Human Capital Formation, Life Expectancy and the Process of Development*

Matteo Cervellati
UPF, Barcelona and University of Bologna

Uwe Sunde
IZA, Bonn and University of Bonn

December 12, 2002

Abstract

This paper presents a microfounded theory of long-term development. We model the interplay between economic variables, namely the process of human capital formation and technological progress, and the biological constraint of finite lifetime expectancy. All these processes affect each other and are endogenously determined. The model is analytically solved and simulated for illustrative purposes. The resulting dynamics reproduce a long period of stagnant growth as well as an endogenous and rapid transition to a situation characterized by permanent growth. This transition can be interpreted as industrial revolution. Historical and empirical evidence is discussed and shown to be in line with the predictions of the model.

JEL-classification: E10, J10, O10, O40, O41

Keywords: Long-term development, endogenous lifetime duration, endogenous life expectancy, human capital, technological progress, growth externalities, industrial revolution

*The authors would like to thank Antonio Ciccone, Adriana Kugler, Matthias Messner, Omar Licandro and Nicola Pavoni, as well as participants at ASSET 2002, Cyprus, for very helpful comments. Financial support from German Research Foundation DFG, and IZA is gratefully acknowledged. Contact address: matteo.cervellati@econ.upf.es, sunde@iza.org. All errors are our own.

1 Introduction

The past two centuries were characterized by widespread and profound changes in human living conditions. For aeons, a more or less stable and unchanged environment prevailed, with a strong preponderance of agriculture and trade of basic goods, rigid social structures with usually a small ruling class, and comparably poor medical conditions. But suddenly within just more than two hundred years, that is just a few generations, the economic environment mutated utterly: the structure of the economy changed completely with industrialization breaking its way, reducing the importance of agricultural activities in favor of the industrial and the service sector. Personal life changed in every dimension to an extent not seen before or after. The traditional social environment ceased to exist, as the vast majority of the population became educated, and acquired knowledge beyond the working knowledge of performing a few manual tasks inherited by previous generations. Literacy, which used to be the privilege of a little elite, became widespread among the population. The process of human capital accumulation accelerated as more and more people acquired the ability to innovate, and to use innovations. On the other hand, the spread of new technologies in turn made it more profitable to acquire knowledge. Also the biological environment sharply changed. Lifetime duration, which had been virtually the same for thousands of years, increased sharply within just a few generations. Mortality fell significantly and fertility behavior changed profoundly, hygienic conditions improved as sanitation became more important and widespread.

Economists have always had a great interest in understanding the reasons and the mechanics of these dramatic changes, in particular against the background of the fact that large parts of the world are still underdeveloped. Several recent contributions address the issue of the economic transition from stagnant, Malthusian regimes to permanent growth, like Komlos and Artzrouni (1990), Goodfriend and McDermott (1995), Tamura (1996), Hansen and Prescott (1998), Lucas (2002), Galor and Weil (2000), Galor and Moav (2002b), Jones (2001) and Jones (2002). The driving forces, which explain the economic transition towards higher growth paths in these models, are technical progress, physical capital accumulation, population growth, and, most importantly, the process of human capital accumulation. The analyzed decision processes also affect fertility behavior and are capable of producing an endogenous demographic transition towards a regime of lower population growth. However, in the light of the changes in personal living conditions that accompanied the economic developments, also lifetime duration played a crucial role in the process of development. But, as some authors like Mokyr (1993) already pointed out, two separate strands of the literature, one about the causes and mechanics of the industrial revolution, and another about the decline in mortality, largely coexist without any obvious connection or compatibility between the two.

Some recent contributions explicitly include mortality or lifetime duration to explain the mechanism of economic development. Kalemli-Ozcan *et al.* (2000) develop a general equilibrium model to study the effects of exogenous changes in mortality on schooling and human capital accumulation. Croix and Licandro (1999), Boucekkine, de la Croix, and Licandro (2002a) and Boucekkine, de la Croix, and Licandro (2002b) consider endogenous growth models in which life expectancy is exogenous and affects the level of schooling, which in turn determines growth. Swanson and Kopecky (1999) present cross-country evidence for the relevance of life expectancy on growth, and develop a model of human capital accumulation in which individuals have a finite lifespan. Reis-Soares (2001) explores the link between life expectancy, educational attainment and fertility choice in the context of long-run development, and presents cross-country evidence for interactions between life expectancy, income, schooling levels and fertility.

All these contributions emphasize that lifetime duration plays a crucial role for human capital investments, which in turn determine growth. Moreover, the empirical evidence they present or cite, strongly supports this view. However, potential reverse effects of development on lifetime duration have been largely neglected in this literature. There is now general agreement in the fields of economic history and demography that economic development and the level of human capital profoundly affects lifetime duration and living conditions. A large body of empirical evidence supports the view that higher levels of development are correlated with longer life expectancy. This evidence suggests that traditionally little education and knowledge about health and means to avoid illness supported the outbreak, propagation and mal-treatment of diseases and ultimately led to high mortality. However, an increasing popular knowledge of the treatment of common diseases and about the importance of hygiene and sanitation, as well as the availability of respective technologies, helped to increase life expectancy somewhat over time (see Mokyr, 1993). There is also evidence for an inverse relation between parents' schooling and child mortality, suggesting that life expectancy increases in parents' human capital. Evidently, a mother's level of education has positive effects on life expectancy of her children (Schultz, 1993). The invention of new drugs, which depends crucially on the human capital involved in research, increased life expectancy (see Lichtenberg, 1998). Blackburn and Cipriani (2002) cite further empirical evidence for the view that life expectancy depends on economic conditions. Moreover, they develop a model with endogenous life expectancy, in which the economy may end up in different development regimes, depending on the initial conditions. Similarly, Kalemli-Ozcan (2002) shows the possibility for multiple equilibria in a model in which individuals decide upon their fertility and the education of their children, once life expectancy is seen as endogenous and depending on income per capita. In Tamura (2002), human capital induces

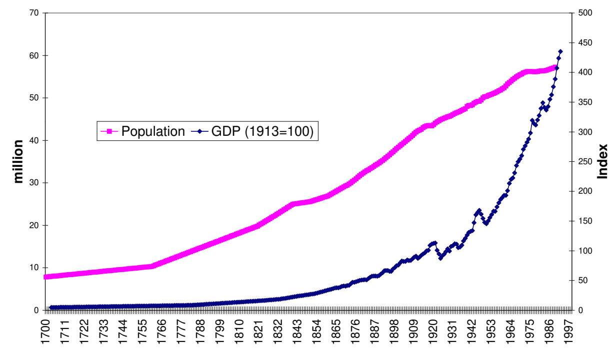

Figure 1: GDP and Population Size in the U.K.

a falling mortality which eventually induces a demographic transition with a reduction in fertility.

Hence, there is little dispute in the literature that life expectancy is a crucial determinant of human capital accumulation and economic development, and that the level of human capital and development in general affects lifetime duration. However, in the context of the early stages of the industrial revolution issues are still largely unsettled. There is still disagreement among economic historians, see Riley (2001) and Easterlin (2002), about whether the onset of increases in life expectancy can be precisely dated for different countries. There is a similar disagreement whether this onset coincided with the beginning of the industrial revolution and the transition to a faster regime of growth, or whether changes in life expectancy preceeded or followed changes in the economic environment. Figure 1 shows the pattern of development of GDP and the size of population for the United Kingdom.[1] A long era of little growth in the size of population as well as output is followed by an acceleration in the development of both variables during the second half of the 18th century. While GDP seems to grow unboundedly ever since, population growth eventually dips after the 1950s, see Maddison (1991) for detailed data for several countries exhibiting these patterns. The increase in population size suggests that the reduction in fertility rates, which is studied by a number of recent contributions, is more than compensated by an increase in lifetime duration: At the same time as the development takes off, from the 18th century onwards, mortality decreased. Boucekkine, de la Croix, and Licandro (2002b) cite evidence from life tables and parish registers from Geneva and Venice, which show that life expectancy as measured at age ten already increased between 1640 and

[1]The data are taken from Maddison (1991) and exclude South Ireland. Missing intermediate values are obtained by linear interpolation. Data for other European countries exhibit similar patterns.

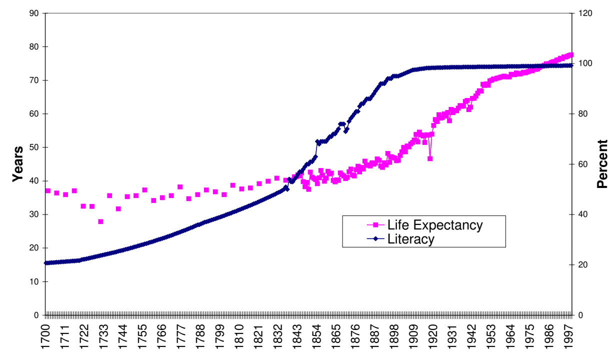

Figure 2: Development of Life Expectancy and Literacy

1740 in these urban centers. Moreover, adult mortality seems to have fallen before child mortality declined substantially. Improvements in knowledge of deseases and hygiene eventually caused average life expectancy at birth as well as at lager ages to increase, as illustrated in Figure 2.[2] Simultaneously to the early developments during the 18th century, literacy began to spread over the population, as is illustrated in Figure 2 by the ability to sign documents (see also Boucekkine, de la Croix, and Licandro (2002b)).[3] According to Maddison (1991), the average education in 1820 was about two years for both sexes. By 1989, this number had risen to over eleven years. This evidence illustrates that not only economic development, as measured by GDP and GDP per capita accelerated substantially during the industrial revolution. The level of human development as such greatly improved as illustrated by life expectancy and education.[4]

The question, which factor was causally responsible for all these profound changes, is still hotly debated. Some authors explain the decline in mortality and the increase in life expectancy by increases in household incomes and technological progress (see e. g. McKeown, 1977). However, this view has been criticized on the basis of the empirical evidence, which suggests that technological (medical) progress took off too late to explain early increases in lifetime duration. Moreover, by and large, the standard of living in terms of income, housing and nutrition of the majority of the population hardly changed before 1850, indicating that this explanation does not tell

[2]Data are taken from Www.Mortality.Org (2002) and Floud and McCloskey (1994).

[3]Data are taken from Cipolla (1969) and Floud and McCloskey (1994) and contain literacy for France due to data limitations for England. The pattern of development was qualitatively similar in both countries with France lagging somewhat behind.

[4]Authors such as Sen (1977,1985) have questioned the use of economic indicators like wealth or consumption as a relevant measure for well-being. Rather, economic resources should be seen as means that allow to improve the individual well being, through better nutrition, health, education etc.

the entire story, see Mokyr (1993). Others, like Boucekkine, de la Croix, and Licandro (2002b) and the references therein, argue that already at the dawn of the industrial revolution mortality declined. They view this decline as as an exogenous event, which in turn is argued to have triggered more investment in human capital and faster growth. Subsequent changes in mortality are again interpreted as endogenous consequences of economic development. However, this line of argument leaves the cause of the industrial revolution essentially unexplained.

The contribution of this paper is to provide a unified framework to analyze the interactions between human capital accumulation, technological progress and lifetime duration in the context of long term development, in which all these relevant processes are determined endogenously. The model to be presented has three basic building blocks. The first block is a micro-founded model of human capital formation in which overlapping generations of heterogeneous individuals decide upon the type and the amount of human capital to acquire during their lives. This optimal choice depends crucially on their life expectancy and the state of technology. The second block is the idea that vintage human capital is the primary engine of economic growth. Human capital affects the state of technology in terms of productivity in the production process. The resulting technical progress makes future investments in human capital more profitable. The third block is motivated by the historical and demographic evidence mentioned above and concerns the effects of the economic and social environment on lifetime duration. In particular, the evolution of life expectancy is endogenously linked to the process of human capital accumulation.

The main mechanism of the model to be presented below can be summarized as follows. Individuals maximize their lifetime utility by their choice of human capital accumulation. This decision shapes the structure of the economy during their lives. In their choice, individuals take their expected lifetime duration as given, and they do not consider the effect of their decision on the life expectancy of future generations. This in turn creates an externality on future human capital decisions. Moreover, the level of human capital created by a generation of individuals affects productivity and therefore the growth potential of the economy in the future. In principle, there is a virtuous cycle of human capital accumulation and growth. However, as long as the biological barrier of low life expectancy is binding, development of the economy will be very slow. The economy is virtually trapped on a slow growth path. Eventually, once average life lifetime duration is high enough and the level of technology is sufficiently advanced to induce large proportions of the population to acquire high quality human capital, growth takes off. A phase of fast development and a profound change in the structure of economy, which can be interpreted as an industrial revolution, starts, and the economy converges within a few generations to a new path with higher growth rates than before. As a consequence of the increase in life expectancy, population size grows even though fertility behavior is unchanged.

The paper is organized as follows. In section 2 we describe the individual problem of human capital accumulation in the face of given technologies, as well as the economic environment, and solve for the intragenerational equilibrium. The intertemporal links between subsequent generations, and the process of dynamic development are presented in section 3. There, we also present the main result, a characterization of long-term development. Section 4contains a simulation of the model to illustrate how it can account for the long-run development experience. Section 5 concludes. All proofs are collected in the appendix.

2 The Model

In this section we analyze the process of human capital formation for a given generation. We introduce the production of final consumption goods and we study the static equilibrium of the economy. The next section deals with the links between generations and analyzes the evolution of the dynamic system.

We start by looking at the individual problem of investing in human capital. The types of human capital at disposal differ in the way they are built up, and in the returns individuals receive from them. The main inputs in building up human capital are individual ability and time spent for education. From the individual point of view, the time available is limited by the expected lifetime duration, which is therefore taken as given by the individual. The same is true for the returns to human capital, which are determined on competitive markets at the aggregate level. The role of real resources as input for the human capital formation process, as well as issues related to capital market development and public provision of education, is neglected in this paper. Instead, we focus on changes in the economic and biological environment creating the necessary and sufficient conditions for large parts of the population to acquire human capital.[5] One distinctive feature of human capital is that it differs inherently between generations, since it is formed in a changing technological environment. Aggregate production in the economy is the outcome of the use of different vintages of technology and human capital. The individual problem is then which type of human capital to acquire and how much of it. The intragenerational equilibrium is characterized by the interplay of individual optimizing behavior and aggregate market conditions.

[5] In the presence of frictions and market imperfections, these conditions might not be sufficient for development, see e.g. Galor, Moav, and Vollrath (2002), and Galor and Moav (2002a).

2.1 Production of Human Capital

The economy is populated by an infinite sequence of overlapping generations of individuals. Generations will be denoted with subscript t. Every generation is born l periods after the birth of the respective previous generation.[6] In order to isolate the development effects related to lifetime duration and human capital accumulation, any links between generations through savings or bequests are excluded. A generation consists of a continuum of agents with population size normalized to one. Individuals face a lifetime duration specific to their generation, whose determinants will be discussed below. Every individual is endowed with ability $a \in [\underline{a}, \overline{a}]$ and abilities are exogenously distributed with a density $f(a)$.[7] Any member of a given generation faces a decision regarding the accumulation of human capital to be described below. Every generation has to build up the stock of human capital capital from zero, since the peculiar characteristic of human capital is that it is embodied in people (even if the production can be easier if the previous generation had a lot of it).[8]

In order to make an income, individuals have to spend their ability and some of their living time to form some human capital. There are several types of human capital, which differ with respect to their production process and the returns they generate. For simplicity, we concentrate on the simple case of only two types of human capital. Along the lines of growth theory, one type of human capital is interpreted as high-quality, and growth enhancing. This type is labelled *theoretical* human capital and is denoted by h. This type of human capital is the primary engine of modern economic growth. It is characterized by a high content of abstract knowledge. This type of human capital is important for innovation and development of new ideas, since abstract knowledge helps to solve a problem never faced before by resorting to known abstract concepts.

The second type is labeled *applied human capital*, denoted by p, and can be interpreted as labor capacity. It contains less intellectual quality, but more manual and practical skills that are important in performing tasks related to existing technologies.[9]

Both types of human capital are produced using time e and individual

[6] Instead of assuming a fixed frequency of births, one could alternatively model the length of the time spell between the births of two successive generations, hence the timing of fertility, as a function of the life expectancy of the previous generation. The case of an economy consisting of non-overlapping, subsequent generations of individuals is a special case of the presented set-up, where l equals the respective previous generation's life duration.

[7] We assume that the *ex ante* distribution of innate ability or intelligence does not change over the course of generations.

[8] This is essentially the idea behind Becker, Murphy, and Tamura (1990).

[9] In the language of labor economics, theoretical human capital could be associated with skilled labor, while applied human capital is associated with unskilled labor.

ability a as inputs: $p = p(e, a)$, and $h = h(e, a)$. These production processes are inherently different. The levels of both types of human capital increase in the time spent in forming them. The main difference lies in the effectiveness of time. To acquire theoretical human capital h, it is necessary to first spend time on the building blocks of the elementary concepts without being productive in the narrow sense. This view of human capital formation is in line with the mastery learning concept as understood by for example Becker, Murphy, and Tamura (1990), which states that learning complicated materials is more efficient when the elementary concepts are mastered. Once the basic concepts are internalized, the time spent on theoretical human capital is very productive. On the contrary, the time devoted to acquire applied human capital p is immediately effective, albeit with a lower overall productivity. Personal ability is relatively more important in acquiring theoretical human capital.

Formally, the following production functions exhibit these different characteristics of the processes of building up human capital:

$$h = \begin{cases} \alpha(e - \underline{e})a & \text{if } e \geq \underline{e} \\ 0 & \text{if } e < \underline{e} \end{cases} \tag{1}$$

and

$$p = \beta e \,. \tag{2}$$

In order to acquire theoretical human capital h, an agent needs to pay a fix cost $\underline{e}$ measured in time units while for applied human capital p the fix cost is smaller and normalized to zero.[10] Any unit of time produces α units of h and β units of p with $\alpha \geq \beta$. Ability is modeled as increasing the production of human capital h per unit of time.

There is just one consumption good in the economy. Strictly speaking, agents face an intertemporal problem of maximizing their lifetime utility. Utility is linear in consumption and there is no discounting. We abstract from life cycle considerations and normalize the discount factor to zero, so agents are indifferent with respect to the date of consumption. In this case, it is sufficient for agents to maximize total lifetime earnings in order to maximize their individual lifetime utility. This is done by choosing optimally the type of human capital to acquire and the time e spent producing it.[11]

Since building up human capital takes time, individuals face an intertemporal trade-off between spending time on building human capital and spending time and using the acquired human capital on working and earning in-

[10] We abstract from other costs of education, like tuition fees etc. Moreover, the fixed cost is assumed to be constant and the same for every generation. Costs that increase or decrease along the evolution of generations would leave the qualitative results of the paper unchanged.

[11] Equivalently, concave utility, discounting and perfect capital markets could be introduced to model lifecycle considerations. Without affecting the main results, these issues are beyond the scope of the current analysis.

come. While accumulating human capital, agents cannot work. This means that agents must optimally decide how to split their expected lifetime between human capital formation and work. We abstract from leisure and learning on-the-job.

This setting is chosen to catch two crucial features of the human capital formation process. The first one is that larger lifetime duration induces individuals to acquire more of any type of human capital. The second feature is that increasing lifetime duration makes theoretical, high quality human capital relatively more attractive for individuals of any level of ability. Any alternative model of human capital formation reproducing these two features would be entirely equivalent for the purpose of this paper. Alternative settings like learning on-the-job could similarly be used to illustrate the importance of lifetime duration for human capital formation.

An agent can either decide to aquire h or p but not both. Formally, in his choice, he takes life expectancy T_t as well as the wages for unit of human capital as given. Denote by $w_h^t(\tau)$ and $w_p^t(\tau)$ the wage rate paid at any moment in time τ to every unit of human capital of type h or p, respectively, aquired by generation t. Consequently, we can express total lifetime utility, i.e. total earnings V, of every individual a of generation t acquiring each type of human capital as:

$$\begin{aligned} V_h^t(e_h, a) &= \int_{e_h}^{T_t} h w_h^t(\tau) d\tau \\ &= \int_{e_h}^{T_t} \alpha(e_h - \underline{e}) a w_h^t(\tau) d\tau \text{ , and} \end{aligned} \tag{3}$$

$$\begin{aligned} V_p^t(e_p, a) &= \int_{e_p}^{T_t} p w_p^t(\tau) d\tau \\ &= \int_{e_p}^{T_t} \beta e_p w_p^t(\tau) d\tau \text{ .} \end{aligned} \tag{4}$$

2.2 Aggregate Production

We consider an economy with multiple sectors of production, in which new technological vintages become available overtime. The stocks of human capital of both types available in the economy at any moment in time, i.e. embodied in all generations alive at that date, are the only factors of production. Wage rates are determined in the macroeconomic competitive labor market and equal marginal productivities. In particular we model, along the line of Hansen and Prescott (2002), a one-good-two-sectors economy.[12]

[12]The focus of the paper is not on the macroeconomic role of demand for different consumption goods, so we assume that each sector produces the same good. Alternatively one could model different sectors as producing differentiated intermediate goods to be

Sectors are structurally different in their intensity of use of different human capital. Denote as $\mathcal{P}$ the sector using p relatively more intensively and $\mathcal{H}$ the sector using h relatively more intensively. Technology is modeled as total factor productivity. Technological process takes place in both sectors in the form of new production technologies characterized by a larger total factor productivity becoming available over time. Technological improvements are modeled as vintages since older production functions are still available in each sector and can potentially be used along with the newest ones.[13] Denote by $A^v_H(\tau)$ and $A^v_P(\tau)$ as the total factor productivites and by $Y^{\mathcal{P}}_v(\tau)$ and $Y^{\mathcal{H}}_v(\tau)$ the production realized in sector $\mathcal{P}$ and $\mathcal{H}$ using vintage technology v at time τ. Then total production at time τ is given by:[14]

$$Y(\tau) = \sum_v Y^{\mathcal{P}}_v \tau) + \sum_v Y^{\mathcal{H}}_v(\tau) \, . \tag{5}$$

Human capital is inherently heterogenous across generations, because individuals acquire their human capital in an environmnent characterized by the availability of different vintages of technologies. Agents of each generation can acquire human capital, allowing them to use technologies up to the latest available vintage. Human capital is thus characteristic for a generation. This implies that a generation's stock of human capital of either type is not a perfect substitute of older and younger generations' human capital and is sold at its own price. Let the respective aggregate amounts of human capital acquired by generation t be $P_t = \int_{\underline{a}}^{\overline{a}} p_t(a) f(a) da$ and $H_t = \int_{\underline{a}}^{\overline{a}} h_t(a) f(a) da$. The istantaneous wage rates are given by:

$$w^t_h(\tau) = \frac{\delta Y(\tau)}{\delta H_t} \, , \text{ and} \tag{6}$$

$$w^t_p(\tau) = \frac{\delta Y(\tau)}{\delta P_t} \tag{7}$$

To make the model analytically tractable, we consider a Cobb-Douglas specification of the production function and we assume that every vintage of human capital fully specializes in the respective latest vintage of technology, so that $t = v$.[15] As a benchmark, we consider the extreme case in which

used in the production of a unique final good.

[13]This means that different technologies of productions are available at any moment in time. If we interpret the different sectors e.g. as agricultural and industrial, the production of corn can then take place using donkeys or modern tractors.

[14]The specification used by Hansen and Prescott (2002) is contained as the special case when only the latest vintage can be used.

[15]In other words, this assumption implies that e.g. a mechanic in the late 20th century knows how to repair a common rail diesel engine, but not a steam engine. However, as will become clear below, vintages build upon the advances of previous vintages, e.g. common rail diesel engines incorporate technological principles that partly derive from the use of steam engines.

every sector uses only one type of human capital. The production function are thus:

$$Y_t^{\mathcal{P}} = A_P^t P_t^\gamma \text{ , and } Y_t^{\mathcal{H}} = A_H^t H_t^\gamma \text{ ,} \tag{8}$$

respectively, with $\gamma \in (0,1)$ and $A_P^v(\tau),\ A_H^v(\tau) \in \mathcal{R}^+$.[16]

2.3 Intragenerational Equilibrium

Consider the decision problem for members of a given generation t of individuals. In the following we omit the corresponding subscripts '$_t$' as long as there is no danger of confusion.

To maximizes his lifetime utility, an agent compares the maximum lifetime utility he can get by acquiring one type of human capital or the other. Consequently, he chooses to acquire p or h depending on whether:

$$V_p^*\left(e_p^*, a, w_p\right) \gtrless V_h^*\left(e_h^*, a, w_h\right),$$

where:

$$e_h^* = \arg\max V_h\left(e_h, a, w_h\right) = \left(T - e_h\right)\alpha(e_h - \underline{e})aw_h \text{ ,}$$

and

$$e_p^* = \arg\max V_p\left(e_p, a, w_p\right) = \left(T - e_p\right)\beta e_p w_p \text{ .}$$

The optimal time investments are then given by:

$$e_h^* = \frac{T + \underline{e}}{2}, \text{ and } e_p^* = \frac{T}{2},$$

respectively.

The individual levels of human capital in the two cases are obtained by substituting the optimal time investments back:

$$h^*\left(T, a\right) = \alpha\frac{T - \underline{e}}{2}a \text{ ,} \tag{9}$$

and

$$p^*\left(T, a\right) = \beta\frac{T}{2} \text{ .} \tag{10}$$

Accordingly, indirect lifetime utilities are given by:

[16] In principle, both sectors could be characterized by different productivity parameters γ_H and γ_P. This case will be illustrated in the simulations below. However, while the main results remain unaffected by asserting a common value to both sectors, it simplifies the analytic tractability of the model considerably. Encorporating both types of human capital in both sectors of production does not alter the results as long as the difference in the relative intensities of their use in the respective sector is maintained and no input is indispensable.

$$V_p^*\left(p^*, a, w_p\right) = \frac{T^2}{4}\beta w_p \ , \tag{11}$$

and

$$V_h^*\left(h^*, a, w_h\right) = \alpha a \frac{\left(T-\underline{e}\right)^2}{4} w_h \ . \tag{12}$$

Obviously, agents with higher ability have a comparative advantage in the accumulation of H and the lifetime utility for those investing in h increases monotonically in the ability parameter.

An agent is indifferent between acquiring h or p if and only if:

$$V_p^*\left(e_p^*, a, w_p\right) = V_h^*\left(e_h^*, a, w_h\right) \ . \tag{13}$$

For every vector of wage rates there is only one level of ability $\widetilde{a}$ for which the indirect utilities are equal:

$$\widetilde{a} = \frac{w_p}{w_h}\left[\left(\frac{\beta}{\alpha}\right)\frac{T^2}{\left(T-\underline{e}\right)^2}\right] \ . \tag{14}$$

Due to the monotonicity of V_h^* in ability, all agents with $a < \widetilde{a}$ will optimally choose to acquire human capital P , while those with ability $a > \widetilde{a}$ will optimally choose to obtain H. Note that, as previously mentioned, all individuals with higher ability than $\widetilde{a}$ choosing to acquire theoretical human capital actually enjoy larger lifetime earnings than those with lower ability than the threshold and thus choosing to invest in applied human capital.

This fact allows us to simplify notation. In what follows, denote by $\lambda(\widetilde{a})$ the fraction of the population acquiring human capital of type p, and by $(1-\lambda(\widetilde{a}))$ the fraction of the population acquiring human capital h.

In fact, these proportions can be written as:

$$\lambda(\widetilde{a}) := \int_{\underline{a}}^{\widetilde{a}} f(a)da \tag{15}$$

$$1-\lambda(\widetilde{a}) := \int_{\widetilde{a}}^{\overline{a}} f(a)da \tag{16}$$

By inspection of equation (14) and since $T-\underline{e} > 0$, one can see that the fraction $1-\lambda(\widetilde{a})$ increases with lifetime duration T, with the relative wage $\frac{w_h}{w_p}$ and with $\frac{\alpha}{\beta}$.

Take for simplicity the case of a uniform distribution of abilities in the interval $[0,1]$. In this case the aggregate levels of theoretical and applied human capital denoted by H and P, respectively, can be explicitly computed as:

$$P(\widetilde{a}) = \int_0^{\widetilde{a}} p\left(T,a\right)da = \widetilde{a}\beta\frac{T}{2} \ , \tag{17}$$

and

$$H(\widetilde{a}) = \int_{\widetilde{a}}^{1} h\left(T,a\right)da = \left(\frac{1-\widetilde{a}^2}{2}\right)\alpha\frac{T-\underline{e}}{2}\,. \tag{18}$$

For computational convenience, we assume uniform distribution of abilities on the support $[0,1]$ in the remainder of the paper, unless noted otherwise.[17]

Factors of production are sold in the competitive market and receive wages equal to their marginal productivity. The resulting instantaneous wage rates as defined in (7) that are compatible with the macroeconomic equilibium are given by:

$$\begin{aligned} w_h &= A_H\gamma H^{\gamma-1}\,, \qquad \text{and} \\ w_p &= A_P\gamma P^{\gamma-1}\,. \end{aligned}$$

Given this setting, we define the intragenerational equilibrium of this economy as follows:

Definition 1. *The intragenerational equilibrium is a vector:*

$$\left\{\{h^*(T,a)\}_{a\in[\underline{a},\overline{a}]}\,,\{p^*(T,a)\}_{a\in[\underline{a},\overline{a}]}\,,H^*,P^*,w_h^*,w_p^*,\tilde{a}^*\right\}$$

such that, for any given T and distribution $f(a)$ we have:

$$h^*\left(T,a\right) = \alpha\frac{T-\underline{e}}{2}a\,, \qquad \forall a\geq\widetilde{a}^* \tag{19}$$

$$p^*\left(T,a\right) = \beta\frac{T}{2} \qquad \forall a<\widetilde{a}^* \tag{20}$$

$$H^* = \int_{\widetilde{a}^*}^{\overline{a}} h^*\left(T,a\right)f(a)da \tag{21}$$

$$P^* = \int_{\underline{a}}^{\widetilde{a}^*} p^*\left(T,a\right)f(a)da \tag{22}$$

$$w_h^* = A_H\gamma H^{*\gamma-1} \tag{23}$$

$$w_p^* = A_P\gamma P^{*\gamma-1} \tag{24}$$

$$\widetilde{a}^* = \frac{w_p^*}{w_h^*}\left[\left(\frac{\beta}{\alpha}\right)\frac{T^2}{(T-\underline{e})^2}\right] \tag{25}$$

The equilibrium system (19) to (25) defines an implicit function in $(\widetilde{a}^*,T)$ linking the equilibrium cut-off level of ability $\widetilde{a}^*$ to lifetime duration T. Since,

[17] In fact, the results can be generated in the model with any uni-modal distribution function of abilities. It is easy to check that the results also go through if a degenerate ability distribution function with just one ability level for all members of the population is assumed. However, the process of how individuals sort into equilibrium would be less clear, since there would be no ability cut-off separating the population, but only a certain decomposition of the population into the two groups required by equilibrium conditions.

$$\frac{w_p^*}{w_h^*} = \frac{A_P \gamma P^{*\gamma-1}}{A_H \gamma H^{*\gamma-1}} = \left[\left(\frac{\alpha}{\beta} \right) \frac{(T - \underline{e})^2}{T^2} \right] \widetilde{a}^* = \frac{w_p^*}{w_h^*} \,. \tag{26}$$

Substituting for $P(\widetilde{a}^*)$, $H(\widetilde{a}^*)$ from Equation (17) and (18), and remembering that the equilibrium concerns any given generation t, we have:

$$\widetilde{a}_t^* \left(\frac{(1 - \widetilde{a}_t^{*2})}{2\widetilde{a}_t^*} \right)^{\gamma-1} = \frac{A_{P.t}}{A_{H.t}} \left(\frac{\beta}{\alpha} \right)^{\gamma} \left(\frac{T_t}{T_t - \underline{e}} \right)^{\gamma+1} \tag{27}$$

This relation between the equilibrium ability threshold for the acquisition of abstract human capital and life expectancy T will be of eminent importance for the analysis of development later on. For notational convenience, reformulate Equation (27) by solving for lifetime expectancy as a function of the ability threshold to get:

$$T(\widetilde{a}^*) = \frac{\underline{e}}{1 - \frac{g(\widetilde{a}^*)}{\Omega}} \,, \tag{28}$$

with

$$g(\widetilde{a}^*) = \frac{(1 - \widetilde{a}^{*2})^{\frac{1-\gamma}{1+\gamma}}}{\widetilde{a}^{*\frac{2-\gamma}{1+\gamma}}} k \,, \tag{29}$$

$$k = 2^{-\frac{1-\gamma}{1+\gamma}} \,, \text{ and}$$

$$\Omega = \left[\left(\frac{A_H}{A_P} \right) \left(\frac{\alpha}{\beta} \right)^{\gamma} \right]^{\frac{1}{1+\gamma}} : . \tag{30}$$

It is easy to see that $g(\widetilde{a}^*) > 0$, $\forall \widetilde{a} \in [0, 1]$. Note that $T(\widetilde{a}^*)$ is defined for all $\widetilde{a}^* \in [\underline{\widetilde{a}}^*, 1]$ with $\underline{\widetilde{a}}^* : g(\underline{\widetilde{a}}^*) = \Omega \Leftrightarrow \lim_{\widetilde{a}^* \to \underline{\widetilde{a}}^*} T(\widetilde{a}^*) = \infty$, and that $\forall \widetilde{a}^* \in [\underline{\widetilde{a}}^*, 1] : 1 - \frac{g(\widetilde{a}^*)}{\Omega} > 0$. The value $\underline{\widetilde{a}}^* > 0$ represents a maximum fraction of the population that would optimally choose to acquire human capital H for any given level of relative productivity $\frac{A_H}{A_P}$. This maximum fraction cannot be exceeded, even if the biological constraint of finite lifetime duration would disappear (i. e. if $T \to \infty$).

There exists a unique pair of expected lifetime duration and ability that satisfies the conditions for an intragenerational equilibrium:

Proposition 1. *There exists exactly one intragenerational equilibrium characterized by the a pair* $(\widetilde{a}^*, T^*)$, *with* $\widetilde{a}^* \in [\underline{\widetilde{a}}^*(\Omega), 1]$ *and* $T \in [\underline{e}, \infty)$, *which satisfies condition* (27).

In this context, it is worth noting that the maximum proportion of the population that would acquire H in the absence of biological constraints, $1 - \underline{\widetilde{a}}^*(\Omega)$, is increasing with the relative productivity of the sector using theoretical human capital intensively, $\frac{A_H}{A_P}$. This observation will prove useful later on and is therefore summarized in:

Lemma 1. *The lower bound on the support of ability thresholds decreases as* Ω *increases, that is* $\frac{\partial \underline{\widetilde{a}}^*(\Omega)}{\partial \Omega} < 0$.

2.4 Properties of the Intragenerational Equilibrium

The equilibrium relation between lifetime duration and the proportion of the population investing in human capital presented in equation (27) will be a crucial determinant of the dynamic system. This relation is determined endogenously for a given generation through the interplay of individual optimizing behavior and aggregate equilibrium conditions. According to the following proposition, the ability cut-off is lower for higher expected lifetime duration, that means that a higher share of the population decides to obtain human capital of type h if they expect to live longer. Moreover, the function $\widetilde{a}^*(T)$, representing the threshold ability defining the proportion of the population acquiring human capital h, is S-shaped: as lifetime duration increases the proportion of population choosing h increases first slowly, then increasingly rapidly, until this increase slows down again as the ability threshold converges to ever lower levels.

Proposition 2. *The cut-off level $\widetilde{a}^*(T)$, which identifies the equilibrium fraction of members of a generation acquiring human capital h, is an increasing, S-shaped function of expected lifetime duration T of this generation, with zero slope for $T \longrightarrow 0$ and $T \longrightarrow \infty$, and exacly one inflection point.*

The full proof is contained in the appendix. The economic meaning behind the S-shape is easier to grasp when looking at the equilibrium relation in the (λ, T)-space. The equilibrium locus can be rationalized as follows. For low lifetime durations, the share of population investing in h is small, and also relatively large increases in average lifetime duration do not change this structure of the economy much. The reason for this is that due to the fixed cost involved with acquiring h, the remaining time to use the acquired h to earn income is too short for a large part of the population to be worth the effort. Once average lifetime duration increases sufficiently, the fixed cost constraint binds for fewer and fewer people, so the structure of the economy changes more rapidly towards a higher fraction of people acquiring h. However, the speed of this structural change decreases as an ever larger share of the population is engaged in h due to decreasing returns in both sectors: Since only few individuals decide to invest in p the relative wage w_h/w_p decreases affecting the individual choice of human capital accumulation.

Having characterized the static behavior of the economy, we now turn to the dynamic process of development.

3 The Process of Economic Development

In the economy described in the previous section, lifetime duration is considered as given from the individual viewpoint. The structure of the economy in every generation is the outcome of individual decisions and depends on average expected lifetime duration. On the other hand, in the long run and from a macroeconomic perspective, lifetime duration is endogenous. Expected lifetime duration is related to the level of development through the structure of the economy. This section models the intertemporal interplay between these two mechanisms characterizing the development path of an economy starting from an initial situation with low average lifetime duration.

The first component of the dynamic system governing the development of the economy is the equilibrium relation between lifetime duration and the proportion of the population investing in human capital. The condition for an intragenerational equilibrium for generation t given in equation (27) in the previous section, defines an implicit relationship between $a(t)$ and $T(t)$, which, for brevity, can be denoted as:

$$a_t = \Lambda(T_t, A_t) \,. \tag{31}$$

3.1 Links Between Generations

The previous section examined the individual education decision of members of a given generation. While life expectancy is a parameter which individuals have to take into account when making their education decisions, we assume that they cannot directly influence it. Rather, expected lifetime duration of children may depend on the level of development and the quantity and quality of human capital of the society at the time of their birth, that is by the level of knowledge acquired by the previous generation.[18] Recent empirical findings show that the level of GDP and literacy are positively correlated with average expected lifetime duration (see Swanson and Kopecky, 1999, and Reis-Soares, 2001). Of course, this effect has also an impact on the children's decision of which type of human capital to acquire, as will become clear below.

We formalize this positive externality of the achievements of a generation for the following generations by making the simple assumption that expected lifetime duration of a given generation t is an increasing function of the fraction of the population of the previous generation $(t-1)$ that acquired theoretical human capital.[19] This can be rationalized by the idea that expected lifetime duration of a generation depends on the level of development

[18] Admittedly, this is only true to a certain extent. Of course, individuals can effectively influence their life expectancy by their life style, smoking habits, drug and alcohol consumption, sports and fitness behavior etc. However, for this they have to know which factors and activities are detrimental and which are advantageous for average life duration. The picture we have in mind is therefore more general: people born in the 18th century did not have medical facilities, or knowledge about health and sanitation comparable to people born in the late 20th century. In our view it is this sort of knowledge that primarily determines life expectancy and mortality, and this knowledge has to be acquired over time and is passed-on from generation to generation.

[19] Clearly, this is just a simplification. Life expectancy might also depend on the share

at the time of its birth:

$$T_t = \Upsilon(\lambda_{t-1}) = \underline{T} + \rho(1 - \lambda_{t-1}) , \tag{32}$$

where $(1 - \lambda_{t-1}) = 1 - \lambda(\widetilde{a}^*_{t-1}) = \int_{\widetilde{a}^*_{t-1}}^{\overline{a}} f(a)da$ is the fraction of generation $(t-1)$ that has acquired human capital of type h. Note that by the definition of λ, life expectancy is a function of the threshold ability level for the decision to acquire general human capital h of the respective generation:

$$T_t = \Upsilon(\widetilde{a}^*_{t-1}) , \tag{33}$$

There is a biological barrier to extending lifetime duration implicitly contained in the specification of equation (32) since by definition of λ the lifetime duration is bounded from above and thus cannot be increased beyond a certain level. We take this as a commonly agreed empirical regularity (see also Vaupel, 1998). The minimum lifetime duration without any human capital of type h is given by $\underline{T}$. The precise functional form of this relation entails no consequences for the main results, and a (potentially more intuitive) concave relationship would not change the main argument.

The second link between consecutive generations is related to total factor productivity and follows the tradition of endogenous growth theory. The level of human capital acquired in a given period increases total factor productivity in subsequent periods. As a consequence, the level of development of an economy exhibits an externality on the subsequent generations. This interpretation is similar to the idea that the stock of ideas transfers into the productivity of future generations suggested by Jones (2001). In the model, we adopt Jones' specification, which is a generalization of the original contribution of Romer (1990). By its nature, theoretical human capital H is relatively more productivity enhancing than practical human capital P. Moreover, the positive effect is stronger in the sector $\mathcal{H}$ that uses theoretical human capital more intensively, since it is the more innovative sector, applying and implementing new and innovative technologies faster. Consequently, total factor productivity (TFP) growth the sector $\mathcal{H}$ is a function of the stock of H and the level of productivity already achieved in this sector.[20] Advances in technology are embodied in the latest vintage of intermediate input $\mathcal{H}$:

$$\dot{A}_{H.t} = \frac{A_{H.t} - A_{H.t-1}}{A_{H.t-1}} = \delta H^{\phi}_{t-1} A^{\chi}_{H.t-1} , \tag{34}$$

where $\delta > 0$, $\phi > 0$, and $\chi > 0$. This can be re-written to:

$$A_{H.t} = \left(\delta H^{\phi}_{t-1} A^{\chi}_{H.t-1} + 1\right) A_{H.t-1} . \tag{35}$$

of the total population that has acquired theoretical human capital at the time of birth of a generation, without qualitatively changing the results.

[20]In the specification used, this function exhibits decreasing returns, while Romer (1990) assumed constant returns. The advantage of the present specification is that it is less rigid and more realistic.

What is important for the argument of the paper is the relative strength of these impacts, so there is no loss in constraining the productivity effect to A_H only. Thus, for simplicity we assume $\dot{A}_{P.t} = 0$ so that total factor productivity in the first sector is constant and can be normalized to 1: $A_{P.t} = A_{P.0} = 1 \; \forall t \in [0, \infty)$.[21] For notational simplicity, we will denote the relative total factor productivity of the two sectors as

$$A_t \equiv \frac{A_{H.t}}{A_{L.t}} \text{ for every } t \in \{0, \infty\} \; . \tag{36}$$

If we assume that the distribution of abilities is uniform, we can substitute $H_{t-1} = \frac{\alpha}{2} (T_{t-1} - e) (1 - \lambda_{t-1})$ from Equation (18) into (35), and obtain an explicit expression for the dynamic evolution of relative productivity:

$$A_t = \left\{ \delta \left[\frac{\alpha}{2} (T_{t-1} - e) (1 - \lambda_{t-1}) \right]^{\phi} A^{\chi}_{.t-1} + 1 \right\} A_{t-1} = F(A_{t-1}, T_{t-1}, \lambda_{t-1}) \tag{37}$$

This specification emphasizes the particular role of theoretical human capital in the accumulation of knowledge, and subsequently for technological progress. The specific functional form has little impact. In fact every, functional specification alternative to (34), which implies a positive correlation between $\dot{A}_t$ and H_t would yield qualitatively identical results. It is also worthwhile noting that the qualitative features of the model are unaltered if technological process is taken to be exogenous, that is if $\dot{A}_t = \varepsilon > 0$.[22]

These inter-generational linkages close the model.

3.2 Dynamics of Development

The solution of the model laid down so far allows to analyze the process of development as an interplay of individually rational behavior and macroeconomic externalities. The static equilibrium relationship (31) holds for every generation, while every generation takes life expectancy along equation (33), and productivity growth according to equation (35) into account. Thus, the development of the economy is characterized by the trajectories of lifetime duration T_t, the fraction of the population acquiring human

[21] In general, both types of human capital can have a positive intertemporal effect on total factor productivity of both sectors, as long as the technological externality is biased towards H-type human capital. In the simulations presented below, we actually allow total factor productivity in the sector using practical human capital intensively to grow according to:

$$A_{P.t} = \left(\delta_P H^{\phi_P}_{t-1} A^{\chi_P}_{P.t-1} + 1 \right) A_{P.t-1} \; .$$

This reflects the historical fact that agricultural productivity also increased as productivity in other sectors went up, e. g. during the industrial revolution, see Streeten (1994).

[22] As will become clearer below, the only consequence of an exogenous change in relative productivity $\dot{A}$ is the missing re-inforcing feedback effect of endogenous technological progress after the industrial revolution.

capital λ_t, and relative productivity A_t. For notational simplicity, denote $\widetilde{a}^*$ simply as a. Taking into consideration the one-to-one relationship between λ_{t-1} and a_{t-1}, the dynamic path is fully described by the infinite sequence $\{a_t, T_t, A_t\}_{t\in[0,\infty)}$, resulting from the evolution of the three dimensional, nonlinear first-order dynamic system derived from equations (31), (33) and (37):

$$\begin{cases} a_t &= \Lambda(T_t, A_t) \\ T_t &= \Upsilon(a_{t-1}) \\ A_t &= F(A_{t-1}, T_{t-1}, a_{t-1}) \end{cases} \tag{38}$$

The development is influenced by the level of human capital of type H accumulated in the past (reflected in the level of relative TFP and the average lifetime duration) and by the current generation, and characterized completely by the respective ability thresholds. The human capital structure of the economy has two effects, one on productivity in aggregate production, which in turn affects relative prices for human capital, and another on the next generation's life expectancy. Both effects concern the main determinants of individual education decisions, and thus affect the structure of human capital accumulation of the subsequent generation, and so on.

The analysis of the dynamic behavior of the economy can be simplified by looking at the dynamic adjustment of human capital and lifetime duration conditional on the value of the relative productivity. We therefore concentrate attention on the properties of the following system, which is conditional on any $A > 0$:

$$\begin{cases} a_t &= \Lambda(T_t, A) \\ T_t &= \Upsilon(a_{t-1}) \end{cases} \tag{39}$$

This system delivers the dynamics of human capital formation and life expectancy for *any given* level of technology. From the previous discussion we know that the first equation of the conditional system is defined for $a_t \in (\underline{a}_t(A), 1]$ and $T \in [e, \infty)$.

In what follows, we denote the S-shaped locus $T_t = \Lambda^{-1}(a_t, A)$ in the space $\{T, a\}$, which results from the intragenerational equilibrium, by $HH(A)$, and the locus $T_t = \Upsilon(a_{t-1})$ representing the intergenerational externality on lifetime duration by TT. Any steady state of the conditional system is characterized by the intersection of the two loci $HH(A\ )$ and TT:

Definition 2. *A dynamic equilibrium of the conditional system given by* (39) *is a vector* $\{a^C, T^C\}$ *with* $a^C \in (\underline{a}(A), 1]$ *and* $T^C \in [e, \infty)$*, which constitutes a steady state solution for the dynamic system* (39) *such that, for any* $A \in (0, \infty)$*:*

$$\begin{cases} a^C &= \Lambda(T^C, A) \\ T^C &= \Upsilon(a^C) \end{cases}$$

We are now in a position to characterize the set of steady states of the conditional system:

Proposition 3. *The conditional dynamic system given by* (39) *can be characterized for any* $A \in (0, \infty)$ *and in the ranges* $a \in (\underline{a}(A), 1)$, $T \in (e, \infty)$ *as follows:*

(i) *There exists at least one steady state.*

(ii) *Any steady state is characterized by a strictly positive amount of both types of human capital:* $H(A) > 0$ *and* $P(A) > 0$.

(iii) *There exist at most three steady states denoted by* $E^H(A) \equiv \{a^H(A), T^H(A)\}$, $E^u(A) \equiv \{a^u(A), T^u(A)\}$ *and* $E^L(A) \equiv \{a^L(A), T^L(A)\}$ *with the following properties:*

   (a) $a^H(A) \leq a^u(A) \leq a^L(A)$ *and* $T^H(A) \geq T^u(A) \geq T^L(A)$*;*

   (b) $E^H(A)$ *and* $E^L(A)$ *are locally stable;*

   (c) $E^u(A)$ *is locally unstable;*

   (d) *if there is only one steady state, then it is globally stable, and it can be labeled as* H *or* L *according to the curvature of* $HH(A)$: $\frac{\partial^2 T(a^H(A))}{\partial a^2} > 0$ *and* $\frac{\partial^2 T(a^L(A))}{\partial a^2} < 0$.

According to this proposition, there exists at least one dynamic equilibrium. Given the S-shape of $HH(A)$, the conditional system exhibits at most three dynamic equilibria, two of which are stable and one unstable. The two dynamic equilibria at the extremes of the support are locally stable while the intermediate one is not. The 'high' equilibrium is characterized by a relatively large fraction of the population acquiring H and large lifetime expectancy, and the locus $HH(A)$ is locally convex in a_t^H. The 'low' equilibrium is characterized by low lifetime duration and correspondingly a little share of the population acquiring H. The locus $HH(A)$ is locally concave in a^L.

Figure 3 illustrates the dynamic system characterizing the economy, when there exist three equilibria. The linear curve TT represents the intergenerational externality of a generation's theoretical knowledge H on the life expectancy of the next generation, as stated by Equation (33). The S-shaped locus $HH(A)$ illustrates pairs of ability thresholds and lifetime durations (a, T) described by Equation (31), for which the static equilibrium conditions are fulfilled. Thus, the intersection of the two curves satisfies the conditions for a dynamic equilibrium.

The analysis of the full dynamic system must account for the evolution of all the variables at the same time. To do this, it is necessary to study the behavior of the relative productivity. The objective of this section is to

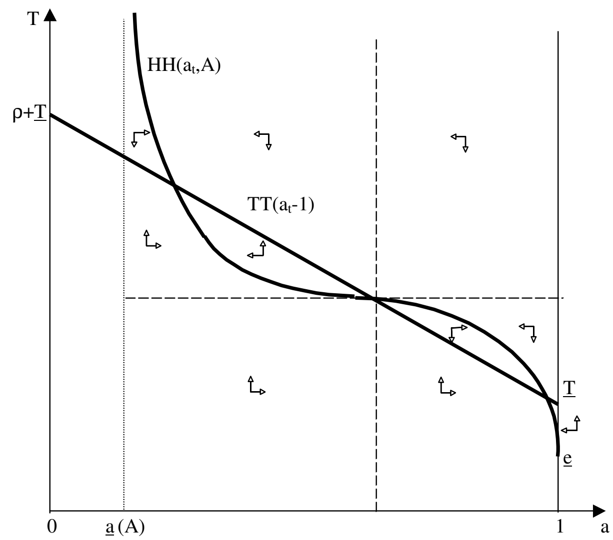

Figure 3: Phase Diagram of the Conditional Dynamic System

give a characterization of the different phases of development of an economy starting from a little productivity A and characterized by a low lifetime expectancy. We argue that these initial conditions once have been historically and empirically relevant for all developed countries in the past and still remain so in most of the underdeveloped countries today. To this end it is sufficient to concentrate attention to the main characteristics of the dynamic evolution of A, while there is no need to characterize its path in detail.

We begin the analysis of development by looking at productivity changes over the course of generations. Human capital H helps in adopting new ideas and technologies, and thus creates higher productivity gains than practical human capital P. This means that in the long run relative productivity A_t will tend to increase, which in turn tends to reinforce the role of theoretical human capital H. This result is summarized by the following lemma:

Lemma 2. *Relative Productivity* A_t *increases monotonically over time with* $\lim_{t\longrightarrow\infty} A_t = +\infty$.

Therefore, productivity increases faster in the sector using theoretical human capital H more intensively, so that this sector becomes relatively more productive over time. As a consequence, H becomes more attractive to acquire. Note that the strict monotonicity of A_t over time depends on the assumption $\dot{A}_{P.t} = 0$. However, this assumption is not necessary for the main argument. What is crucial is that relative productivity will *eventually* be increasing once a sufficiently large fraction of the population acquires H. In the simulations below, we allow $\dot{A}_{P.t} > 0$ starting from large $A_{P.0}$ and small $A_{H.0}$. Relative productivity A_t is, in that case, initially decreasing, reflecting the larger innovative dynamics of sector P in during early stages of development, but since H is relative more important than P for technological progress in any sector in the long run, A_H leapfrogs A_P. Therefore, A_t is eventually increasing and keeps increasing from this point on. The qualitative prediction is totally unchanged, with the only difference that in early stages of development the high productivity in the P sector reinforces the tendency to acquire P and, in this way, delays massive human capital acquisition even further.

As A_t increases, the fraction of the population investing in H also increases. Lifetime expectancy necessary to make an agent of ability a indifferent between acquiring any kind of human capital tend to decrease and the locus $HH(A)$ shifts down for any a (excluding the extremes):

Proposition 4. *The life expectancy required for any given level of ability to be indifferent between acquiring h or p decreases, as relative productivity A increases: the locus $HH(A)$ is such that $\frac{\partial T(a,A)}{\partial A} < 0, \ \forall \ a \in (0,1)$.*

Then, according to the proposition, the more productive theoretical human capital becomes relatively to applied human capital, the less restrictive is the fixed cost requirement of acquiring it, as the break-even of the investment in education is attained at a lower age.

We are now prepared to analyze full dynamic solution of the system. Given the results so far, permanent productivity growth implies that for a given life expectancy the ability threshold for becoming theoretically educated decreases, inducing a higher fraction of the population to acquire theoretical human capital. Of course, this has feedback effects through the externalities of this increase of the aggregate stock of theoretical knowledge in the economy on life expectancy and productivity of the subsequent generation.

We focus on an non-developed economy in which life expectancy at birth is low, as for example during the middle ages.[23] Since the relative productivity A is low, investing in h is relatively costly for large part of the population as the importance of the fix cost for education $\underline{e}$ is large. This means that the concave part of the $HH(A)$-locus is large and the conditional system is characterized by a dynamic equilibrium of type $\left\{a^L(A), T^L(A)\right\}$, exhibiting low life expectancy and a little class of individuals deciding to acquire theoretical human capital, as the ability threshold is very high at a^L. This situation is depicted in panel *(1)* of Figure 4.

In early stages of development, both the relative productivity gains, as well as the effect on the ability threshold are relatively small. Consequently,

[23] As will become clear below, starting from this point is without loss of generality. However, even though the model is also capable of demonstrating the situation of developed economies, the main contribution lies in the illustration of the transition from low to high levels of development.

also the feedback effects on lifetime duration and productivity are close to negligible, but just not quite negligible. Over time, productivity growth makes investing in h easier for everybody as h becomes relatively more valuable, and life expectancy increases slowly. Graphically, the non-linear locus HH of pairs of (a, T) satisfying intragenerational static equilibrium shifts downwards over time and the importance of the concave part decreases.

After a sufficiently long period of this early stage of development, the non-linear locus HH exhibits a tangency point, and eventually three intersections rather than one with the linear locus TT of pairs of (a, T) of the intergenerational externality on life expectancy. From this point onwards, in addition to E^L, also steady states of type E^u and E^H with lower ability thresholds emerge. The intermediate equilibrium is locally unstable, and the economy remains trapped in the area of attraction of the L-type equilibria, as there is no possibility to attain the high life expectancy required for the economy to settle into a H-type equilibrium. This situation is depicted in panel *(2)* of Figure 4.

As generations pass, the dynamic equilibrium induced by initially low life expectancy moves along TT. The consecutive downward shifts of $HH\,(A)$, however, eventually lead to a situation in which the initial dynamic equilibrium lies in the tangency of the two curves, as shown in panel *(3)* of Figure 4. In the neighborhood of this tangency, the static equilibrium locus lies below the linear curve, such that the equilibrium is not anymore stable. Already the following generation faces a life expectancy that is high enough to induce a larger fraction to acquire human capital than in the previous generation. At this point a unique E^H steady state exists, as is shown in panel *(4)* of Figure 4. A period of extremely rapid development is triggered, during which life expectancy virtually explodes, and the human capital structure of the population changes dramatically towards theoretical, h-type education. This phase of rapid change in general living conditions and the economic environment reflects what happened during the industrial revolution.

This phase of fast development lasts for a few consecutive generations. Relative Productivity A is eventually sufficiently large to render investing in h optimal for the majority of individuals. The reason is that the individual fix cost $\underline{e}$ is relatively low for all individuals endowed with at least some low level of ability.

In later stages, after the transition to the high conditional dynamic equilibrium, steady but small increases in life expectancy and in the fraction of the population acquiring human capital are observed. Eventually, the economy ends up in a series of dynamic equilibria characterized by high expected lifetime duration and a low ability requirement for the adoption of theoretical knowledge, E^H. Life expectancy and the share of the population acquiring Human capital h keep increasing. However, the extent of this late growth is very moderate. Life expectancy converges slowly to some (biologically determined) upper bound $\rho + \underline{T}$, which is never achieved.

Once the majority of the population is theoretically educated, ever further technological progress cannot sustain the growth in highly innovative human capital, and living conditions improve less and less rapidly. Some fraction of the population will always acquire applied knowledge, as the ability threshold never reaches zero. This is what happened after the dramatic changes during industrial revolution, and what still happens today. The following section presents a simulation of the model, which illustrates the evolution of the main variables of the model.

In the following proposition, we summarize this process of development. The evolution of the system is given by the infinite sequence of ability thresholds, life expectancies and relative productivities $\{a_t, T_t, A_t\}_{t\in[o,\infty)}$, starting in a situation of an undeveloped economy:

Proposition 5. (Development Path of the Economy) *Consider an undeveloped economy with initially A_0 being small such that, without loss of generality, the conditional system (39) is characterized by a unique steady state of type E^L as formalized in Proposition 3. The solution of the dynamic system* (38) *exhibits the following features:*

(i) *There exists a unique $t^1 \in [0,\infty)$ such that $\forall t < \left(t^1 - 1\right)$ the conditional system (39) is characterized by a unique equilibrium $E^L(A_t)$: $a^L(A_t) > a^L(A_{t+1})$ and $T^L(A_t) < T^L(A_{t+1})$.*

(ii) *At $t = t^1$, the conditional system exhibits two steady state equilibria: $E^u(A_{t^1})$ and $E^L(A_{t^1})$. The economy remains situated in the area of attraction of the conditional steady state $E^L(A_{t^1})$.*

(iii) *There exists a unique $t^2 \in \left(t^1,\infty\right)$ such that $\forall t > t^1 \wedge t < t^2$ the conditional system is characterized by three steady states: $E^H(A_t)$, $E^u(A_t)$ and $E^L(A_t)$ with the economy situated in the area of attraction of $E^L(A_{t^1})$: $a^L(A_{t+1}) < a^L(A_t)$ and $T^L(A_{t+1}) > T^L(A_t)$.*

(iv) *At $t = t^2$, the conditional system displays two steady state equilibria: $E^H(A_t)$ and $E^u(A_t)$.*

(v) *For any $t > t^2$, the conditional system (39) is characterized by a series of unique and globally stable equilibrium of type $E^H(A_t)$ with: $a^H(A_{t+1}) < a^H(A_t)$ and $T^H(A_{t+1}) > T^H(A_t)$.*

It is important to note that the actual trajectory of the system depends on the initial conditions and cannot be precisely identified in general. Proposition 5 in fact states that the system moves period by period in the area of attraction of the locally stable conditional state E^L during phases (i) to (iv). In phase (v), the system converges to a series of globally stable steady states E^H.

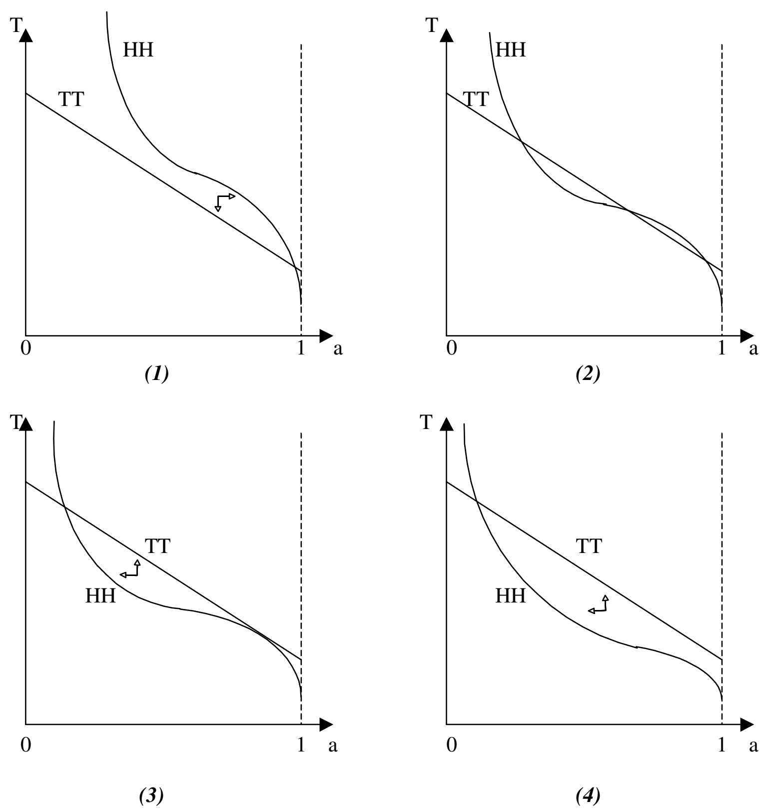

Figure 4: The Process of Development

In historical terms, the model therefore exemplifies the different stages of development. Europe could be thought of as being trapped in a sequence of E^L equilibria during ancient times and the middle ages. At some point during the late 18th century development took off, as the multiplicity of equilibria vanished, and the economies were no longer trapped in the bad equilibrium with low human capital and low life expectancy. Living conditions changed dramatically, and one could think that European economies today are in dynamic E^H equilibria. However, one could also think that e. g. African economies are still trapped today in dynamic equilibria characterized by low life expectancy and little theoretical knowledge (like literacy).

According to this model, an industrial revolution was inevitable, and its timing depended on the particular parameters and the initial conditions. This feature of the model depends on the type of technological progress that is in line with the tradition of endogenous growth theory. Essentially, technological progress is the accumulation of knowledge over time as in Romer (1990). An alternative view of technological progress with stochastic elements, as destruction of knowledge, forgetting and non-continuous, periodic improvements, could imply different predictions about the inevitability of the industrial revolution. For example, one could easily introduce random shocks affecting life expectancy and/or the stock of theoretical human capital in the economy, representing events exogenous to the economic system such as wars. These might prolong or even completely prevent the economic and biological transitions characterized in this paper. In this sense, an industrial revolution would not be inevitable anymore, but due to reasons that lie beyond the mechanisms described here. Different views about the structure of technological progress clearly would also imply different conclusions about the scope for development enhancing policies.

4 A Simulation of the Development Process

This section presents a simulation of the model to illustrate its capability to replicate this phenomenon. We simulate the model using parameters reflecting empirical findings where possible. However, note that these simulations do not claim utmost realism, and we do not calibrate and fine-tune the model in order to achieve an optimal fit with real world data. Rather, the simulations are meant as an illustration of the workings of the model. Table 1 contains the values of the parameters and initial conditions used for the baseline specification of the model.

Table 1: Parameter Values for Simulation

$\alpha =$	0.5;	$\delta_P =$	0.05;	$\rho =$	75.0;	$A_P(0) =$	1.6;
$\beta =$	0.5;	$\phi_H =$	0.95;	$\underline{e}=$	15.0 ;	$\widetilde{a}(0)=$	0.9911;
$\gamma =$	0.6;	$\phi_P =$	0.95;	$\underline{T}=$	25.0 ;		.
$\delta_H =$	0.11;	$\chi =$	0.75;	$A_H(0) =$	1.0 ;		.

Marginal productivity of time spent in education, given a specific level of ability, is assumed to be the same in the production of both types of human capital. The macro-economic returns to human capital production are decreasing in both sectors (γ). In the simulation we assume that TFP is growing with the stock of theoretical human capital of the preceeding generation, H_{t-1} in both sectors, albeit at a faster rate in the sector using theoretical human capital more intensively ($\delta_H > \delta_P$). Both sectors exhibit the same extent of decreasing returns to this stock of human capital. We assume the total scope of extending life expectancy by research, medical inventions and the like as 75 years (ρ). The baseline life expectancy is 25 years, which is in line with Streeten (1994) who cites evidence that average life expectancy in central Europe was even lower than 25 before 1650. This means that even if the entire population would engage in accumulating theoretical knowledge, a life expectancy of 100 years could not be exceeded. The minimum requirement of lifetime with respect to accumulating theoretical human capital is 15 years. Moreover, we start initially from a situation in which total factor productivity is 1.6 times higher in the applied human capital sector. This reflects the fact that at this point in time already a large number of generations has acquired applied knowledge that has increased TFP over time. Finally, we assume that in the first period of the simulation 0.89 percent of the population pay the fixed cost in terms of time spent for education and accumulate theoretical human capital. We simulate the economy over 250 generations. If one wanted to directly test the predictions for the industrial revolution, the simulation period comprises roughly a horizon from 1000 to 2250, interpreting every 5 years as the beginning of a new generation.

Simulation results for life expectancy and the fraction of the population acquiring theoretical human capital are depicted in Figure 5. Life expectancy remains at a low level for many generations, before at a certain point (around 1760) a period of rapid growth in average lifetime duration begins. In fact, even before this period of rapid growth life expectancy is increasing over time, but with very little increments over the generations. However, then life expectancy increases from mid-20 to over 60 within just a few generation. Eventually, the increments decrease, and the growth of life expectancy slows down again, but never actually stops or gets nearly as small as before the transition. Accordingly, just when life expectancy starts

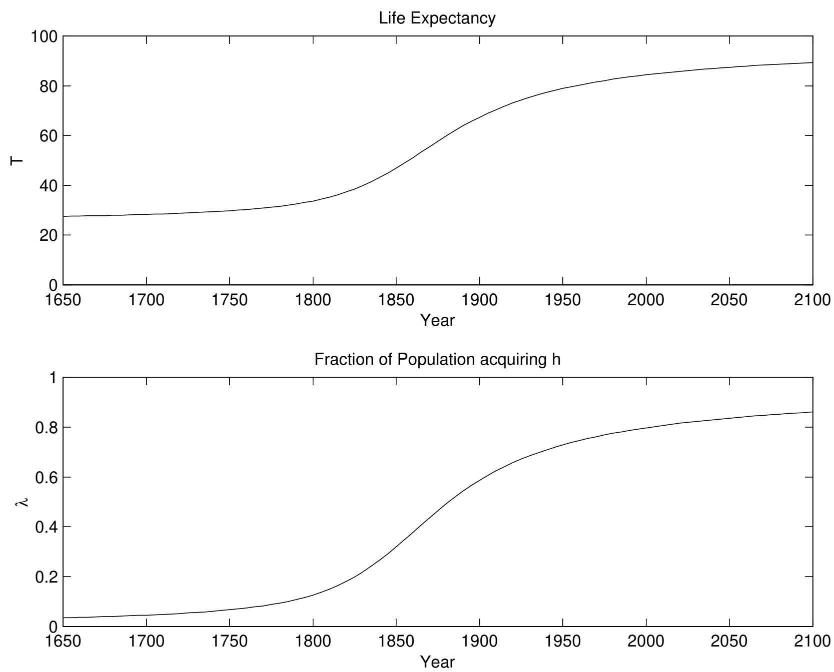

Figure 5: Simulation of Life Expectancy T and the Proportion of the Population with Theoretical Education, λ

to take off, the social structure of the economy starts changing rapidly, as ever larger proportions of the population acquire theoretical human capital. This is reflected in a rapid decrease of the ability threshold for abstract education. However, also this evolution slows down from its initial rapidness, as the share of educated people exceeds roughly three quarters of the population. When more than 90 percent accumulate theoretical knowledge, this fraction hardly grows anymore. Nevertheless, due to the permanent growth in TFP, the aggregate stock of theoretical human capital keeps increasing, even after the transition, albeit at a somewhat slower rate.

Simulation results for TFP, aggregate income and population size are shown in Figure 6. TFP in the theoretical sector is about ten times higher after a time comparable to 250 years since the beginning of the industrial revolution. Also the stock of applied knowledge increases further, thanks to the externality of theoretical human capital on TFP also in the P-sector, and is about three to four times higher after 250 years since industrialization started. Aggregate income grows only very slowly before the industrial revolution. But then it virtually explodes, and keeps growing rapidly, even when growth in life expectancy and the fraction of theoretically educated ebbs away.

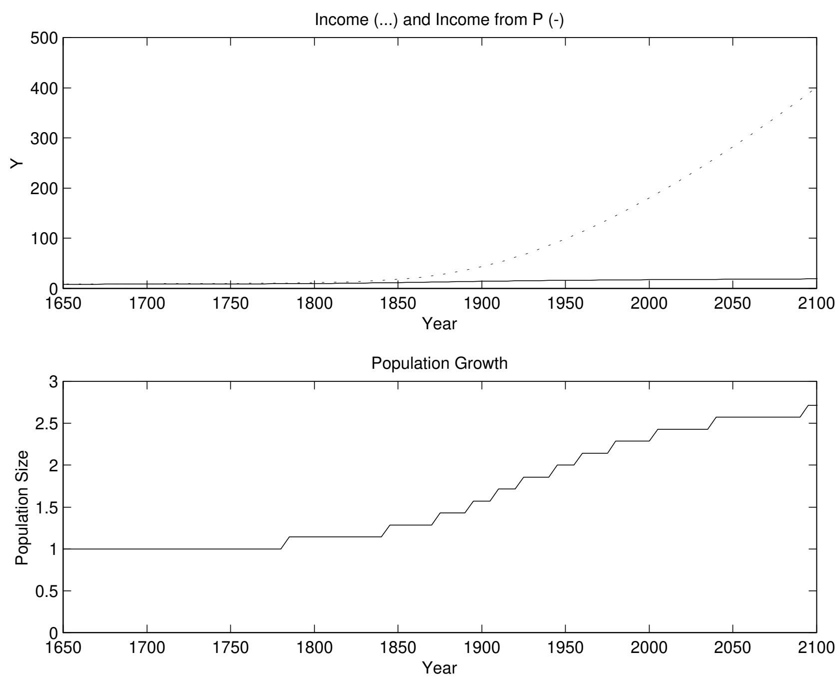

Figure 6: Simulation of Income, Productivity and Population Size

The assumption of fixed frequency of reproduction while life expectancy is endogenous implies that, at the same time, several generations can populate the economy. The number, and thus population size, depends on the respective life expectancies the generations are endowed with. Hence, the size of the population can grow, even though individual fertility behavior is assumed to be constant and the same throughout generations. In the simulation, the frequency in which new generations are born is taken to be five years. Thus, children do not build on the knowledge of their parents, but on the knowledge of their immediately preceeding generation, whose in turn depends on that of older generations, etc. The development of the size of the population along the dynamic equilibrium path is illustrated in the lower panel of Figure 6 Holding fertility constant, the population size almost triples as life expectancy increases in the process of economic development. Population growth slows down again once the growth in life expectancy slows down.[24] A final observation is the endogenous structural transition illustrated by the decline in the share of income generated using applied knowledge and the inverse increase in the income share of theoretical knowledge in Figure 7.

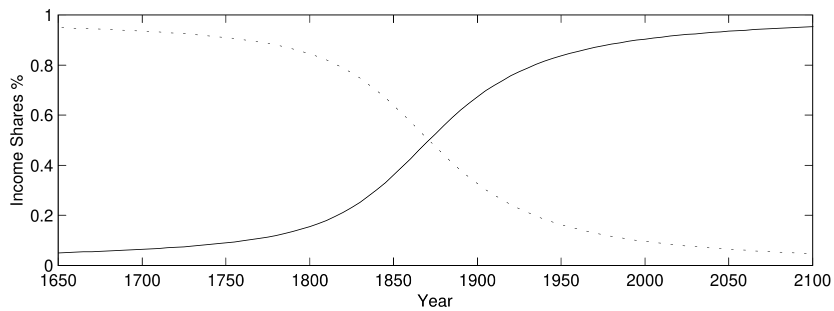

Figure 7: Simulation of Income Shares Y_P/Y (...) and Y_H/Y (-)

5 Concluding Remarks

One major puzzle, which economic explanations for the industrial revolution have to address, is the apparently long stagnancy of economic conditions and life expectancy, which is suddenly followed by a period of fast and dramatic changes in both these dimensions. Previous contributions model the experience of the industrial revolution as the transition from a primitive, stagnant to a developed regime exhibiting permanent growth. What eventually triggered this rapid transition is the topic of a lively discussion within the profession. This paper offers an explanation which is not based on the existence of different regimes of the economy, but interprets long-

[24]The non-smooth, jagged development of the population size follows from the fact that the number of populations alive at each point in time is an integer.

term development including the experience of the industrial revolution as the continuous evolution of the dynamic system of the economy.

We present a simple microfoundation of human capital accumulation that allows to explain patterns in long-term economic development, explicitly taking complex interactions between economic, social and biological factors into account, and model economic development and changes in life expectancy as endogenous processes. An implication of this view is that even during the apparently stagnant environment before the industrial revolution, economic and biological factors affected each other. There is thus no need to explain a change in regimes, or a driving shock that triggered the transition.

Life expectancy is the crucial state variable in the individual education decision. In turn, this education decision has implications for the education decision of future generations, both through life expectancy and productivity changes. This means that advances in technological progress, human capital formation and lifetime duration reinforce each other. However, the peculiarity of human capital is that every generation has to acquire it anew. But the costs for human capital formation are prohibitively high for large parts of the population when the level of development is still low and expected lifetime duration is short. However, at a certain point in time the entire system is sufficiently developed so that the positive feedback loop has enough momentum to overcome the retarding effects of costs for human capital formation. The resulting development path exhibits an S-shape, with a long period of economic and biological stagnation, followed by a relatively short period of dramatic change in living conditions and the economic and social environment. The mechanism presented in this paper is able to reproduce the observed patterns of long-term economic development without the need of relying on some exogenous events and strict temporal causalities. By simulating the model for illustration purposes, we show that the long-run behavior of income, income growth, productivity, lifetime duration, population size and other key indicators of development implied by the model is in line with empirical evidence and stylized facts.

A Appendix

Proof of Proposition 1:

Proof. Consider Equation (28). For notational simplicity, denote $\widetilde{a}^*$ simply as a, $\underline{e}$ as e. By standard calculus,

$$T'(a) = \frac{e\frac{g'(a)}{\Omega}}{\left[1 - \frac{g(a)}{\Omega}\right]^2} < 0\,, \tag{40}$$

since $g'(a) = -\frac{g(a)}{1+\gamma}\left[\frac{2-\gamma-\gamma a^2}{2a^2(1-a^2)}\right] < 0$, $\forall a \in [0,1]$. Therefore, we conclude that for a given set of parameters $\gamma \in (0,1)$, $A = \frac{A_H}{A_P}$, for every $a \in [\underline{a}, 1]$ there is one and only one $T > 0$ such that (28) is satisfied. □

Proof of Lemma 1:

Proof. The claim follows from the definition of $\underline{\widetilde{a}}^*(\Omega)$, and the fact that $g(\widetilde{a}^*)$ is strictly decreasing in $\widetilde{a}^*$. □

Proof of Proposition 2:

The intuition of the proof proceeds as follows: We solve equilibrium condition (27) for T as a function of $\widetilde{a}^*$ and investigate the behavior of this function. Due to the fact that $T(\widetilde{a}^*)$ is strictly monotonically decreasing within the admissible support the function is invertible within this range of support. We then show that there exists one and only one $\widetilde{a}^*$ for which the second derivative of this function equals zero. Since the condition for the second derivative to equal zero cannot readily be solved for $\widetilde{a}^*$, we decompose it into two components and show that one is strictly monotonically increasing within the support while the other is strictly monotonically decreasing, such that there must exist one and only one $\widetilde{a}^*$ for which the condition is satisfied by the intermediate value theorem. But if $T(a)$ has a single inflection point and is invertible, also $a(T)$ has a single inflection point and is therefore S-shaped.

Proof. Consider again Equation (28). We use the notational shorthands as in proof of Proposition 1. Using standard calculus, one can now show that:

$$T'(a) = \frac{e\frac{g'(a)}{\Omega}}{\left[1 - \frac{g(a)}{\Omega}\right]^2}\,, \tag{41}$$

and

$$T''(a) = \frac{e\frac{g''(a)}{\Omega}\left[1 - \frac{g(a)}{\Omega}\right] + 2e\frac{\left[g'(a)\right]^2}{\Omega^2}}{\left[1 - \frac{g(a)}{\Omega}\right]^3}\,. \tag{42}$$

Due to the fact that $T'(a) < 0$ $\forall a \in [\underline{a}, 1]$, we note that the function $T(a)$ is invertible in the range $a \in [\underline{a}, 1]$ of the support. Note also that $T(a) \geq \underline{T}$ $\forall a \in [0,1]$, so the inverse function $a(T)$ is strictly monotonically decreasing for all positive T.

It will prove useful to substitute a^2 with b and to re-write $g(a) \equiv h(b) = \frac{(1-b)^{\frac{1-\gamma}{1+\gamma}}}{b^{\frac{2-\gamma}{2(1+\gamma)}}}k$, $g'(a) \equiv h'(b)$, and $g''(a) \equiv h''(b)$, where $\underline{b} = (\underline{a})^2$. Thus define $T(a) =$

$\mathcal{T}(b)$, so the derivatives $T'(a) = \mathcal{T}'(b)$ and $T''(a) = \mathcal{T}''(b)$ can be re-written in terms of b:

$$\mathcal{T}'(b) = -\frac{[\mathcal{T}(b)]^2}{e}\frac{h'(b)}{\Omega}$$

Existence of an inflection point can already be inferred from a closer examination. Since

$$h'(b) = -\frac{k}{2(1+\gamma)}(1-b)^{\frac{-2\gamma}{1+\gamma}} b^{\frac{-1}{1+\gamma}}(2-\gamma-\gamma b) = -\frac{k}{1+\gamma}h(b)B(b) < 0 \quad \forall b \in [\underline{b}, 1]$$

(where $B(b) = \frac{1-\gamma}{1-b}\frac{2-\gamma}{2b}$), we know that also $\mathcal{T}'(b) < 0 \quad \forall b \in [\underline{b}, 1]$. Moreover, one immediately sees that $\lim_{b\to 1} h'(b) = -\infty \Leftrightarrow \lim_{a\to 1} T'(a) = -\infty$, such that T has infinitely negative slope at both boundaries of the admissible support, suggesting that there must exist at least one inflection point. From these arguments it is also clear that the slope of the inverse function, $a'(T)$, converges to zero at both boundaries of the support.

Analysis of the second derivative $\mathcal{T}''(b)$ allows to show existence and uniqueness of an inflection point. In particular, $\mathcal{T}''(b) = 0$ requires:

$$\begin{aligned}
h''(b)\left(1-\frac{h(b)}{\Omega}\right) &= -\frac{2}{\Omega}\left(h'(b)\right)^2 \\
\Leftrightarrow \frac{kh(b)}{1+\gamma}\left[\frac{B^2(b)}{1+\gamma} - B'(b)\right]\left(1-\frac{h(b)}{\Omega}\right) &= -\frac{2kh(b)}{\Omega(1+\gamma)}B^2(b) \\
\Leftrightarrow \left(\frac{-1}{1+\gamma} + \frac{B'(b)}{B^2(b)}\right) &= \frac{2k}{\Omega(1+\gamma)}\left(\frac{h(b)}{1-\frac{h(b)}{\Omega}}\right). \qquad (43) \\
(LHS) &= (RHS)
\end{aligned}$$

Noting that

$$\frac{B'(b)}{B^2(b)} = \frac{-2\gamma b^2 + 4b(2-\gamma) + 2(\gamma-2)}{(2-\gamma-\gamma b)^2},$$

one finds that

$$\frac{\partial\left(\frac{B'(b)}{B^2(b)}\right)}{\partial b} = \frac{8(2-\gamma)(1-\gamma)}{(2-\gamma-\gamma b)^3} > 0, \forall \gamma \in (0,1), b \in [0,1].$$

This implies that the LHS of the condition for an inflection point ($\mathcal{T}''(b) = 0$), equation (43), is strictly monotonically increasing in b. Furthermore, applying calculus one can also verify that the RHS of condition (43) is strictly monotonically decreasing in b on the support $[0,1]$:

$$\frac{\partial\left(\frac{h(b)}{1-\frac{h(b)}{\Omega}}\right)}{\partial b} = \frac{h'(b)}{\left(1-\frac{h(b)}{\Omega}\right)^2} < 0, \forall b \in [\underline{b}, 1].$$

In order to ensure that there is a value of b for which (43) is satisfied, it remains to be shown that the value of the LHS is smaller than that of the RHS for $b = \underline{b}$ and larger for $a = b = 1$. Noting that $LHS(b=1) = \frac{-1}{1+\gamma} + \frac{1}{1-\gamma} > 0$ and that $RHS(b=1) = 0$ since $h(1) = 0$, one sees that the latter claim is true. The facts that $h'(b) < 0 \; \forall b \in [0,1]$, and that $\lim_{b \downarrow \underline{b}} h(b) = \infty$ indicate that $h(b)$ exhibits a saltus at $b = \underline{(b)}$. Since $LHS(0) = \frac{-1}{1+\gamma} - \frac{2}{2-\gamma} < 0$ and due to the fact that the LHS is strictly monotonically increasing $\forall b \in [0,1]$, the values of LHS and RHS can only be equal for one single value of a. These arguments are illustrated in Figure 8. This

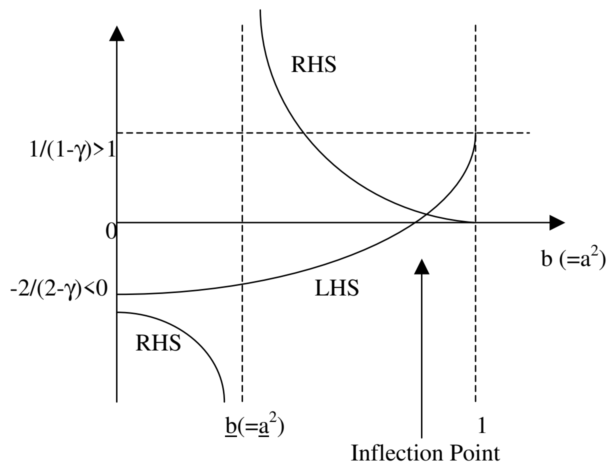

Figure 8: Existence and Uniqueness of an Inflection Point

means that there exists one and only one level of $b \in [\underline{b}, 1]$ such that $T''(b) = 0$. From the fact that the function is invertible in this range of the support, and since there is a one-to-one relationship between a and b, we conclude that the function $a(T)$ also exhibits exactly one inflection point.

□

Proof of Proposition 3:

Proof. Note: As long as there is no danger of confusion, we suppress the subscripts '$_t$' for generation t for notational convenience (e. g. $T_t(a_t) = T(a)$, etc.).

Ad *(i)*: Existence of a dynamic equilibrium for the conditional system. Recall that the locus TT is linear with slope $-\rho$ and values $T(a = 0) = \underline{T} + \rho$ and $T(a = 1) = \underline{T}$. From the proof of Proposition 2 we know that, for any $A > 0$, the locus $HH(A)$ is such that $\lim_{a \downarrow \underline{a}(A)} T_t(a, A) = \infty$, and that its value is monotonically decreasing $\forall a > \underline{a}(A)$. Hence, if the value of this non-linear relation at $a = 1$ is smaller than that of the linear relation of the intergenerational externality, there must exist at least one intersection by the intermediate value theorem. However, note that $T(1) = \underline{e} \; \forall t$, and that by assumption $\underline{e} < \underline{T}$. That means the fixed cost for theoretical education is always lower than any minimum life expectancy, otherwise theoretical education would never be an alternative, not even for the most able individual in the world. Hence a dynamic equilibrium exists for every generation t.

Ad *(ii)*: From the proof of *(i)* and noting that any steady state is characterized by an interior solution with $a < 1$, since $T(a = 1) = \underline{T} > \underline{e}$, which in turn implies that $H_t > 0$ and $P_t > 0$ for any $t > 0$.

Ad *(iii)*: The claims follow from Proposition 2: We know that $HH(A)$ has always a unique turning point and takes values above and below TT at the extremes

$\underline{a}(A)$ and 1. Hence the two curves can intersect *at most* three times, while they intersect at least once by *(i)*. Claim *(*a) follows from the negative slopes of both loci that allow to rank steady states. Claims *(b)* and *(c)* are true since, in the extreme equilibria E^H and E^L, $HH(A)$ intersects TT from above, which means that the system is locally stable, while the opposite happens in the intermediate equilibrium E^u, since $HH(A)$ must cut TT from below. Thus E^u is locally unstable. Claim *(d)* follows from the fact that if only one steady state exists it must be stable since $HH(A)$ starts above TT and ends below so it must cut from above. The concavity/convexity of $HH(A)$ in the stable equilibria is used to identify them since in case of multiplicity one must be in the concave and the other in the convex part. ☐

Proof of Lemma 2

Proof. By assumption, $\delta > 0$, $\phi > 0$, and $\chi > -1$ in equation (35), such that $\dot{A}_{H.t} > 0$, and $A_{H.t} > A_{H.t-1}$ $\forall t$. A_{t-1} and A_t are linked in an autoregressive way, and equation (37) is of the form $A_t = (c_{t-1}+1)A_{t-1} = d_{t-1}A_{t-1}$, where $d_{t-1} = \delta H_t^{\phi} A^{\chi}_{.t-1} + 1 > 1$ for any t, since from Proposition 3 $H_t > 0$ for any t and $\delta > 0$. This means that the process is positive monotonous and non stationary. Starting with any $A_0 > 0$ we can rewrite $A_t = \left(\prod_{i=1}^{t} d_{i-1}\right) A_0$, where $\left(\prod_{i=1}^{t} d_{i-1}\right) > 1$ and $\lim_{t\longrightarrow\infty}\left(\prod_{i=1}^{t} d_{i-1}\right) = \infty$.

Note: If there is TFP growth also in the P-sector, it is sufficient for the argument to hold to assume that $\delta > \delta_P \geq 0$, $\phi \geq \phi_P$ and $\chi \geq \chi_P$ in equation (35) and footnote 21. Then, the relative increment to TFP each period is larger in the H-sector, and the claim holds for identical initial values. For higher initial values of A_P it only holds after sufficiently many periods (generations) have passed. ☐

Proof of Proposition 4

Proof. As in the proof of Proposition 2, solve equation (27) for $T(a)$ to get:

$$T_t(a_t) = \frac{e}{1 - \frac{g(a_t)}{\Omega_t}} . \tag{44}$$

The claim follows by partial derivation of equation (44), $\frac{\partial}{\partial \Omega_t} T_t(a_t) = -\frac{g(a_t)e}{(\Omega_t - g(a_t))^2} < 0$ $\forall a_t \in [\underline{a_t}, 1]$. ☐

Proof of Proposition 5:

Proof. Claims *(i)* and *(ii)*: Consider Equation (44), and denote denote the derivative $HH'(A) = \frac{\partial T_t(a_t, A_t)}{\partial a_t}$. From proposition 2 we know that $HH'(A)$ is U-shaped and takes infinite value at the extremes of the support $\{\underline{a}(A_t), 1\}$, with a unique minimum corresponding to the inflection point of $HH(A_t)$. Existence of at least one dynamic equilibrium has been shown in Proposition 3. Hence, there exists an initial level A_0 sufficiently small such that only one equilibrium $E^L(A_0)$ exists. To see that this equilibrium is of type L, note that if $A_0 \longrightarrow 0$, then the loci $HH(A)$ and TT (which has slope equal to ρ) can cross only once for a level a close to 1 since: $\lim_{A_0\longrightarrow 0} \underline{a}(A_0) = 1$. Also along $HH(A)$ the $\lim_{A_0\longrightarrow 0} \frac{\partial T_0}{\partial a_0} = \infty > \rho$; thus $\forall a_0 \in [\underline{a}(A_0), 1]$ and $HH'(A_0)$ has value always larger than ρ. This is illustrated in Figure 9. Remember that $\frac{\delta T_t(a_t,A_t)}{\delta a_{A_t}} < 0\ \forall a_t \in [\underline{a}(A_t), 1]$. As time passes, relative productivity A_t increases, as shown in Lemma 2. Hence, $\forall t > 0$ the locus $HH(A_t)$ lies below the locus $HH(A_0)$ while the locus TT is unchanged. Both curves necessarily intersect for $a^L(A_{t+1}) < a^L(A_t)$.

For illustrative purposes, continue with the Proof of Claim *(v)*: The locus $HH(A)$ becomes L-shaped as A gets large, and exhibits infinite value at $a = 0$ and value $\underline{e}$ elsewhere. The $\lim_{A\longrightarrow\infty} T(a, A) = \underline{e},\ \forall a \in (0, 1]$. Since $\underline{a}(\infty) = 0$, and $T(\underline{a}, A) = \infty$ for any A, then $T(0, \infty) = \infty$. This means that, eventually there will be a unique equilibrium with a_t^H close to zero. Existence of another equilibrium can be ruled out by contradictions since it would imply that $HH(A_t) > \rho$ for some $a < a_t^H$ which is impossible due to the previous limits. Hence, from a certain generation onwards there exists a unique high type equilibrium E^H.

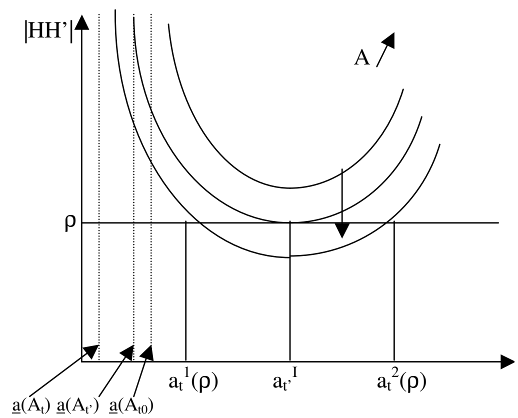

Figure 9: Emergence of Multiple Equilibria

Proof of Claims *(iii)* and *(iv)*: Obviously, the transition from a unique dynamic equilibrium E^L to a unique dynamic equilibrium E^H, the system must pass a phase characterized by multiple steady state equilibria. As generations pass we have from (41):

$$\frac{\partial}{\partial \Omega_t}\left|\frac{\partial T_t(a_t, A_t)}{\partial a_t}\right| = \frac{-\Omega_t + g(a_t) - 1}{\left(1 - \frac{g(a_t)}{\Omega_t}\right)^3 \Omega_t^3} g'(a_t) e \frac{1}{\Omega_t} < 0\ \forall a_t \in [\underline{a}_t(A_t), 1]\,. \tag{45}$$

Also $\lim_{A\longrightarrow\infty}\left(\frac{\partial T_t(a_t,A_t)}{\partial a_t}\right) = 0\ \forall a \neq 1$ and $a \neq 0$. Hence, as shown in Figure 9, the locus $HH'(A)$ shifts down monotonically. Its value eventually converges to zero *in the interior* of the bounded support as $A \longrightarrow \infty$. From Proposition 2 we know that the locus $HH'(A)$ has a unique minimum corresponding to the inflection point of $HH(A)$. Denote the ability threshold of the inflection point as a_t^I for any t.

Thus the full system passes from a series of unique equilibria E^L, characterized by concavity of $HH(A)$ to a series of unique equilibria E^H, characterized by convexity of $HH(A)$. By continuity the dynamic equilibrium must, at the certain point be characterized by the tangency between the two loci in the inflexion point of HH (so that $HH(A)$ is locally concave and convex).

Since as a consequence of Lemma 2 A grows monotonically and without bound as generations go by, and due to the resulting downward shift of $HH'(A)$, there exists a $t' < \infty$ such that: $\frac{\partial T_{t'}(a_{t'}^{I}, A_{t'})}{\partial a_{t'}} = HH'(a_{t'}, A_{t'}) = \rho$. For any $t > t'$ there exist two levels of a: $a_t^1(\rho) < a_t^I < a_t^2(\rho)$ such that: $\frac{\partial T_t(a_t^1(\rho), A_t)}{\partial a_t} = HH'(a_t^1(\rho), A_t) = \frac{\partial T_t(a_t^2(\rho), A_t)}{\partial a_t} = HH'(a_t^2(\rho), A_t) = \rho$. Hence, the locus $HH(A)$ is convex in $a_t^1(\rho)$ and concave in $a_t^2(\rho)$. Also, as the locus $HH'(a, A)$ shifts down, $\frac{da_t^1(\rho)}{dt} < 0$ and $\frac{da_t^2(\rho)}{dt} > 0$ (see also Fig. 9). Finally, note that $\lim_{A \longrightarrow \infty} a_t^1(\rho) = 0$ and $\lim_{A \longrightarrow \infty} a_t^2(\rho) = 1$ since $HH'(\cdot, \infty)$ takes value zero in the interior of the support as shown previously.

By continuity, as the locus $HH(a, A)$ shifts downward, there exists a t^1 such that: $\frac{\partial T_{t'}(a_{t1}^1(\rho), A_{t1})}{\partial a_{t'}} = HH'(a_{t^1}, A_{t^1}) = \rho$ and that $HH(a_{t^1}, A_{t^1}) = TT$. At t^1, a new locally unstable steady state emerges in which $a_{t^1}^1(\rho) = a_{t^1}^u$, but the system remains trapped in the locally stable low equilibrium. For any $t > t^1$ the locus $HH(A)$ shifts further downwards and exhibits three intersections with locus TT. Now define $\Delta(t) = a_t^2(\rho) - a_t^L \quad$ with $\Delta(t^1) > 0$. Now note that $a_t^2(\rho) < a_t^L$ for any $t > t'$, and that $\frac{\partial a_t^2(\rho)}{\partial t} > 0$ and $\lim_{A \longrightarrow \infty} a_t^2(\rho) = 1$. From the proof of Claims (i) and (ii) above we know that $\frac{\partial a_t^L}{\partial t} < 0$. Hence, $\frac{\partial \Delta}{\partial t} < 0$ and therefore there $\exists$ $t^2 > t^1$: $\Delta(t^2) = 0$ such that: $a_{t^2}^2(\rho) = a_{t^2}^L = a_{t^2}^u$.

During the life of generation t^2 there exist only two dynamic steady state equilibria with $E_{t^2}^L = E_{t^2}^u$. For this generation, $HH(a_{t^2}, A_{t^2})$ is tangent to TT so that $HH(a_{t^2}, A_{t^2})$ is both concave and convex in the equilibrium. However, due to condition (45), the static equilibrium locus HH keeps shifting down, and hence and $\frac{\partial a_t^2(\rho)}{\partial t} < 0$ also for $t > t^2$. From this point onwards the concave part of $HH(A)$ lies entirely below TT, and only one dynamic equilibrium E^H remains. For any $t > t^2$ the dynamic system converges, at a decreasing speed, to a sequence of E_t^H equilibria where $a^H(A_{t+1}) < a^H(A_t)$ and $T^H(A_{t+1}) > T^H(A_t)$ since the locus $HH(a_{t+1}, A_{t+1})$ lies always below the locus $HH(a_t, A_t)$ while the locus TT remains unchanged. The dynamic equilibrium for $t > t^2$ is characterized by much lower levels of a and much higher levels of T than before t^2. As a consequence, the adjustment involves rapid reductions in a and increases in T for a few generations after t^2, but the growth rate eventually slows down as the system gets closer to E^H.

Note: The dynamics of the full system necessarily follows the stages described in the proposition and the actual path just depends on the initial value of A_0. If A_0 is large enough for the conditional system to display three steady states, then the convergence to one of the stable ones would depend just on the initial value of lifetime duration. If lifetime duration is long enough to bring the system to the E^H equilibrium right from the beginning, then we are considering an economy that already escaped the low life expectancy equilibrium. Clearly the empirically relevant initial conditions suggest to choose initial T_0 in the area of attraction of E^L. $\square$

References

Becker, G. S., K. M. Murphy, and R. Tamura (1990): "Human Capital, Fertility and Economic Growth," *Journal of Political Economy*, 98(1), 12–37.

Blackburn, K., and G. P. Cipriani (2002): "A Model of Longevity, Fertility and Growth," *Journal of Economic Dynamics and Control*, 26, 187–204.

Boucekkine, R., D. de la Croix, and O. Licandro (2002a): "Vintage Human Capital, Demographic Trends, and Endogenous Growth," *Journal of Economic Theory*, 104(2), 340–375.

——— (2002b): "Early Mortality Declines at the Dawn of Modern Growth," *EUI Working Paper*, ECO 2002/11.

Cipolla, C. (1969): *Literacy and Development in the West*. Penguin Books, Middlesex, England.

Croix, D. d. l., and O. Licandro (1999): "Life Expectancy and Endogenous Growth," *Economics Letters*, 65, 255–263.

Easterlin, R. A. (2002): "Review of James C. Riley: "Rising Life Expectancy: A Global History"," *Economic History Services*, URL: http://www.eh.net/bookreviews/library/0448.shtml(Feb. 12).

Floud, R., and D. McCloskey (1994): *The Economic History of Britain Since 1700*. Cambridge University Press, Cambridge.

Galor, O., and O. Moav (2002a): "Das Human Kapital," *mimeo*.

Galor, O., and O. Moav (2002b): "Natural Selection and the Origin of Economic Growth," *Quarterly Journal of Economics*, 117(4), 1133–92.

Galor, O., O. Moav, and D. Vollrath (2002): "Divergence and Overtaking: Land Abundance as a Hurdle for Education Reforms," *mimeo*.

Galor, O., and D. N. Weil (2000): "Population, Technology, and Growth: From Malthusian Stagnation to the Demographic Transition and Beyond," *American Economic Review*, 90(4), 807–828.

Goodfriend, M., and J. McDermott (1995): "Early Development," *American Economic Review*, 85(1), 116–133.

Hansen, G., and E. Prescott (2002): "From Malthus to Solow," check reference.

Hansen, G. D., and E. C. Prescott (1998): "Malthus to Solow," *NBER Working Paper*, 6858.

Jones, C. I. (2001): "Was an Industrial Revolution Inevitable? Economic Growth Over the Very Long Run," *mimeo.*

——— (2002): "Sources of U.S. Economic Growth in a World of Ideas," *American Economic Review*, 92(1), 220–239.

Kalemli-Ozcan, S. (2002): "Does Mortality Decline Promote Economic Growth?," *Journal of Economic Growth*, 7(4), 411–439.

Kalemli-Ozcan, S., H. E. Ryder, and D. N. Weil (2000): "Mortality Decline, Human Capital Investment, and Economic Growth," *Journal of Development Economics*, 62, 1–23.

Komlos, J., and M. Artzrouni (1990): "Mathematical Investigations of the Escape from the Malthusian Trap," *Mathematical Population Studies*, 2(4), 269–287.

Lichtenberg, F. R. (1998): "Pharmaceutical Innovation, Mortality Reduction, and Economic Growth," *NBER Working Paper*, 6569.

Lucas, R. E. (2002): *Lectures on Economic Growth.* Harvard University Press, Cambridge, Massachusetts.

Maddison, A. (1991): *Dynamic Forces in Capitalist Development: A Long-Run Comparative View.* Oxford University Press, Oxford.

McKeown, T. (1977): *The Modern Rise of Population.* Butler and Tanner, London.

Mokyr, J. (1993): "Technological Progress and the Decline of European Mortality," *American Economic Review*, 83(2), 324–330.

Reis-Soares, R. (2001): "Life Expectancy, Educational Attainment, and Fertility Choice," *mimeo.*

Riley, J. C. (2001): *Rising Life Expectancy: A Global History.* Cambridge University Press, New York.

Romer, P. M. (1990): "Endogenous Technological Change," *Journal of Political Economy*, 98(5), S71–S102.

Schultz, P. T. (1993): "Mortality Decline in the Low-Income World: Causes and Consequences," *American Economic Review*, 83(2), 337–342.

Sen, A. (1977): "On Weights and Measures: Informational Constraints in Social Welfare Analysis," *Econometrica*, 45(7), 1539–72.

——— (1985): *Commodities and Capabilities.* North-Holland, Amsterdam.

Streeten, P. (1994): "Human Development: Means and Ends," *American Economic Review*, 84(2), 232–237.

Swanson, C. E., and K. J. Kopecky (1999): "Lifespan and Output," *Economic Inquiry*, 37(2), 213–225.

Tamura, R. (1996): "From Decay to Growth: A Demographic Transition to Economic Growth," *Journal of Economic Dynamics and Control*, 20(6-7), 1237–1261.

——— (2002): "Human Capital and Economic Development," *Federal Reserve Bank of Atlanta working paper*, 2002-5.

Vaupel, J. W. (1998): "Demographic Analysis of Aging and Longevity," *American Economic Review*, 88(2), 242–247.

Www.Mortality.Org (2002): "Human Mortality Database," provided by University of California, Berkeley, and Max Planck Institute for Demographic Research, downloaded Nov. 11, 2002.