

# Growth and Longevity from the Industrial Revolution to the Future of an Aging Society

David de la Croix

Dept. of econ. & CORE,  
Univ. cath. Louvain

Thomas Lindh

Institute for Future  
Studies

Bo Malmberg

Institute for Future  
Studies & dept. of Human  
Geography, Stockholm  
University

June 2006

---

<sup>1</sup>De la Croix acknowledges financial support from the Belgian French speaking community (Grant ARC 03/08-235 “New macroeconomic approaches to the development problem”) and the Belgian Federal Government (Grant PAI P5/21, “Equilibrium theory and optimization for public policy and industry regulation”).

Contact addresses: Université catholique de Louvain, Department of Economics, Place Montesquieu 3, B-1348 Louvain-la-Neuve, Belgium. Institute for Futures Studies, Box 591, SE-101 31 Stockholm, Sweden.

E-mail: [delacroix@ires.ucl.ac.be](mailto:delacroix@ires.ucl.ac.be), [bo.malmberg@framtidsstudier.se](mailto:bo.malmberg@framtidsstudier.se), [thomas.lindh@framtidsstudier.se](mailto:thomas.lindh@framtidsstudier.se).

We are grateful to Fati Shadman and participants to seminars at IRES, European Central Bank, PAA 2006 (Population Association of America) and International Symposium for Forecasters, for useful comments on an earlier draft.

## Abstract

Aging of the population will affect the growth path of all countries. To assess the historical and future importance of this claim we use two popular approaches and evaluate their merits and disadvantages by confronting them to Swedish data. We first simulate an endogenous growth model with human capital linking demographic changes and income growth. Rising longevity increases the incentive to get education, which in turn has ever-lasting effects on growth through a human capital externality. Secondly, we consider a reduced-form statistical model based on the demographic dividend literature. Assuming that there is a common DGP guiding growth through the demographic transition, we use an estimate from post-war global data to backcast the Swedish historical GDP growth. Comparing the two approaches, encompassing tests show that each of them contains independent information on the Swedish growth path, suggesting that there is a benefit from combining them for long-term forecasting.

*Journal of Economic Literature Classification numbers:* Demographic Transition, Life Expectancy, Education, Income Growth.

*Key Words:* J11, O41, I20, N33.

CORE Discussion paper 2006/64

# Introduction

Population aging is an inescapable consequence of lower mortality and fertility. This process will affect all countries on earth, starting with the most developed ones. The consequences of aging for future income growth are of prime importance for the conduct of economic policy, but they are still largely unknown. To shed light on this issue we investigate whether demographically based models can be used to account for past income growth and to forecast future economic growth using population projections.

Although demographic variables depend on economic development, they are still to a large extent predetermined. This demographic inertia is exploited in demographic projections to yield forecasts that are substantially more reliable than any projection of economic variables. Therefore, using demographic projections as independent variables to forecast economic growth is a promising avenue.

There are two traditions to analyze the interaction between demographic trends and long-run growth prospects. The first one consists in building theoretical models to achieve a consistent view of the mechanisms that can drive the growth process, either qualitatively (Galor and Weil (2000), Lagerlöf (2003)) or quantitatively (Boucekkine, de la Croix, and Licandro (2003)). The second tradition has an agnostic view of the mechanisms actually in place; it analyzes the empirical relationships between demographic variables and growth in income per capita in recent data, and extrapolates growth rates on the basis of demographic projections (see Bloom and Williamson (1998)). Both approaches, however, share the idea that a decline in mortality may serve as a trigger for modern economic growth.

We believe that a good model for long-term forecasting should be able to shed light on the history of growth since the Malthusian stagnation to modern growth, through the industrial revolution and the demographic transition. We will therefore confront both approaches to Swedish long-term data. Looking at Sweden is particularly relevant, not only because the Swedish demographic transition is very typical, but also because excellent demographic data are available from the mid 18th century and onwards. Estimates of per capita GDP stretches back to the 18th century too.

In a first step, we use these long-term data to calibrate a demographically based growth model so as to reproduce the take-off process and the rise in growth rates from stagnation prior to the eighteenth century to 2% growth in the twentieth century. The main mechanisms at work are that rises in life expectancy increase the incentive to get education, which in turn has ever-lasting effects on growth through a human capital externality and there is a scale effect from active population on growth.

In the second step, we consider a demographically-based statistical growth model estimated on global post-war growth data to study whether it can account for the long-term growth process that can be observed in the Swedish data. The global model estimates show a drift in the most productive activity period with life expectancy. The peak productivity shifts from around 30 years of age when life expectancy is low to an age around 50 for actual life expectancies in developed and emerging economies. The model is then used to backcast Swedish economic growth back to 1750 making use of the long-term

demographic data that we have available. The backcast shows that the statistical model can account not only for recent changes in per capita income but also for the long-term process of Swedish economic development since the mid 19th century.

Both approaches are used to forecast income growth in Sweden over the period 2000-2050. Given this purpose, we treat demographic variables as exogenous throughout this paper. An assessment of the performance of a combination of the forecasts show that this leads to smaller expected forecast errors, one reason being that any model specific bias is corrected by the combination (Diebold and Lopez 1996).

Our conclusion is that the Swedish case provides a valuable test-ground that allows an evaluation of both theoretical and statistical approaches to demographically based models of long-term economic growth. Our analysis highlights that the correlation between mortality decline, age structure change and income growth is not just a statistical artifact in recent data. This relation also conforms with the theoretically expected effect of mortality change on human capital accumulation and productivity. Our results suggest that a universal and highly regular process connects demographic change and economic development. Provided this connection remains intact in the future reliable methods for long-term forecasts of GDP using demographic projections can be developed.

The paper is organized as follows. Section 1 presents the Swedish demographic transition from 1750 to 2050, which will serve as input in our two models. Section 2 details the theoretical model, its main theoretical implications, and its calibration to data. Section 3 describes the statistical model, its estimation on World data, and the robustness checks which have been carried on. Section 4 compares both simulated growth patterns with actual Swedish data and looks at the properties of the two models to forecast income growth beyond 2000. Section 5 concludes.

# 1 Long-term Trends in Sweden

## Population

Already in 1749, Sweden established a public agency with a responsibility for producing population statistics. These statistics were based on population records kept by the parish priests of the Swedish Lutheran church. Thanks to this effort we have access to detailed data of high quality on how mortality and fertility changed as Sweden developed from a poor agricultural country in the 18th century into a rich, highly industrialized country in the 20th century (Hofsten and Lundström 1976).

Figures 1-3 presents how some key mortality indicators in Sweden have developed during the last 250 years. Figure 1 shows the probability of dying before age 10. Figure 2 gives expected remaining years of life for people that have reached 65 years of age. Figure 3 shows the probability for men of surviving to age 65, given that they survived to age 10. All the graphs are based on time series with annual data.

By 1850, childhood mortality had been improved considerably. Around 1750, 40 % of all children died before age 10. By 1850, this figure was down to about 25 %. This is

Figure 1: Mortality under 10

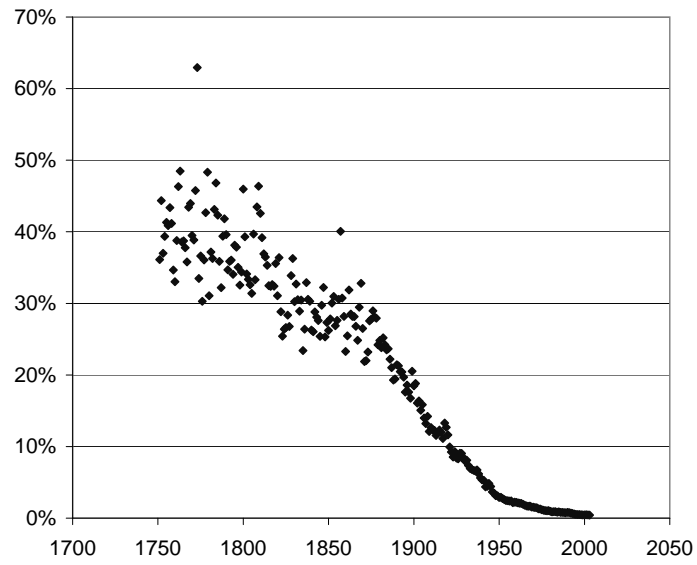


Figure 2: Life expectancy at age 65

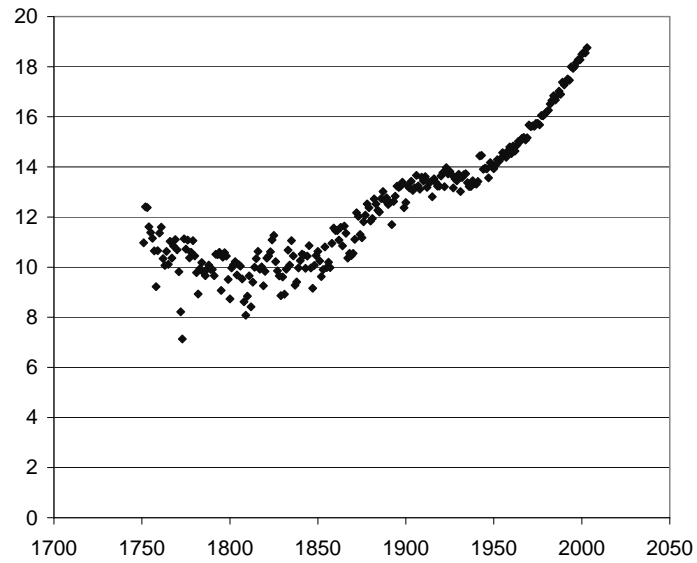
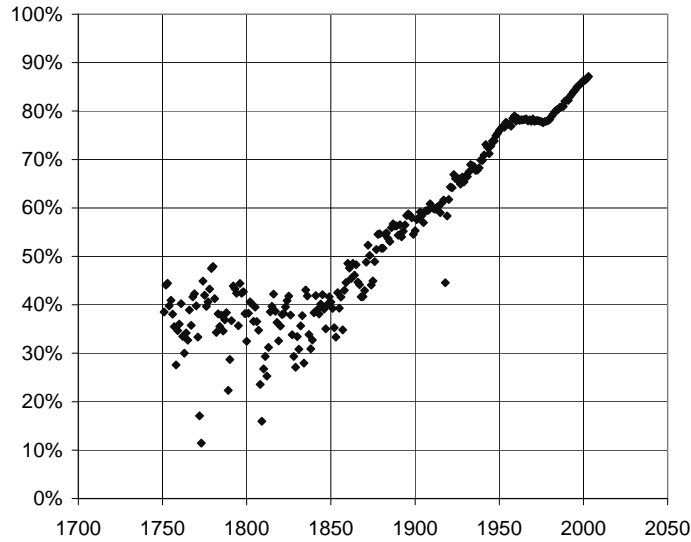


Figure 3: Male survival from age 10 to age 65



still a high number, but the large decline had generated an acceleration in population growth.

After 1850 the age pattern of mortality improvements changed as adult survival began to improve quite significantly.

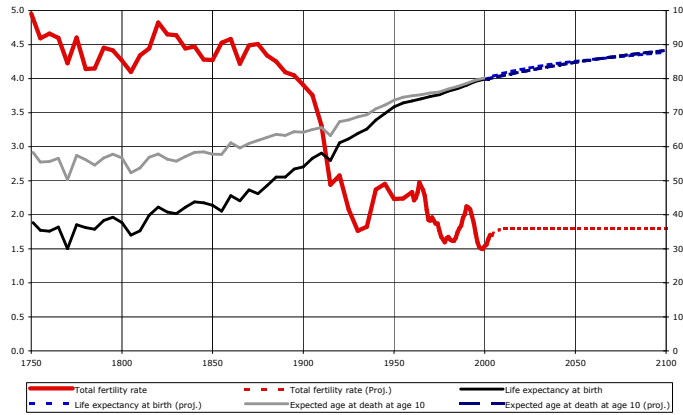
From 1870, a period of long-term, essentially uninterrupted improvement in survival for all age groups was initiated. Below-10 mortality fell from about 25% in 1870 to under 0.5% in 2000. Male survival from age 10 to age 65 increased from about 40% in 1850 to 87% in 2000. Life expectancy at age 65 increased from about 10 years in 1850 to almost 19 years in 2000.

Figure 4 summarizes the mortality changes by two measures: life expectancy at birth and remaining life expectancy at age 10. As can be seen in this graph, increases in adult life expectancy lag behind life expectancy at birth. When adult life expectancy slowly starts to increase around 1825, there has already been a quite substantial increase in life expectancy at birth. The time horizon of Figure 4 extends to 2100, also showing the assumptions on mortality and fertility we have used in the forecasts of the Swedish economy.

In addition, Figure 4 also shows changes in the Swedish total fertility rate after 1750. As can be readily seen from this graph, a clear downward trend in Swedish fertility did not materialize until the last quarter of the 19th century. That is, at a time when mortality had been declining for almost a century.

The long term trend in mortality and fertility has led to a total transformation of the Swedish age structure (Malmberg and Sommestad 2000). This is illustrated in Figure 5. Here the population has been divided into five twenty-years age brackets: 0-19, 20-39, 40-59, 60-79 and 80+. Declining mortality and fertility leads to a change in the age

Figure 4: Summary of the demographic transition



structure from a population dominated by children and young adults to a population where all twenty-year age brackets except the oldest have about the same size.

The expansion of the 0-19 age group is concentrated to the 1820-1920 period; the young adult groups expands 1840-1940; the middle-aged population multiply fast between 1870 and 1970; whereas the expansion of the 60-79 groups is concentrated to the 20th century. The 80+ group, on the other hand, starts to expand rapidly only after 1970. Since it is well-known that the economic behavior of individuals undergoes profound changes from childhood, youth, early adult years, into middle age and during old age, these shifts in the age structure can be expected to have substantial economic effects (Lee 2003).

## Education

Constructing long time-series for education is made difficult by the successive reforms of the Swedish educational systems. The solution chosen here is to take the current system as a starting point and to assign earlier educations to the categories in use today. Since the 1970s Swedish education is divided into three levels: extended primary education, upper secondary education, and tertiary education. Extended primary education comprises grade 1-9, that is, primary and lower secondary education. Upper secondary school includes both theoretical and vocational educations. For the post-1870 period, data on educational enrollment are available in the official Swedish statistics. Pre-1870 data are based on calculations made by Sjöstrand (1961) and Aquilonius and Fredriksson (1941). As can be seen in Figure 6, the expansion of Swedish education has been a four-step process. The first step was an expansion of primary and lower secondary education that took place from the mid-19th century to the early 20th century. The second step was the expansion of upper-secondary education. This expansion accelerated after the First World War and continued up to about 1980. The third step was the post-1940 expansion of extended primary education. Part of this expansion was due to an increase in the cohort-size following a baby-boom in the 1940s, but the extension of compulsory education from six to nine years was also an important factor. The fourth step has been

Figure 5: Changing age structure

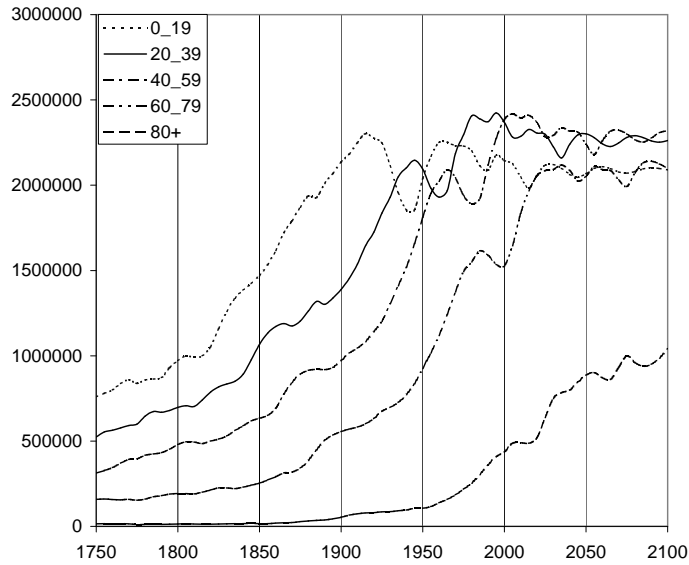


Table 1: Years of education

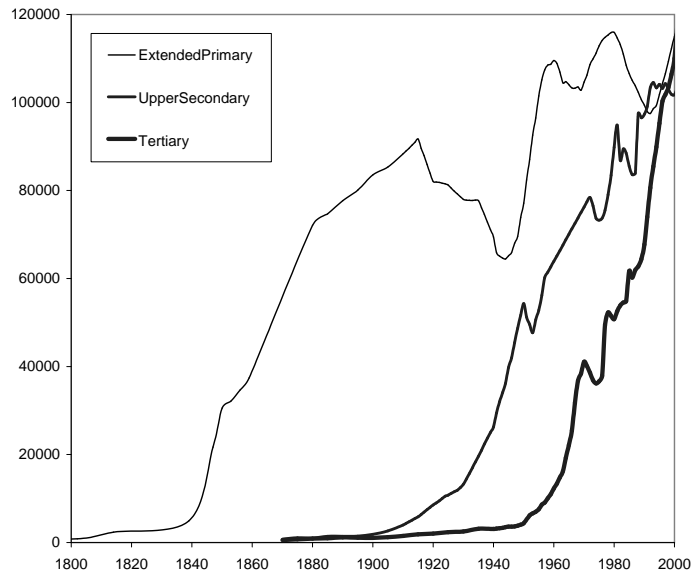
year	Mean length of education for birth cohort	Mean length of education of those aged 25 years
1820	0.53	0.26
1870	5.57	4.22
1913	5.72	5.64
1950	10.81	7.17
1980	13	11.88

the expansion of tertiary education after 1950. This expansion was particularly fast in the 1960s and the 1990s.

In Table 1 we present data on the mean length of education per birth cohort. For cohorts born before 1930 these figures are obtained by adding the yearly, age-specific enrollment rates that result from dividing the number of enrolled per grade with the mean cohort size in the relevant age bracket. For cohorts born between 1930 and 1976, mean length of education is as observed in 2004 in Statistics Sweden (2005), Swedish Register of Education. The figures for cohorts born after 1976 are based on the assumption that growth in the observed mean length of education will continue until it reaches 13 years for the cohort born in 1980 and then remains constant.



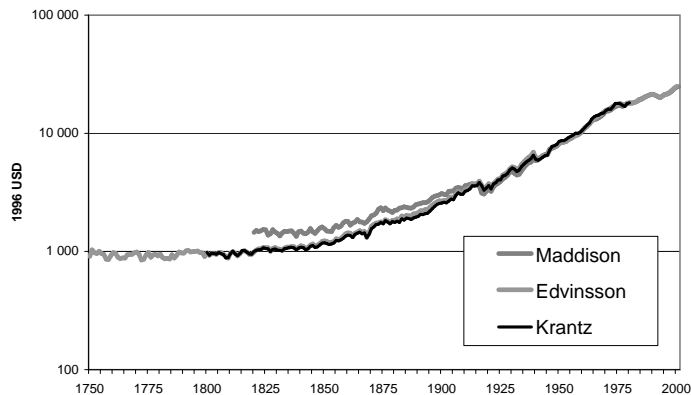
Figure 6: Enrollment history



## Income

Historical estimates of GDP per capita in Sweden are available from several sources. Back to 1861 they all build on work done by Lindahl, Dahlgren, and Kock (1937) but lately these estimates have been extended backwards, e.g. Krantz (1997) brings the annual estimates back to 1800 and Edvinsson (2005) all the way back to 1720. Maddison (2003) also has published an estimate—which is based on previous estimates by Krantz. As is clear from Figure 7 the Maddison estimate differs considerably from those of the two Swedes. Maddison gets the level of real GDP per capita about 50 percent higher in 1820 than the other two estimates.

Figure 7: Historical log GDP per capita estimates in 1996 USD



Apart from the difference in levels that arise in the turbulent early 1900s the series do actually agree rather well on the growth process in the 19th century. Up to the 1820s we have stagnation in per capita income, then a slight rise in income after the Napoleonic wars is discernible, but at a very modest level averaging around half a percent a year. There is an increasing growth trend though and after 1850 average growth rates start to exceed the 1 percent level. After a crisis around the 1870s which also sparked off a very substantial emigration out of the country, mainly to the United States, this growth takeoff gains strength again to rise above the 2 percent level in the early 20th century. Apart from temporary setbacks connected to the World Wars and later oil crises the long-run averages have remained at those levels ever since.

From a level of around 1000 (in 1996 USD) a year in per capita income in the stagnant period (or close to 1500 according to Maddison) Swedish GDP per capita has today reached 25 000. Depending on which estimate we prefer regarding the initial level this is twenty or twenty five times the original subsistence level.

## 2 The Endogenous Growth Model

We first propose a theoretical model to achieve a consistent view of the link between demographic changes and income growth over the 250 years for which data are available.

The transition from stagnation to growth has been the subject of intensive research in the growth literature in recent years. Galor and Weil (2000) propose a unified theory of economic growth in which the inherent Malthusian interaction between technology and population, accelerated the pace of technological progress through rising population density, and ultimately brought about an industrial demand for human capital. Human capital formation and thus further technological progress triggered a demographic transition, enabling economies to convert a larger share of the fruits of factor accumulation and technological progress into growth of income per capita.

Boucekkine, de la Croix, and Licandro (2002) argue that this picture should be completed to account for the specific effect of mortality on the incentive to accumulate human capital. They show in Boucekkine, de la Croix, and Licandro (2003) that the very first acceleration of growth can be related to early drops in adult mortality. In a recent paper, Boucekkine, de la Croix, and Peeters (2007) develop a quantitative theory that argues that the effect of population density of human capital formation prior to the Industrial Revolution was a major force in the process of industrialization. They provide foundations for the effect of population density on human capital formation in the transition from stagnation to growth. The increase in population density made the establishment of schools profitable, stimulating human capital formation (and thereby technological progress) and economic growth.

We use a model adapted from papers by de la Croix and Licandro (1999) and Boucekkine, de la Croix, and Licandro (2002). In this model, demographic variables are exogenous and influence income growth rates through human capital accumulation. Of course, demographic changes are not able to explain the whole pattern of development, and we

shall use the model to measure the changes we need in the other variables to reproduce convincingly the growth of Sweden over two hundred and fifty years.

## 2.1 The Model

*Demographic Structure.* Time is continuous and at each point in time there is a continuum of generations indexed by the date at which they were born. Each individual has an uncertain lifetime. The unconditional probability for an individual belonging to the cohort  $t$  of reaching age  $a \in [0, A(t)]$ , is given by the survival law

$$m(a, t) = \frac{\alpha(t) - e^{\beta(t)a}}{\alpha(t) - 1}, \quad (1)$$

with both functions  $\alpha(t) > 1$  and  $\beta(t) > 0$  being continuous. This two-parameter function is much more realistic than the usual one-parameter function used for example in Kalemli-Özcan, Ryder, and Weil (2000): like the actual survival laws, it is concave, reflecting the fact that the death probability increases with age. It also allows to define a maximum age  $A(t)$  that an individual can reach as

$$A(t) = \frac{\log(\alpha(t))}{\beta(t)}. \quad (2)$$

Assuming that the initial size of a newborn cohort is  $N(t)$ , its size at time  $z > t$  is

$$N(t) m(z - t, t), \text{ for } z \in [t, t + A(t)]. \quad (3)$$

The mortality processes  $\alpha(t)$ ,  $\beta(t)$  and the process for births  $N(t)$  are considered as exogenous in the model in conformity with the purpose of forecasting, as noted in the Introduction. For given  $(\alpha(t), \beta(t), N(t))$  we can easily compute life expectancy at all ages, and sizes of any population group. The unconditional life expectancy associated to (1) is

$$\Lambda(t) = \frac{\alpha(t) \log(\alpha(t))}{(\alpha(t) - 1)\beta(t)} - \frac{1}{\beta(t)}. \quad (4)$$

An increase in life expectancy can arise either through a decrease in  $\beta$  or an increase in  $\alpha$ . These two shifts do not lead to the same changes in the survival probabilities. When  $\alpha$  increases, the improvement in life expectancy relies more on reducing death rates for young and middle-age agents. When  $\beta$  decreases, the old agents benefit the most from the drop in death rates, which has an important effect on the maximum age.

The size of total population at time  $t$  is given by

$$P(t) = \int_{t-\bar{A}(t)}^t N(z) m(t - z, z) dz, \quad (5)$$

where  $\bar{A}(t)$  is the age of the oldest cohort still alive at time  $t$ , i.e.,  $\bar{A}(t) = A(t - \bar{A}(t))$ . The birth rate can be written as  $N(t)/P(t)$ .

*Production Technology.* There is a unique material good, the price of which is normalized to 1, that can be used for consumption. The production function is linear in the stock of human capital:  $Y(t) = H(t)$ . Hence, firms employ the whole labor force to produce as long as the wage per unit of human capital is lower or equal to one. The equilibrium in the labor market thus implies that the wage per unit of human capital is constant through time and equal to one, i.e.,  $w(t) = 1$ , for all  $t$ . The normalization of human capital productivity to unity does not affect the equilibrium.

*The Households' Problem.* An individual born at time  $t, \forall t \geq 0$ , has the following expected utility:

$$\int_t^{t+A(t)} c(t, z) m(z-t, t) e^{-\theta(z-t)} dz, \quad (6)$$

where  $c(t, z)$  is consumption of generation  $t$  member at time  $z$ . The pure time preference parameter is  $\theta$ .

We assume the existence of complete annuity markets. This assumption is equivalent to one with no annuity markets, but with a redistribution of the wealth of the deaths to the persons of the same generation. The inter-temporal budget constraint of the agent born at  $t$  is:

$$\int_t^{t+A(t)} c(t, z) R(t, z) dz = h(t) \int_{t+T(t)}^{t+F(t)} R(t, z) dz. \quad (7)$$

$R(t, z)$  is the contingent price paid by a member of generation  $t$  to receive one unit of the physical good at time  $z$  in the case where he is still alive. By definition,  $R(t, t) = 1$ . The left-hand side is the actual cost of contingent life-cycle consumptions. The right-hand-side is the actual value of contingent earnings. The individual enters the labor market at age  $T(t)$  with human capital  $h(t)$ , and earns a wage  $w(z) = 1$  per unit of human capital.  $F(t)$  is the age until which individuals can work. It can be interpreted either as the age above which the worker is not able to work any longer, or as a mandatory retirement age.

Human capital accumulation depends on the time spent on education,  $T(t)$ , and on the average human capital  $\bar{H}(t)$  of the society at birth, and on a technology parameter  $\mu$ :

$$h(t) = \frac{\mu}{\eta} \bar{H}(t) T(t)^\eta, \quad (8)$$

The presence of  $\bar{H}(t)$  introduces the typical externality which positively relates the future quality of the agent to the cultural ambience of the society (through for instance the quality of the school). This formulation amounts to linking the externality to the output per capita, which is another way of reflecting the general quality of a society. The parameter  $\eta \in ]0, 1[$  is the elasticity of income to years of schooling.

The problem of the representative individual of generation  $t$  is to select a consumption contingent plan and the duration of his education in order to maximize his expected utility subject to his inter-temporal budget constraint, and given the per capita human

capital and the sequence of contingent wages and contingent prices. The corresponding first order necessary conditions for a maximum are

$$m(z - t, t) e^{-\theta(z-t)} - \lambda(t)R(t, z) = 0 \quad (9)$$

$$\eta T(t)^{\eta-1} \int_{t+T(t)}^{t+F(t)} R(t, z) dz - T(t)^\eta R(t, t + T(t)) = 0, \quad (10)$$

where  $\lambda(t)$  is the Lagrangean multiplier associated to the inter-temporal budget constraint. Since  $R(t, t) = 1$  and  $m(0, t) = 1$ , we obtain from equation (9)  $\lambda(t) = 1$ . Using this in (8), we may rewrite contingent prices as

$$R(t, z) = m(z - t, t) e^{-\theta(z-t)}. \quad (11)$$

Equation (11) reflects that, with linear utility, contingent prices are just equal to the discount factor in utility, which includes the survival probabilities.

The first order necessary condition for the schooling time is (10). The first term is the marginal gain of increasing the time spent at school and the second is the marginal cost, i.e., the loss in wage income if the entry on the job market is delayed.

From (10) and (11) the solution for  $T(t)$  should satisfy:

$$T(t) m(T(t), t) e^{-\theta T(t)} = \eta \int_{T(t)}^{F(t)} m(a, t) e^{-\theta a} da, \quad (12)$$

where the right hand side represents the discounted flow of wages per unit of human capital. Notice that optimal schooling does not depend on the efficiency of education  $\mu()$ , because  $\mu()$  affects symmetrically opportunity costs and benefits, but aggregate human capital will depend on it.

*Aggregate Human Capital.* The productive aggregate stock of human capital is computed from the human capital of all generations currently at work:

$$H(t) = \int_{t-\bar{F}(t)}^{t-\bar{T}(t)} e^{n(z)z} m(t - z, z) h(z) dz, \quad (13)$$

where  $t - \bar{T}(t)$  is the last generation that entered the job market at  $t$  and  $t - \bar{F}(t)$  is the oldest generation still working at  $t$ . Then,  $\bar{T}(t) = T(t - \bar{T}(t))$ , and  $\bar{F}(t) = F(t - \bar{F}(t))$ . Accordingly, the size of active population is:

$$P^A(t) = \int_{t-\bar{F}(t)}^{t-\bar{T}(t)} N(z) m(t - z, z) dz, \quad (14)$$

The average human capital at the root of the externality (8) is obtained by dividing the aggregate human capital by the size of the population given in (5):

$$\bar{H}(t) = \frac{H(t)}{P(t)}. \quad (15)$$

The dynamics of human capital accumulation can be obtained by combining (8) with (13) and (15). To evaluate  $H(t)$ , for  $t \geq 0$ , we need to know initial conditions for  $H(t)$ , for  $t \in [-\bar{A}(0), 0[$ .

## 2.2 Some Theoretical Results

In Boucekkine, de la Croix, and Licandro (2002) some interesting properties of the theoretical model have been derived. Let us provide the intuition for four of them.

**Property 1** *A rise in life expectancy increases the optimal length of schooling.*

A key property of the model is that a decrease in the death rates, or equivalently, an increase in life expectancy induces individuals to study more. This prediction is consistent with the joint observation of a large increase in both life expectancy and years of schooling during the last 150 years.

**Property 2** *When demographic variables are constant through time, income grows at a constant rate.*

There is thus a balanced growth path. The value of the long-run growth rate is a function of various factors, such as the efficiency of education  $\mu$ . Notice that the income of an individual does not grow over time; growth in the economy is linked to the appearance of new generations. Hence, the objective function of an individual is always finite.

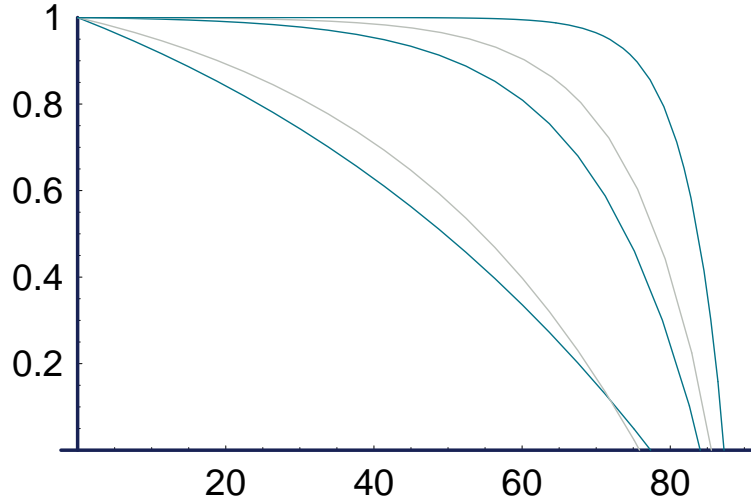
**Property 3** *A rise in life expectancy  $\Lambda$  has a positive effect on economic growth for low levels of life expectancy and a negative effect on economic growth for high levels of life expectancy.*

Intuitively, the total effect of an increase in life expectancy results from combining three factors: (a) agents die later on average, thus the depreciation rate of aggregate human capital decreases; (b) agents tend to study more because the expected flow of future wages has risen, and the human capital per capita increases; (c) the economy consists more of old agents who did their schooling a long time ago. The two first effects have a positive influence on the growth rate but the third effect has a negative influence. Notice that the two last effects are still effective even if there is a fixed retirement age (which does not change with life expectancy) or if we had assumed that human capital becomes fully depreciated after a given age. This is due to the fact that a rise in life expectancy reduces the probability of dying during the activity period.

**Property 4** *There is a growth-maximizing fertility rate  $N/P$ .*

The model generates an interior  $N/P$ -maximizer for the per capita growth rate. More precisely, the functional relation between these two variables is hump-shaped. This feature relies on the vintage nature of the economy. Indeed, when fertility is relatively low, the share of retired workers in the population is relatively high. Increasing fertility thus increases the active population and the growth rate. However, when fertility is very high, the students are the main group in the population. Lowering fertility would then increase the active population. In the two extreme cases, very low and very high fertility, the size of active population compared to total population is small which depresses growth. There is thus a level of fertility which maximizes the activity rate. To this rate there corresponds a growth-maximizing demographic growth rate.

Figure 8: Model's survival functions in 1750, 1860, 2000, 2050 and 2200



### 2.3 Calibration

We calibrate the demographic processes of the model on the population data presented in Section 1. In order to focus on adult mortality, we disregard the huge swings affecting infant mortality in the 17th, 18th and early 19th centuries. Accordingly, we will consider that the birth date in our model corresponds to age 10 in the data. One decision variable is affected by this time shift, the schooling time,  $T(t)$ . If the birth date is 10, one can legitimately argue that the true schooling time is not  $T(t)$ , but  $T(t) + T_0$ , where  $T_0$  is the time spent at school before 10. In our empirical assessment, we take into account this crucial aspect and replace  $T(t)$  with  $T(t) + T_0$  in the model. More precisely, we set  $T_0 = 4$ , which means that the representative individual has already cumulated four years of education at birth.

To calibrate the model, the exogenous processes  $\alpha(t)$ ,  $\beta(t)$ , and  $N(t)$  should be made explicit. We assume that all these processes follow a polynomial function of time. Polynomials of order 3 are sufficient to capture the main trends in the data. For the survival function processes  $\alpha(t)$  and  $\beta(t)$ , the parameters of the polynomial are chosen by minimizing the distance between the model's life expectancies at age 10 to 80 with their empirical counterparts. The implied survival laws for various years are reported in Figure 8.

The parameters of the process for  $N(t)$  are chosen so that the distance between the share of the age groups 10-15, 15-30, 30-50, 50-65, and 65+ in total total population  $P(t)$  and the observed levels is minimized.

For the risk free interest rate, we choose 3% per year, which sets  $\theta = 0.03$ .<sup>1</sup> The effective

---

<sup>1</sup>Robustness analysis shows that the value of  $\theta$ , provided it remains small, does not influence the characteristics of our simulations.

retirement age  $F(t)$  is set constant to 53 (that is age 63). This number is in accordance with the estimate of the effective retirement age by Blondal and Scarpetta (1997) on the recent past. Since we do not have more information on its historical value, we keep it constant through time.

A value for the elasticity of income to schooling could be drawn from the estimations of the wage equations (see the discussion in de la Croix and Doepke (2003)) which yield a value  $\eta = 0.6$ . This value, though, is correct only for the recent years; it is very likely that the return to education was lower in the past. We therefore use the model to compute an elasticity that is in line with the actual length of education reported in Section 1. This amounts to solve equation (12) for  $\eta$  after having replaced  $T(t)$  by the observations. Once we have computed a series for  $\eta$ , we smooth it by estimating a polynomial of order three in time, to eliminate the short/medium run variations we are not interested in. In Table 2 the evolution of  $\eta$  is reported.

Table 2: Elasticity of income to schooling

year	$\eta$
1800	0.18
1850	0.22
1900	0.31
1950	0.41
2000	0.50

Hence, the observed increase in educational attainment cannot be entirely explained by higher longevity; one also needs an increase in the return to schooling. If mortality decrease had been the only factor, assuming a constant  $\eta = 0.5$ , schooling would have increased by only 1.75 years, which represents 20% of the total increase of 9 years (from 4 to 13). The other 80% needs to be explained by other factors increasing the return to education, such as skill-biased technical progress, public funding of education etc.

Another parameter that is likely to have changed over two centuries is the parameter  $\mu$  affecting the efficiency of education. Here we want to reflect the idea of “population-induced” technical progress as in Galor and Weil (2000), Lagerlöf (2003), and Boucekine, de la Croix, and Peeters (2007). This assumption is meant to capture a positive effect in more dense populations of transmission of skills and knowledge, i.e. in regions with shorter geographical distance between people. To calibrate the the changes in this process, we follow Lagerlöf (2006) by assuming that population exerts a positive effect only in a certain range. We therefore calibrate a process of the form:

$$\max [\mu_0, \mu_1 + \min[\mu_2 P_t^A, \mu_3]] .$$

The parameters are chosen so that the distance between the growth of income per capita along the balanced growth path in 1760, 1835, 1865, 1895, 1925, 1955 and 1985 and the growth rate estimated by Maddison over the corresponding 30 year periods is minimized. The implied level of efficiency is plotted as a function of time in Figure 9. We observe



Figure 9: The efficiency level  $\mu$  of education over time.

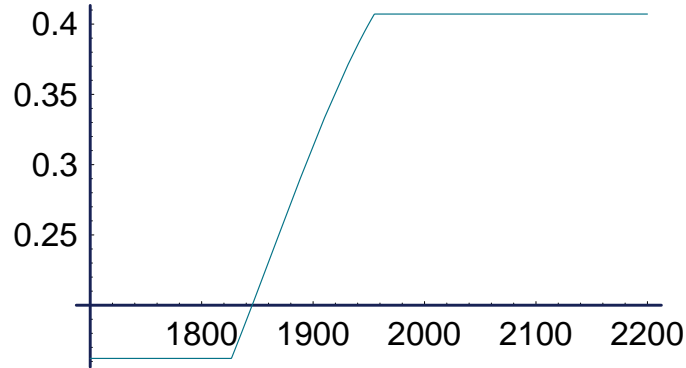


Table 3: Simulated annual growth rate of income per capita

year	growth rate	year	growth rate
1800	0.288	2050	2.122
1850	0.684	2100	1.917
1900	1.660	2150	1.765
1950	2.203	2200	1.698
2000	2.203	2250	1.627

that the efficiency of education starts to rise around 1820 once population passes a given threshold; in the middle of the twentieth century, productivity stops increasing, and the scale effect does not play any role further.

## 2.4 Simulation results

We run our simulation assuming that the economy was on a balanced growth path prior to 1750. Then we use the method proposed by Boucekkine, Licandro, and Paul (1997) to solve models with differential-difference equations. The simulation covers the period 1750-2300. The simulation and projection results of income per capita are analyzed in depth in Section 4. We limit here the discussion to the results in terms of growth rates presented in Table 3.

Growth starts from very low levels in the eighteenth century, accelerates during the nineteenth century, reaches a maximum in 1963 with 2.293% and then declines. It does not revert to the pre-industrial level however, since permanent changes in the return to education and in productivity have occurred. In 2250, the model yields a growth rate similar to the one in the beginning of the twentieth century. This can be seen as the cost of aging.

Figure 10: Income per capita (logs): baseline and constant mortality scenario

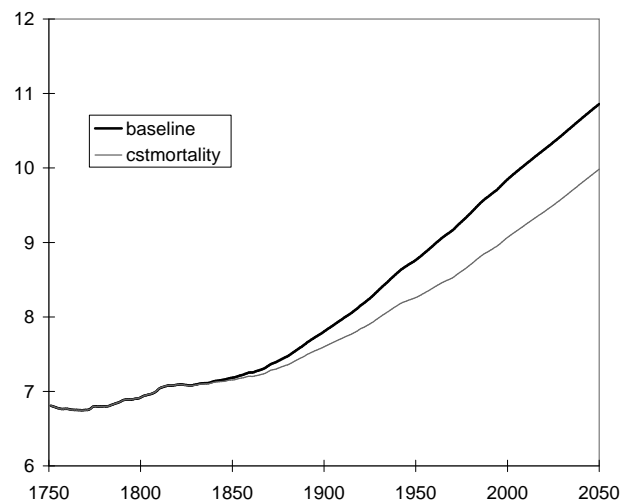
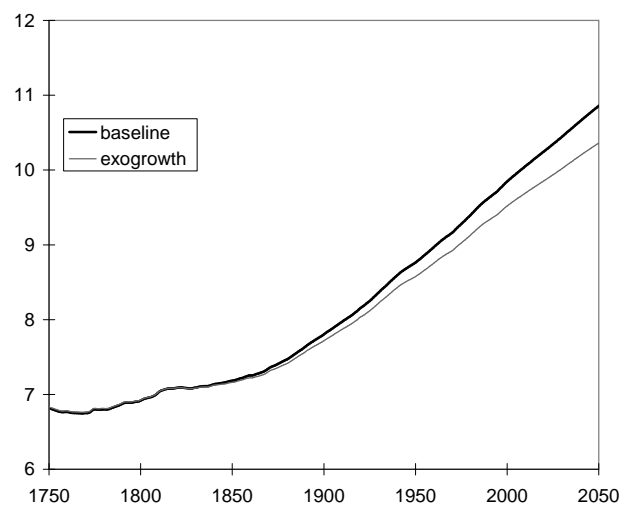


Figure 11: Income per capita (logs): baseline and exogenous growth



Before moving to the demographic dividend model, we address two specific questions. a) What would have happened if longevity stayed at the 1750 level ? b) Is the endogenous growth assumption crucial for our results ?

To answer the first question, we run a counterfactual experiment by keeping the parameters  $\alpha$  and  $\beta$  at their 1750 level. This has several implications. First, incentives to invest in education are reduced, and the length of schooling will never go beyond 11 years (7+4) instead of 13 in the baseline simulation. Second, active population will be much lower, implying that the effect of scale on productivity is now smaller. Third, the age structure of the population is modified, with fewer old persons at all dates. The total effect of these factors is to depress income per capita compared to the baseline simulate after 1870. Figure 10 compares the baseline simulation with the counterfactual experiment. In 2000, GDP per capita would be 55% lower if mortality was at its 1750 level.

The role of endogenous growth can be understood by changing very slightly the specification of equation 8. Assume instead that human capital follows:

$$h(t) = \frac{\mu(t)}{\eta(t)\rho} \bar{H}(t)^\rho T(t)^{\eta(t)} \quad (16)$$

With  $0 < \rho < 1$ , the externality is weaker, and, in the absence of exogenous technical progress, the long-run growth rate of the economy is zero. Figure 11 shows the resulting income per capita compared to the baseline if we set  $\rho = .98$  without changing anything else. In the beginning it makes little difference, but as time passes, the improvements in mortality and in technology yield less persistent results. In 2000, GDP per capita is lower than the baseline by 28%. For the future, the fact that  $\rho$  is below 1, even by a very small amount, matters a lot. The gap to the baseline is widening sharply after 2000 and income per capita with  $\rho = .98$  will stop growing asymptotically (In 2300, the end of our simulation, it is still growing at 1% per year).

### 3 The Demographic Dividend Model

The above section has demonstrated that an endogenous growth model can be used to mimic the effect of observed shifts in mortality, fertility, and age structure on the Swedish long-term per capita income growth. This gives a strong theoretical underpinning to the proposition that demographic change is a key element in the economic development process. Endogenous growth models incorporating demographic elements are not, however, the only approach to assess the importance of demographic factors in the analysis of economic growth. A more direct approach has been to incorporate demographic measures such as life expectancy and measures of age structure in Barro-type, cross-country growth regression. In general, demographic variables have been shown to have strong and significant effects on per-capita income growth in these estimations. A relevant question, therefore, is if a dividend model would be able to account for long-term economic growth in countries that experienced the demographic transition already in the 19th and early 20th century?

One way to answer this question would be to estimate a time series regression on the long-term demographic and GDP data of a country with an early demographic transition. There are many problems to that approach however: structural breaks, persistence in both dependent and independent variables, and high collinearity among independent variables. It is therefore a risk that the demographic effects on income growth are drowned by such problems. Therefore, the approach taken here is, instead, to employ an empirical dividend model that has been estimated on modern, global, cross-country panel data to backcast Swedish per capita income growth back to 1750. The backcast can then be compared to the available empirical estimates of Swedish per capita GDP as well as to the simulated path of per capita income presented in the preceding section.

Such a backcast, using a model estimated on a different data set, does in fact represent an out-of-sample test of the stability of the original model. An evaluation of the backcasting-performance of the dividend model, therefore, will not only cast light on the possible importance of demographic factors in Swedish economic development. It also shows whether the information from currently developing countries is useful for describing historical experience. Furthermore, such historical stability would add credibility to long-term forecasts by increasing the likelihood of continued stability.

### 3.1 A Cross-Country Estimation

At least three arguments underscore the importance of age structure for per capita income. One is the savings argument. In countries with high child-dependency rates, savings rates will be low and this may lead to low productivity if domestic capital formation is constrained by savings. This argument was first put forward by Coale and Hoover (1958). Second, a high dependency rate implies a low worker per capita ratio and this should lead to a lower per capita income in a direct way by a pure accounting effect. Kelley and Schmidt (2005) summarize this argument and review much of the demographic dividend literature up to date. Third, as demonstrated by Lindh and Malmberg (1999) on OECD data age structure within the working-age population is also of importance.

The dividend model here uses levels of per capita GDP instead of growth rates as dependent variable and age shares instead of age group growth rates as explanatory variables. That is, a level regression is used instead of a first-difference estimation. An argument for using a level-specification is evidence showing co-integration between GDP and age structure in a OECD sample (Österholm (2004)). This implies that standard least square estimates of coefficients are superconsistent. If the cointegration assumption does not hold a potential problem is that an estimation using non-stationary time-series can result in a spurious regression. However, in a panel context, this problem is substantially ameliorated (Phillips and Moon (1999)). More important, our intention is to use the regression results for out-of-sample backcasting. Failure to produce a successful backcast immediately expose any spurious regression problem.

Dividend models, typically, also includes life expectancy as one of the explanatory variables. First, increasing life-expectancy is likely to increase savings by increasing the

risk for survival into old age dependency. Second, higher life expectancy increases the expected return of education. The theory in the previous section has focused on this latter mechanism and left out physical capital. However, a capital externality with life-cycle saving would work in a rather similar way and we will not here attempt to identify these effects separately. From the theory we would expect that increases in the share of the active population are associated with higher income, and more so the higher levels life expectancy reaches, up to the point where the increase in life expectancy mainly increases the retired population. For a single time series changing life-expectancy is reflected directly in the age structure and we would not expect to be able to identify separate effects from increasing longevity and growing shares of elderly in the population. In the country panel with observations from different stages of the demographic transition the effects of different age shares on GDP will shift over time as increasing length of education and increasing longevity pushes the effective working ages upwards. By interacting age group shares with life-expectancy this expected gradual drift in age effects can be compensated and age effects are estimated as varying with life-expectancy.

### 3.2 Data and Specification of Model

The details of the cross-country estimation together with extensive regression diagnostics have been presented in Lindh and Malmberg (2004). The presentation here, therefore, will concentrate on the main features of the model.

Our economic data are taken from Penn World Table Version 6.1, Center for International Comparisons at the University of Pennsylvania (CICUP), October 2002. We use 111 countries which had coherent data for at least the period 1961-1996 using the variable RGDPCH (the chain indexed PPP-adjusted real GDP estimate) which is available for many countries since 1950. We deleted countries with shorter time series both because we wanted to maintain a reasonably balanced panel and because we know from time series estimation that too short time series are unreliable when estimating the correlations to age structure. Demographic variables, stretching from 1950 up to the end of the 1990s are from UN World Population Prospects (2000) from which we also have consistent projections up to 2050.

Our estimation model allows for a panel regression with the logarithm of per capita GDP,  $y$ , as dependent variable predicted by the independent demographic variables: the logarithms of age shares,  $a$ , and the logarithm of life expectancy at birth,  $t$  being the time period, and including interaction terms between life expectancy,  $e_0$ , and age shares to catch the upward drift in the economically active period of the life cycle:

$$y_{it} = \alpha \log e_{0it} + \sum_{k=0-14}^{65+} (\beta_k + \gamma_k \log e_{0it}) a_{kit} + \eta_i + \nu_t + \varepsilon_{it} \quad (17)$$

We allow for country-specific intercepts through  $\eta_i$  and  $\nu_t$  accounts for time-specific effects. A potential problem is that life expectancy is highly correlated with age structure,

especially the size of older groups, and more seriously with the interaction terms themselves. However, checking the correlation matrix it turns out that log life expectancy is more strongly correlated (about 0.8) with log GDP per capita than with any of the age share variables. Recursive estimation proves that the parameter estimates are robust and stable given that the time series dimension is long enough.

The demographic variables are lagged one step to ensure that they are pre-determined, but these variables are highly persistent so this does not make much difference. Endogeneity therefore may still be an issue for any structural interpretation of the coefficients but for the purpose of forecasting, or in this case backcasting, bias has to be traded for precision anyway. Using instrumental variables thus becomes less of a choice even if we had been able to find good instruments.

Without the interaction terms GDP per capita would be described by a Cobb-Douglas index of age shares and life expectancy to capture "technological change". This is thus essentially a standard production function specification where we use population shares as substitutes for production factor intensities and the interaction with life expectancy to catch cross effects with human and physical capital, or if you like knowledge capital. The logarithmic form ameliorates problems with heteroskedasticity and also makes it possible to include the whole distribution of age shares in the fixed effect estimation, since the exact linear dependence of the full set of age group shares is broken.

Based on previous work (Lindh and Malmberg (1999)) an aggregation of age groups to children 0-14, young adults 15-29, mature adults 30-49, middle aged 50-64, and old age 65+ is known to work well in growth equations for the OECD without running into collinearity problems. This corresponds roughly to the age intervals in which humans are first dependent on parents, second finding their place in adult life and forming a family, third raising their family, fourth preparing for retirement and fifth retiring. The limits for these functional groups are, of course, not exact. They vary both with time and culture, as well as the institutions that transmit and govern the economic effects of each age group. Nor do we expect effects to be uniform within the limits. This specification is thus a pragmatic approximation for estimating growth effects from the continuous age distribution. The age distribution in turn proxies for the actual functional changes in behavior and resources over the life cycle which are the real causes for the GDP effects.

### **3.3 Parameter Estimates of the Dividend Model**

To simplify the out-of-sample tests below, we have only used data up to 1996 in the estimations reported in Table 4. The estimates show that life expectancy is positively correlated with per capita income. The estimates of interaction effects also indicate that the basic hypothesis is valid; life expectancy modifies the correlation with demographic age structure by shifting life phases. In Figure 12 we visualize this shifting pattern of age elasticities on income that is implied by the interaction model. The effect of young and mature adults decreases with life expectancy while the negative effects of dependents tend to decrease also. This shifts the hump of the life cycle pattern upwards and makes it flatter and less pronounced as life expectancy rise. This might indicate that increasing

Table 4: Estimates of Equation (17)

	$\alpha, \beta_k$	$\gamma_k$
$\log e_0$	<b>5.412</b> (2.38)	
$\log a_{0-14}$	-5.45 (3.01)	1.062 (0.69)
$\log a_{15-29}$	3.704 (2.25)	-0.872 (0.52)
$\log a_{30-49}$	<b>3.800</b> (1.79)	<b>-0.831</b> (0.42)
$\log a_{50-64}$	0.251 (1.02)	0.017 (0.24)
$\log a_{65+}$	<b>-7.597</b> (0.65)	<b>1.873</b> (0.16)
	$\bar{R}^2$	0.964

Note: Model with fixed time and country effects. Bold face indicates that the estimates are significantly different from zero on the 5 percent level. Standard errors are adjusted for the unbalanced panel.

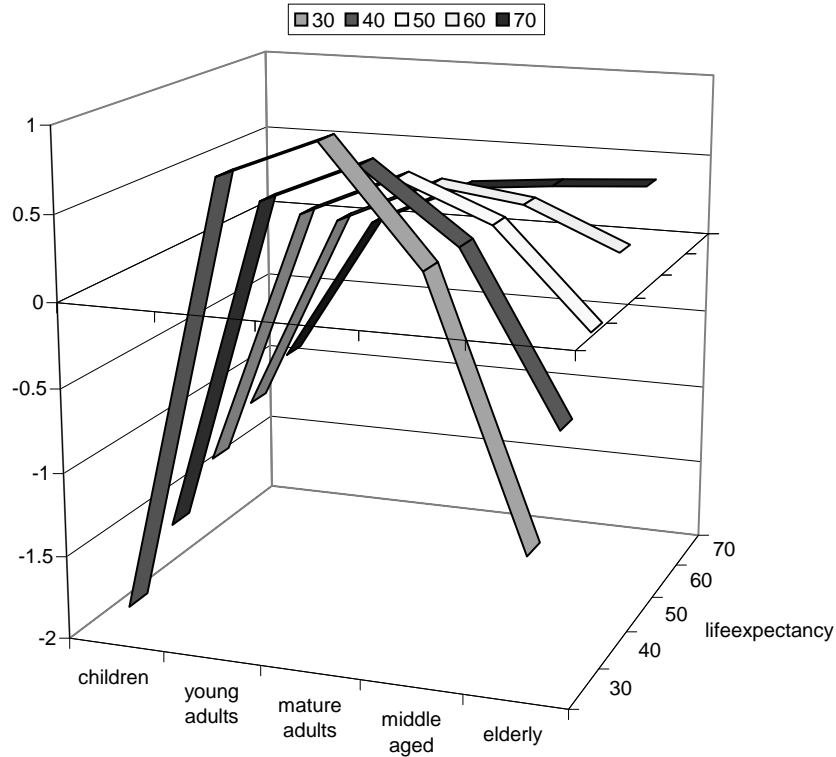
length of education in low mortality populations reduces positive effects from young adults while increasing them for middle aged and even perhaps for the elderly. The latter conclusion is highly uncertain due to the collinearity issue between children and elderly although it is intriguing that we actually see a trend in that direction.

At low and medium levels of life expectancy the age effects on per capita income are dominated by the balance between children and young adults. Child-rich populations tend to be poor whereas countries with declining child dependency and an expanding young adult population enjoy rising per capita income. At higher levels of life expectancy it is instead a high share of middle age adults (30-49 and 50-64 years old) that ensure good economic prospects.

## 4 Long-term Growth in Sweden: an Evaluation

To generate a Swedish projection for the period 1751-2050 we use demographic data taken from the Human Mortality Database (Berkeley and Rostock). We have updated with the latest estimates (fall 2004) up to 2003 and projections up to 2050. Using the

Figure 12: The pattern of shifting age elasticities from the heterogeneous model



coefficients of the interaction model in Table 4 columns 2 and 3 it is then trivial to make the projection. However, since mortality data and thus life expectancy is very volatile in the historical data we have used the same smoothed series that were used in the simulation part. The confidence intervals have been computed in the standard way only taking account of parameter uncertainty and assuming a normal distribution.<sup>2</sup>

Figure 13 shows the projections both backwards and forwards with 95 percent prediction intervals and compared to two of the historical estimates presented in Section 2.<sup>3</sup> The historical estimates and the simulated curve fall (barely) within the confidence intervals but the 19th century projection part underestimates growth in the second half of that century and misses part of the high growth take-off, whether that takes place in the 1860s, as indicated by Maddison, or in the 1890s, as Edvinsson shows. There is slightly less curvature in the simulation path compared both to the projection from the world sample and historical estimates and actual data, at least in the latter part of the period. Within the period 1870 to 1950 the projections of both models are very close to historical estimates data and quite well centered within the prediction bands. Since the simulation has been calibrated on the Maddison estimates it is not surprising to see that there is

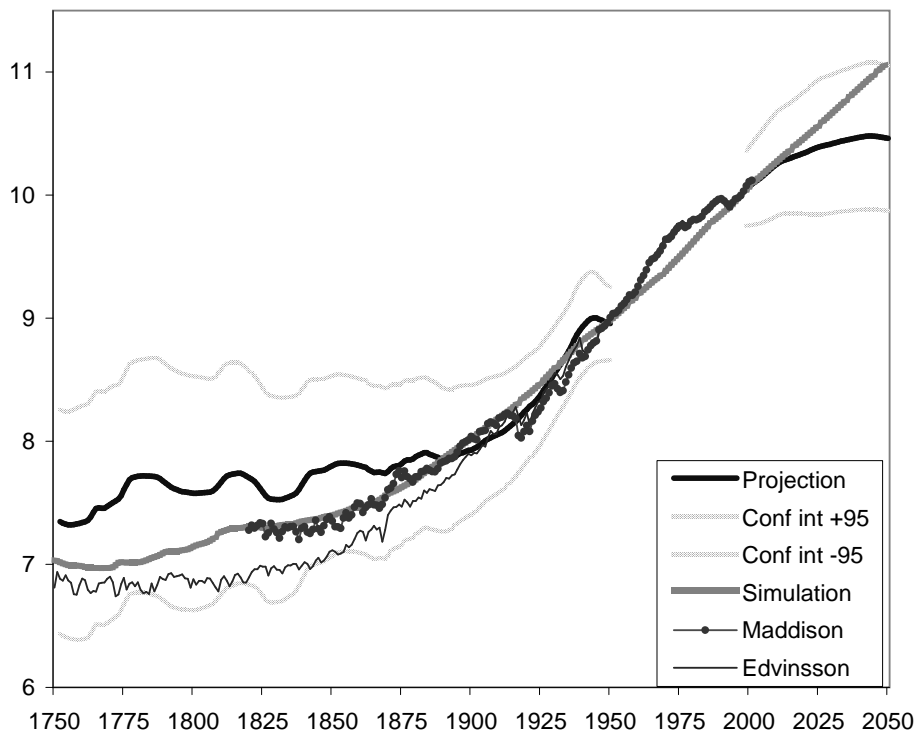
<sup>2</sup>The projections in the graphs below were therefore adjusted to fit exactly to the Penn data at the endpoints of the interval, both forward and backwards by ratio linking in 1951 and 1998.

<sup>3</sup>We use two historical estimates to better grasp the uncertainty surrounding GDP per capita before World War II.



a good agreement to these. Also note that the forecast series and the simulation path agrees quite well on the future path of GDP per capita up till the second decade of the 21st century when the projection starts to indicate stagnation while the simulated GDP per capita due to its built-in endogenous growth continues to rise. Let us now look closer at the period 1820-1950.

Figure 13: Comparisons over the period 1750-2050

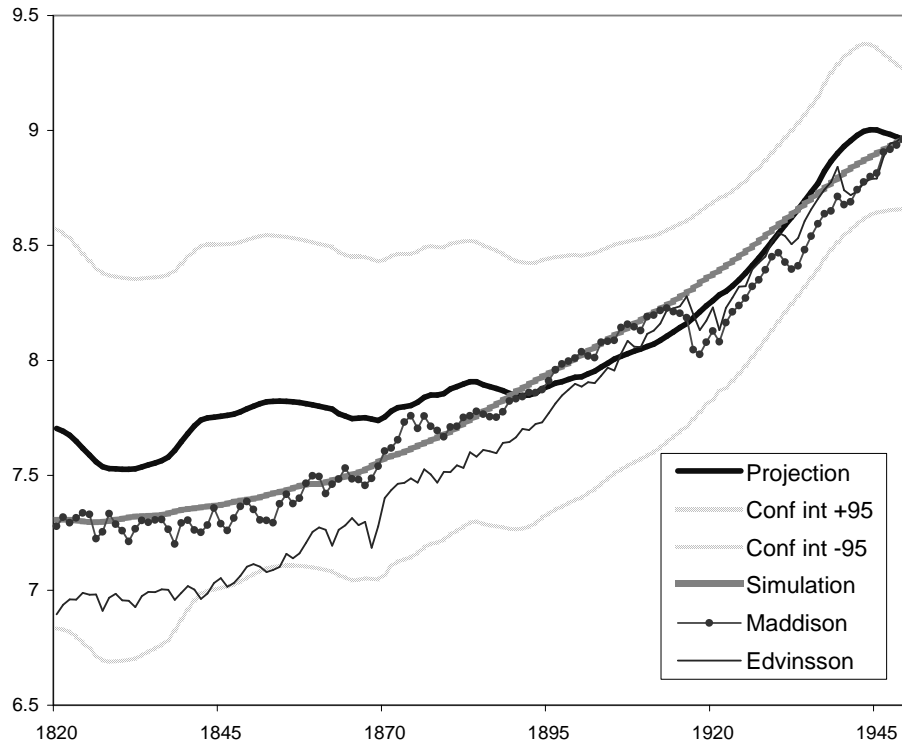


#### 4.1 Comparisons over the period 1820-1950

In Figure 14 we have zoomed in on the period 1820-1950. The simulated GDP per capita is generally slightly higher than both the backcast and the historical estimates from the 1890s to the 1930s. The backcast level in turn is generally higher than the historical estimates in the period encompassing the World Wars and the Interwar period but on the whole it agrees with the Maddison estimates fairly well down to around 1870 but is higher than Edvinsson's estimates before 1900. The relative deviations for different periods are shown in Table 5.

The order of the relative differences between the backcast and Maddison back to 1870 is actually about the same as the differences between the Edvinsson and Maddison estimates but increases somewhat as we go further back in time. In view of the immense changes in available technology that has taken place between the late 19th century and

Figure 14: Same data as in previous figure but focusing on the period 1820-1950.



the late 20th century one might have expected that the relation between income level and demographic structure in developing countries in Africa and Asia would look rather different than it did in Sweden 100 years earlier. But the relation seems to have been essentially the same at least back to the end of the 19th century. This strongly suggests that the economic development associated with the demographic transition is a universal process allowing us to treat global data as stemming from a common non-stationary DGP.

We do not have any final answer why the backcast underestimates rates of growth in the 19th century but the simulation model suggests that the demographic dividend model does not account for the booster effect of the size of the active population on the endogenous growth mechanism through human capital. In the forecast period the human capital externality also may be behind the simulation diverging upwards from the demographically based projection. On the other hand the simulation model does not directly account for negative child dependency effects, ignores infant mortality and predicts that an aging workforce will be less productive.

If the projection is the more relevant model we could therefore conjecture that the simulated path have missed some negative influences. The increase in fertility that Swedish population projections assume together with the rejuvenation of the workforce that will take place when the baby boomers from the 1940s retire are the main influences that slows down the growth rate in the projection model.

Table 5: Relative deviations from Maddison as benchmark

	Edvinsson	Simulated series	Back projection
1820-1850	-0.261	0.044	0.426
1851-1870	-0.203	0.040	0.423
1871-1890	-0.174	-0.025	0.127
1891-1910	-0.114	0.015	-0.068
1911-1930	0.060	0.184	0.065
1931-1950	0.060	0.116	0.184

If decreasing returns to  $\bar{H}$  are introduced in the simulation model, simulated GDP per capita turns downwards getting much closer to the econometric forecast. However, it turned out to be impossible to mimic the historic GDP per capita evolution at the same time. While this is no conclusive proof we interpret it as an indication that an externality in human capital that generates endogenous growth helps to explain the historic development.

## 4.2 Can Forecast and Simulation be Combined?

The discussion above suggests that the two models may incorporate different sets of demographic information useful for forecasting. Thus a combination of the forecasts may be useful. Since the seminal paper by Bates and Granger (1969) a vast literature on the combination of forecasts has been generated. This literature suggests that considerable reductions of forecast errors can be achieved by using weighted averages of forecasts rather than any particular forecast per se. The question of why that works has not yet been fully answered but, as pointed out by Diebold and Lopez (1996), all forecasting models are in practice misspecified simplifications of the actual DGPs that generate real data and there is therefore a potential benefit to be had by pooling different biases. A recent evaluation by Hendry and Clements (2001) discuss the potential explanations and confirm that if indeed the correct conditional expectation of a weakly stationary process is known then combination of forecasts is ineffective in reducing forecast errors.

Mis-specification, inefficient use of all available information or non-stationarities then appear as pre-conditions for achieving gains from combination. In practice all three of these sources for combination gains are abundant. More specifically in this case we use two very different approaches. The empirical information used to calibrate the simulation model overlaps only slightly in the second half of the 20th century with the global data used to estimate the econometric model and only for one country out of 111. And in fact the Maddison estimate of GDP per capita differs substantially from

that of PWT. By assumption the cross-country information on countries early in their development provides information also on the historical time series of Sweden. On the other hand we lack information on the mechanisms in the economy that generates this process, but some of this information has presumably been built into the simulation model. The crucial question is then whether the two models encompass each other? Or do they both contain independent information useful for predicting actual GDP per capita?

Armstrong (1989) in a succinct summary of the then available studies of forecast combination concludes that combination of forecasts seems to be most useful for long-range forecasting and also that combination seems to yield the most gain when the methods used are very different. The simulation and the projection forward predict very similar growth rates over the near future and also backwards. It is difficult to evaluate their performance in actual forecasting out-of-sample, especially since we would not expect them to track annual variation in GDP per capita.<sup>4</sup>

As a first step we assume we can treat the series as weakly stationary and check whether the combination works. In Table 6 we report the results from regressions of the following form

$$y_t = \alpha_0 + \beta_s \hat{y}_{t,s} + \beta_p \hat{y}_{t,p} + \varepsilon_t \quad (18)$$

where the left hand side are GDP per capita measures while  $\hat{y}_{t,s}$  and  $\hat{y}_{t,p}$  denote the simulated series and the projected series, respectively. As dependent variable we use both the longer Edvinsson series that has not been used in either estimation or calibration and the shorter Maddison series that has been used in calibration but differs from the PWT data that were used for estimation. Since we are testing the general agreement the projection series used in the test has not been level shifted in 1950 and 1996 as in the graphs.

This equation generates unbiased combination forecasts as a by-product even if the individual forecasts are biased (see Granger and Ramanathan (1984)) and is often referred to as the simple linear combining method by e.g. Deutsch, Granger, and Teräsvirta (1994) if  $t$  is in the future while Diebold and Lopez consider the same equation as a test of whether the two forecasts encompasses each other or not. If the coefficient vector is  $(0, 1, 0)$  the simulation would be said to encompass the projection model while if it is  $(0, 0, 1)$  then the projection encompasses the simulation model. For any other values neither model encompasses the other and both forecasts contain useful information about the DGP in question. In our case, of course, neither of the projections are forecasts estimated from a given time series in the usual sense, nor are the historical time series

---

<sup>4</sup>It is not quite clear what we would mean by an out-of-sample test in either of the two cases. Should data not used in the calibration of the simulation model constitute a test sample, and if so over what period would we need to evaluate a model designed to run over several centuries? The projection on the other hand has a clearly defined period of data that the equation is estimated on, but it is not clear what status we would give considerable deviations in one individual country within the close future. Since such deviations should occur we have little guidance for an out-of-sample test using only Swedish data.

Table 6: Encompassing tests for different models and series.

	$\alpha_0$	$\beta_s$	$\beta_p$	<b>Maddison</b>		Test (0, 1, 0)	Test (0, 0, 1)	$\bar{R}^2$	DW
				AR	MA				
OLS	-943.1 (215.9)	0.618 (0.234)	0.514 (0.242)			42.036 (0.000)	196.8 (0.000)	0.985	0.054
AR	-694.6 (836.4)	0.594 (0.129)	0.504 (0.140)	0.975 (0.017)		4.680 (0.004)	8.286 (0.000)	0.999	1.312
ARMA	-669.0 (678.6)	0.662 (0.159)	0.436 (0.172)	0.956 (0.023)	0.440 (0.069)	2.709 (0.047)	7.348 (0.000)	0.999	1.991
<b>Edvinsson</b>									
OLS	-1212.8 (226.6)	0.588 (0.228)	0.573 (0.235)			164.058 (0.000)	1278.3 (0.000)	0.987	0.062
AR1	-1048.2 (519.8)	0.669 (0.115)	0.468 (0.125)	0.970 (0.016)		6.170 (0.000)	15.65 (0.000)	0.999	1.151
ARMA	-1030.180 (434.757)	0.722 (0.142)	0.422 (0.154)	0.949 (0.021)	0.439 (0.059)	5.019 (0.002)	16.299 (0.000)	0.999	1.918

Note: In the OLS case the test is  $\chi^2$  in the other models F-tests. The full length of the series has been utilized. Maddison's series 1820-2001 and Edvinsson's 1750-2000.

observed data in the usual sense. Nevertheless it is of interest to see whether the projections contain information that is helpful in predicting the available GDP estimates. Indeed we find that to be the case.

In Table 6 the upper half tests against the Maddison series, first using ordinary least squares (though the standard errors are corrected for heteroskedasticity and autocorrelation by the Newey-West method). Clearly the result indicates that both models contain information valuable for the prediction and that they are not encompassing. Using the much longer Edvinsson series as dependent variable yields approximately the same results. Correcting for serial correlation in the residuals both by AR and ARMA corrections does not change the basic impression of that.

In Table 7 we report the mean square relative errors using the full length of the comparison series. When the serial correlation is modeled we achieve unbiased forecasts that reduce the errors quite substantially. When comparing the OLS combination to the

Table 7: Mean square relative errors for forecasts and combined forecasts.

Compared to Maddison					
No of obs	Simulation	Projection	Comb OLS	Comb AR	Comb ARMA
182	0.116	0.312	0.144	0.037	0.036
Compared to Edvinsson					
No of obs	Simulation	Projection	Comb OLS	Comb AR	Comb ARMA
249	0.194	0.657	0.169	0.041	0.042

simulation there is no reduction against the Maddison but against the Edvinsson series there is still a reduction notwithstanding that we know that the projection deviates strongly for more than a hundred years in the beginning of the series. It seems rather obvious that we could make the case for both the approaches being useful even stronger by restricting the tests to a period starting in the later half of the 19th century.

There may, however, be some stationarity problems apart from autocorrelation since all the series are rather obviously trended in a non-linear way. Then our tests here would not be valid since the parameter estimates in that case would not have a standard distribution. Diebold and Lopez recommend that in such cases one should instead use encompassing tests from a specification due to Fair and Shiller (1990). Instead of estimating equation (18) that directly compares the level series one can compare whether the changes in GDP per capita are predicted by the predicted changes.

$$y_{t+k} - y_t = \alpha_0 + \beta_s (\hat{y}_{t+k,s} - y_t) + \beta_p (\hat{y}_{t+k,p} - y_t) + \varepsilon_t \quad (19)$$

Then the risks of spurious regression due to high persistence is removed by the differentiation and the tests are reliable tests for whether both series actually provide information for the prediction.

For obvious reasons neither the simulation nor the projection contain much information for predicting business cycle noise. The simulation is a long-term model with no allowance whatsoever for business cycle movements. The projection also ignores business cycles by conditioning on slow-moving demographic variables. We, therefore, redo the encompassing tests successively increasing the horizon. As we let  $k$  increase the predictive information in the forecasted changes also increases. This is due to the business cycle errors starting to cancel out when the horizon grows.

In Table 8 we report a sample of results from different horizons using an autoregressive model. The tests for encompassing reject the hypothesis that any of the series encompass the other. The results are very consistent over this broad range of horizons (and in fact the combination coefficients are not too dissimilar for every choice of forecasting period). This holds irrespective of which of the historical estimates we have as dependent variable.

Table 8: Results from first-order autoregressive model of GDP per capita changes on forecasted changes. Only the coefficients for the simulated change and the projected change are reported.

Horizon of forecast	2 years	5 years	10 years	20 years	40 years
Dep: Maddison					
$\beta_s$	0.577	0.450	0.382	0.507	0.455
$\beta_p$	0.461	0.438	0.419	0.456	0.463
$F$ -test (0, 1, 0)	0.011	0.004	0.002	0.017	0.020
$F$ -test (0, 0, 1)	0.001	0.000	0.000	0.001	0.004
$\bar{R}^2$	0.787	0.944	0.979	0.994	0.998
Dep: Edvinsson					
$\beta_s$	0.540	0.618	0.485	0.778	0.311
$\beta_p$	0.344	0.391	0.357	0.369	0.377
$F$ -test (0, 1, 0)	0.004	0.017	0.003	0.007	0.000
$F$ -test (0, 0, 1)	0.000	0.000	0.000	0.000	0.000
$\bar{R}^2$	0.794	0.946	0.980	0.994	0.998

Thus we conclude that a combination of the approaches really is useful and will reduce forecasting errors given some control for the serial correlation in the errors. Both models in fact contribute useful information for predicting the Swedish GDP per capita as far as we can judge from their ability to predict the historical estimates. As the weights in the linear combination are roughly not significantly different from each other in most cases a forecast somewhere in between the simulation path and the demographic projection should have a lower forecast error and be more likely than either of the paths suggested by the individual models.

## 5 Conclusion

In this paper we have presented new evidence supporting the idea that demographic change is a key determinant of long-term growth in per capita income. The analysis has used the case of Sweden as a kind of laboratory to test two different approaches, since this is a country for which high-quality demographic and economic data are available from 1750 and onwards. The first piece of evidence is a formal model showing how increasing population scale, the extension of effective working life, declining fertility and induced increases in education can account for the acceleration of Swedish per capita income growth rates after 1750. The second piece of evidence is an empirical model

based on the demographic dividend approach showing that shifts in mortality and age structure can account not only for the growth in per capita income across 111 countries during the last 40 years, but also for the long-term increase in Swedish per capita income from the early 19th century and onwards.

In our view these two approaches are complementary. The formal model has its base in the endogenous growth literature analyzing the role played by mortality decline, fertility change and induced human capital accumulation in the acceleration of economic growth rates. What we have shown here is that an endogenous growth model can reproduce the Swedish long-term per capita income growth as the result of observed shifts in mortality, fertility and education. This gives a firm theoretical underpinning to the proposition that demographic change and human capital responses are key elements in the economic development process.

The empirical model, instead is rooted in the "demographic dividend" literature that focuses on how declining dependency rates have boosted economic growth in countries that have experienced a decline in fertility during the post-war period. Here we have shown that a demographic dividend model estimated on modern data can be used to successfully track the long-term growth experience of a country that experienced the demographic transition much earlier than the developing countries of today. That a model estimated on modern data can be used to successfully backcast economic growth is evidence of a considerable stability in the empirical relationship between demographic structure and growth in per capita income.

The complementarities between the two approaches are further demonstrated by the encompassing analysis presented in section 4. According to the results from this analysis, both models contain useful information about the data generating process. This suggests that a forecast based on a combination of the two models can be expected to perform better than a forecast using just one of the two.

Thus, our analysis of the Swedish growth process has not only considerably strengthened the argument that the fundamental shifts in the human conditions that are associated with the demographic transition are fundamental also for the process of modern economic growth. This leads to a proposal for further research on long term economic forecasts to be based on a combination of formal modeling and more traditional econometric methods.

## References

- Aquilonius, K, and V Fredriksson. 1941. *Svenska folkskolans historia*. Stockholm.
- Armstrong, J. S. 1989. "Combining Forecasts - the End of the Beginning or the Beginning of the End." *International Journal of Forecasting* 5 (4): 585–588.
- Bates, J. M., and Clive W. J. Granger. 1969. "The Combination of Forecasts." *Operations Research Quarterly* 20:451–468. reprinted in T. C. Mills (ed.)(1999), *Economic Forecasting*, Edward Elgar, Cheltenham.



- Blondal, Sveinbjørn, and Stefano Scarpetta. 1997. "Early Retirement in OECD Countries: The Role of Social Security Systems." *OECD-Economic-Studies* 0 (29): 7–54.
- Bloom, David E., David Canning, and Bryan; Graham. 2003. "Longevity and Life-Cycle Savings." *Scandinavian Journal of Economics* 105 (3): 319–338 (September).
- Bloom, David E., and Jeffrey G. Williamson. 1998. "Demographic Transitions and Economic Miracles in Emerging Asia." *World Bank Economic Review* 12 (3): 419–455 (September).
- Boucekkine, Raouf, David de la Croix, and Omar Licandro. 2002. "Vintage human capital, demographic trends and endogenous growth." *Journal of Economic Theory* 104:340–375.
- Boucekkine, Raouf, David de la Croix, and Omar Licandro. 2003. "Early Mortality Declines at the Dawn of Modern Growth." *Scandinavian Journal of Economics* 105 (3): 401–418 (September).
- Boucekkine, Raouf, David de la Croix, and Dominique Peeters. 2007. "Early Literacy Achievements, Population Density and the Transition to Modern Growth." *Journal of the European Economic Association*.
- Boucekkine, Raouf, Omar Licandro, and Christopher Paul. 1997. "Differential-difference equations in economics: On the numerical solution of vintage capital growth models." *Journal of Economic Dynamics and Control* 21:347–362.
- Coale, A.J., and E.M. Hoover. 1958. *Population Growth and Economic Development in Low-Income Countries*. Princeton: Princeton University Press.
- de la Croix, David, and Matthias Doepke. 2003. "Inequality and growth: Why differential fertility matters." *American Economic Review* 93 (4): 1091–1113.
- de la Croix, David, and Omar Licandro. 1999. "Life expectancy and endogenous growth." *Economics Letters* 65 (2): 255–263.
- Deutsch, M., C. W. J. Granger, and T. Teräsvirta. 1994. "The Combination of Forecasts Using Changing Weights." *International Journal of Forecasting* 10 (1): 47–57.
- Diebold, Francis X, and Jose A Lopez. 1996. "Forecast Evaluation and Combination." In *Statistical methods of finance*, Volume 14 of *Handbook of Statistics series*. New York and Oxford: Elsevier, North-Holland.
- Edvinsson, Rodney. 2005. *Growth, Accumulation, Crisis*. Volume 41 of *Stockholm Studies in Economic History*. Stockholm: Almqvist & Wiksell International. PhD thesis.
- Fair, Ray C., and Robert J. Shiller. 1990. "Comparing Information in Forecasts from Econometric Models." *American Economic Review* 80 (3): 375–389 (June).
- Galor, Oded, and David Weil. 2000. "Population, technology, and growth: from the Malthusian regime to the demographic transition and beyond." *American Economic Review* 90 (4): 806–828.
- Granger, Clive W.J., and Ramu Ramanathan. 1984. "Improved Methods of Combining Forecasts." *Journal of Forecasting* 3:197–204.

- Hendry, DF, and MP Clements. 2001. "Pooling of forecasts." *Econometrics Journal* 7:1–31.
- Hofsten, E, and H Lundström. 1976. *Swedish population history: main trends from 1750 to 1970*. Stockholm: LiberFörlag.
- Kalemli-Özcan, Sebnem, Harl Ryder, and David Weil. 2000. "Mortality Decline, Human Capital Investment, and Economic Growth." *Journal of Development Economics* 62 (1): 1–23.
- Kelley, Allen C., and Robert M. Schmidt. 2005. "Evolution of Recent Economic-Demographic Modeling: A Synthesis." *Journal of Population Economics* 18:275–300.
- Krantz, Olle. 1997, March. "Swedish Historical National Accounts 1800-1990 – Aggregated Output Series." manuscript.
- Lagerlöf, Nils-Petter. 2003. "From Malthus to modern growth: can epidemics explain the three regimes?" *International Economic Review* 44 (2): 755–777.
- . 2006. "The Galor-Weil Model Revisited: A Quantitative Exercise." *Review of Economic Dynamics* 9:116–142.
- Lee, Ronald. 2003. "The Demographic Transition: Three Centuries of Fundamental Change." *Journal of Economic Perspectives* 17 (4): 167–190 (Fall).
- Lindahl, Erik, Einar Dahlgren, and Karin Kock. 1937. *National Income of Sweden 1861-1930, Part One and Two*. Volume III of *Wages, Cost of Living and National Income of Sweden, 1860-1930*. London: P.S. King & Son, Ltd.
- Lindh, Thomas, and Bo Malmberg. 1999. "Age Structure Effects and Growth in the OECD, 1950-90." *Journal of Population Economics* 12 (3): 431–449 (August).
- Maddison, Angus. 2003. *The World Economy: Historical Statistics*. Paris: OECD Development Centre.
- Malmberg, Bo, and Lena Sommestad. 2000. "The hidden pulse of history: population and economic change in Sweden: 1820-2000." *Scandinavian Journal of History* 25:131–146.
- Österholm, Pär. 2004. "Estimating the Relationship between Age Structure and GDP in the OECD Using Panel Cointegration Methods." Technical Report, Uppsala University, Department of Economics.
- Phillips, Peter C.B., and Hyungsik R. Moon. 1999. "Linear Regression Limit Theory for Nonstationary Panel Data." *Econometrica* 67 (5): 1057–1111 (September).
- Sjöstrand, W. 1961. *Pedagogikens historia*. Lund: Gleerup.
- United Nations Population Division. 2000. *World Population Prospects: The 2000 Revision*. UN Department of Economic and Social Affairs.