Institute for
FUTURES STUDIES

Lagerlöf, Nils-Peter

From Malthus to Modern Growth: Can Epidemics Explain the Three Regimes?

Paper presented at the conference: ”Population Dynamics and the Macro Economy” organized by Harvard Center for Population and Development Studies and IFS. September 2000.

Arbetsrapport/Institutet för Framtidsstudier; 2001:1
ISSN 1652-120X
ISBN 91-89655-14-1

From Malthus to Modern Growth: Can Epidemics Explain the Three Regimes?

Nils-Petter Lagerlöf*
Department of Economics
Concordia University
1455 de Maisonneuve Blvd. W.
H3G 1M8 Montreal QC
Canada
E-mail: nippe@alcor.concordia.ca

February 6, 2001

*This paper, and an earlier version, have benefited from comments from Debasis Bandyopadhyay, Bill Cardak, Alexia Fürnkranz-Prskawetz, Paul Klein, Tomas Kögel, Johan Lagerlöf, Greg LeBlanc, Hans Lundström, Jim McIntosh, Olivier Morand, Rachel Ngai, Mike Sadler, Avi Simhon, and participants at a seminar at the University of Melbourne; the Midwest Macro Meetings 2000 in Iowa; a workshop at the Max-Planck Institute for Demographic Research in Rostock, Germany; the SED Meetings 2000 in Costa Rica; the Econometric Society World Meetings 2000 in Seattle; and a workshop at Harvard University financed by the Swedish Institute for Future Studies.

Abstract:

These are the stylized facts of long-run economic and demographic development, as described by Galor and Weil (AER 1999, 2000): Under an initial Malthusian Regime the growth rates of population and per-capita income are both low. Then follows a Post-Malthusian Regime, with higher growth rates of both population and per-capita income. Finally, the economy transits into a Modern Growth Regime, with falling population growth rates, but accelerated growth rates of per-capita income. This paper models the transition through all these three regimes endogenously. The model also captures the empirical regularity of a simultaneous fall in the level and the volatility of death rates, and the fact that death rates fell before birth rates. Throughout time, we let epidemic shocks hit the economy at a constant rate. However, with rising human capital the impact of these shocks is mitigated. For many generations the economy is stuck in a Malthusian Regime with volatile and high death rates. Sooner or later it experiences a phase of relatively mild epidemics. Mortality declines, enabling population and human capital to simultaneously start growing: a Post-Malthusian Regime. Once human capital growth has taken off, epidemic shocks have smaller impact. Finally comes a stage where parents start having fewer children, and instead invest more in their education: a quality-quantity switch. This triggers faster growth in human capital. The economy enters the Modern Growth Regime.

1. Introduction

> *"This entire sequence of events from 1894 to 1921 occurred under the eyes of professionally sensitized medical teams whose job was to find out the best control to plague. [...] Without such study and the prophylactic measures that followed, the twentieth century might well have been inaugurated by a series of plagues reaching completely around the earth, with death tolls dwarfing those recorded from the age of Justinian and the fourteenth century, when the Black Death ravaged Europe and much of the rest of the Old World. [...] Doctors and public health officers, in fact, probably forestalled epidemics that might have checked or even reversed the massive world-wide growth of human population that distinguishes our age from all that have gone before."*
>
> William H. McNeill (1976, pp. 156-158), speculating about the possible consequences of a series of plague epidemics in China, one of which reached Canton and Hong Kong in 1894, but never spread to the rich world.

Throughout history the levels of human population and its living standards have remained fairly constant. The sudden increase in income that came with the industrial revolution 200 years ago is a very recent phenomenon, and so is the initial rise — and subsequent fall — in population growth that followed it.[1]

This observation provokes some questions. What caused the industrial revolution in the first place, and why did it take so long to occur? Was it inevitable? Could it have happened at a different point in time, and/or in a different region of the world? These issues are addressed in a recent growth literature. Within

[1] This follows from logical reasoning: had income before the industrial revolution grown at rates similar to what we see today, income levels 1000 years ago, say, would have been so low that no one could reasonably have survived. Thus, income levels growing at present rates must be quite a recent phenomenon [cf. e.g. Kuznets (1971, pp. 23-27), and Pritchett (1997)].

one unified framework, various authors try to explain both the stagnant levels of income and population for long periods of time, the recent and sudden rise in income growth, and the associated changes in population growth.

The economic and demographic development of the Western World the last couple of millennia has been described by Galor and Weil (1999, 2000) as passing through three distinct stages, or regimes. The first is called the *Malthusian Regime*. Here population and living standards are constant, or grow very slowly. The relationship between per-capita income and population growth is positive: small increases in income lead to increased population growth. This regime prevails from around the first century A.D. (or even earlier) until around the late 18th century. Thereafter, the economy transits into what Galor and Weil call a *Post-Malthusian Regime*, with increased growth rates of *both* population and per-capita income. The relationship between income and population growth remains positive, as in the Malthusian Regime. The final stage of development is the *Modern Growth Regime*, starting towards the end of the 19th century. Here per-capita income growth accelerates even further, whereas population growth declines, reflecting a negative relationship between the two.

Our paper sets up a growth model which endogenously generates the transition through these three regimes. One ingredient is the presence of mortality shocks, called epidemics. However, what differs between the three regimes is not the shocks per se — they are generated by the same exogenous distribution throughout time; what is changing is the way society *responds to* the shocks. This makes our model able to explain two other facts about the demographic transition: firstly that death rates fall before birth rates, and secondly that not only the level of death rates fall, but also their volatility.

We use a two-period overlapping-generations model. Parents decide how many children to have, and spend time caring for them. Part of that time builds up children's human capital. Parents thus face the well-known trade-off between spending more time (higher quality) on fewer children (lower quantity), or vice versa.

A fraction of the children die before adulthood. As mentioned, this size of this fraction is subject to epidemic shocks. We assume that the impact of these epidemics is *less severe* if human capital is abundant. This captures the effect of advances in medical skills, and knowledge about how diseases are spread, as suggested by the quotation above. At the same time we assume that the impact of an epidemic becomes *more severe* if population density is high, because diseases then spread more easily. In fact, the very establishment of epidemic diseases among humans is closely related to the rise in population density that came with the birth of agriculture and cities [McNeill (1976), Diamond (1999)]. Also, up until at least the end of the 19th century, urban mortality rates in Northwest Europe were higher than those in rural areas [Kunitz (1983)].

However, in our model population density is not only a bad thing: we allow for a "scale effect" so that the productivity in human-capital production increases with population density. This captures a positive effect from the growth and establishment of cities on learning and the spread of knowledge. This assumption is suggested by e.g. Becker, Glaeser and Murphy (1999), and supported by the empirical findings of e.g. Glaeser and Mare (1994). See also Johnson (2000).

The Malthusian Regime in our model is a locally stable steady-state where death and birth rates are both high, and population roughly constant. Moreover, mortality is highly volatile, increasing dramatically in periods of big epidemic shocks. In periods with mild shocks population expands. This makes the impact of the next epidemic worse, equilibrating population back to its Malthusian state. With luck, the economy can escape this trap if it experiences a sustained phase of mild shocks. The resulting population expansion raises productivity in human capital production. More human capital in turn mitigates the effects of future epidemics. In the beginning of such a take-off the economy is still vulnerable: a severe shock can wipe out a big part of the population, reduce human capital productivity, and push society back to the Malthusian trap. This could be what happened to Europe during the Black Death 1346-50.

However, once human capital growth gets a chance to take off the economy becomes increasingly resistant to shocks. While new diseases arrive at an unaltered rate, more abundant human capital mitigates their impact. Society learns new and better ways to stop epidemics, as illustrated by the quotation in the beginning of this section. Mortality rates fall and flatten out. This is consistent with data from e.g. Sweden (a country with relatively good statistics from its demographic transition). As shown in Figure 1.1, as death rates fell, they also became less volatile. The same pattern holds for the rest of Europe.[2] Despite lower mortality, in our model birth rates are still unchanged at this stage, so population expands. This is the Post-Malthusian Regime.[3]

All through the Malthusian and Post-Malthusian Regimes, parents invest nothing in their children's education. However, with growing population and human capital productivity, parents at some stage find it optimal to invest time-costly education in their offspring. This makes children more costly to rear, inducing a fall in birth rates: the standard quality-quantity switch. As mortality cannot fall further (at least not below zero), population growth starts falling. Simultaneously, human capital growth jumps to even higher levels, as education constitutes an input in human capital production. The economy enters the Modern Growth Regime.

Earlier attempts to jointly model the industrial revolution and the demographic transition are relatively recent. An incomplete list as of today should probably include Kremer (1993), Goodfriend and McDermott (1995), Hansen and Prescott (1998), Lucas (1998), Doepke (1999), Galor and Weil (2000), Galor and Moav (2000a,b), and Tamura (2000).[4] In most of these papers, as in ours, endo-

[2] See Livi-Bacci (1997, p. 118), Maddison (1982, pp. 50-52), and Easterlin (1996, pp. 7-9).

[3] Birth rates in the Malthusian and Post-Malthusian Regimes are constant in our model, rather than volatile as in Figure 1.1. See Section 3 for a discussion of this.

[4] These papers build on earlier work by Becker and Barro (1988), Barro and Becker (1989), and Becker, Murphy and Tamura (1990), who were (among) the first to model a quality-quantity

geneity of human-capital investment and/or fertility are the key elements driving the dynamics.

However, none of these papers allow explicitly for mortality (in some cases not even fertility). An early paper with growth and mortality is Ehrlich and Lui (1991). In their model an exogenous fall in mortality raises the return to human-capital investment, and pushes the economy onto a sustained growth path. As human-capital growth takes off, parents substitute quantity for quality (reduce birth rates), leading to a demographic transition. Kalemli-Ozcan, Ryder, and Weil (2000) model similar mechanisms in a continuous time framework. Our model differs in the channels through which the mortality decline promotes growth. Rather than directly raising the returns to human-capital investment, falling mortality causes a population expansion and a subsequent rise in the return to education via the scale effect in human-capital production. Once productivity in human-capital production reaches a critical level a non-negativity constraint on education time ceases to bind, and the quality-quantity switch sets in. This is also what causes the delayed fall in births rates, after death rates have started to fall. Moreover, unlike Ehrlich and Lui (1991) and Kalemli-Ozcan et al. (2000), we model mortality endogenously.[5]

The common approach to endogenize mortality is to postulate that mortality depends negatively on some other variable, which measures living standards, or economic development. This other variable may be per-capita income, as in Kalemli-Ozcan (2000b)[6]; consumption, as in Jones (1999); or human capital, as

trade-off in children into an endogenous growth framework.

[5]Another long-run growth model with exogenous, but explicit, mortality is Morand (2001).

[6]Kalemli-Ozcan's (2000b) model differs from ours, and most earlier work, by explicitly letting each child face a probability of dying: the fraction of children dying is thus uncertain to the parent (whereas we let it be known). If marginal utility of children is convex, a fall in the death probability of each child (i.e. child mortality) lowers the precautionary demand for children, and leads to an associated increase in education. See also Kalemli-Ozcan (2000a), Pörtner (2001), and Tamura (2001).

in our model, Morand (2000), and Tamura (2001). One novelty of our approach is the balance between two counteracting mortality factors, population density and human capital, and the way they amplify and mitigate epidemic shocks.

The remainder of this paper is organized as follows. Section 2 sets up the model, describing production and consumption (Section 2.1), mortality (Section 2.2), and human capital (Section 2.3). The utility function is specified in Section 2.4, where the optimal behavior of the agents is also derived. Section 2.5 presents the dynamics and draws the phase diagrams: first when education time is constrained to zero (Section 2.6), and then when education time operative (Section 2.7). Section 2.8 describes how to interpret the full transition from a Malthusian to a Modern Growth Regime in terms of the phase diagrams, and Section 2.9 shows how the model can be calibrated and simulated to generate the same transition. Section 3 ends with a concluding discussion.

2. The Model

Consider an overlapping generations model, where agents live for two periods: as children, and adults. In childhood, agents make no decisions, but require time to rear and educate. The time required by each child has a fixed component, which the parent cannot influence, and a component which the parent can choose freely, subject to a non-negativity constraint. The latter is called education time. Both types of time constitute inputs in the production of the children's human capital, but education time is more productive. Some children die; others enter adulthood, taking as given their human capital stocks, and decide how much to consume, how many children to rear, and how much time to spend educating their own children.

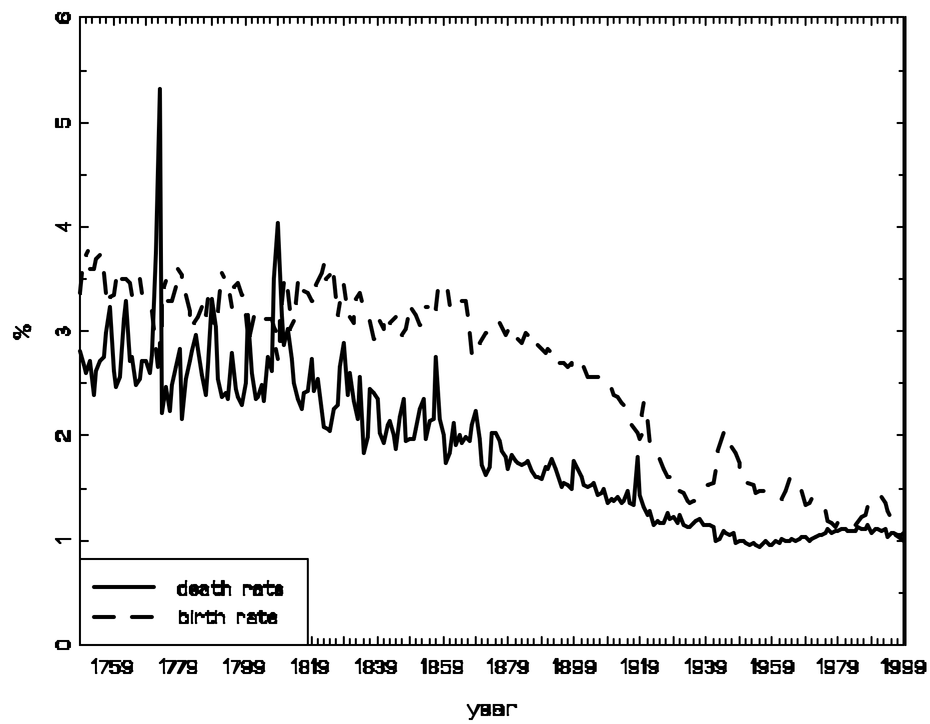

Figure 1.1: Death and birth rates for Sweden 1749-1999. Source: Statistics Sweden (1999), and Hans Lundström at Statistics Sweden.

2.1. Consumption and production

Letting all production take place in the household sector, period t output of the consumption good is given by

$$Y_t = Dl_t(L + H_t), \tag{2.1}$$

where Y_t is output, D is a productivity parameter, l_t is time input in the consumption goods sector, and $L+H_t$ denotes the time-augmenting human capital. Nature equips every agent with L units of human capital, and H_t measures the human-capital component inherited from parents. When there is no risk of confusion, we shall refer to H_t as human capital.

Childhood consumption is set to zero, so goods produced are spent on consumption by the adult:

$$Y_t = C_t, \tag{2.2}$$

where C_t is the adult's consumption.

2.1.1. Time

The budget constraint for time is given by

$$1 = l_t + (v + h_t)B_t, \tag{2.3}$$

where B_t is the number of born children (or births), and $v + h_t$ is the time spent on each child. Each adult has an endowment of time equal to 1. The number of children is continuous and non-negative. We abstract from gender heterogeneity, so each individual parent reproduces herself.

The component v is a fixed time cost of rearing one child, which can be interpreted as the time required to nurse the child just enough to keep her alive.

Parents are effectively constrained to undertake this nursing time.[7] The component h_t measures time spent educating each child. This is a choice variable to the parent, but subject to a non-negativity constraint.

2.2. Mortality

Some fraction of the births B_t die before reaching the second phase of life. This fraction is denoted $1 - T_t$, which we shall refer to as mortality. We assume that mortality increases when the economy is hit by certain shocks denoted ω, which can be interpreted as epidemics. These are *i.i.d*, with a lower and an upper bound, denoted $\underline{\omega}$ and $\overline{\omega}$, respectively. Whereas the distribution of these shocks is constant over time, the effects of them need not be. In 18th-century Europe death rates were very volatile, increasing dramatically in periods of epidemics and famines. Towards the second half of the 18th century, as the industrial revolution took place, these mortality peaks flattened out. See Livi-Bacci (1997, p. 118), Maddison (1982, pp. 50-52), and Easterlin (1996, pp. 7-9). Our model can explain this pattern.

We assume that the severeness of epidemics (1) increases with population density, but (2) falls with human capital. Assumption (1) captures the idea that diseases spread more easily in densely populated areas. This seems to have empirical support. As described by Diamond (1999, Ch. 11), so-called crowd diseases (which tend to spread as epidemics) occurred together with large and dense human populations. They first began to appear with the rise of agriculture around 10,000 B.C. (as the number people a given amount of land could support increased), and accelerated with the rise of densely populated cities somewhat later.

Assumption (2) captures the mortality reducing effects of improved medical skills, and knowledge about how diseases spread. An illustration is given by the quotation in the beginning of the introduction. Also, an improved skill level in

[7] Equivalently, utility could equal minus infinity if the parent spends less than v units of time on each child.

the economy probably helped to improve governing in the newly formed European nation states. According to Easterlin (1996, p. 7) early reductions in mortality were connected with an "improved ability of central administrations to isolate entire regions from epidemics and to contain subsistence crises." See also Kunitz (1983). Similarly, the more recent AIDS epidemic has hit poor countries harder, where the levels of human capital are lower, and government institutions often weaker.

Let P_t denote the (adult) population size in period t. With land fixed, P_t also measures population density. To keep the understanding of the model simple, we let mortality depend on the ratio $H_t/P_t \equiv R_t$, i.e.,

$$T_t = T\left(R_t, \omega\right). \tag{2.4}$$

We assume:

Assumption 1. $T : \mathbb{R}_{++} \times [\underline{\omega}, \overline{\omega}] \rightarrow (0,1)$ *satisfies*

$$\frac{\partial T\left(R_t, \omega\right)}{\partial \omega} < 0, \tag{2.5}$$

$$\frac{\partial T\left(R_t, \omega\right)}{\partial R_t} > 0, \tag{2.6}$$

$$\frac{\partial^2 T\left(R_t, \omega\right)}{\partial \omega \partial R_t} > 0 \quad \forall\ R_t \geq \overline{R} > 0, \tag{2.7}$$

$$\lim_{R_t \rightarrow \infty} T(R_t, \omega) = 1 \quad \forall\ \omega \in \left[\underline{\omega}, \overline{\omega}\right], \tag{2.8}$$

$$\lim_{R_t \to 0} T(R_t, \omega) = 0 \quad \forall\, \omega \in [\underline{\omega}, \overline{\omega}]\,. \tag{2.9}$$

The interpretation goes as follows. (2.5) says that an increase in ω (an epidemic) raises mortality (lowers the survival rate T_t); (2.6) says that mortality falls with the human-capital stock, and rises with population density (recall $H_t/P_t \equiv R_t$); (2.7) says that the effect of an epidemic shock on mortality is smaller if human capital is abundant and/or population not so dense (given that R_t exceeds some positive level $\overline{R}$, explained soon); (2.8) says that if human capital grows at a *faster* sustained rate than population, the mortality rate approaches zero; (2.9) says that if human capital grows at a *slower* sustained rate than population, the mortality rate approaches 100%.

The requirement that $R_t \geq \overline{R}$ in (2.7) is made because when R_t approaches zero (infinity) the marginal negative impact on $T(R_t, \omega)$ of an increase in ω must be negligible, since $T(R_t, \omega)$ then equals zero (unity) for any ω. [See (2.8) and (2.9).] Put differently, the marginal impact of ω-shocks on mortality must be greatest (in absolute terms) at some intermediate level of R_t, i.e. $\overline{R}$.[8] However, for intuitive understanding, we can pretend that $R_t \geq \overline{R}$ always holds. In the parametric example in Section 2.9 the cross-derivative in (2.7) will be positive in the relevant interval.

2.3. Human capital

The production function for human capital takes the form

$$H_{t+1} = A(P_t)\,[L + H_t]\,(\rho v + h_t)\,, \tag{2.10}$$

[8] To see this, the reader may draw a diagram with $\partial T(\cdot)/\partial\omega$ on the vertical axis, and R_t on the horizontal. From (2.8) and (2.9), we know that $\partial T(\cdot)/\partial\omega = 0$ at $R_t = 0$ and when $R_t \to \infty$. In between, $\partial T(\cdot)/\partial\omega < 0$. Say it has some unique minimum, which we can think of as $\overline{R}$. Then, for $R_t \geq \overline{R}$, it must hold that $\partial T(\cdot)/\partial\omega$ is increasing in R_t, i.e., $\partial^2 T(\cdot)/\partial\omega\partial R_t > 0$.

where $A(P_t)$ is a productivity parameter (which depends on population density, but is taken as fixed by each atomistic agent; see below), and ρv measures the direct inheritance of human capital from one generation to the next, where $\rho \in (0,1)$. This is a novelty compared to e.g. Becker et al. (1990): the time spent nursing each child, v (which the parent is effectively forced to undertake), adds to the human capital of the child, but not as efficiently as the time spent educating the child, h_t. We can think of this as capturing the fact that talking and interacting with a child teaches it certain language skills, and a basic acquaintance with customs and social conventions. This automatic spill-over of skills will be driving growth in human capital at early stages of economic development, when $h_t = 0$.

We assume a type of "scale effect": that the effectiveness with which one generation transmits human capital to the next $[A(P_t)]$ is increasing with population density. (Recall that land is fixed, so population size P_t also measures population density.) Our assumption is meant to capture a positive effect on learning, or transmission of skills and knowledge, in regions with shorter geographical distance between people, i.e. *cities*.

A similar assumption is made by Becker et al. (1999), who sketch a model with an urban human capital sector, and a rural goods production sector. In their model, productivity may be decreasing with population size in the consumption goods sector, whereas in the human capital sector population size has a positive effect on productivity.

There are many ways to explicitly model such improved learning, or knowledge spill-overs, in cities. Glaeser (1999) focuses on imitation, whereby agents' skills are improved through random contact with other skilled people. Another explanation for why cities are good for learning could be scale economies in schooling and research, as suggested by Johnson (2000). Whereas we do not model cities or urbanization explicitly, we have similar mechanisms in mind: a more dense population facilitates contacts between people, and the growth of educational institutions, and thereby the transmission of skills from one generation to the next.

Regardless of theoretical explanations, some direct empirical evidence in the so-called New Economic Geography literature does support the basic argument. Glaeser and Mare (1994) show that the wage gap between cities and rural areas is due to a better ability of cities to generate human capital growth, rather than e.g. ability bias.

The importance of small distances between people for the spread of ideas is also illustrated by Jaffe et al. (1993), who find that citations of patents are more common in the same geographical areas (state, country, or metropolitan area) as the original patent. It seems reasonable that geographical distance should have been even more important in historic times, in the absence of telephones, modern transportation, and internet technology.

Leaving a more precise formulation of the relationship between population and human capital transmission for future work, we simply black-box the relationship.

We assume the following:

Assumption 2. $A : \mathbb{R}_{++} \to (\widehat{A}, A^*)$ *satisfies*

$$A'(P_t) > 0, \tag{2.11}$$

$$\lim_{P_t \to \infty} A(P_t) \equiv A^* < \infty, \tag{2.12}$$

$$\lim_{P_t \to 0} A(P_t) \equiv \widehat{A} > 0. \tag{2.13}$$

The interpretation goes as follows. From (2.11) human-capital productivity increases with population density. From (2.12), increasing population density does not raise human-capital productivity indefinitely, but only up to the level A^*. This rules out indefinitely increasing human capital growth rates. From (2.13), as population goes to zero human-capital productivity approaches something positive ($\widehat{A}$) rather than becoming zero. This ensures that the economy cannot "die out."

2.4. Preferences

An adult agent active in period t (referred to as agent t) maximizes a utility function given by

$$U_t = \ln(C_t) + \alpha \ln(B_t T_t) + \alpha\delta \ln(L + H_{t+1}) \tag{2.14}$$

subject to the above budget constraints (2.1), (2.2), (2.3) and (2.10). We assume that $\delta \in (\rho, 1)$ (to guarantee the existence of an interior solution), and that $\alpha > 0$. In (2.14) the second term measures the utility of quantity of surviving children. With logarithmic utility it does not matter if we specify preferences over surviving children $B_t T_t$, or births B_t, since $\ln(B_t T_t) = \ln(B_t) + \ln(T_t)$ and T_t depends only on things which are taken as given by agent t (H_t, P_t, and ω). Thus, $\ln(T_t)$ is just an additive constant in the maximization problem.

The third term measures the utility of quality, which is here simply given by the total human capital of the offspring $(L + H_{t+1})$. This formulation could be interpreted as a reduced form of an old-age security motive for rearing children, as in Nishimura and Zhang (1992), and Lagerlöf (1997).

Alternatively, one could set up a "dynastic" utility function, where the parent cares about the total welfare of the children, but that would complicate matters due to the non-negativity constraint on h_t, which would make the value function non-differentiable. [A related problem has been pointed out by Tamura (1996, Footnote 11).]

Letting the weight on quality be written as $\alpha\delta$ is simply for notational convenience.

Maximizing (2.14) over the number of births (B_t) tells us that

$$B_t = \left(\frac{\alpha}{1+\alpha}\right)\frac{1}{v + h_t}, \tag{2.15}$$

i.e., time spent on children, $(v + h_t)B_t$, is a constant fraction of the unit time endowment, following from the logarithmic utility.

There is an alternative interpretation of mortality in the model. Say there is no child mortality, so all children born also enter the second phase of life. Let T_t instead denote the remaining length of life for agent t as she enters adulthood. This cannot exceed the length of the period, which is normalized to unity, so the restrictions in Assumption 1 still apply. The budget constraint for time in (2.3) is now written $T_t = l_t + (v + h_t)B_t$, and the optimal number of births of agent t (which also equals the number of surviving children) becomes identical to (2.15) except that the unit time-endowment is replaced by T_t. Thus, the number of surviving children is still $B_t T_t$. With this interpretation, falling mortality (increased T_t) expands the amount of time available for the parent, and thus the number of births. This formulation is the same as that of Morand (2000). In the remainder of this paper we shall instead stick to the child-mortality interpretation.

The first-order condition for h_t tells us that

$$h_t \geq \frac{1}{1-\delta}\left[v(\delta - \rho) - \frac{L}{A(P_t)(L + H_t)}\right]. \tag{2.16}$$

The weak inequality in (2.16) follows from the non-negativity constraint on h_t: if the right-hand side of (2.16) is negative $h_t = 0$ (and the inequality is strict); otherwise (2.16) holds with equality. Recall that we are assuming $1 > \delta > \rho$, so that sustained growth in H_t, and/or P_t, must make h_t operative at some stage (see below).

As long as education time is operative (not constrained to zero) it increases with the marginal return to human capital investment $[A(P_t)(L+H_t)]$. This hinges on the biological endowment of human capital (L) being positive, since — with logarithmic utility — the return would otherwise not affect the optimal amount invested. Here, an increase in the return decreases the "discounted" value of L (the value of L expressed in time). This causes a positive relationship between the return to education $[A(P_t)(L+H_t)]$ and optimal education time (h_t).

2.5. Dynamics

An important ingredient in our model is the stochastic epidemic shock ω, but to understand the mechanisms driving the results it is useful to start by thinking of economies with ω being constant over time: either high or low. It turns out that a high-ω country can be stuck in a Malthusian steady state, whereas a low-ω country will follow a path leading to an industrial revolution. Once we have seen that, we can start to think about the dynamics of an economy which randomly switches between ω being high and low. This is examined in Section 2.8.

In period t, P_t adult individuals each have $B_t T_t$ surviving children who become adults in the next period. Thus P_t evolves dynamically according to

$$P_{t+1} = P_t B_t T_t. \tag{2.17}$$

We distinguish between two situations: one where education time is constrained to zero $(h_t = 0)$, and one where it is operative $(h_t > 0)$. The first case prevails during the Malthusian, and (as the dynamics evolve) Post-Malthusian Regimes; the second during the Modern Growth Regime. From (2.16) we see that $h_t > 0$ if

$$A(P_t)(L+H_t) > \frac{L}{v(\delta-\rho)}, \tag{2.18}$$

and $h_t = 0$ otherwise. In Figures 2.1 and 2.2 the non-negativity constraint on h_t is binding in the area below the dashed curve; above it is not binding. The dashed curve need not be shaped the way we have drawn it, but its slope is negative.

2.6. The dynamic system when $h_t = 0$

Consider first the case where education time is constrained to zero. Setting $h_t = 0$ in the human capital production function (2.10), we can write

$$H_{t+1} = A(P_t)\,[L + H_t]\,\rho v. \tag{2.19}$$

The dynamics of P_t is given by (2.17): $P_{t+1} = P_t B_t T_t$. The birth rate B_t is given by (2.15), with $h_t = 0$. The survival rate T_t is given by in (2.4). This gives:

$$P_{t+1} = P_t B_t T_t = P_t \left(\frac{\alpha}{1+\alpha}\right) \frac{T\left(\frac{H_t}{P_t}, \omega\right)}{v}. \tag{2.20}$$

(Recall again that $R_t = H_t/P_t$.) We now have a two-dimensional system of difference equations: (2.19) and (2.20). We analyze it in the phase diagrams in Figures 2.1 and 2.2. In a standard fashion, we begin by deriving the loci along which human capital and population are constant.

2.6.1. The $(\Delta H_t = 0)$-locus when $h_t = 0$

Imposing $H_{t+1} = H_t$ in (2.19) we can write the $(\Delta H_t = 0)$-locus as

$$P_t = A^{-1}\left(\frac{H_t}{\rho v\,[L + H_t]}\right). \tag{2.21}$$

When P_t exceeds the right-hand side of (2.21) H_t is growing over time, and when P_t is less than the right-hand side of (2.21) H_t is falling. When H_t is sufficiently small, it will always grow over time, since $A(P_t)$ is bounded from below [see (2.13)]. This implies that the $(\Delta H_t = 0)$-locus has a horizontal intercept.

Clearly, the $\Delta H_t = 0$ locus slopes upwards since

$$\frac{dP_t}{dH_t}\,\Big|_{\Delta H_t = 0} = \frac{L}{\rho v A'(P_t)}\left(\frac{1}{L + H_t}\right)^2 > 0. \tag{2.22}$$

Figures 2.1 and 2.2 show the $(\Delta H_t = 0)$-locus sloping upwards (below the dashed curve, where $h_t = 0$), and with a horizontal intercept. (We have drawn it concave, but it need not be.) As shown by the arrows, human capital is growing above the $(\Delta H_t = 0)$-locus, and falling below it.

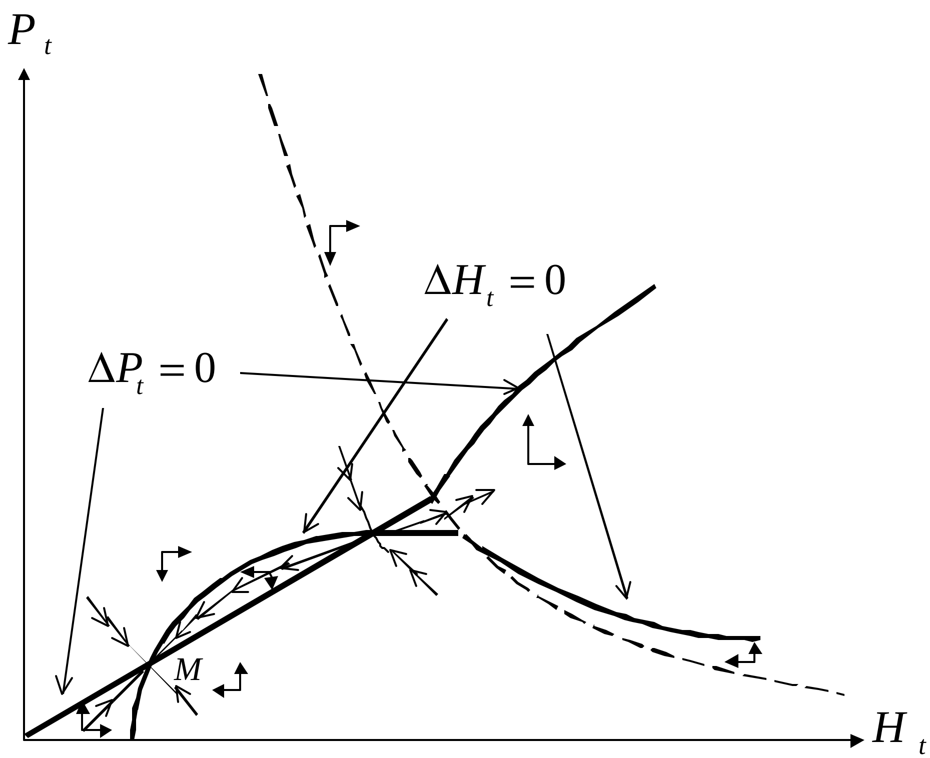

Figure 2.1: Below the dashed line $h_t = 0$, and above $h_t > 0$. This diagram illustrates the case when ω is high (a severe epidemic shock). There is a locally stable Malthusian steady-state equilibrium (point M), as well as a path which eventually leads to an industrial revolution when the $(h_t = 0)$-locus is intersected.

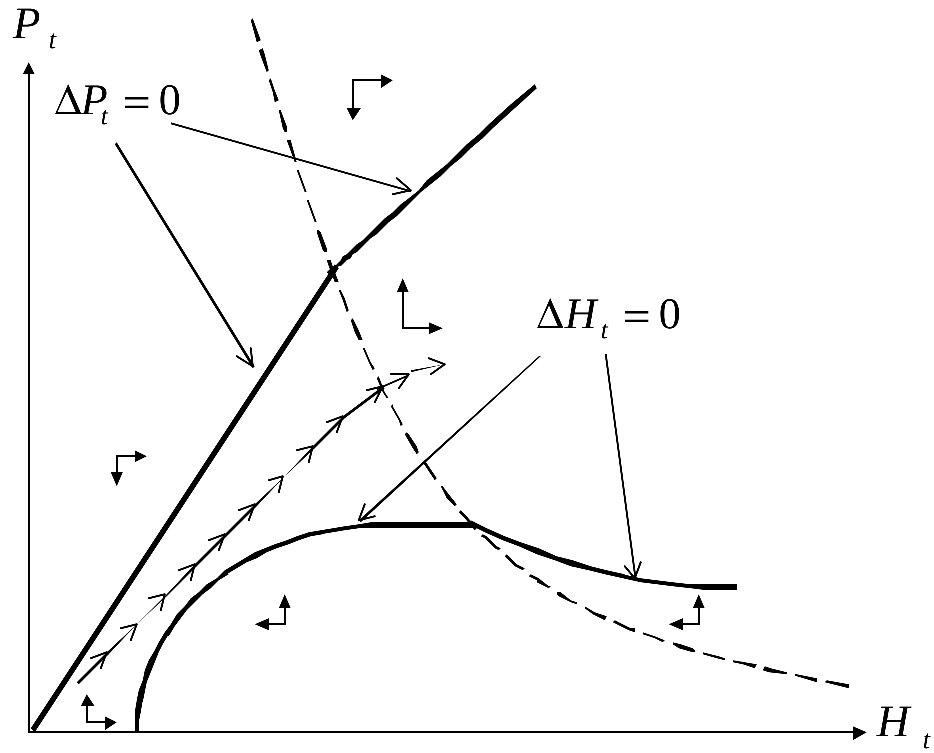

Figure 2.2: Below the dashed line $h_t = 0$, and above $h_t > 0$. This diagram illustrates the case when ω is low (no, or small, epidemic shock). The economy will always follow a path which intersects the $(h_t = 0)$-locus at some stage, leading to an industrial revolution.

2.6.2. The $(\Delta P_t = 0)$-locus when $h_t = 0$

At a given epidemic shock ω, (2.20) tells us that the population is constant when the ratio $R_t = H_t/P_t$ is such that $[\alpha/(1+\alpha)]\,T\,(R_t, \omega)\,/v = 1$. This gives the $(\Delta P_t = 0)$-locus as a straight line (below the dashed curve, where $h_t = 0$), as drawn in Figures 2.1 and 2.2. [The linearity of the $(\Delta P_t = 0)$-locus comes from our specification of the survival function $T(R_t, \omega)$, which depends only on $R_t = H_t/P_t$. With a more general form the $(\Delta P_t = 0)$-locus would still be sloping upwards.] As shown by the arrows, population grows in the region below the $(\Delta P_t = 0)$-locus, and falls above it.

The slope of the $(\Delta P_t = 0)$-locus depends on the realized shock ω. A high ω (a severe epidemic) implies that the ratio of human capital to population must be relatively high for population to be constant, compared to when ω is low. Therefore, a high ω makes the $(\Delta P_t = 0)$-locus flatter. [Since we have P_t on the vertical axis, the slope equals P_t/H_t, i.e., the inverse of $R_t = H_t/P_t$.] Another way of seeing the same thing is to note that a rise in ω shrinks the region in which population is growing.

2.7. The dynamic system when $h_t > 0$

Above the dashed curve in Figures 2.1 and 2.2, education time is operative, and the dynamics of the human capital stock is derived from the human capital production function (2.10), substituting for the optimal h_t, as given by (2.16) holding with equality. After some algebra, this gives

$$H_{t+1} = \frac{v\delta(1-\rho)A(P_t)\,[L + H_t] - L}{1-\delta}. \tag{2.23}$$

Similarly, after substituting the expression for optimal h_t in (2.16) into the expressions for the birth rate in (2.15), some algebra tells us that

$$B_t = \left(\frac{\alpha}{1+\alpha}\right)\frac{(1-\delta)A(P_t)\left[L+H_t\right]}{v(1-\rho)A(P_t)\left[L+H_t\right]-L} \equiv B(H_t,P_t). \tag{2.24}$$

From (2.24) and (2.17) the population dynamics can be written

$$P_{t+1} = P_t B(H_t,P_t)T(H_t/P_t,\omega), \tag{2.25}$$

where $B(H_t,P_t)$ is defined in (2.24).

We now have a new two-dimensional system of difference equations: (2.23) and (2.25). To analyze the dynamics, we start by deriving the loci along which P_t and H_t are constant.

2.7.1. The ($\Delta H_t = 0$)-locus when $h_t > 0$

Setting $H_{t+1} = H_t$ in (2.23), and solving for P_t, we can write the ($\Delta H_t = 0$)-locus as

$$P_t = A^{-1}\left\{\frac{L+(1-\delta)H_t}{v\delta(1-\rho)\left[L+H_t\right]}\right\}. \tag{2.26}$$

The slope of this locus is negative:

$$\frac{dP_t}{dH_t}\Big|_{\Delta H_t=0} = \frac{-1}{A'(P_t)}\left(\frac{1}{v\delta(1-\rho)\left[L+H_t\right]}\right)^2 \delta^2 v(1-\rho)L < 0, \tag{2.27}$$

as we have drawn it in the diagrams in Figures 2.1 and 2.2, where we also note that the locus converges to $A^{-1}\left(\frac{(1-\delta)}{v\delta(1-\rho)}\right)$ as H_t goes to infinity (which may, or may not, be positive, as we have drawn it). Positions above this locus imply growing levels of human capital, and positions below imply falling levels.

It is easy to verify that the ($\Delta H_t = 0$)-locus in the region where $h_t > 0$, must coincide with the corresponding locus when h_t is constrained to zero, at exactly the point where these loci intersect the dashed curve in Figures 2.1 and 2.2. If the unconstrained choice of h_t equals exactly zero, the economy behaves as if $h_t \geq 0$ was binding.

2.7.2. The $(\Delta P_t = 0)$-locus when $h_t > 0$

From (2.25) we see that population is constant when

$$B(H_t, P_t)T(H_t/P_t, \omega) = 1. \tag{2.28}$$

This implicitly defines the $(\Delta P_t = 0)$-locus. First note that it must coincide with the corresponding $(\Delta P_t = 0)$-locus for the case when $h_t = 0$ exactly on the dashed line, where the constraint $h_t \geq 0$ is on the border of being binding. In other respects, the precise shape of the $(\Delta P_t = 0)$-locus depends on the form of the functions $A(P_t)$ and $T(H_t/P_t, \omega)$. The way it is drawn in Figures 2.1 and 2.2 is one example.

However, we can say something about how this locus slopes asymptotically. Letting H_t go to infinity in (2.24) we get

$$\lim_{H_t \to \infty} B(H_t, P_t) = \left(\frac{\alpha}{1+\alpha}\right) \frac{1-\delta}{v(1-\rho)} \equiv B^*. \tag{2.29}$$

Sustained growth in human capital will eventually cause the birth rate to converge to B^* in (2.29). [Note that $A(P_t)$ is bounded from below by $\widehat{A}$ in (2.13), so $A(P_t)\,[L + H_t]$ will go to infinity if H_t does.] Another way of deriving B^* is to let H_t go to infinity in the expression for education time in (2.16), and then substitute the resulting expression into the expression for the birth rate in (2.15). Asymptotically, the slope of the $(\Delta P_t = 0)$-locus is thus defined by the ratio P_t/H_t which makes $T(H_t/P_t, \omega) = 1/B^*$. Since $T(H_t/P_t, \omega) < 1$, B^* must exceed one for there to exist some asymptotically constant ratio H_t/P_t along which $\Delta P_t = 0$. We therefore assume:

Assumption 3. $\left(\frac{\alpha}{1+\alpha}\right) \frac{1-\delta}{v(1-\rho)} \equiv B^* \geq 1.$

Section 3 discusses the implications of $B^* < 1$.

2.8. The full transition through the three regimes

To understand the dynamics, we can think of an economy which has a constant level of ω over time: either high (constantly hit by severe epidemics), or low (with epidemics virtually absent). Consider first the phase diagram in Figure 2.1. This illustrates the situation of an economy where ω is high. In the lower part of the diagram (below the dashed line, where $h_t = 0$), the $(\Delta P_t = 0)$-locus intersects the $(\Delta H_t = 0)$-locus twice. The low intersection, marked M, constitutes a locally stable steady-state equilibrium: if the economy is situated near M it will converge to M (of course, given that ω is kept constant). This is a Malthusian equilibrium, in the sense of e.g. Galor and Weil (1999, 2000). The mechanism which equilibrates the economy is the increased mortality effect of a larger population: if population density increases, so that the economy ends up above the $(\Delta P_t = 0)$-locus, the mortality rate goes up, and the population is thus reduced until the economy goes back to M. In Figure 2.1 there is also a path with growing levels of population and human capital. An economy following that path will eventually intersect the dashed line. That is where education time (h_t) becomes operative, so that human capital starts growing faster and the birth rate falls: the economy experiences an industrial revolution and a demographic transition.

Consider next the case where epidemic shocks are virtually absent (ω is small). This is illustrated in Figure 2.2. Here the $(\Delta P_t = 0)$-locus never intersects the $(\Delta H_t = 0)$-locus. There is no Malthusian trap. The only path is the one leading to an industrial revolution.

Now think of an economy which randomly switches between ω being high and low. Having long been in a state with frequent epidemics (high ω) the economy is situated somewhere near the Malthusian trap, point M in Figure 2.1. Then ω suddenly falls. The $(\Delta P_t = 0)$-locus becomes steeper, as in Figure 2.2. This enables population and human capital to start growing. As more human capital builds up mortality falls and population expands. Higher population density raises human capital productivity $[A(P_t)]$, and so the cycle feeds back into itself. In the absence of any further epidemics, the economy would be heading safely for an industrial revolution as it intersects the dashed line. But consider what happens if the economy is hit by a new epidemic wave (ω rises again). The ($\Delta P_t = 0$)-locus then becomes flatter again, as in Figure 2.1.

If the economy has reached a position beyond the threshold saddle path in Figure 2.1 it will not contract back to the Malthusian trap. As it continues to accumulate more human capital it can become safe from even the worst epidemic shocks, as the impact of these shocks falls with the level of human capital. (This hinges upon human capital growing faster than population; see below.) If the economy has *not* reached a safe position when the new epidemic comes, it will contract back to the Malthusian trap. The mortality increase could be extra severe, due to the higher population density [see (2.7)]. At the same time, the human capital accumulated will help mitigate this effect.

For an industrial revolution (and the associated demographic transition) to take place the economy must be spared from epidemic shocks long enough. This explains how it can be stuck in a Malthusian trap for very long, and then suddenly escape. Once it escapes, it will first experience population and human capital growing simultaneously, as in Figure 2.2 in the area below the dashed line. This is the *Post-Malthusian Regime* of Galor and Weil (1999, 2000). Note that in the Post-Malthusian Regime population growth comes from falling mortality rates. Birth rates remain unchanged, since education time (h_t) is constrained to zero [see (2.15)].

When the trajectory intersects the dashed line the economy transits into what Galor and Weil (1999, 2000) call the *Modern Growth Regime*. This is characterized by an increased growth rate of human capital, and a fall in birth rates and population growth rates. Both effects come from education time becoming operative (h_t becoming positive). Human capital grows faster because education is an input in human capital production [see (2.10)]. At the same time, increased education time makes children more expensive, which induces a fall in birth rates [see (2.15) again]. This is a standard quality-quantity switch.

The positive growth effects of education time becoming operative has some empirical support. In Britain education started to increase at the same time as growth rates did. The average years of schooling started to increase with the cohort born around 1800, and was later followed by sharp increases in the proportion of the male population with a university degree, starting with the cohort born in the 1880's. [See Matthew, Feinstein and Odling-Smee (1982, Ch. 4).]

As long as human capital grows at a faster rate than population our model predicts that the impact of the epidemic shocks tend to vanish over time, as H_t/P_t goes to infinity [see (2.8)]. This is consistent with data: during the demographic transition, not only the level of death rates fell, but also their volatility. Note that, in our model and probably in the real world too, the distribution of the epidemic shocks is constant over time. Only the way in which the mortality rates respond to the shocks change.

Using the production function for human capital (2.10), and letting population go to infinity in (2.12), the (gross) growth rate of H_t in the Modern Growth Regime becomes

$$\gamma_H^* \equiv \frac{\delta v A^*(1-\rho)}{1-\delta}. \tag{2.30}$$

The growth rate of population is given by $B(H_t, P_t)T(H_t/P_t, \omega)$ [see (2.25)]. If H_t/P_t goes to infinity, and the mortality rate $1 - T(H_t/P_t, \omega)$ goes to zero, the population growth rate will asymptotically become identical the birth rate B^* in (2.29). Does it actually hold that H_t/P_t goes to infinity? To make this conjecture true we must assume that $\gamma_H^* > B^*$, i.e.,

Assumption 4. $\frac{\delta v A^*(1-\rho)}{1-\delta} > \left(\frac{\alpha}{1+\alpha}\right)\frac{1-\delta}{v(1-\rho)}$.

Note that Assumptions 3 and 4 together imply that human capital exhibits sustained growth ($\gamma_H^* > 1$).

We need to make sure that all the assumptions we have made so far can hold simultaneously. To do this the next section sets up functional forms for $T(\cdot)$ and $A(\cdot)$ that satisfy Assumptions 1 and 2. We then calibrate the growth rates of H_t and P_t to realistic levels for modern economies (which also satisfy Assumptions 3 and 4), and simulate the model with epidemic shocks.

2.9. Simulations

We let the parametric form for $A(P_t)$ be

$$A(P_t) = A^* - \widetilde{A} + \widetilde{A}\left(\frac{P_t}{\eta + P_t}\right), \tag{2.31}$$

where $\eta > 0$ is an exogenous parameter, and $A^* - \widetilde{A} > 0$ corresponds to $\widehat{A}$ in (2.13). Clearly, (2.31) satisfies Assumption 2.

We let the parametric form for $T(H_t/P_t, \omega)$ be given by

$$T(H_t/P_t, \omega) = \left(\frac{H_t}{\omega P_t + H_t}\right)^{\sigma}, \tag{2.32}$$

where $\omega > 0$ is the stochastic epidemic shock, uniformly distributed on $[\underline{\omega}, \overline{\omega}]$, and $\sigma > 0$ is an exogenous parameter. Clearly, (2.32) satisfies Assumption 1.

As numerical values we choose those listed in Table 2.1. These are chosen to satisfy all assumptions made so far. A number of further restrictions were imposed to shape the ($\Delta P_t = 0$)- and ($\Delta H_t = 0$)-loci the way they are drawn in Figures 2.1 and 2.2. The values of $\underline{\omega}$ and $\overline{\omega}$ were chosen to make an escape from the Malthusian trap possible for $\underline{\omega}$, but not for $\overline{\omega}$.

We also set the parameter values to give growth rates of output per capita and population in the Modern Growth Regime in the vicinity of what we see in most modern in economies today. Letting each period correspond to 35 years,

ρ	δ	α	v	D	L	A^*	$\widetilde{A}$	$\underline{\omega}$	$\overline{\omega}$	σ	η
0.9982	0.9987	0.0389	0.025	1	1	82.11	44.16	0.0496	0.0679	1	10000

Table 2.1: Parameter values

human capital growth (which can be interpreted as per-capita income growth) is calibrated to 3% per annum. The population growth rate is calibrated to 0.25% per annum.[9]

The simulated annual growth rates of human capital and population are shown in Figure 2.3. It replicates the three regimes of Galor and Weil (1999, 2000). The first 120-130 generations experience low growth rates of both population and human capital, with the population growth rate being very volatile. This is the Malthusian Regime. Then both population and human capital simultaneously start growing faster. Population growth also starts becoming less volatile. This is the Post-Malthusian Regime. Then the economy enters its final phase, the Modern Growth Regime: population growth drops to 0.25%, and human capital growth spurts to 3% (as we have calibrated them). As seen, this happens simultaneously with an increase in education time (h_t).

Figure 2.4 shows the timing of the different components of the demographic transition: the annual birth and death rates.[10] First mortality falls, while fertility stays constant; then fertility falls, and mortality stabilizes. The phase where death rates fall below birth rates is associated with the high population growth phase in Figure 2.3. The picture is clearly consistent with the stylized facts about the demographic transition described by e.g. Livi-Bacci (1997, Ch. 4).

One striking feature is the simultaneous fall in the volatility of mortality. This

[9] Using the numerical values in Table 2.1 and the expression for human-capital growth in the Modern Growth Regime in (2.30) gives (approximately) $(1.03)^{35}$, and similarly for the population growth rate.

[10] The annual death rate is computed as $1 - T_t^{1/35}$ (i.e., one minus the annual survival rate), and the annual birth rate as $B_t^{1/35} - 1$.

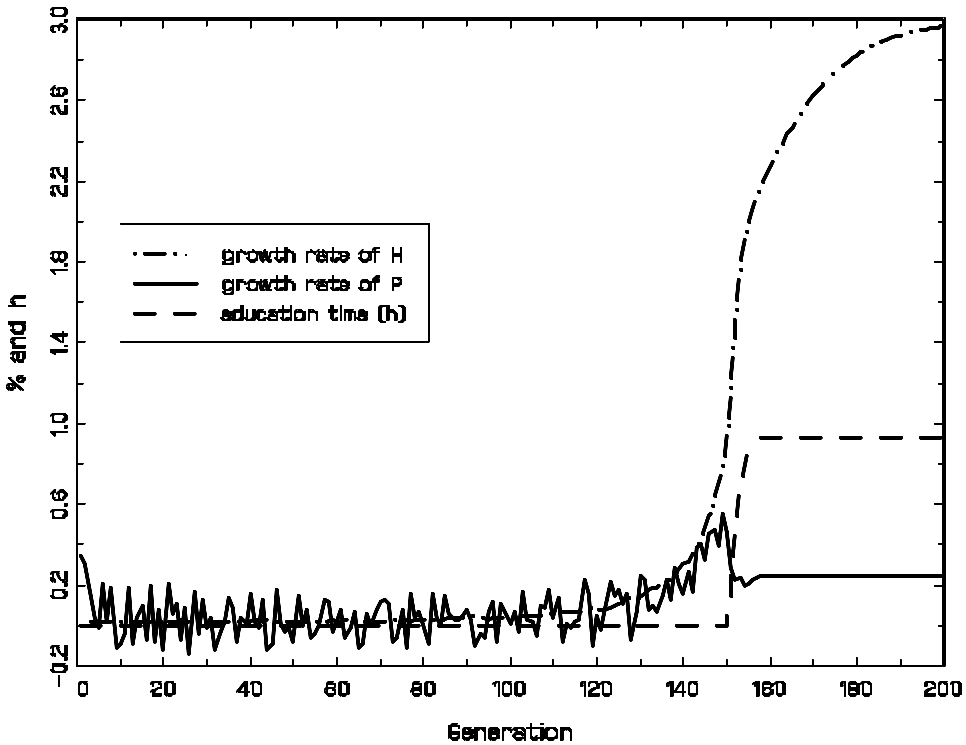

Figure 2.3: The three regimes simulated.

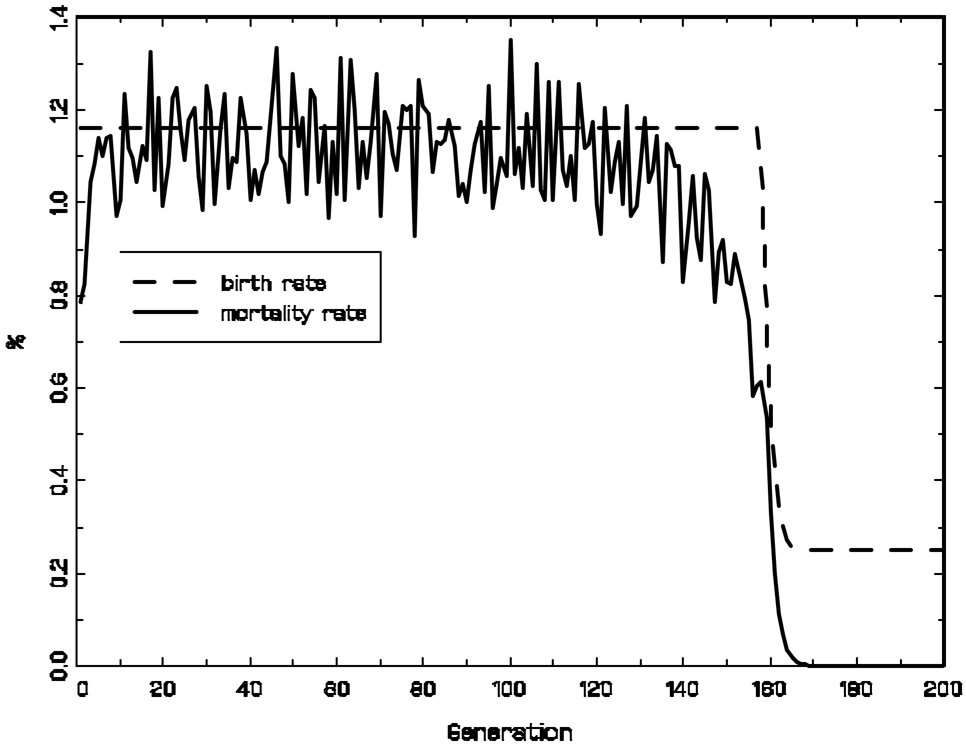

Figure 2.4: Simulated birth and death rates.

is what we see in the data for e.g. Sweden, as shown in Figure 1.1.[11] It also holds for 18th-century continental Europe, where death rates increased dramatically in periods of epidemics and famines. Towards the second half of the 18th century, as the industrial revolution took place, the mortality peaks flattened out. For a further discussion of these facts, see Livi-Bacci (1997, p. 118), Maddison (1982, pp. 50-52), and Easterlin (1996, pp. 7-9).

The levels of the birth and death rates are considerably lower in our simulation, compared to the Swedish figures in Figure 1.1. However, the data in Figure 1.1 refer to the total number of births and deaths, divided by total population. The death rates for children aged between 5 and 15 were much lower, around 0.6—1.3% in the 1750's. Infants and very old people had higher death rates, which is what pulls the numbers up in Figure 1.1. [See Statistics Sweden (1999).] From the view point of our model, 5—15 is probably the age group we should look at, since that is when schooling decisions are made.

3. Conclusions

According to Galor and Weil (1999, 2000), on its way from stagnation to modern growth, the Western World has passed through three stages: a *Malthusian Regime*, with almost constant levels of population and per-capita income, and a positive relationship between per-capita income and population growth; a *Post-Malthusian Regime*, with positive growth rates of both population and per-capita income, and still a positive relationship between per-capita income and population growth; and finally a *Modern Growth Regime*, with even higher growth rates of per-capita income, but lower growth rates of population, and the relationship between per-capita income and population growth being negative. Galor and Weil (1999, 2000) call for a model which can explain all three regimes in a unified framework. That

[11] However, in contrast to what we see in the Swedish data, the birth rate is flat in our model. See the discussion in Section 3 below.

is what our model does.

The following assumptions drive our results: we allow advances in medical skills to reduce the impact on mortality of epidemics; we let population density have the opposite effect, i.e. increasing the impact on mortality of epidemic shocks; and we allow population density to have a positive effect on productivity in the human capital sector. We also cite empirical evidence supporting these assumptions.

Our model is capable of generating the transition through all the three regimes endogenously. Through all regimes, the epidemic shocks hit the economy at the same rate. However, with rising human capital the impact of them is mitigated. The results are most easily seen in the simulated example in Figures 2.3 and 2.4. For many generations the economy is stuck in a Malthusian equilibrium with volatile and high death rates. Sooner or later it experiences a phase of relatively mild epidemics, so that mortality declines, enabling population and human capital to simultaneously start growing. This is the Post-Malthusian Regime. Once human capital growth has taken off, the impact of further epidemics dies out. Then at some stage a quality-quantity switch makes parents rear fewer children, and start investing in their children's education. This also triggers faster growth in human capital. The economy enters the Modern Growth Regime.

Let us go back to the questions posed in the beginning of this paper. Is the industrial revolution inevitable? In our model: yes and no. Given that it is *feasible* to escape the Malthusian trap by experiencing sufficiently many periods with sufficiently mild epidemic shocks, the economy must escape at some stage. It is just a matter of time and luck; the right conditions will arrive in finite time. The industrial revolution is in that sense inevitable (as is the Post-Malthusian Regime preceding it). However, when simulating the model, we find that it can take very many generations to get a growth take-off — even when we choose parameter values that make this theoretically certain at some stage.

So why did the industrial revolution start in Europe, and not the Americas or

Australia? According to Jared Diamond (1999), in his best-seller "Guns, Germs, and Steel," epidemics are indeed part of the answer. [See also McNeill (1976).] When colonizing these lands, Europeans brought with them crowd diseases against which the native inhabitants were less resistant. The reason why Europeans had developed so many more crowd diseases is in part its frequent trade contacts across the Eurasian continent (see below). In comparison, the densely populated areas in the Americas (the Andes, Mesoamerica, and the Mississippi Valley) were not connected at all. But more important, according to Diamond, was the fact that the Americas and Australia had fewer domesticated animals than Europe. It is from such livestock which many deadly microbes have jumped the species barrier onto humans. The low number of domesticated animals in the Americas and Australia is in turn due to a lack of animals suited for domestication. In terms of our model, one could say that the native Americans were originally quite lucky. Thanks to an exogenously smaller endowment of animals to domesticate they were spared from epidemics. Of course, when the Europeans arrived the situation changed, as the natives lacked immunity against the new European diseases. However, had Columbus never set sail in the late 15th century maybe we could have seen an industrial revolution in the Americas.

But why did the industrial revolution not start in China or India? From the point of view of our model we could think of Eurasia as one economy, especially from around 1300, when Eurasian population density was sufficiently high to facilitate contacts and travels across the continent (by land, initiated by the Mongol Empire, and by sea, initiated by the Europeans). This increased the spread of ideas (which can be thought of as the endogenous effect from population density on human capital productivity in our model). However, it also turned the continent into one gigantic germ pool, easing the spread and worsening the impact of epidemics (which can be thought of as the endogenous effect from population density on mortality in our model). The Black Death in Europe 1346-50 is the most well-known example. Still, there were some local epidemiological differences across

Eurasia. India was worse off than China and Europe due to its climate being more suitable for diseases to develop and spread. [See McNeill (1976, pp. 106-107).] This may have made the necessary population and human-capital growth take-off harder to sustain. China did not have a disadvantageous climate, but may simply have been hit by worse shocks. China's population was almost constant throughout the 17th century, associated with a documented increase in epidemics. During the same time period, Europe (in particular Britain) experienced comparatively mild epidemics, and higher population growth. [See McNeill (1976; pp. 228-229 for China and p. 142 for Britain).]

The birth rate in our model is constant (up until the increase in education time), whereas in the data it is volatile (though not as volatile as the mortality rate; see Figure 1.1). One way to extend our model to include birth rate volatility could be to allow for shocks to output (e.g. the parameter D in our model). These could correspond to famines and bad harvests. One would then also have to specify a consumption goods cost of having children, because with only a time cost and logarithmic utility (as in our present setting) any shock to output would be neutral to birth rates, since the income and substitution effects cancel out. A 10% fall in income means a 10% fall in the time cost of children, and would leave the birth rate unchanged [see (2.15)]. We believe formulating a model which can explain both birth and death rate volatility should be an interesting task for future research.

Assumption 3 ensures that the birth rate asymptotically approaches something positive as human capital grows. With a zero mortality rate a positive birth rate is required for positive sustained population growth. But birth rates are very low in many European countries today, even below replacement levels. What happens if we allow for such low birth rates in our model? In terms of the phase diagrams, it can be seen that if B^* falls below unity, the $(\Delta P_t = 0)$-locus will eventually start sloping downwards and intersect the horizontal axis. This would either lead to population dying out, while an "imaginary" human capital stock still kept growing [recall that human-capital productivity goes to $\widehat{A}$ when P_t goes to zero; see (2.13)]; or it would make H_t and P_t stabilize at some non-growing levels. The first scenario we should think has no empirical relevance; the second contradicts our belief that the last few century's sustained growth will never die out. We impose Assumption 3 to ensure that none of this happens, and let readers 1000 years from now be the judges of whether, or not, this was a good guess.

References

[1] Barro, R.J., and G.S. Becker, 1989, Fertility choice in a model of economic growth, Econometrica 57, 481-501.

[2] Becker, G.S., and R.J. Barro, A reformulation of the economic theory of fertility, Quarterly Journal of Economics 103, 1-25.

[3] Becker, G.S., E.L. Glaeser, and K.M. Murphy, 1999, Population and Economic Growth, American Economic Review 89, 145-149.

[4] Becker, G.S., K.M. Murphy and R. Tamura, 1990, Human capital, fertility, and economic growth, Journal of Political Economy 5, 12-37.

[5] Diamond, J., 1999, Guns, germs, and steel, W.W. Norton & Company, New York.

[6] Easterlin, R.A., 1996, Growth triumphant, The University of Michigan Press, Ann Arbor.

[7] Ehrlich, I. and F.T. Lui, 1991, Intergenerational trade, longevity, and economic growth, Journal of Political Economy 99, 1029-1059.

[8] Galor, O., and O. Moav, 2000a, From physical to human capital accumulation: inequality and the process of development, manuscript Brown University and Hebrew University.

[9] − − −, 2000b, Natural selection and the origin of economic growth, manuscript Brown University and Hebrew University.

[10] Galor, O., and D. Weil, 1999, From Malthusian stagnation to modern growth, American Economic Review 89, 150-154.

[11] −−−, 2000, Population, technology, and growth: from the Malthusian regime to the demographic transition and beyond, American Economic Review 90, 806-828.

[12] Glaeser, E.L., 1999, Learning in cities, Journal of Urban Economics 46, 254-277.

[13] Glaeser, E.L., and D.C. Mare, 1994, Cities and skills, Hoover Institution Working Paper E-94-11.

[14] Hansen, G.D., and E.C. Prescott, 1998, Malthus to Solow, NBER Working paper 6858.

[15] Haines, M.R., 1995, Disease and health control through the ages, in: J.L Simon (ed.), The state of humanity, Blackwell, Oxford.

[16] Jaffe, A.B., M. Trajtenberg, and R. Henderson, 1993, Geographic localization of knowledge spillovers as evidences by patent citations, Quarterly Journal of Economics 108, 577-598.

[17] Johnson, G., 2000, Population, food, and knowledge, American Economic Review 90, 1-14.

[18] Jones, C., 1999, Was the industrial revolution inevitable? Economic growth over the very long run, manuscript, Stanford University.

[19] Kalemli-Ozcan, S., 2000a, A stochastic model of mortality, fertility, and human capital investment, manuscript, University of Houston.

[20] ———, 2000b, Does mortality decline promote economic growth?, manuscript, University of Houston.

[21] 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.

[22] Kremer, M., 1993, Population growth and technological change: one million B.C. to 1990, Quarterly Journal of Economics 108, 681-716.

[23] Kunitz, S.J., 1983, Speculations on the European mortality decline, Economic History Review 36, 349-364.

[24] Kuznets, S. 1971, Economic growth of nations: total output and production structure, Belknap Press, Cambridge, Massachusetts.

[25] Lagerlöf, N.-P., 1997, Endogenous fertility and the old-age security hypothesis: a note, Journal of Public Economics 64, 279-286.

[26] Livi-Bacci, M., 1997, A concise history of world population, second edition, Blackwell, Oxford.

[27] Lucas, R.E., 1998, The industrial revolution: past and future, book manuscript prepared for the Kuznets Lectures.

[28] Maddison, A., 1982, Phases of capitalist development, Oxford University Press, Oxford.

[29] Matthews, R.C.O., C.H. Feinstein, and J.C. Odling-Smee, 1982, British Economic Growth 1856-1973, Clarendon Press, Oxford.

[30] McNeill, W.H., 1976, Plagues and peoples, Anchor press, Doublebay.

[31] Morand, O.F., 2000, Health and the process of economic development, manuscript, University of Connecticut.

[32] −−−, 2001, Population, technology, and growth, manuscript, University of Connecticut.

[33] Nishimura, K. and J. Zhang, 1992, Pay-as-you-go public pensions with endogenous fertility, Journal of Public Economics 48, 239-258.

[34] Pörtner, C.C., 2001, Children as insurance, Journal of Population Economics 14.

[35] Pritchett, L. 1997, Divergence, big time, Journal of Economic Perspectives 11, 3-17.

[36] Statistics Sweden, 1999, Population development in Sweden in a 250-year perspective, Demografiska Rapporter 1999:2.

[37] Tamura, R., 1996, From decay to growth: a demographic transition to economic growth, Journal of Economic Dynamics and Control 20, 1237-1261.

[38] −−−, 2000, Human capital and the switch from agriculture to industry, manuscript, Clemson University.

[39] −−−, 2001, Human capital and economic development, manuscript, Clemson University.

Former Working Papers:

- **Arbetsrapport/Institutet för Framtidsstudier; 2000:1**
  Malmberg, Bo & Lena Sommestad. *Heavy trends in global developments. Idea platform for MISTRA's future strategy.*
- **Arbetsrapport/Institutet för Framtidsstudier; 2000:2**
  Malmberg, Bo & Lena Sommestad. *Tunga trender i den globala utvecklingen.* Uppdrag för Stiftelsen för Miljöstrategisk forskning (MISTRA).
- **Arbetsrapport/Institutet för Framtidsstudier; 2000:3**
  Lee, Ronald, Mason, Andrew & Timothy Miller. *From Transfers to Individual Responsibility: Implications for Savings and Capital Accumulation in Taiwan and the United States.*
- **Arbetsrapport/Institutet för Framtidsstudier; 2000:4**
  Mayer, David. *On the Role of Health in the Economic and Demographic Dynamics of Brazil, 1980-1995.*
- **Arbetsrapport/Institutet för Framtidsstudier; 2000:5**
  Galor, Oded & Omer Moav. *Natural Selection and the Origin of Economic Growth.*
- **Arbetsrapport/Institutet för Framtidsstudier; 2000:6**
  Malmberg, Bo & Lena Sommestad. *Four Phases in the Demographic Transition, Implications for Economic and Social Devleopment in Sweden, 1820-2000.*

ISSN 1652-120X
ISBN 91-89655-14-1