The Economic and Demographic Transition, Mortality, and Comparative Development*

Matteo Cervellati
University of Bologna
IZA, Bonn

Uwe Sunde
University of St.Gallen
IZA, Bonn
CEPR

November 11, 2011

Abstract

This paper presents a unified growth theory of the economic and demographic transition. The dynamic evolution of adult longevity, child mortality, fertility and the education composition of the population is endogenously determined. The paper provides a quantitative investigation of the interdependencies between demographic and economic variables in generating the transition, and of the role of mortality differences for comparative development. The model is calibrated to data from Sweden and evaluated against historical time series and cross-country panel data. The quantitative analysis documents that the unified growth framework can rationalize both the patterns in historical time series data and contemporaneous cross-country panel data, including the bi-modal distribution of the endogenous variables across countries. The quantitative results suggest that differences in exogenous baseline mortality might explain a substantial part of the observed differences in the timing of the take-off across countries and the density distribution of the main variables of interest.

JEL-classification: E10, J10, J13, N30, O10, O40

Keywords: Economic and Demographic Transition, Adult Mortality, Child Mortality, Quantitative Analysis, Unified Growth Model, Heterogeneous Human Capital, Comparative Development, Development Traps.

*The authors wish to thank Graziella Bertocchi, Francesco Billari, Javier Birchenall, Michele Boldrin, Raouf Boucekkine, Michael Burda, Nancy Chau, Matthias Doepke, David de la Croix, Oded Galor, Nezih Güner, Marcus Hagedorn, Omar Licandro, Omer Moav, Giovanni Prarolo, Alexia Przkawetz, Rodrigo Soares, Vegard Skirbekk, Chiara Strozzi, Nico Voigtlaender, Joachim Voth, and Romain Wacziarg, as well as participants at the Workshop on Demographics, Health and Economic Development at Stanford University, the SAET in Vigo, the SED in Vancouver, the VfS Council for Population Economics, the VfS annual conference in Graz, the UN Development Division in New York, UCLA Anderson, the University Carlos III, Madrid, the University of Cologne, the University of Modena, University Milano Bicocca, University of Munich, the University of Salerno, the Vienna Institute of Demography, the Bank of Italy, University Tor Vergata in Rome, the Einaudi Institute for Economics and Finance Rome, the LSE-UCL development seminar, London, Queen Mary University, London, University of Uppala, University of Helsinki, the ESPE Conference in London, the Rags-to-Riches Workshop at Universitat Pompeu Fabra, Barcelona, the Simposio d'Analisi Economico in Zaragoza, the IZA Workshop on Demographic Change, the ASSA meeting in San Francisco, and the Macro Workshop of the Verein für Socialpolitik for helpful discussions and comments. We are also particularly grateful to David de la Croix, Vegard Skirbekk for sharing their data. Matteo Cervellati: m.cervellati@unibo.it, Uwe Sunde: uwe.sunde@unisg.ch.

1 Introduction

From the mid 19th century, European countries experienced fundamental changes in living conditions. Within a few generations mortality and fertility rates dropped to unprecedented levels. At roughly the same time of this demographic transition, an economic transition lead to widespread education and sustained income growth after a stagnant development during the entire previous history. The stylized patterns are very similar across countries and times, including countries that entered their demographic and economic transition much later than the European forerunners.[1] Nonetheless, by 1970 about half of all countries in the world had not yet experienced the onset of the transition, and in 2000 still 40 percent of these countries were trapped in underdevelopment.[2] Several important questions regarding these long-run development patterns still remain open: What are the underlying forces behind the different dimensions of economic and demographic development? Why have some countries developed early on, others only with a delay, and why do many countries still remain trapped in poor living conditions today? What is the role of mortality given that today's underdeveloped countries are predominantly located in areas with a high exposure to infectious diseases?

This paper addresses these questions by providing a unified theory of the economic and demographic transition that allows for a systematic quantitative investigation of the role of the mortality environment for long-run growth. The early contributions in the unified growth literature have modeled the economic transition as the result of a fertility transition that occurred in response to technological change in combination with an education-fertility trade-off, or as resulting from an acceleration in education following a reduction in mortality.[3] We build a prototype unified growth theory by considering an occupational choice framework in which individuals decide about acquiring skilled or unskilled human capital, as well as about their

[1]Demographers such as Kirk (1996) notice that also "in non-European countries undergoing the demographic transition in the mid 20th century, the regularities are impressive". See also Lee (2003), and Galor (2005).

[2]In these countries life expectancy at birth was less than 55 years in 1970, average total fertility was around six children per woman and the share of population with completed secondary education was less than one out of five. Demographers conventionally identify the onset of the demographic transition with life expectancy at birth increasing above 50-55 years and a sustained drop in fertility, see Chesnais (1992).

[3]The works by Galor and Weil (2000), Kögel and Prskawetz (2001), Galor and Moav (2002), and Doepke (2004) investigate the role of fertility for long-term development. Contributions by de la Croix and Licandro (1999), Kalemli-Ozcan, Ryder and Weil (2000), and Boucekkine, de la Croix and Licandro (2002, 2003) study the role of mortality reductions in homogenous human capital models, while Cervellati and Sunde (2005) consider heterogeneous human capital. Chakraborty (2004) investigates the role of mortality for growth through savings decisions. See also the surveys by Galor (2005) and Doepke (2007).

fertility in terms of number and quality of children. Higher adult life expectancy crucially affects the type of human capital acquired and the fertility choice by inducing a larger share of individuals to acquire skilled human capital.[4] Endogenous improvements in life expectancy and skill-biased technological change trigger an economic take-off to sustained growth and eventually lead to a lower net fertility since skilled workers have fewer children.[5]

The theoretical framework is applied in a quantitative analysis of the role of exogenous differences in adult life expectancy for long-run comparative development. The model is calibrated targeting data moments for Sweden, the textbook case of long-run economic and demographic transition. The dynamic simulation of the calibrated model is compared to actual time series data from Sweden for the period 1750-2000 and to cross-country panel data for the period 1960 to 2000. The simulated model features a non-linear path of long-run development in all central dimensions, including education, fertility, longevity and income per capita. The patterns of this development path match closely the historical time series.[6] At the same time, the simulated data also documents the ability of the the unified growth framework to match the patterns of comparative development in cross-sectional panel data.[7] Despite the non-linear development dynamics, the simulated data display monotonic and almost linear correlations between the equilibrium share of educated individuals and all central variables. When interpreted in a cross-sectional perspective the simulated data match the correlations emerging from cross-country panel data for 1960 and 2000. The calibrated model also reproduces the well documented concave relationship between income per capita and life expectancy that is known as "Preston Curve". A further implication of the model, which has not been investigated in the unified growth literature, is

[4]The model essentially nests the unified growth frameworks studied by Galor and Weil (2000) and Cervellati and Sunde (2005) and models child and adult mortality as in Soares (2005). Alternative frameworks to study the role of mortality for fertility include Boldrin and Jones (2002), Lagerlöf (2003), Tertilt (2005), Doepke (2005), Tamura (2006), Kalemli-Ozcan (2009), Chen (2010), and Jones and Schoonbroodt (2010).

[5]The role of differential fertility has been studied by de la Croix and Doepke (2003, 2004) and Doepke (2004), who abstract from mortality. The focus on the role of mortality for fertility and education for the demographic transition is shared with de la Croix and Licandro (2007), who consider investments in health.

[6]The calibrated model also matches the duration of the transition and delivers a drop in gross as well as net fertility as consequence of an increase in life expectancy. This effect has been difficult to rationalize in previous quantitative studies (see Doepke, 2004 and 2005, and Kalemli-Ozcan, 2003 and 2009).

[7]Previous quantitative studies of unified growth models, including Doepke (2004) and Lagerlof (2003, 2006) have concentrated on matching long-run time series patterns. Eckstein, Mira, and Wolpin (1999), Kalemli-Ozcan (2002), de la Croix, Lindh, and Malmberg (2008), and Bar and Leukhina (2010a, 2010b) also provide quantitative investigations of the change in demographic variables for long-run development, albeit not within a unified growth framework. To our knowledge, this paper provides the first attempt to study the quantitative role of life expectancy and the implications of a unified growth theory also from a cross-country perspective.

a hump-shaped relationship between life expectancy and subsequent changes in the education composition of the population. This prediction is also shown to be in line with patterns in cross-country data.

The calibrated model is used to study the quantitative role of mortality by simulating a counterfactual economy that differs from the benchmark calibration only in terms of the exogenous mortality environment (which is calibrated targeting data moments for the highest mortality countries in 2000 rather than for Sweden in 1800).[8] The results document that cross-country differences in baseline longevity can result in substantial delays of the economic and demographic transition.[9] Differences in baseline longevity are shown to leave both the cross-sectional relationships and the Preston Curve essentially unaffected, however. Finally, the quantitative analysis of an artificial world composed of countries that differ in terms of baseline longevity but that are otherwise identical delivers cross-country distributions of all variables of interest that match the actual worldwide distributions in 1960 and 2000 well.[10]

The analysis delivers novel insights to the underlying mechanics of the long-run development process and of comparative development patterns. The results document that the observed patterns in time series, cross-section correlations, and in the cross-country distributions of the variables of interest can all be generated by the same non-linear development process. The findings also suggest that differences in the mortality environment across countries can explain a large part of the observed differences in the timing of the take-off across countries, and hence in the cross-country distribution of the main variables of interest. At the same time, exogenous differences in longevity leave the cross-sectional correlations essentially unaffected. This can help explaining why the empirical role of mortality differences for long-run development is difficult to identify with linear estimation frameworks.[11] In summary, the quantitative results support

[8]The fact that unlike in most other unified growth models the acceleration in income the present model is not generated by the transition from corner to interior solutions allows for smooth comparative statics on the main parameters of interest such as baseline mortality, and facilitates the derivation of cross-sectional predictions.

[9]This finding is in line with recent results by Chakraborty, Papageorgiou, and Sebastian (2010) that suggest that the disease environment can have an important impact on economic development.

[10]The distributions are bi-modal in 1960 and rather uni-modal in 2000 as most countries have experienced the onset of the transition. The bi-modality of the cross-country distribution in the simulation is due to the acceleration in the changes of all variables during the transition to the balanced growth. The quantitative analysis helps rationalizing earlier findings of twin peaks in fertility, in the world income distribution (Azariadis and Stachurski, 2005, for an overview), and in life expectancy (Bloom and Canning, 2007).

[11]Consistent with this view, Cervellati and Sunde (2011a, 2011b) document that accounting for the non-monotonic effects of life expectancy can reconcile the contradictory empirical findings for the growth effect of longevity by Acemoglu and Johnson (2007) and Lorentzen et. al (2008).

the hypothesis that underdeveloped countries follow very similar development processes as the Western forerunners, although with a substantial delay, and suggest that the observed differences in extrinsic mortality environments across countries might explain a substantial part of this delay.

The paper is organized as follows. Sections 2 and 3 introduce the model and derive the analytical predictions while Section 4 presents the quantitative analysis. Section 5 provides a discussion and Section 6 concludes. All proofs are relegated to the Appendix.

2 The Model

This section presents the theoretical framework. The model is presented using the specific functional forms that are applied in the calibration of the model in Section 4, even though the functional form assumptions are not needed for the analytical results in Section 3. Functional forms are specified in line with the previous literature and the available evidence, and to minimize the number of parameters.

2.1 Set up

The economy is populated by a discrete number of generations of individuals denoted by $t \in \mathbb{N}^+$. There are two relevant subperiods in the life of an individual: childhood and adulthood. The duration of childhood is denoted by $K_t = k$ while the duration of adulthood is denoted by T_t. Each individual of generation t survives to age k with probability $\pi_t \in (0, 1)$. Surviving children become adults, survive with certainty until age $k + T_t$, and then die. The variable T_t therefore represents both life expectancy at age k and the maximum duration of adulthood.[12] In the model, T_t is a summary statistic of the effective time available during adulthood. An alternative interpretation of T_t would be as a "health augmented" time endowment of adults.

Reproduction is asexual and takes place at age $m \geq k$, which therefore represents the length of a generation. A generation of adults consists of a mass of agents of size $N_{t+1} = N_t \pi_t n_t$ where n_t is the average (gross) fertility of the parent generation. Every individual of generation t is endowed with an innate ability $a \in [0, 1]$, which is randomly drawn from a distribution $f(a)$ that

[12]This modeling of child survival and adult longevity follows Soares (2005). Empirically, the observed increases in adult life expectancy T are mainly due to the increase in survival probabilities at intermediate ages, rather than maximum lifetime horizon, see, e.g. Boucekkine, de la Croix and Licandro (2003). Considering a deterministic adult life expectancy (that is, a perfectly rectangular survival probability distribution) is without loss of generality but simplifies the description of the model substantially by allowing to abstract from age dependent mortality rates and uncertainty. In the quantitative analysis, $1 - \pi_t$ corresponds to child mortality, and T_t to life expectancy at age five (so that $k = 5$). Assuming a constant death rate before age 5, life expectancy at birth is $\pi_t(5 + T_t)$.

does not change over the course of generations. For the calibration of the model we assume a (truncated) normal ability distribution with mean μ and standard deviation σ.

2.2 Preferences and Production

During childhood individuals are fed by their parents and make no choices. Those individuals that survive childhood make decisions about their own education and their fertility at the beginning of their adulthood to maximize their (remaining) lifetime utility. Individuals derive utility from own consumption, c, and the quality, q, of their (surviving) offsprings πn. As in Soares (2005), the lifetime utility of an individual i of generation t is additively separable and given by,

$$\int_0^{T_t} \ln c_t^i(\tau)\, d\tau + \gamma \ln\left(\pi_t n_t^i q_t^i\right)$$

where $\gamma > 0$ is the weight of the utility that parents derive from their surviving children relative to their own adults lifetime consumption. Following Soares (2005), we set the subjective discount rate and interest rates equal to zero and assume perfect capital markets, which implies that individuals perfectly smooth consumption within their adult period of life, $c_t^i(\tau) = c_t^i$ for all τ. These assumptions allow us to abstract from life cycle considerations and to focus on the long term evolution of the economy. The utility can therefore be expressed as,

$$U(c_t^i, \pi_t n_t^i q_t^i) = T_t \ln c_t^i + \gamma \ln\left(\pi_t n_t^i q_t^i\right) \tag{1}$$

The key feature of this formulation is that individuals can smooth consumption over their adult life, but they cannot perfectly substitute the utility from their own consumption with utility derived from their offsprings.[13]

The only inputs of production are *skilled* human capital, denoted by s, and *unskilled* human capital, denoted by u. We follow the vintage growth literature that treats human capital as inherently heterogenous across generations since individuals acquire their skills in environments characterized by the availability of a particular vintage of technology. The respective aggregate stocks of human capital of both types represent the only factors of production that are used in a constant returns to scale technology to produce a consumable commodity as

$$Y_t = A_t\left[(1-x_t)\left(H_t^u\right)^\eta + x_t\left(H_t^s\right)^\eta\right]^{\frac{1}{\eta}}\ , \tag{2}$$

[13] Assuming that parents derive utility directly from the quality of their children, as in Becker and Lewis (1973), allows us to study the change in the quantity-quality trade-off in the simplest way. For a more articulated analysis of the quantity-quality trade-off that explicitly considers that parents care about the utility of their children see Becker and Barro (1988), Becker, Murphy, and Tamura (1990) and Jones, Schoonbroodt, and Tertilt (2011) (that also present a comprehensive survey of the literature).

where $\eta \in (0,1)$, H_t^u and H_t^s are the aggregate levels of human capital of each type supplied by generation t, $x_t \in (0,1)$ is their relative productivity, and A_t reflects total factor productivity (TFP). Generation t only operates the technological vintage t, which is characterized by relative productivity of the two types of human capital, x_t. The production function (2) is a specialized (CES) formulation of the vintage production function by Chari and Hopenhayn (1991). As in Boucekkine, de la Croix, and Licandro (2002) and Cervellati and Sunde (2005), the vintage of technology is linked to generation-specific knowledge in terms of skilled and unskilled human capital. The returns to human capital of the two types are determined in general equilibrium on competitive markets and equal marginal productivity,

$$w_t^s = \frac{\partial Y_t}{\partial H_t^s} \quad , \quad w_t^u = \frac{\partial Y_t}{\partial H_t^u} \, . \tag{3}$$

While the vintage structure is not needed for the main mechanism, it greatly simplifies the analysis since the optimal choices of acquiring human capital by generation t do not depend on the optimal choices of the (unborn) generations of workers that will enter in the labor market in the future.[14]

The level of human capital acquired by each individual is increasing in the level of innate ability, a, $h^j(a)$ with $dh^j(a)/da \geq 0$ for each $j = \{u, s\}$. Individual ability is relatively more important in producing skilled human capital. As studied below, this delivers a natural equilibrium sorting of the population into skilled and unskilled. For simplicity, we make the assumption that ability only matters for skilled human capital. An individual with ability a acquires $h^s(a) = e^{\alpha a}$ units of human capital if he decides to become skilled, and $h^u(a) = e^{\alpha\mu}$ if he decides to be unskilled.[15] An individual that decides to become skilled, respectively unskilled, pays a fix cost, measured in term of adult time, of $\underline{e}^s > \underline{e}^u \geq 0$.[16]

[14]Vintage models that relax the assumption that human capital is perfectly homogenous across different age cohorts have many convenient technical properties for investigating long term development. The vintage assumption greatly simplifies the computation of the general equilibrium of our model since the dynamic system does not involve a forward looking component in the general equilibrium share of skilled workers. See also Boucekkine, de la Croix, and Licandro (2011) for an extensive discussion of the vintage human capital literature.

[15]Assuming that individuals with average ability can produce the same quantities of skilled or unskilled human capital is a natural benchmark. This normalization also implies that the average quantity of human capital of each type would be the same if acquired by the entire population (so that there are no scale effects associated with the acquisition skilled human capital during the process of development). The relative importance of ability for the level of skilled and unskilled human is irrelevant for the results, however. As discussed below, irrespective of the relative productivity of ability for the two skills, the economy passes from an equilibrium where almost all individuals are unskilled to an equilibrium where almost all individuals are skilled during the process of development.

[16]More complex skills may involve more complex and costly processes of skill acquisition and maintenance. The

As in Galor and Weil (2000), raising a child involves a time cost $r_t = \widetilde{r}_t + \underline{r}$ where $\underline{r} > 0$ is a fix time cost that needs to be spent and $\widetilde{r}_t \geq 0$ is the extra time that can be spent voluntarily in addition.[17] The time spent increases the child's quality according to,

$$q_t\left(\underline{r}, \widetilde{r}_t, g_{t+1}\right) = \left[\widetilde{r}_t \delta \left(1 + g_{t+1}\right) + \underline{r}\right]^{\beta} \tag{4}$$

where $g_{t+1} = (A_{t+1} - A_t)/A_t$, $\beta \in (0,1)$, and $\delta > 0$. The functional form (4) implies a complementarity between technical progress and the effectiveness of the extra time invested in children's (the quality time $\widetilde{r}_t$). As discussed in more details below, this formulation captures in the simplest way that faster technological progress increases the incentives to invest more time in raising children, as in Galor and Weil (2000).

2.3 Mortality and Technological Change

Adult Life Expectancy and Child Survival. A large body of evidence documents that environmental factors, in particular macroeconomic conditions, are crucial determinants of individual health. Child and adult mortality appear to be affected by the macro environment in different ways, however. The evidence reported by Cutler et al. (2006) suggests that human capital is more important for adult longevity than per capita income since adult longevity depends on the ability to cure diseases and is related to the level of medical knowledge. Better living conditions in terms of higher incomes, but also in terms of access to water and electricity, are more important for increasing the survival probability of children, see Wang (2003) for a survey.

In line with this evidence, we consider a differential impact of human capital and income on adult and child mortality. Adult longevity of generation t is assumed to be increasing in the share of skilled individuals in the parent generation,

$$T_t = \Upsilon\left(\lambda_{t-1}\right) = \underline{T} + \rho \lambda_{t-1} \tag{5}$$

where $\underline{T}$ is the baseline longevity that would be observed in the economy in the absence of any skilled human capital, and $\rho > 0$ reflects the scope for improvement. Since $\lambda \in (0,1)$, the maximum level of adult longevity is given by $\overline{T} = \underline{T} + \rho$.

---

crucial feature for the mechanism is that workers who decide to be skilled face a lower effective lifetime that is available for market work during their adulthood. To capture this feature in the simplest way, we follow Cervellati and Sunde (2005) and consider a fix cost in terms of time.

[17] Both increase children's quality but with different relative intensity. The cost $\underline{r}$ can be interpreted as the minimum investment required for the children to have a chance to survive to adulthood and may include feeding (or dressing) the child. The extra investment $\widetilde{r}_t$ can be interpreted as quality time that is not needed to insure the survival like e.g. talking, playing or reading a book with the child.

The child survival probability π_t depends on living conditions at the time of birth reflected by per capita income,

$$\pi_t = \Pi\left(\lambda_{t-1}, y_{t-1}\right) = 1 - \frac{1-\underline{\pi}}{1+\kappa\lambda_{t-1}y_{t-1}} \tag{6}$$

with $\kappa > 0$ and where $1 > \underline{\pi} > 0$ is the baseline child survival that would be observed in an economy with $\lambda_{t-1}y_{t-1} = 0$. Larger total income Y_{t-1} improves the probability of children reaching adulthood while population size N_{t-1} deteriorates living conditions and reduces child survival rates.[18]

Technology. Technological progress is reflected in the emergence of a new vintage of technology characterized by TFP, A_t, and a higher relative weight of skilled human capital in the production process, x_t. In the tradition of Nelson and Phelps (1966), we consider technological progress that is biased towards high-skill intensive production and that depends on the available skilled human capital. The relative productivity of skilled human capital in production, x_t, increases with the share of skilled workers in the previous generation, λ_{t-1}, and with the scope for further improvement, $1 - x_{t-1}$,

$$\frac{x_t - x_{t-1}}{x_{t-1}} = X\left(\lambda_{t-1}, x_{t-1}\right) = \lambda_{t-1}(1 - x_{t-1}) \,. \tag{7}$$

For any λ_t, improvements are smaller as x_t converges to its upper limit at $x = 1$.

Finally, improvements in total factor productivity, A_t, are increasing with the share of skilled workers in the previous generation,[19]

$$g_{t+1} = \frac{A_{t+1} - A_t}{A_t} = G\left(\lambda_t\right) = \phi\lambda_t \quad , \quad \phi > 0 \,. \tag{8}$$

3 Analytical Results

This section derives the analytical results starting from the optimal fertility and education decisions in partial equilibrium. We then characterize the intra-generational general equilibrium and the dynamic evolution of the economy overtime.

[18] Considerable evidence documents the negative effect of population density and urbanization on child mortality, especially during the early stages of the demographic transition, see e.g. Galor (2005).

[19] This specification is also in line with endogenous growth models such as Aghion and Howitt (1992) if ϕ is interpreted as average size of an innovation and the labor involved in research is increasing in λ_t.

3.1 Individual Decision Problem

The total time available during adulthood is limited by adult longevity T_t, or by some exogenous limit to the number of years in the labor market (e.g., due to retirement), $R > 0$.[20] The effective time available for productive activities during adulthood is therefore bounded from above by $\overline{T}_t = \min\{T_t, R\}$. An individual i with education $j = \{u, s\}$ cannot use more than the available time and cannot spend more than the total earnings for total consumption, so that

$$\overline{T}_t \geq l_t^i + \underline{e}^j + \pi_t n_t^i r_t^i \,, \tag{9}$$

and

$$l_t^i w_t^j h_t^j(a) \geq T_t c_t^i \,, \tag{10}$$

where l_t^i is the total time spent working. Given the utility function (1) both constraints will be binding at the optimum. Combining (9) and (10) delivers the budget constraint conditional on being skilled or unskilled, $j = \{u, s\}$,

$$T_t c_t^i = \left(\overline{T}_t - \underline{e}^j - \pi_t n_t^i r_t^i\right) w_t^j h_t^j(a) \tag{11}$$

The problem of an individual i with ability a born in generation t is to choose the type of human capital he wants to acquire, $j \in \{u, s\}$, the number of children, n_t^i, and the time invested in raising each child, r_t^i, so as to maximize utility (1) subject to (11). This is equivalent to maximizing

$$T_t \ln\left[(1/T_t)\left(\overline{T}_t - \underline{e}^j - \pi_t n_t^i r_t^i\right) w_t^j h_t^j(a)\right] + \gamma \ln\left(\pi_t n_t^i q_t^i\right) \,. \tag{12}$$

Optimal Fertility and Time Spent in Children. The optimization problem is strictly globally concave and the first order conditions uniquely identify the optimal fertility level and the time spent raising children conditional on the type of human capital.

Lemma 1. *For any $\left\{w_t^j, T_t, \pi_t, g_{t+1}\right\}$, the optimal fertility of an individual acquiring human capital $j = \{u, s\}$ is given by,*

$$n_t^{ij} = \frac{\gamma\left(\overline{T}_t - \underline{e}^j\right)}{\left(T_t + \gamma\right) r_t^{ij} \pi_t} \tag{13}$$

where r_t^{ij} is given by,

$$r_t^{ij} = r_t^* = max\left\{\underline{r}, \frac{1 - \left[1/\left(\delta(1 + g_{t+1})\right)\right]}{1 - \beta}\underline{r}\right\} \tag{14}$$

[20] The assumption of a limit R, which may be due to a compulsory retirement or some effective limitation to labor force participation at old ages, is not needed for the main results but adds a realistic feature for the analysis of the quantitative role of bounds to productive life when longevity increases to old ages. In the quantitative analysis, the parameter R is calibrated exogenously to match the effective retirement age.

Fertility decreases with child survival (since individuals only care about surviving children) but increases with life expectancy (through a positive income effect) as long as $T_t < R$. Further increases in longevity above R (so that $\overline{T}_t = R$) lead to a reduction in fertility, however. The reason is that a longer old age life increases the marginal utility of lifetime income since more income is needed to deliver a constant consumption over the life cycle, thereby lowering fertility as a longer expected time in retirement increases the weight on private consumption. Fertility is decreasing with the time invested in children, in line with a standard quantity-quality trade-off. The optimal time spent raising each child does not depend on the type of human capital acquired by parents and displays the same features as the mechanism of Galor and Weil (2000). When technical progress is too low parents may optimally decide not to invest any extra time in raising their children beyond the minimum level, so that $r_t^{ij} = \underline{r}$. If a positive extra time is invested in raising children then a larger level technological progress g_t increases r_t^* and reduces optimal fertility for unskilled and skilled individuals.[21] The quantity-quality trade-off and the optimal investment in children's quality does not depend on adult longevity and child mortality.[22]

Optimal Type of Skills. Agents with higher ability have a comparative advantage in acquiring skilled human capital. For any vector of wages there exists a unique ability threshold for which the indirect utilities from acquiring the two types of human capital are equal.

Lemma 2. *For any* $\{w_t^s, w_t^u, T_t, \pi_t\}$ *there exists a unique* $\widetilde{a}_t$ *implicitly defined by*

$$\frac{h_t^s(\widetilde{a}_t)}{h_t^u} = \left(\frac{\overline{T}_t - \underline{e}^u}{\overline{T}_t - \underline{e}^s}\right)^{\frac{T_t+\gamma}{T_t}} \frac{w_t^u}{w_t^s} \tag{15}$$

such that all individuals with $a \leq \widetilde{a}_t$ *optimally choose to acquire unskilled human capital* $j = u$ *while all individuals with* $a > \widetilde{a}_t$ *acquire skilled human capital* $j = s$.

Consequently, for any distribution of abilities $f(a)$, there is a unique share of individuals λ_t that optimally acquire skilled human capital which is given by,

$$\lambda_t := \int_{\widetilde{a}_t}^1 f(a)\, da \tag{16}$$

From (15) and (16), the share of skilled workers λ_t is increasing in the relative wage w_t^s/w_t^u, decreasing in $\underline{e}^s$, increasing in adult longevity T_t, and is unaffected by child mortality π_t.

[21] From (14), there is a $\underline{g} > 0$ defined by $r_t^*(\underline{g}) = \underline{r}$ such that for any $g_{t+1} > \underline{g}$ then $r_t^* > \underline{r}$ and $dr_t^*/dg_{t+1} > 0$.

[22] See also Doepke (2005), Moav (2005), Hazan and Zoabi (2006) and Jones, Schoonbroodt, and Tertilt (2011).

The Effects of Mortality on Education Composition and Fertility. As a consequence of Lemma 2, the average fertility in the population is given by

$$n_t^* = N\left(T_t, \lambda_t, \pi_t\right) = \frac{\gamma}{\left(T_t+\gamma\right) r_t^* \pi_t}\left[(1-\lambda_t)(\overline{T}_t - \underline{e}^u) + \lambda_t(\overline{T}_t - \underline{e}^s)\right] \tag{17}$$

where r_t^* is given by (14).

Gross fertility is decreasing in π_t through a *substitution effect* but net fertility is independent of π_t. From (15), child mortality does not affect individuals' choices regarding their own skill acquisition choices either. The effect of adult longevity on fertility is more complex. From (13), higher adult longevity T_t increases gross fertility as long as $T_t < R$ due to a positive *income effect*, but decreases gross fertility when $T_t \geq R$ as the income effect turns negative. In addition, a higher T_t reduces fertility by a *differential fertility* effect since it increases the share of skilled workers (reflected by a higher λ_t) who, everything else equal, have fewer children.[23] If this composition effect is large enough then both average gross and net fertility can decrease following an increase in adult life expectancy even if $T_t < R$. A further indirect effect arises through the externality of the share of skilled individuals on growth. Although the quantity-quality trade-off is not directly affected by adult longevity and child mortality, parents substitute quantity for quality in the face of technological progress, which depends on λ_t. The effect of T_t on the skill composition therefore reduces fertility also by indirectly affecting the future parental investments in the quality of children.

3.2 Intra-Generational General Equilibrium

From (3), the ratio of competitively determined wages is

$$\frac{w_t^u}{w_t^s} = \frac{1-x_t}{x_t}\left(\frac{H_t^s}{H_t^u}\right)^{1-\eta} . \tag{18}$$

The aggregate levels of the two types of human capital supplied by generation t are given by,

$$H_t^u = N_t \int_0^{\widetilde{a}_t} h_t^u f(a)da \qquad \text{and} \qquad H_t^s = N_t \int_{\widetilde{a}_t}^1 h_t^s(a) f(a)da\,. \tag{19}$$

Since there is a one-to-one relationship between the share of skilled workers λ_t and the threshold ability $\widetilde{a}_t$, H_t^u is decreasing in λ_t. For any $\{T_t, \pi_t, x_t\}$, the general equilibrium of generation t is characterized by a λ_t that is compatible with the vector of market wages. From (15), λ_t is increasing in w_t^s/w_t^u while the market wage ratio (18) is decreasing in λ_t. Hence, there is a

[23]This differential fertility effect is in line with a large body of evidence from different countries and epochs, see Bengtsson and Dribe (2006), Soares (2006), and Skirbekk (2008) for a meta-analysis of the demographic literature.

unique equilibrium share λ_t^* which is obtained by substituting (19) and the wage ratio (18) into (15). This equilibrium share is implicitly characterized by the ability threshold $\widetilde{a}_t$ in

$$\frac{h_t^u \left(\int_{\widetilde{a}_t}^1 h^s(a)f(a)da\right)^{1-\eta}}{h^s(\widetilde{a}_t)\left(\int_0^{\widetilde{a}_t} h_t^u f(a)da\right)^{1-\eta}} = \frac{x_t}{1-x_t}\left(\frac{\overline{T}_t - \underline{e}^s}{\overline{T}_t - \underline{e}^u}\right)^{\frac{T_t+\gamma}{T_t}} \tag{20}$$

as shown in the Appendix. Equation (20) implicitly identifies a unique equilibrium share of skilled workers as function of adult longevity and technology, which is denoted by

$$\lambda_t^* = \Lambda(T_t, x_t)\,. \tag{21}$$

Proposition 1. *For any generation t with $\{T_t \in (\underline{e}^s, \infty), \pi_t \in (0,1), x_t\}$ there exists a unique λ_t^*, characterized by* (21)*, and a unique equilibrium vector, $\left\{H_t^{j*}, w_t^{j*}, n_t^{j*}\right\}$ for $j = u, s$, for which* (15)*,* (18) *and* (19) *jointly hold. The equilibrium share of skilled individuals λ_t^* is an increasing function of T_t, with slope zero for $T \searrow \underline{e}^s$ and $T \nearrow \infty$.*

The key state variables affecting λ_t^* are adult longevity T_t and the relative importance of the different types of human capital in the production function, x_t. An increase in T_t leads to an increase in the share of skilled individuals λ_t^*. The effect of T_t on λ_t^* is non-linear, however. For low T_t the share of population investing in skilled human capital is small due to the fix cost $\underline{e}^s > \underline{e}^u$, which prevents a large part of the population from receiving sufficient lifetime earnings when becoming skilled. When T_t is low the locus $\Lambda(T_t, x_t)$ is convex and large increases in T_t are needed induce a significant fraction of individuals to acquire skilled human capital. At the opposite extreme, if T_t is very large, the locus $\Lambda(T_t, x_t)$ is concave making large improvements in T_t necessary to induce further increases in λ_t due to the decreasing returns to human capital of either type, which drive down the relative wage w^s/w^u.[24] To shorten notation we denote by λ_t the equilibrium share of skilled workers in the following.

3.3 Dynamics

Adult longevity T_t, given in (5), is a function of the share of skilled individuals in the parents' generation. The evolution of x_t is characterized by (7). The share of skilled, λ_t is determined by the intra-generational equilibrium implied by Proposition 1. The total factor productivity, A_t,

[24] That λ_t^* is flat for $T_t = \underline{e}^s$ and $T_t = \infty$ does not depend on the presence of retirement as shown in the Appendix. Characterizing analytically the second derivative of $\lambda(T_t)$ is not possible at this level of generality. That there is only one inflection point, so that $\lambda(T_t)$ is increasing and s-shaped, in the parametrization used in the calibration in Section 4 can be shown numerically and can be established analytically when imposing simplifying assumptions on the ability distribution, see Cervellati and Sunde (2008).

evolves as in (8). Child survival probability, π_t, evolves according to (6) and also depends on y_{t-1} and, therefore on T_{t-1}, x_{t-1}, λ_{t-1}, and A_{t-1}. Fertility is determined in (17). The dynamic path of the economy is therefore given by a sequence $\{T_t, x_t, \lambda_t, A_t, \pi_t, n_t\}$ for $t = [0, 1, .., \infty)$, which results from the evolution of the nonlinear first-order dynamic system,

$$\begin{cases} T_t &= \Upsilon(\lambda_{t-1}) \\ x_t &= X(x_{t-1}, \lambda_{t-1}) \\ \lambda_t &= \Lambda(T_t, x_t) \\ A_t &= A_{t-1}\left(1 + G(\lambda_{t-1})\right) \\ \pi_t &= \Pi(T_{t-1}, x_{t-1}, \lambda_{t-1}, A_{t-1}) \\ n_t &= N(T_t, \lambda_t, \pi_t) \end{cases} \tag{22}$$

The system is block recursive. Baseline longevity $\underline{T}$ and the past level of the share of skilled workers, λ_{t-1}, determine adult longevity T_t, which in turn affects the current share of skilled workers and technological change. Productivity A_t and child mortality π_t only depend on past levels of the variables and do not affect the evolution of the dynamic system (22) in terms of the crucial state variables T_t, λ_t and x_t. There are no scale effects in the dynamics so that the level of fertility, n_t, does not affect their evolution either.

The Economic and Demographic Transition. From (7) and (8), the endogenous skill biased technical change leads to a monotonic increase in the importance of skilled human capital and total factor productivity overtime. Also, from (8), the growth rate of technology increases with the share of individuals acquiring skilled human capital and is bounded from above.

Lemma 3. *The TFP, A_t, and the relative productivity of skilled human capital x_t increase monotonically over generations with $\lim_{t\longrightarrow\infty} x_t = 1$, $\lim_{t\longrightarrow\infty} A_t = +\infty$ and $\lim_{t\longrightarrow\infty} g_t = \phi$.*

The dynamic evolution of the economy, given by the system (22), exhibits an endogenous economic and demographic transition along the long-run development path.

Proposition 2. [Economic and Demographic Transition] *Considering a sufficiently low x_0, the development path of the economy is characterized by:*
(i) An initial phase with few individuals acquiring skilled human capital, $\lambda \simeq 0$, low longevity, $T \simeq \underline{T}$, large child mortality $\pi \simeq \underline{\pi}$, slow income growth, and gross fertility given by,

$$n \simeq \gamma \frac{\underline{T} - \underline{e}^u}{(\underline{T} + \gamma)\, \underline{r}\ \underline{\pi}}\ . \tag{23}$$

(ii) A final phase of balanced growth in income per capita, with large life expectancy, $T \simeq \overline{T}$, low child mortality $\pi \simeq 1$, almost the entire population acquiring h^s human capital, $\lambda \simeq 1$ and

$$n \simeq \gamma \frac{min\{\overline{T}, R\} - \underline{e}^s}{(\overline{T} + \gamma)\, \overline{r}} \; . \tag{24}$$

where $\overline{r}$ is obtained from (14) evaluated at $g_{t+1} = \phi$ and $\overline{T}$ from (5) evaluated at $\lambda = 1$.

From Proposition 1, the equilibrium share of skilled workers $\lambda_t = \Lambda(T_t, x_t)$ is an increasing but non-linear function. The shape of this equilibrium locus depends on the relative productivity of the different types of human capital, x_t, which affects relative wages. The lower x_t, the flatter the function Λ and the lower the equilibrium share λ for any given T_t. The skill-biased technological change increases x_t across generations, which makes the function (21) successively steeper. From (7) the importance of skilled human capital increases over the course of generations, although initially the improvements are small. This process involves reinforcing feedbacks between increases in human capital, and mortality reductions and technological progress. The economy eventually converges endogenously to a sustained growth path, maximal adult longevity, minimal child mortality, and virtually the entire population acquiring skilled human capital. The dynamics of fertility, n_t, along the transition process results from (17) given the realized levels of T_{t-1}, x_{t-1}, λ_{t-1}, and A_{t-1}. The interaction between adult longevity and the share of skilled therefore determines the timing of the transition to the balanced growth path and may affect the patterns of comparative development across countries, whereas fertility and child mortality do not affect the dynamics of the economy. In the dynamic system (22) all variables are characterized by interior solutions although the speed of their dynamics changes vary over time until the balanced growth path is reached.

4 Quantitative Analysis

We calibrate the model to match data moments for Sweden in 2000 and around 1800. The simulated data generated by the calibrated model are then compared to the historical time series for Sweden over the period 1750-2000 in order to investigate the fit of long-term development dynamics, as well as to cross-country panel data for the period 1960-2000 to analyze the relevance for cross-sectional patterns of comparative development. Finally, we perform controlled variations in baseline adult longevity to study the quantitative role of mortality for the timing of the take-off from quasi-stagnation to sustained growth, for comparative development, and for the worldwide distribution of the variables of main interest.

4.1 Calibration

The calibration of the model requires setting the values of fifteen parameters that characterize the utility and production function $\{\gamma, \eta\}$, technological progress ϕ, adult life expectancy $\{\underline{T}, \rho\}$, child survival $\{\underline{\pi}, \kappa\}$, skill acquisition $\{\underline{e}^u, \underline{e}^s, \alpha\}$, the distribution of ability $\{\mu, \sigma\}$, and the quality of children $\{\beta, \underline{r}, \delta\}$. In addition, as discussed above we allow for the possibility that individuals retire at some exogenously given age R, which also needs to be pinned down. Finally, the age at reproduction m (corresponding to the length of one generation) and two initial conditions for technology, A_0 and x_0, need to be specified. For a given set of parameters and initial conditions, A_0 and x_0, the evolution of all the variables of interest are endogenously determined by the model along the development path in all periods $t = \{0, 1, .., \infty\}$.

The parameters m, R, and η, as well as the initial condition for x_0 are set exogenously. The other parameters and initial conditions are set endogenously to match data moments for Sweden in 2000 (ϕ, α, μ, σ, γ, A_0), or to match data moments for Sweden in 1800 and 2000 ($\underline{e}^u$, $\underline{e}^s$, β, $\underline{r}$, δ, $\underline{T}$, ρ, $\underline{\pi}$, κ). To study of the role of mortality for comparative development, we calibrate an alternative scenario of baseline adult longevity, $\underline{\underline{T}}$, that reflects the worst mortality environment across countries. This calibration targets data moments for pre-transition countries with the highest observed adult mortality in 2000.[25] Finally, we simulate an artificial world composed of 113 countries that are identical in all parameters except for baseline adult longevity, which is distributed in the interval $\{\underline{\underline{T}}, \underline{T}\}$.

The data moments used as targets, the data sources and the calibrated parameters are summarized in Table 1.

Benchmark Calibration. The duration of a generation, the age of retirement and the parameter of the production function are set exogenously.

Length of a generation. Across countries the average age (of women) at first birth before the demographic transition is approximately 20 years.[26] A twenty year frequency also allows for a direct match of the simulated data across different generations with cross-country panel data without the need of interpolation. We therefore set $m = 20$.

Age of retirement. The average effective retirement age was around 64 in Sweden in 2000.[27]

[25] To explore the role of the endogenous cost of raising children in explaining fertility differences across countries, we also consider an alternative calibration of the quantity-quality trade-off by targeting data moments for pre-transitional countries with the highest recorded fertility in 2000.

[26] Mean age at first birth in Sweden around 1800 was slightly higher, see Dribe (2004), while age at first birth is still below 20 in pre-transitional countries in Africa nowadays, see Mturi and Hinde (2007).

[27] The data source is OECD, see `http://www.oecd.org/dataoecd/3/1/39371913.xls`.

Since R is the number of years before retirement at age $k = 5$, we set $R = 59$.

Production Function. The elasticity of substitution between skilled and unskilled workers is typically taken to be $1/(1-\eta) = 1.4$ in the literature, see Acemoglu (2002), which implies setting $\eta = 0.285$.

The remaining parameters are set endogenously by matching model moments to data moments for Sweden which is the prototypical example of the economic and demographic transition and represents a natural benchmark for the calibration and for testing the quantitative fit of the model in terms of long-term development patterns. The demographic and economic data for Sweden are available since the mid 18th Century and are of comparably high quality.

From Lemma 3, $\lambda_\infty \simeq 1$. The data on education document that the enrolment shares in Sweden have essentially reached 100% in primary and lower secondary education after 1980 and 1995, respectively.[28] Consistent with this evidence, we assume that the transition to the balanced growth path is essentially completed by 2000 and in the calibration we take 2000 to be the year in which λ takes a value arbitrarily close to 1.[29] The determination of some parameters requires the solution of a system of simultaneous equations that target data moments in 2000 and before the onset of the transition, which in the case of Sweden occurred in the first decades of the nineteenth century. In these cases we use data moments for Sweden in the period around 1800 with a target for the share of skilled workers of $\lambda = 0.1$, which roughly corresponds to the enrolment rates in early 19th Century Sweden.[30]

Technological Progress. The parameter governing the evolution of TFP, ϕ, is set to match the average annual growth rate of income per capita on the balanced growth path after the transition (which equals the growth rate of technological change). The average growth rate in Sweden over the period 1995-2010 has been about 2.4 percent per year.[31] This implies targeting a growth factor of 1.61 over a twenty-year period. Given the function (8), and with $\lambda = 1$ along the balanced growth path, we set $\phi = 0.61$.

Human Capital and Ability Distribution. The calibration of the time cost associated with the different skills requires setting the values of $\{\underline{e}^u, \underline{e}^s\}$. These parameters are set by targeting

[28]See also Ljungberg and Nilsson (2009) and de la Croix et al. (2008).

[29]Specifically, in the simulation, 2000 corresponds to the first generation for which λ exceeds 0.999.

[30]The available data sources provide slightly heterogeneous information on enrolment rates in the early 19th Century Sweden, with estimates ranging from about 5 to about 15 percent, see de la Croix et al. (2008) and Ljungberg and Nilsson (2009). The precise value of λ before the transition is of little importance, however. The results for alternative parameters obtained by targeting levels of λ up to 0.3 are essentially the same.

[31]See, e.g., the ERS Dataset (`www.ers.usda.gov`) or historical statistics from the Bank of Sweden (`www.historicalstatistics.org/`).

the average years of schooling in 2000 and in 1800. The average years of schooling in Sweden for the cohort age 25-35 was 12 years in 2000, see Lutz et al. (2007), while the earliest available data suggest around 1 year of schooling on average before or around the onset of the transition.[32] Targeting one year (or less) of schooling on average before the transition, and 12 years of schooling on average for 2000 implies setting $\underline{e}^s = 12$ and $\underline{e}^u = 0$.

The parameter α, relating to the importance of ability for individual human capital, and the moments of the ability distribution $\{\mu, \sigma\}$ are calibrated by targeting the income distribution in Sweden by 2000. By 2000, the income distribution in the model results from the income distribution of skilled workers, since $\lambda \simeq 1$. We estimate the income distribution using micro data from the ECHP dataset for individual incomes of full-time employees aged 25 to 45, which corresponds to the two last cohorts in the dynamic simulation, and equivalently to the two first generations with $\lambda = 1$ in the data. The income distribution is approximately log-normal between the 5th and 95th percentile of the data, with slightly thicker tails. The distribution of log incomes has mean 9.7, standard deviation 0.4, and the lowest and highest observed log-incomes are 6.7 and 12.8, respectively, which implies a maximum spread of 6.1.[33]

The individual (per period) income earned by a skilled worker is given by $w_t^s \cdot e^{\alpha a}$, which implies that individual log income is given by $\ln w_t^s + \alpha a$. The assumption of a normal distribution of ability (truncated to lie within a finite interval) and the exponential production function of human capital together imply that for $\lambda = 1$ the distribution of income in the model is also approximately log-normal with thicker tails due to the truncation. With $a \in [0, 1]$, the observed difference between the lowest and the highest income in the data, (12.8-6.7) is matched by setting

[32]The estimates are slightly lower when referring to the entire population alive in Sweden in 2000 since older cohorts are included (for instance 11.4 in the data of Barro and Lee, 2001 and 11.5 years in Ljungberg and Nilsson (2009)). Regarding pre-transitional education levels, the estimates differ somewhat more. Ljungberg and Nilsson (2009) report 1.03 years of schooling in the total Swedish population aged 15-65 in 1870, and 0.1 average standard school years of the population aged 7-14 around 1810-1820, considering absenteeism and length of school years.

[33]Incomes are converted US-$ using an average exchange rate of 9 Kroner for one US-$ in 2000. The moments of the income distribution for the age cohort 25-65 are essentially the same, with the lowest, mean, and highest levels of log income being 6.7, 9.7, and 12.8, respectively, and with a standard deviation of 0.41. The ECHP data are based on surveys and refer to total net income from work, which might explain the small differences between the log income per capita from macro data, which is approximately 10 in 2000, and the mean log income from the micro data that is about 9.7. The relevant data moments extracted from this data set are broadly consistent with other data sources based on register data and alternative surveys for gross earnings, see, Domeij and Floden (2010). The data moments are also close to the ones typically used for the calibration of dispersion in permanent incomes in other OECD countries. For instance, Erosa et al. (2011) match a variance of log permanent earnings in the US of 0.36. Robustness checks show that the results are fairly insensitive to varying the dispersion.

$\alpha = 6.1$. Matching the other data moments therefore requires setting $\alpha\mu = 3$ and $\alpha\sigma = 0.4$, which for $\alpha = 6.1$, implies $\mu = 0.49$ and $\sigma = 0.066$.[34]

Adult longevity. The baseline mortality parameter, $\underline{T}$, and the parameter linking adult life expectancy to human capital, ρ, are calibrated targeting the levels of adult longevity in 2000 and 1800. Life expectancy at age five in Sweden was approximately 76 in 2000 and 48 around 1800.[35] With these targets, the parameters of the function (5) are set to $\underline{T} = 45$ and $\rho = 31$.

Child survival probability. Child mortality in Sweden fluctuated around one third in the period 1760-1800 and was about 0.004 in 2000. Targeting a child survival probability 0.67 and 0.996 for 1800 and 2000, respectively, and using condition (6) delivers a baseline child survival $\underline{\pi} = 0.5$ and $\kappa = 0.005$.[36]

Preferences. The parameter γ is calibrated by targeting gross fertility $n = 1$ along the balanced growth path, which is also equivalent to targeting the net reproduction rates approximately at replacement levels, with child survival at $\pi = 0.996$.[37] The time spent in raising children is determined endogenously in the model and changes overtime with the growth rate of income and technology. We set a target for the number of years spent raising a child in 2000 of $r = 5$ in line with the estimates by Haveman and Wolfe (1995).[38] Solving for γ from (13) with $\lambda = 1$, $\pi = 0.996$, $R = 59$, and $r = 5$ as the respective data moments for 2000 delivers $\gamma = 9$.[39]

[34] It is worth noting that the distribution of cognitive ability (or IQ), which is generally measured in the literature as a truncated normal with mean 100 and standard deviation 15, see, e.g., Neisser et al. (1996), would imply a very similar parametrization when normalized for a support $a \in [0, 1]$, with $\mu = 0.5$ and $\sigma = 0.075$.

[35] The average of life expectancy at age five in the period 1760-1840 was 48.38, in the period 1790-1810 it was 48.06 (Human Mortality Data Base available at `http://www.mortality.org/`). Similar figures are documented for England, France and Italy, see Woods (1997) and Bideau et al. (1997) and Lewis and Gowland (2007). Also note that, as discussed below, in 2000 child mortality is around 0.004, which explains the convergence of life expectancy at 5 plus five years of 80.74, and of life expectancy at birth of 80.45.

[36] The data on child survival are available at: `http://www.mortality.org`. The levels of income per capita needed for the computation of the parameters are taken from the historical statistics from the Bank of Sweden (`www.historicalstatistics.org/`), converted to US-$ using an average exchange rate in 2000 of 9 Kroner for one US-$. The income levels used for the calibration of condition (6) are 22,717 and 884 US-$, which correspond to the GDP per capita of Sweden in 2000 and 1800, respectively, in US-$ per 2000.

[37] Total fertility rates (TFR) in Sweden were on average 1.8 children per woman over the period 1980-2000, with substantial fluctuations. In 1990, the TFR was 2.13, whereas in 2000 it was 1.54 (World Development Indicators). These figures suggest that a gross fertility of 1 (which would correspond to a TFR of 2) along the balanced growth path is a reasonable target. Targets in the range from 0.75 to 1.1 deliver very similar results, however.

[38] This is equivalent to setting a target for the share of work life that is spent in raising a child is about 15 percent which is in line with Doepke (2004) and de la Croix and Doepke (2003).

[39] Given the utility function (1) the weight of children relative to own lifetime consumption changes with T_t, as in Soares (2005). For $\gamma = 9$ the relative weight of children compared to per period consumption, γ/T_t, drops

Production function of children's quality. To calibrate the parameters of the function that determines the quality of children as outcome of parental investments, (4), we use the optimal time investment by parents in children, (14), and the minimum growth rate of technology $\underline{g}$ for which parents spend no extra time in raising children (as characterized in footnote 21). The three parameters $\{\beta, \underline{r}, \delta\}$ are calibrated targeting the levels of gross fertility for Sweden in 1800 and 2000, and the growth rate of technology around the period of the exit from the corner solution of zero investments in children's quality. Gross fertility in Sweden in 1800 and 2000 was $n = 2.3$ and $n = 1$, respectively. A noticeable drop in gross fertility occurs around 1900.[40] In the absence of reliable information on parental investments in quality time we take 1900 to be the period of the exit from the corner solution of the quantity-quality trade-off. We therefore target the growth level of productivity in 1900 to set the level of $\underline{g}$. Estimates of TFP and income per capita growth around 1900 vary between 0.7 and 1.7 percent per year.[41] An average of 1.2 percent delivers a corresponding growth over a 20-year generation of 0.27. With these targets we get $\{\beta = 0.23, \underline{r} = 4.7, \delta = 3.54\}$.[42]

Sweden, like the other European countries, displays pre-transitional fertility levels that are particularly low in a cross-country perspective.[43] The average total fertility rates of the highest fertility countries was around 7, or above, in 2000 as compared to about 5 for pre-transitional Europe.[44] To explore the role of the cost of raising children for the high fertility countries, we calibrate an alternative ("low fertility cost") quantity-quality function. Changing the target for the pre-transitional fertility to $n = 3.5$ and re-calibrating the parameters accordingly delivers $\{\beta' = 0.75, \underline{r}' = 3, \delta' = 1.06\}$.

Initial Conditions and Time Conventions. To simulate the model, we finally need to set initial conditions for the importance of skilled human capital in production, x_0, and for TFP, A_0. The level of x_0 is a free parameter that only affects the timing of the take-off in the

from around 0.18 before the transition to around 0.12 in the steady state.

[40]The data are from Keyfitz and Flieger (1968) and World Development Indicators. The clear drop in fertility after 1900 is apparent also the time series depicted in Figure 3 below.

[41]The largest estimates are based on indexed data and include land, see Krantz and Schön (2007), Schön (2008) and Greasley and Madsen (2010).

[42]As an alternative calibration that does not rely on information about the growth rate of technology during the transition, one can also use information on the share of skilled in terms of λ around 1900 and compute the growth rate that is implied by (8). According to estimates by Ljungberg and Nilsson (2009) average years of schooling for the cohort aged 7-14 was around 4 in 1900. Given $\{\phi = 0.61, \underline{e}^u = 0, \underline{e}^s = 12\}$ as set above, this implies targeting a level of $\underline{g} = 0.2745$, which delivers essentially the same parameters for $\{\beta, \underline{r}, \delta\}$.

[43]The comparatively low fertility levels in Europe compared to other regions is well documented and the reasons are a matter of debate, see, e.g., Voigtlaender and Voth (2010).

[44]Data are from World Development Indicators, see also Figure 6 below.

simulation. Choosing x_0 sufficiently low, the simulation starts in the phase of stagnation which can be made arbitrarily long, to deliver a later onset of the transition in terms of the number of generations that are required until the onset takes place. Recall that the time axis is set with reference to the convergence to the post-transitional balanced growth path (in terms of λ converging to 1) in 2000. This implies that the choice of x_0 essentially determines the beginning of time in the calibration. We set $x_0 = 0.04$, which implies that the simulation covers the period from 0 A.D. until 2000 A.D.[45] The initial level of TFP is a shift parameter that is set in order to match a target level of log GDP per capita equal 10.03 in 2000.[46] This implies setting $A_0 = 15$.

The relevant parameters for the timing of the onset of the transition, with the notable of exception of $\underline{T}$, are calibrated targeting data moments for Sweden in 2000. The parameters governing child mortality and fertility, according to (6) and (14), are calibrated targeting moments in 1800 as well as in 2000. Since the system (22) does not involve any scale effect the calibration of these parameters only affects levels of the child survival probability and fertility, but not the dynamic evolution of the central variables (share of skilled workers, adult longevity and technological progress) and the timing of the take-off.

Cross-country differences in baseline life expectancy. The calibrated model is used to investigate the quantitative role of the mortality environment for comparative development. Sweden (and generally European countries) have a comparably favorable mortality environment which is reflected in a relatively low exposure to infectious diseases, whereas the less developed countries of today are often located in areas with a harsher mortality environment. A permanently higher exogenous exposure to infectious diseases implies faster aging and lower life expectancy under similar (economic) living conditions.[47]

As alternative scenario, we therefore target a life expectancy at age five at 45 years (compared to 48 years reflecting Sweden around 1800 just before the transition). This target is in line with the lowest available measure in 2000 and life expectancy at age five is not much higher in several Sub-Saharan African and Latin American countries.[48] Retaining a target of 76 years for

[45] This parametrization also implies that the income share of unskilled human capital in total production is larger than 99.9% at the beginning of the simulation, and still above 95% in 1800 just before the transition.

[46] The data are from `www.historicalstatistics.org`.

[47] A higher exposure to diseases leads to a faster deficit accumulation and earlier death, see, e.g., Mitnitski et al. (2001) and Searle et al. (2007). Research based on the investigation of skeletons documents that adult longevity during the Mesolithic period was lower in more difficult mortality environments, see Boldson and Paine (2000).

[48] The data source is UN Population Statistics (`www.unstats.un.org`). Data on life expectancy at five for earlier periods are missing for many countries, including most Sub-Saharan Africa countries in 1960. Alternatively, the available information on child mortality and life expectancy at birth in 1960 can be used to derive an estimate

life expectancy at age five on the balanced growth path, and using condition (5), this implies setting a $\underline{\underline{T}}$=40 and $\overline{\rho} = 36$ (rather than $\underline{T}$=45 and $\rho = 31$ as in the benchmark calibration).[49]

The last element of the calibration that is needed to simulate a cross-sectional distribution of the variables of interest is the world-wide distribution of baseline mortality. We simulate a world of 113 countries that only differ in terms of their baseline adult longevity, which lies between the two extremes $\left[\underline{\underline{T}}, \underline{T}\right]$. To create a meaningful distribution, the calibration is based on historical data about differences in disease prevalence across 113 countries constructed by Murray and Schaller (2010). These data have been collected from historical sources from the early 20th century and therefore reflect extrinsic mortality across the world before major health innovations occurred in most countries. The calibration exploits information on whether a particular infectious disease was detected in a country (i.e., not on the spread of the disease or the number of infected cases, which were potentially endogenous to development already in the 19th century).[50] The frequency distribution of the counts of pathogens for all countries of the world is used as distribution of baseline adult longevity within the support $[40, 45]$. The resulting distribution, depicted in Figure 1 in terms of a kernel density plot, is modestly skewed.[51]

---

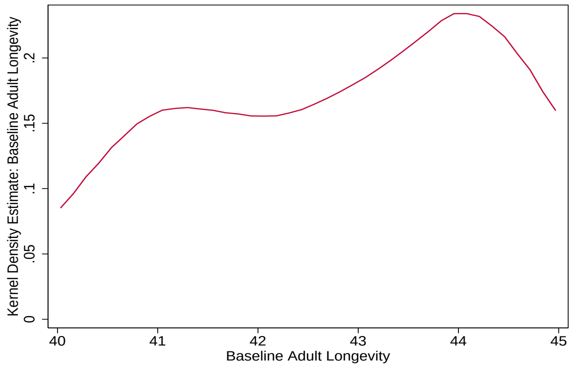
of life expectancy at age five. This delivers a very similar target for the highest mortality countries. In 1960 life expectancy at birth was as low as 33 years in some countries like Afghanistan, and child mortality one third. Assuming a constant death rate below the age of 5, these numbers imply a life expectancy at age five between 44 and 45 years. In some countries, like Swaziland life expectancy at birth is just above 30 years still today (CIA World Factbook). This suggests that 45 is possibly a conservative estimate of baseline adult longevity in the worst conceivable mortality environment.

[49] For robustness, we also replicated the analysis by keeping $\rho = 31$. All the quantitative results are unchanged and the only difference is a lower adult longevity along the balanced growth path for high mortality countries.

[50] For each pathogen we construct a binary indicator of whether or not a disease has been present at severe or epidemic levels at least once in the history up to the early 20th century. The diseases include leishmanias, schistosomes, trypanosomes, leprosy, malaria, typhus, filariae, dengue, and tuberculosis. Six of these diseases fall into the class of multi-host vector-transmitted diseases, which are particularly difficult to prevent or eradicate even today because the pathogens survive in multiple hosts (both humans and animals), and which are bound to specific transmission vectors, like mosquitos, which require a particular geographical habitat. The endemicity of the class of multi-host vector-transmitted diseases is fairly insensitive to economic development and globalization, and thus an informative measure of cross-country differences in the extrinsic mortality environment, see Smith et al. (2007). Cervellati, Sunde, and Valmori (2011) document the health relevance of the number these pathogens in terms of predicting life expectancy and the likelihood of outbreaks of epidemics.

[51] The frequency of simulated countries with baseline longevity $\underline{T} = 45$ corresponds to the frequency of countries with the lowest observed number of multi-host vector-transmitted diseases ever diagnosed (which includes Sweden). Conversely, the frequency of simulated countries with baseline longevity $\underline{\underline{T}} = 40$ corresponds to the frequency of countries with the highest number of multi-host vector-transmitted pathogens (which include several Sub-Saharan African countries). The distribution on the full support $(40, 45)$ is created by a linear intrapolation of the frequency distribution of the counts of multi-host vector-transmitted diseases on a grid of 0.25 diseases. The

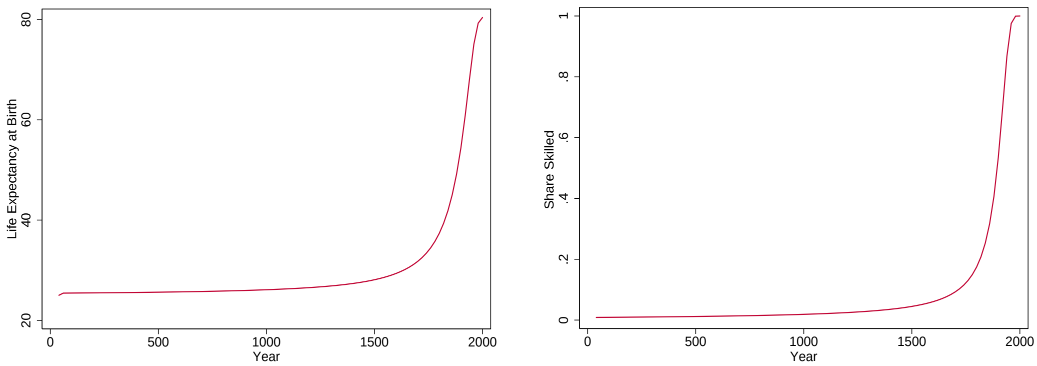
Figure 1: Synthetic Distribution of $\underline{T}$



4.2 Results

Time Series Analysis. Figure 2 depicts the simulated data for the equilibrium share of individuals acquiring skilled human capital and of life expectancy at birth that is obtained from the benchmark calibration of the model. The figure plots the dynamic evolution of these variables over the entire simulation period and illustrates the lengthy phase of slow development that is followed by a transition to the balanced growth path.

Figure 2: Long-Run Development: Simulation of benchmark calibration



(a) Life Expectancy at Birth

(b) Share Skilled λ

Although the transition appears sudden in a long term perspective, it actually takes place over a time horizon of about 200 years. Figure 3 restricts attention to the period 1750-2000 and compares the simulated data to the corresponding time series of historical data from Sweden.[52]

figure plots the resulting distribution of baseline longevity for the 113 countries of the Murray-Schaller data.

[52]Life expectancy and fertility data are taken from the Human Mortality Database (`http://www.mortality.org`), Keyfitz and Flieger (1968) (up to 1960) and World Development Indicators

Figure 3(a) reports the evolution of life expectancy at birth (T0) and life expectancy at age five (plus five years, T5), while panel (b) reports the evolution of child mortality rates. The calibration targeted life expectancy at age five as well as child mortality at two points in time (in 1800 and 2000). The model performs well in matching the evolution of adult longevity over the entire period, both in terms of levels and in terms of the duration of the transition. Also life expectancy at birth (which was not targeted) is matched well. Figures 3(c) and (d) plot the share of skilled individuals, λ, against the primary school enrolment rate and against the (shorter) series of average school years, respectively. Neither data series constitutes a perfect empirical counterpart for λ, but both reflect the education acquisition in the population. The model dynamics resemble the evolution of the enrolment rates in primary education and tend to lead slightly the dynamics of average school years. Given that the model does not account for institutional changes, like compulsory schooling legislation or school systems, the model's dynamics fit the data well.

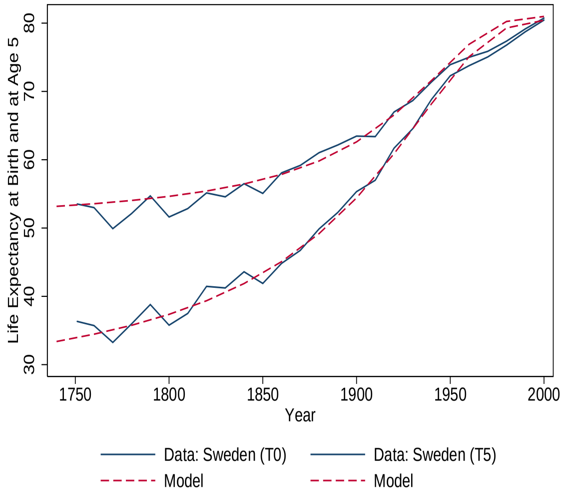
Figure 3(e) depicts gross and net fertility. The model was calibrated by targeting three moments that are apparent in this figure: the levels of gross fertility before and after the transition (1800 and 2000) as well as the exit from the corner solution of zero additional investment in child quality around 1900. The simulation matches not only the initial and terminal levels and the timing of the drop in gross fertility, however, but also the duration, the level and the dynamic evolution of net fertility, which were not explicitly targeted in the calibration. The presence of differential fertility allows to explain the eventual reduction in net fertility following the reduction in mortality, which has been difficult to rationalize in models based on the quantity-quality trade-off (Kalemli-Ozcan, 2003, and Doepke, 2005). The calibration of the quantity-quality function implies that the endogenous cost of raising children are actually very similar before the onset of the transition and on the balanced growth path, with levels of 4.7 and 5, respectively. This implies that the change in the quantity-quality trade-off is small and that the observed drop in gross and net fertility is mainly due to the differential fertility effect and the negative income effect that emerges when life expectancy reaches old ages.[53]

Figure 3(f) depicts evolution of income per capita. Although the technological progress and

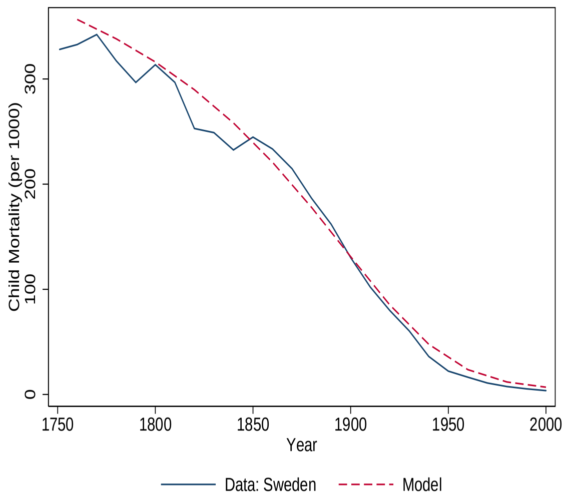
(after 1960), respectively. The Data for GDP, population and GDP per capita is provided by the internet portal for historical Swedish statistics, `www.historia.se` and the Swedish Central Statistical Office, `www.scb.se`, see also footnote 36. The data on schooling are from de la Croix, Lindh, and Malmberg (2008) while the data on average years of schooling are from Ljungberg and Nilsson (2009).

[53] In fact, even if one where to assume a cost of raising children fixed at post-transition levels the time series of fertility for the benchmark would be essentially unchanged. This is not the case for the quantity-quality function calibrated targeting data moments for the high fertility countries, however, as discussed below.

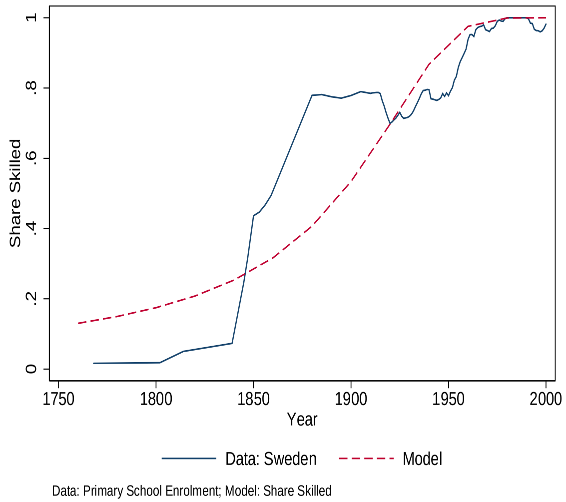
Figure 3: Long-Run Development: Simulation of Benchmark Calibration of the Model and Historical Data for Sweden 1750-2000



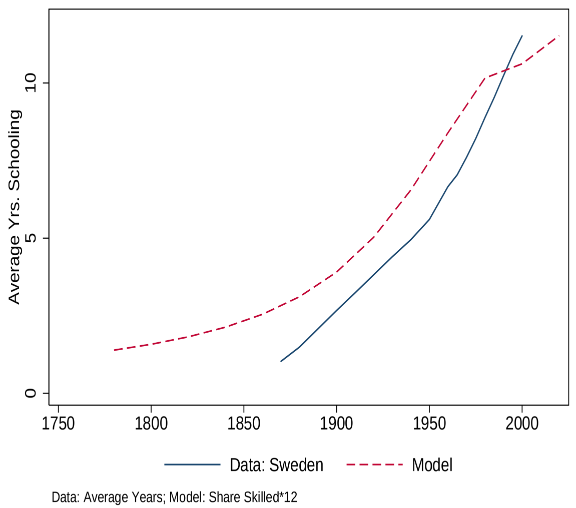
(a) Life Expectancy at Birth



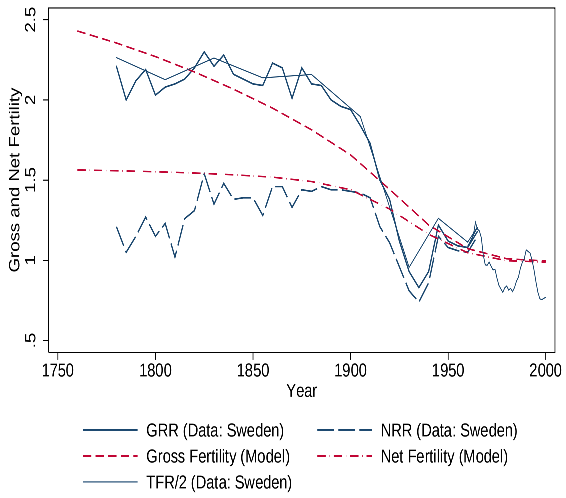
(b) Child Mortality Rate



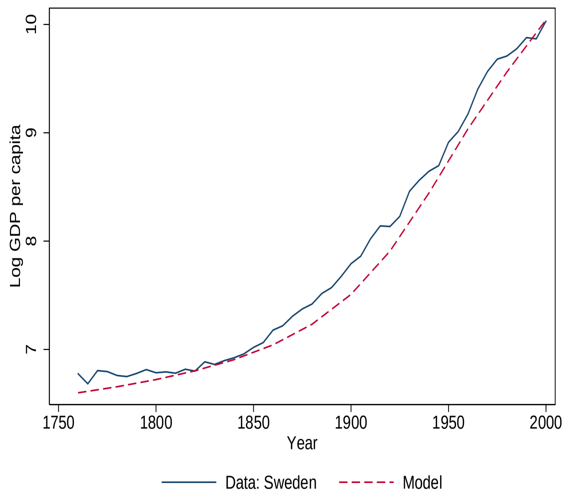
(c) Primary School Enrolment and λ



(d) Average Years of Schooling



(e) Gross and Net Reproduction Rates



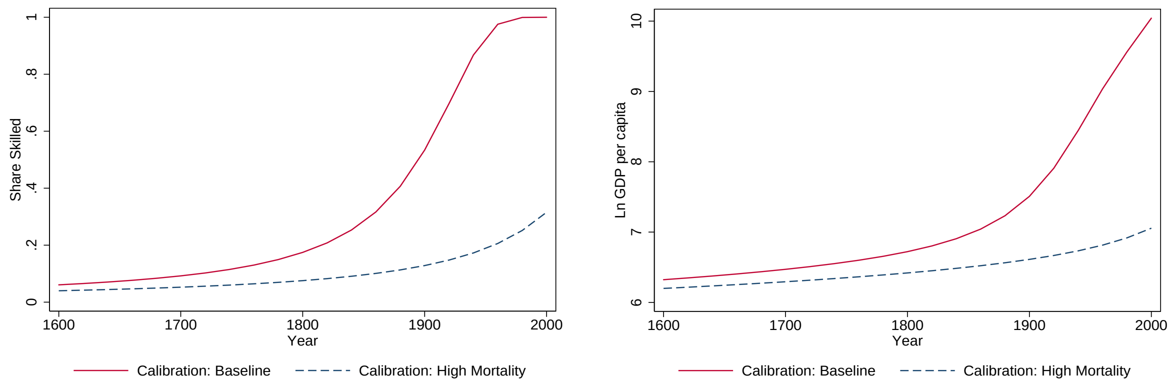
(f) log GDP per capita

the level of initial technology were calibrated to match the growth rate and level of income per capita in 2000, the evolution of income per capita matches closely the data series over the entire period, including the pre-transitional level and the acceleration during the transition.

Mortality and Comparative Development. From the system (22), a lower baseline adult longevity implies a lower equilibrium share of skilled for a given level of development in terms of technology and education of the previous generation, and therefore a delayed take-off. In contrast, a lower baseline child mortality does not affect human capital and technological progress (and therefore the timing of the transition). To investigate the quantitative importance of differences in mortality for comparative development, we replicate the quantitative analysis with the baseline adult longevity re-calibrated for the countries with the highest baseline mortality, while keeping the remaining parameters of the benchmark calibration unchanged. This counterfactual exercise isolates the role of adult longevity by simulating the same model that has been calibrated for data moments of Sweden and investigating the effects of changes in the baseline longevity to levels that reflect those of the highest mortality countries in sub-Saharan Africa.

Figure 4, plots the time series of the share of skilled individuals and income per capita for the benchmark calibration and contrasts them with the counterfactual simulation with $\underline{T} = 40$ rather than $\underline{T} = 45$. The take-off is delayed by about 140 years, or 7 generations.

Figure 4: The Role of Lower Baseline Longevity for Comparative Development



(a) Share Skilled

(b) log Income per capita

The results suggest that differences in extrinsic mortality environment that are compatible with the observed differences in pre-transitional adult longevity of about three years (48 years in Sweden in 1800 and 45 years in sub-Saharan Africa in 2000) might explain an important part of the cross-country differences in comparative development by substantially delaying the timing of the take-off to sustained growth.

Cross-Country Analysis. The analysis so far has focused on a given economy and its development over time (or across generations). We now turn to the assessment of the calibrated model's ability to account for comparative development patterns in spite of the fact that the model was calibrated to the historical evolution of Sweden. If the mechanism driving the transition process is generally valid one would expect that, at each point in time, different countries are in different phases of their (otherwise similar) development process. This conjecture is investigated by comparing the data generated by the dynamic simulation of the calibrated model to panel data for all countries for which we have empirical counterparts of the variables of interest over the period 1960-2000. Figure 5 presents the data generated by the very same calibration of the model for the period 1750-2000 (as depicted in Figure 3), but plotted against the key variable driving the transition in the model, the share of skilled workers λ, at the respective point in time, rather than as time series. As empirical counterpart of λ across countries we use the data from Barro and Lee (2001) and consider the share of the total population with some formal education, generated as one minus the fraction of the population with "no schooling education" in the total population.[54]

Figure 5 panels (a), (b) and (c), plot the simulated data for life expectancy at birth, child mortality and income per capita against λ. These simulated data, which have been generated with the benchmark calibration, are plotted together with corresponding cross country data for 1960 and 2000. The cross-sectional interpretation of the calibrated data fits the cross-country data patterns quantitatively well. For each level of λ the predicted adult longevity and child mortality roughly correspond to the respective average levels in the data. The relation appears also remarkably stable over the 40-year horizon and no substantial shift can be seen in the data pattern from 1960 to 2000. The results also document that the simulated data for the alternative, high mortality, calibration are essentially identical to the benchmark calibration and the correlations between λ. Only the cross-sectional relationship between life expectancy at birth and education is slightly different. This further supports the hypothesis that, apart from

[54] All results are qualitatively and quantitatively very similar using alternative measures like the fraction of total population with at least completed lower secondary education, or the fraction restricted to different age cohorts such as, e.g., age 20-24 years from Lutz, Goujon, and Sanderson (2007). Detailed results are available upon request. We use the Barro-Lee data as benchmark since they are used more frequently and go back to 1960. The other data sources are Human Mortality Database (`www.mortality.org`), the UN Population Statistics (different historical volumes of the UN Demographic Yearbook, `www.unstats.un.org`), the World Development Indicators (`http://data.worldbank.org/data-catalog/world-development-indicators`).

the timing of the take-off, the different countries experience a very similar development process. The joint consideration of Figures 3, 4 and 5 suggests that differences in baseline mortality may be relevant to explain the delay in comparative development, but their effect is difficult to detect with cross-country panel data.

Figure 5(d) plots the share of skilled against the value of the same variable 40 years (two generations) earlier. In the data, this corresponds to plotting the share of educated individuals in 2000 against that in 1960. The calibration performs comparably better for countries with a larger lagged share of educated individuals while it underestimates the improvements in education for countries with low λ in 1960. This provides a first indication that, compared to Sweden or other European countries for the same level of initial share of educated individuals, the developing countries appear to have experienced an acceleration in education acquisition in the last forty years. Also in this case higher baseline mortality leaves the structural evolution of education overtime completely unchanged.

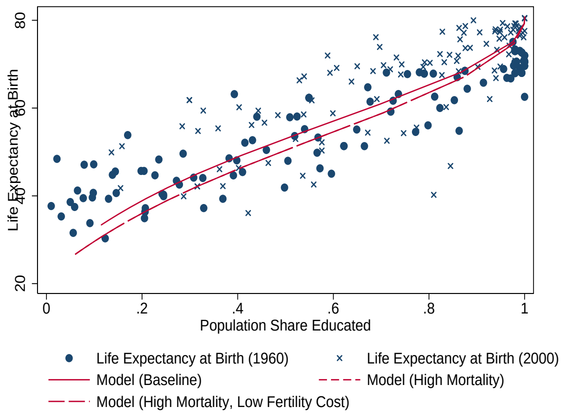
Figure 6(a) and (b) present the respective results for gross and net fertility.[55] The benchmark model matches the fertility levels for the more developed countries (the ones with a relatively large λ) that have undergone the demographic transition around, or shortly after, the period of the demographic transition in Sweden. The benchmark model underestimates the fertility levels for pre-transitional countries with low levels of λ, which in the data correspond to underdeveloped high mortality countries, however. The match between simulated model and data is substantially better for the alternative parametrization of the quantity-quality trade-off that was calibrated targeting data moments for the highest fertility countries.[56]

The non-linearity of the equilibrium locus Λ, characterized in Proposition 1, implies that the changes in λ are largest in the intermediate range where the slope is steepest. Hence, the increase in the share of skilled workers is relatively large in countries with intermediate levels of adult longevity, but relatively small in pre-transitional and post-transitional countries. Furthermore, from the dynamic simulation of Figure 3, the countries with largest increases in λ in the short run display lower increases in the future due to the convergence process, while the countries with lowest longevity display the largest overall improvements. The model therefore predicts

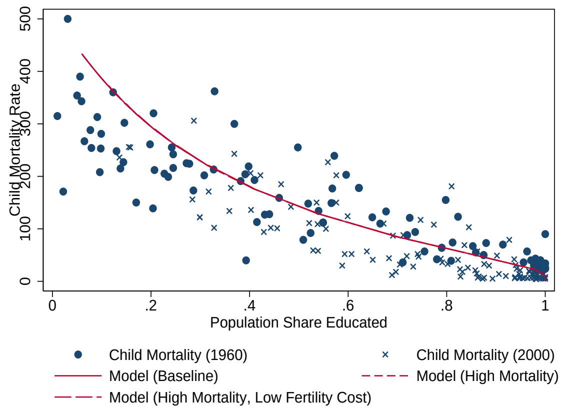
[55]Since reproduction in the model is asexual, the level n refers to the gross reproduction rate (the number of daughters for each woman). In order to compare this number to the data on total fertility rates, we multiply the gross reproduction rate n by two.

[56]The kink in the simulated data corresponds to the exit from the corner solution of the extra time invested in children. Recall that the cross-sectional patterns depicted in Figure 5 are unaffected by the actual calibration of the quantity-quality trade-off and the kink in fertility because the dynamic system (22) is block recursive and does not involve any scale effect.

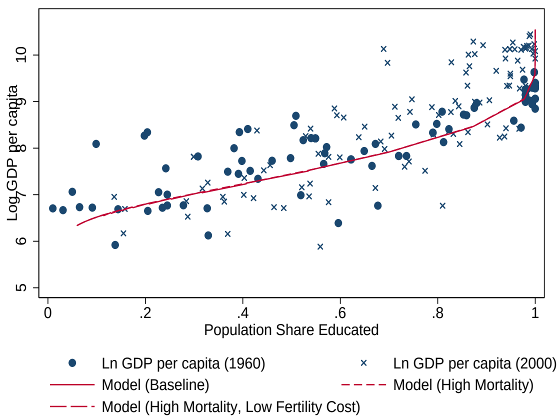
Figure 5: Education, Mortality and Income [Simulation and Data (1960 and 2000)]



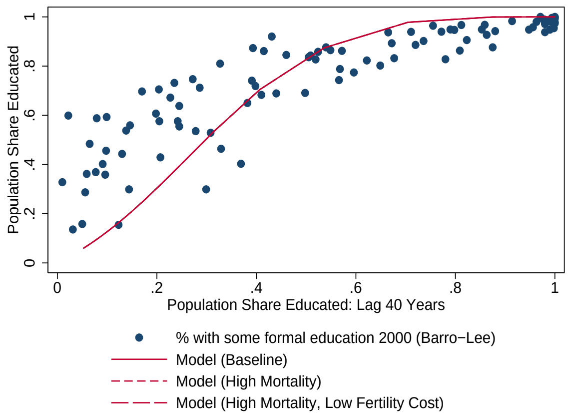
(a) Life Expectancy at Birth



(b) Child Mortality



(c) Log GDP per capita

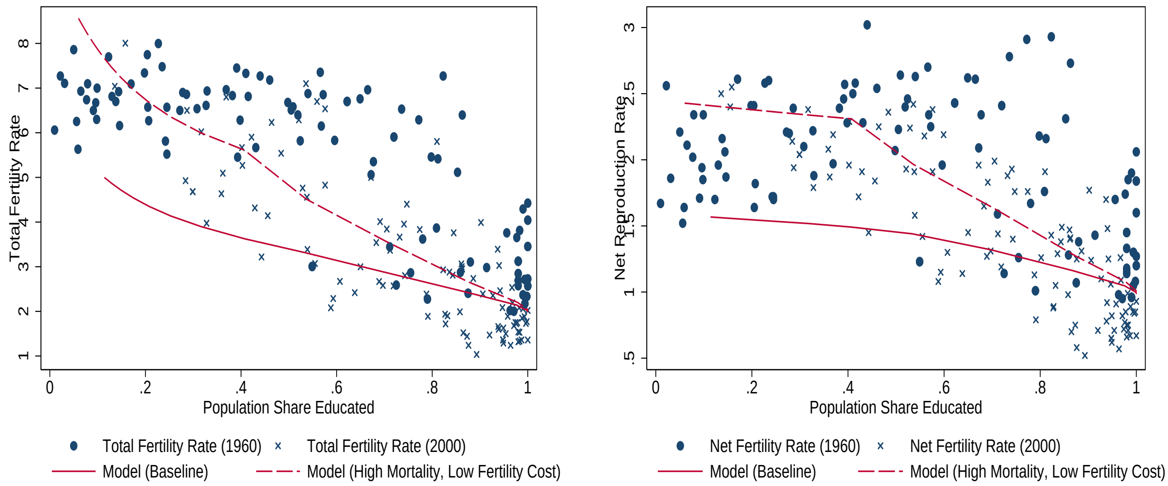


(d) λ 1960 and 2000

a non-monotonic relationship between life expectancy and changes in λ at each point in time. Panels (a) and (b) of Figure 7 depict the relationship between the level of life expectancy in 1960 and the change in the share of individuals with no formal education over the following twenty and forty years in the data (including a quadratic regression line), and compares them to the respective data from the benchmark calibration. The model matches the data well but somewhat underestimates the improvements in the change in education in countries with lower initial life expectancy. This again suggests that, compared to the historical experience of Sweden, education improvements in the poorest countries were comparatively large in the period 1960-2000.

Starting with Preston (1975), demographers have reported a strong correlation of health,

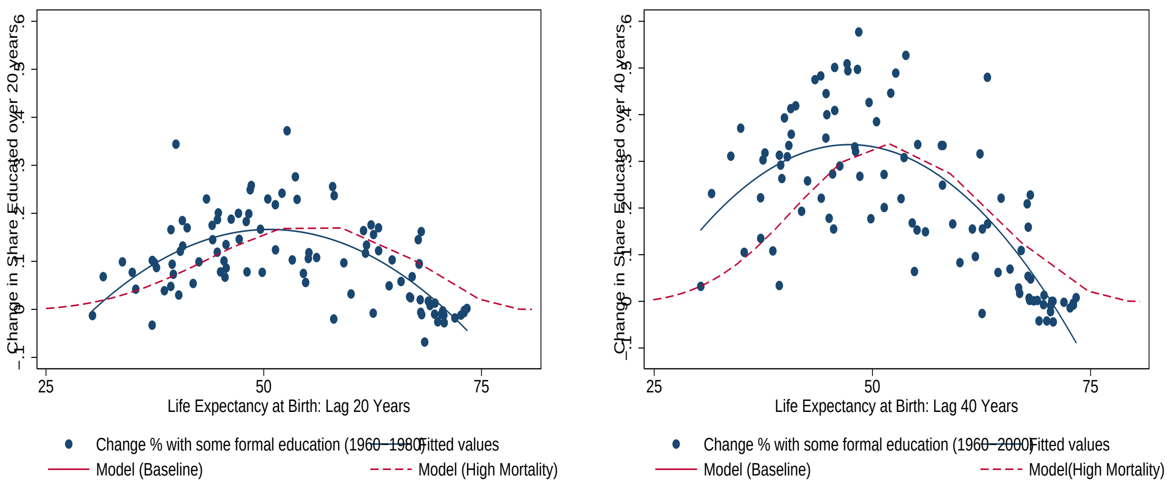
Figure 6: Education and Fertility [Simulation and Data (1960 and 2000)]



(a) Total Fertility Rate

(b) Net Reproduction Rate

Figure 7: Life Expectancy and Changes in Education Composition [Simulation and Data]



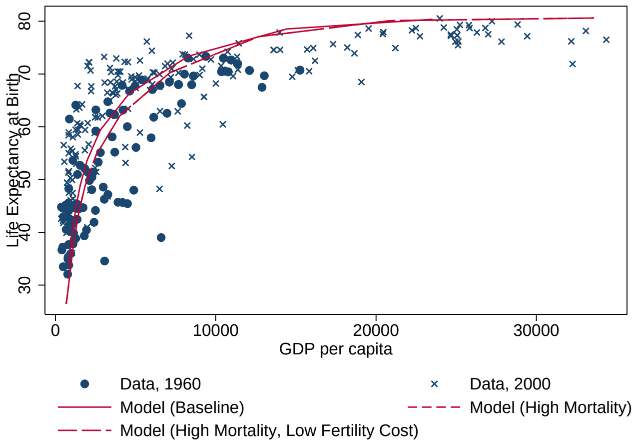
(a) Change in λ over 20 years

(b) Change in λ over 40 years

measured as life expectancy at birth, and income, which becomes weaker as incomes increase, see, e.g., Figure 1 of Cutler, Deaton, and Lleras-Muney (2006). The model also generates such a relationship that is consistent with the empirical Preston-Curve, as illustrated in Figure 8.

Cross-Country Distributions. Figure 2 illustrates that the dynamic evolution of the economy is characterized by a very long period of slow development followed by a (comparatively)

Figure 8: The Preston Curve



rapid transition to a sustained growth path. This feature holds irrespective of the level of baseline adult longevity and of the actual timing of the take-off. A direct implication is that even if different countries have a different timing of the take-off, at each point in time relatively few countries should be observed during the transition (since for most of its history each country is either pre-transitional or post-transitional). Consequently, one would expect the cross-sectional distribution of all variables of interest to display two modes corresponding to the mass of countries that are still pre-transitional or on the balanced growth path, as characterized in Proposition 2.[57] While intuitive, this cross-sectional implication of the non-linear development process has not been pointed out and investigated in the existing unified growth literature.

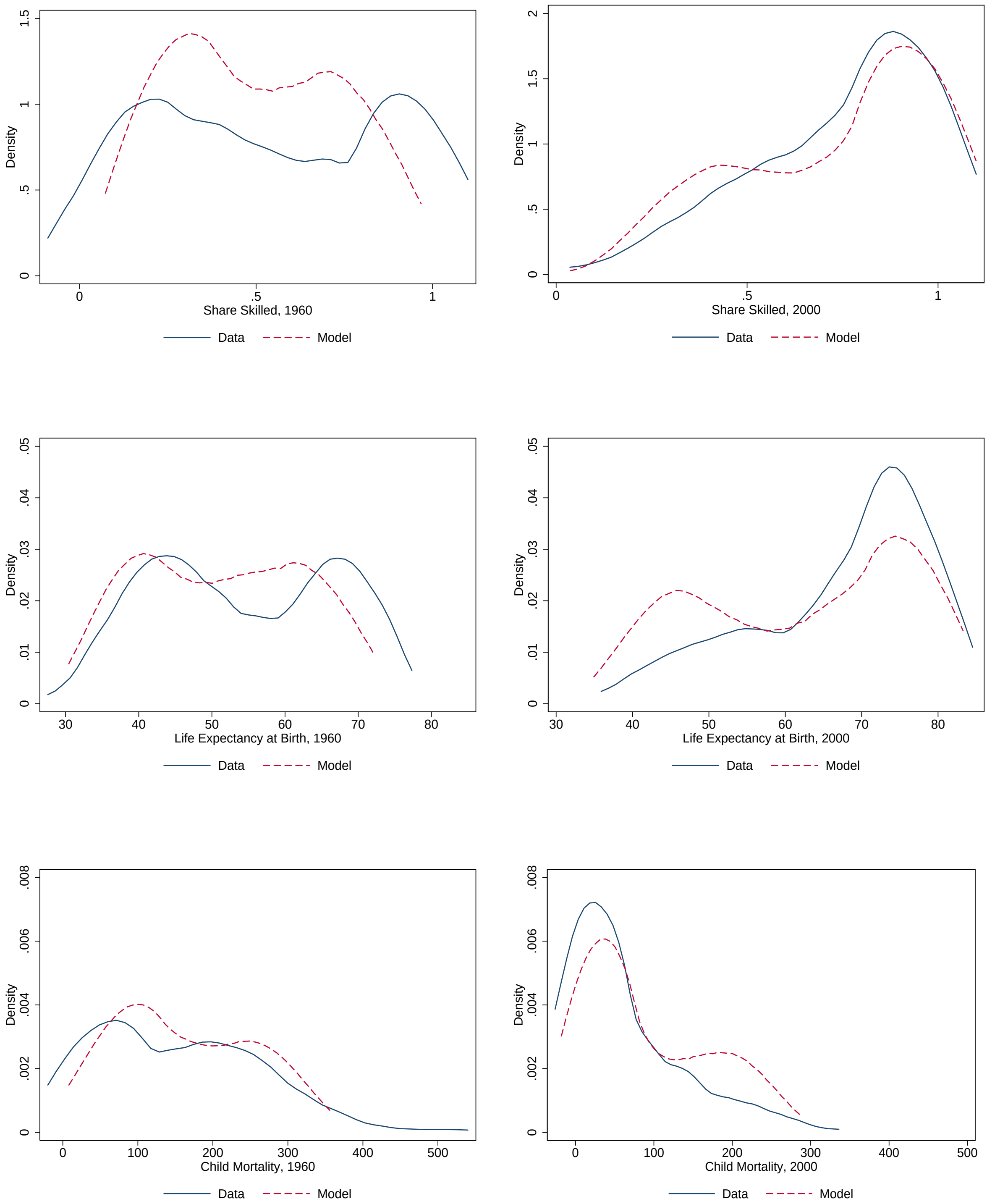
We simulate the dynamic evolution of the 113 countries differing in terms of their baseline adult longevity in the range $[40, 45]$ using the calibrated distribution described above. The data obtained from this artificial world are then pooled, used to estimate the cross-country distribution of all variables of interest in 1960 and in 2000, and compared to the corresponding distributions obtained from cross-county data.[58]

Figure 9 plots the simulated distributions of education, life expectancy at birth and child mortality for the years 1960 and 2000, and contrasts them to the respective distributions of the actual cross-country data by ways of kernel density estimates. Notice that no data moment of

[57]The precise shape of the distribution depends on the actual distributions of the underlying variables, like baseline mortality, that drive the delay in the take-off. Nonetheless, the bi-modality should be detectable as long as sufficiently many countries are still pre-transitional.

[58]For comparability, the distributions of real data are based on a homogenized sample of 90 countries, for which information on the share of skilled individuals, life expectancy at birth, child mortality, total fertility rate, and the net reproduction rate is available for 1960 and 2000. The results are similar when using unrestricted samples for the different variables. Details are available upon request.

Figure 9: Distributions: Education and Mortality [Simulation and Data (1960 and 2000)]



these distributions was explicitly used as target for the calibration of the model. The results can therefore be used to judge the ability of the model to fit the data. For all variables the expected bi-modality is apparent in 1960 both in the simulated and the actual data, while the distributions tend to be more unimodal by 2000 (when most countries have undergone the transition).[59] The simulated data match the patterns of the actual data in terms of the support, the location of the modes, and the shape of the distribution.

Most of the countries display fertility patterns resembling the high fertility countries, rather than Europe. We therefore simulate the artificial world by considering as benchmark the parametrization of the quantity-quality function that was calibrated targeting data moments for these countries. Figure 10 presents the results for total fertility rates and net reproduction rates for 1960 and 2000. The simulation fits the data by roughly capturing the peaks at low and high levels of fertility, as well as the shape of the distribution and its change over the 40-year horizon.[60]

Finally, Figure 11 depicts the world-wide distribution of incomes per capita, which the model fits reasonably well especially for 2000. When interpreting the figure, it should be kept in mind, however, that the model is limited in capturing the world income distribution by construction, since the model is calibrated to Sweden as the most developed county, and since the model does not allow for income spillovers or leapfrogging.[61]

The counterfactual exercise of comparing an artificial world (in which all countries are identical except for the baseline adult longevity) to the actual data suggests a potentially important quantitative role of differences in mortality environment for comparative development patterns. This role has been difficult to identify empirically. The quantitative results presented above suggest that differences in baseline mortality alone can potentially explain a substantial share of the observed cross-country differences in the distribution of the variables of main interest.

[59]The bi-modality of the simulated distribution is not due the actual calibrated distribution of baseline mortality. We have performed the exercise also considering a uniform distribution of baseline mortality. The bi-modality still emerges for 1960.

[60]The actual calibration of the function (14) is irrelevant for the kernel distributions of all variables apart from gross and net fertility. Unreported kernel distributions generated with the calibration for Sweden display a similar fit to the actual data for the most developed countries, but considerably underestimate the location of the peak for high fertility. This also implies that differences in the cost of raising children across countries are potentially more important for the cross-country differences in pre-transitional fertility levels than differences in mortality.

[61]Recall that the model only accounts for a human capital driven production process and does not consider other determinants of cross country income differences, like e.g. differences in physical capital, natural resources or institutions that have been shown to be empirically relevant, nor does it consider possible cross-country spill-overs or transfers of technology and innovations. In addition, while the sample for GDP corresponds to the 90 countries used for the density plots in Figures 9 and 10, the sample for 1960 only contains 72 countries due to data availability, and is therefore not perfectly comparable.

Figure 10: Density Distributions of Fertility [Simulation and Data (1960 and 2000)]

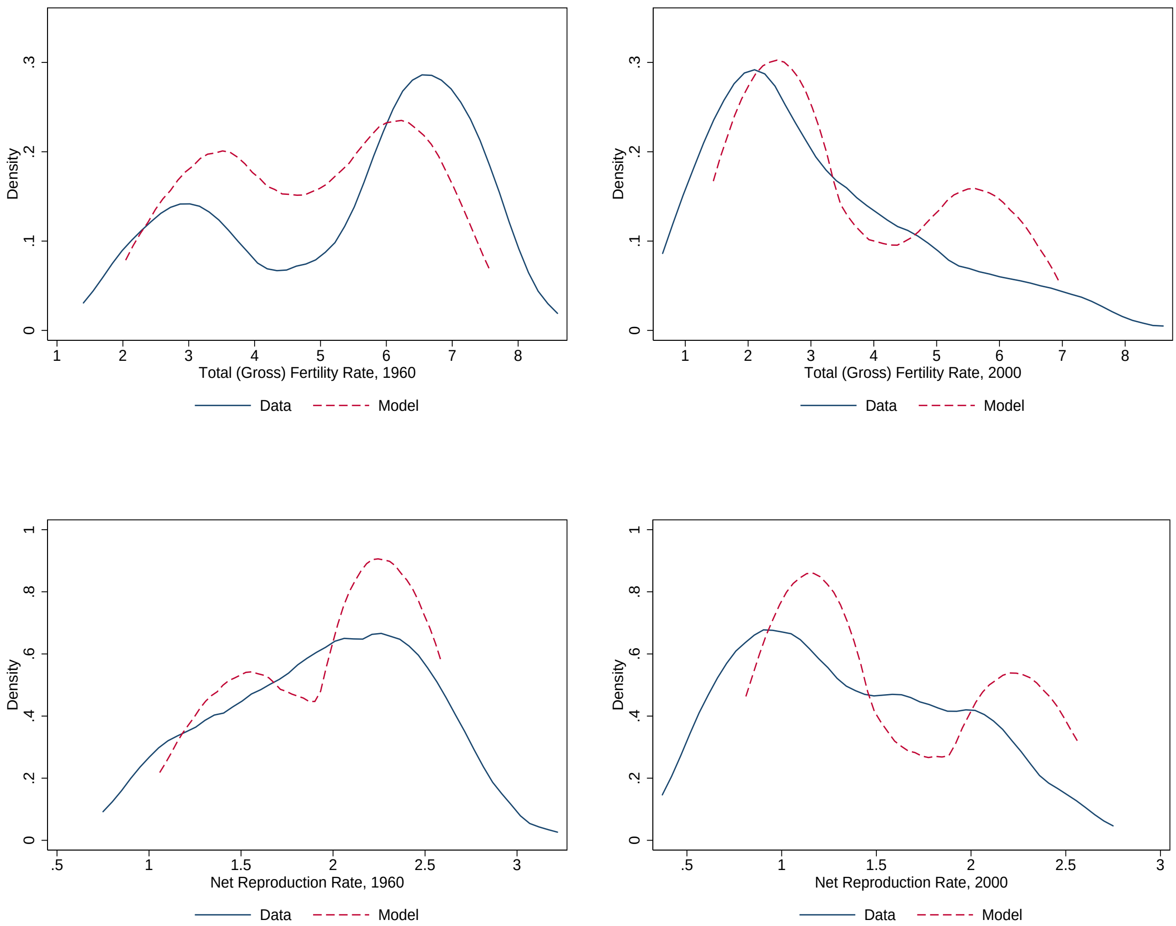
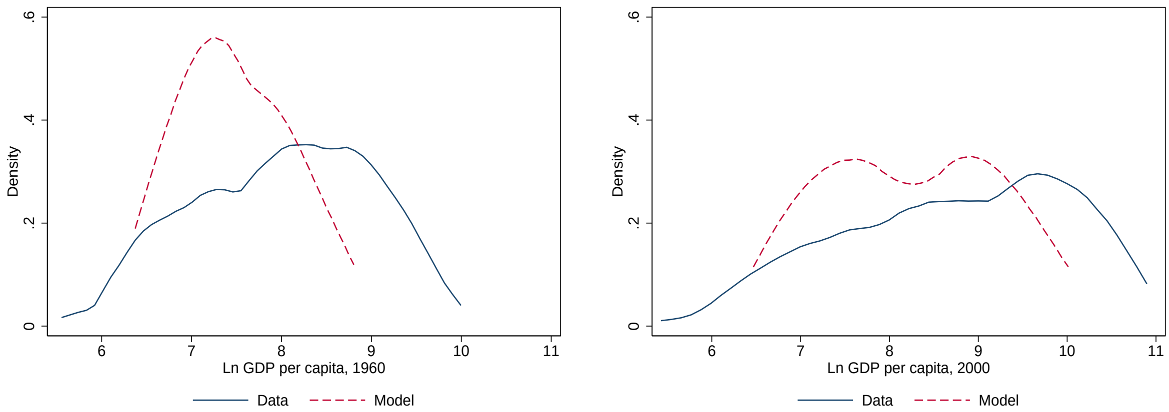


Figure 11: Density Distribution of Income per Capita [Simulation and Data (1960 and 2000)]



5 Discussion

We briefly comment on some relevant assumptions, possible extensions and complementary channels that could help improving the qualitative predictions and the quantitative fit.

Differential Fertility. The existence of a fertility differential by education that emerges from the model is one of the most robust stylized facts in demography. Skirbekk (2008) presents a meta-analysis of fertility studies and finds a strong negative relationship between education and fertility in all available studies in demography. The historical evidence suggests, however, that before 1750 higher social status (or income, wealth, or social class) was often associated with higher fertility.[62] The possibility to observe a reversal in differential fertility can be rationalized by the existence of subsistence levels in consumption. The negative fertility differential by education would emerge in the model after the onset of the transition.

Differential Mortality. Despite being small compared to the changes in average mortality, differential mortality related to education has been observed in the last decades in countries that have completed the demographic transition (Preston and Elo, 1995). The consideration of differential mortality would reinforce the predictions on the role of adult life expectancy for the incentives to acquire skilled human capital.[63] The share of skilled workers would matter not only for the change in average fertility but also for the change in average longevity.

Complementary Channels of the Fertility Transition. The model investigates the endogenous drivers of the fertility patterns in the long run.[64] The quantitative analysis suggests a relatively small role of the endogenous cost of raising children in explaining the patterns of the fertility transition in Sweden. The reason is that, as discussed above, the endogenous cost of raising children appears to have been already quite high before the transition.[65] The changing cost of child raising appears key to explain the dynamic patterns of the high fertility (non-European) countries, however. While reproduction is asexual in the model, in reality most of the time cost of raising children is provided by mothers at least before the demographic transition. Doepke et al. (2007) and Falcão and Soares (2008) explicitly study the role of female education and labor force participation during the demographic transition. Studying the differential participation

[62] Clark and Cummins (2009) discuss evidence suggesting that the rich appear to have been less constrained in their fertility in pre-transitional England.

[63] Technically, the implicit characterization of the equilibrium threshold of ability in (21) would imply a steeper equilibrium locus Λ for any level of technology. A mark-up in adult longevity for individuals acquiring skilled human capital is equivalent to a reduction in the parameter measuring the fixed cost $\underline{e}^s$. Hence, introducing such a mark-up leaves the overall properties of the dynamic system unaffected.

[64] A different question concerns the understanding of fluctuations in fertility patterns in terms of baby booms and baby busts, see Greenwood et al. (2005), Doepke et al. (2007), Murphy et al. (2008), and Jones and Schoonbroodt (2011).

[65] Voigtlaender and Voth (2010) argue that the low European fertility has its roots in the increased female labor force participation in the pastoral sector that followed the black death in the fourteenth century and that has contributed to an increase the cost of raising children in medieval Europe.

of females in (non-farming) labor markets could help explaining the cross-country differences in the cost of raising children. Finally, we have considered an exogenous age of reproduction although postponement is likely to be relevant for the dynamics of the fertility transition. The main quantitative effect of the consideration of fertility postponement would be related to the existence of a tempo effect leading to a temporarily stronger fertility differential associated with education acquisition.

Health and Labor Supply. Hazan (2009) documents that higher life expectancy and education were associated with a lower lifetime labor supply in data from the U.S. over the past 150 years. The predictions of the model are compatible with this evidence. As Galor and Moav (2000), we consider an occupational choice model where skilled and unskilled human capital are paid different wages. Increases in T_t reduce the relative cost for becoming skilled, and thus imply an increase in the share of skilled. The different wage accruing to skilled human capital implies that becoming skilled is associated with a larger total lifetime income and utility, despite the lower effective lifetime labor supply as consequence of the larger cost in terms of time, $\underline{e}^s$. Consequently, an increase in T_t induces individuals to become skilled even though it implies a discretely lower effective work life for these individuals.

Triggers of the transition. Skill-biased technical change increases the importance of human capital monotonically, and deterministically, over the course of generations. We consider skill-biased technological change for simplicity, but this assumption is not necessary for the main argument as long as productivity eventually increases enough to trigger the transition, i.e., to induce a sufficiently large fraction of the population to acquire skilled human capital.[66] Other variables might also trigger the transition, with potentially important implications for development policies. For instance, the incentives for the acquisition of skilled human capital in the model depend on the relative effectiveness of the time invested in acquiring education, which may be affected also by public schooling, as studied by Galor et al. (2009), and endogenous investments in health, as studied by de la Croix and Licandro (2007).[67]

Sources of Stagnation. The slow development before the take-off is due to the fact that it is not optimal to acquire skilled human capital until the returns are sufficiently large, despite the continuous technical progress. The dynamic evolution of the economy does not depend on scale effects and the theory abstracts from the presence of fixed factors of production, like land. As

[66] A monotonic technical change may not be realistic if, for instance, appropriate human capital is needed to innovate or adopt innovations, see Aiyar, Dalgaard, and Moav (2008).

[67] The relative importance of human capital in production and the productivity of education in producing human capital are isomorphic in inducing a larger fraction of skilled individuals λ for any T.

discussed by Weil and Wilde (2009), increases in longevity may reduce income per capita before the onset of the fertility transition if the size of the population increases more than total income in the presence of congestion effects. While not affecting the main mechanism, considering congestion effects would qualify the predictions about the role of longevity before the transition and could increase the quantitative fit of the world distribution of per capita income.

The Malthusian phase is typically modeled as a corner solution of the dynamic system that is the consequence of, e.g., subsistence consumption requirements, and the take-off (sometimes called post-Malthusian phase) is related to the exit from this corner solution. Abstracting from such subsistence effects allows performing smooth comparative statics on the role of the parameters, most notably baseline adult longevity, which is key for the cross-sectional analysis.[68] The consideration of corner solutions could nonetheless help improving the quantitative fit of the model in the time series dimension in terms of fertility levels at the onset of the transition, as shown by de la Croix and Doepke (2003) and de la Croix and Licandro (2007).

6 Concluding Remarks

This paper has proposed a unified theory of the economic and demographic transition. The dynamic equilibrium path is characterized by the endogenous evolution of mortality, fertility, education and income overtime. The model is calibrated to historical data for Sweden and matches closely the historical time series data. When analyzed from a cross-sectional perspective, the calibrated model can account for correlation and distribution patterns of the demographic and economic variables observed in cross country panel data for the period 1960-2000. The quantitative findings support the hypothesis that all countries follow a similar development process, which is characterized by a long period of stagnation, a rapid take-off, and a convergence to a balanced growth path, even though countries differ substantially in terms of the timing of the take-off. The results suggest that the differences in mortality environments across countries can explain the delay in development of those countries that are permanently exposed to harsher disease environments. The non-monotonic relationship between life expectancy and education along the development path also suggests an explanation for the bi-modality of the cross-country distribution of many demographic and economic variables.

The findings are relevant for the debate on the role of geography and of the mortality environment for long-run development. The quantitative results show that an inherently non-

[68]Even small changes in baseline parameters change the dynamic evolution of the system, which would not necessarily be the case in the presence of corner solutions in the main state variables.

linear dynamic development process can generate remarkably stable and essentially linear cross-sectional relationships between demographic and economic variables, such as correlations of education with mortality, fertility or income per capita, even among countries with very different levels of development or at different points of their development process. These cross-country patterns hide important non-linearities, including bi-modal cross-country distributions of the variables of interest. This can explain why the role of cross-country differences in the mortality environment for comparative development was difficult to detect in linear estimation frameworks that exploit cross-country variation.

The quantitative findings suggest some interesting directions for further research. The analysis has concentrated on explaining patterns of long-run and comparative development within a unified model that restricts attention to a single economy. Cross-section implications were derived in a world consisting of countries that are identical except their baseline mortality. The analysis of comparative development could be extended beyond extrinsic mortality differences to compare the quantitative role of alternative possible channels like, for instance, the role of institutions and other relevant cross-differences for explaining the delay in development. Finally, while instructive regarding the main mechanism, the analysis has completely abstracted from spill-overs and interactions between countries at different stages of development. Studying the role of such interactions, for example in terms of transfers of productive technology or medical knowledge, appears a promising direction for future research with potentially relevant implications also for development policies.

References

Acemoglu, D. (2002): "Directed Technical Change," *Review of Economic Studies*, 69(4), 781–809.

Aghion, P., and P. Howitt (1992): "A Model of Growth Through Creative Destruction," *Econometrica*, 60(2), 323–351.

Aiyar, S., C.-J. Dalgaard, and O. Moav (2008): "Technological Progress and Regress in Pre-Industrial Times," *Journal of Economic Growth*, 13(1), 125–144.

Azariadis, C., and J. Stachurski (2005): "Poverty Traps," in *Handbook of Economic Growth*, ed. by P. Aghion, and S. Durlauf, chap. 2. Elsevier Science B.V., Amsterdam.

Bar, M., and O. Leukhina (2010a): "Demographic Transition and Industrial Revolution: A Macroeconomic Investigation," *Review of Economic Dynamics*, 13(4), 424–451.

——— (2010b): "The Role of Mortality in the Transmission of Knowledge," *Journal of Economic Growth*, 15(4), 291–321.

Barro, R. J., and J. Lee (2001): "International Data on Educational Attainment: Updates and Implications," *Oxford Economic Papers*, 53(3), 541–563.

Becker, G. S., and R. J. Barro (1988): “A Reformulation of the Economic Theory of the Family,” *Quarterly Journal of Economics*, 103(1), 1–25.

Becker, G. S., and H. Lewis (1973): “On the Interaction Between the Quantity and Quality of Children,” *Journal of Political Economy*, 81(2), S279–288.

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

Bengtsson, T., and M. Dribe (2006): “Deliberate Control in a Natural Fertility Population: Southern Sweden, 1766-1864,” *Demography*, 43(4), 727–746.

Bideau, A., B. Desjardins, and H. Perez-Brignoli (1997): *Infant and Child Mortality in the Past*. Clarendon Press, Oxford.

Bloom, D. E., and D. Canning (2007): “Mortality Traps and the Dynamics of Health Transitions,” *PNAS, Proceedings of the National Academy of Sciences*, 104(41), 16044–16049.

Boldrin, M., and L. E. Jones (2002): “Mortality, Fertility, and Saving in a Malthusian Economy,” *Review of Economic Dynamics*, 5(4), 775–814.

Boldson, J. L., and R. R. Paine (2000): “The Evolution of Human Longevity from the Mesolithic to the Middle Ages: An Analysis Based on Skeletal Data,” in *Exceptional Longevity: From Prehistory to the Present*, chap. 3. Odense University Press, Odense.

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

——— (2003): “Early Mortality Declines at the Dawn of Modern Growth,” *Scandinavian Journal of Economics*, 105, 401–418.

——— (2011): “Vintage Capital Growth Theory: Three Breakthroughs,” in *Frontiers of Economic Growth and Development*, ed. by de La Grandville.

Cervellati, M., and U. Sunde (2005): “Human Capital, Life Expectancy, and the Process of Development,” *American Economic Review*, 95(5), 1653–1672.

——— (2008): “The Economic and Demographic Transition, Mortality, and Comparative Development,” *University of St.Gallen Discussion Paper*, 2008-21.

——— (2011a): “Disease and Development: The Role of Life Expectancy Reconsidered,” *Economics Letters*, 113, 269–272.

——— (2011b): “Life Expectancy and Economic Growth: The Role of the Demographic Transition,” *Journal of Economic Growth*, 16, 99–133.

Cervellati, M., U. Sunde, and S. Valmori (2011): “The Distribution of Infectious Diseases and Extrinsic Mortality Across Countries,” *Mathematical Population Studies, forthcoming.*

Chakraborty, S. (2004): “Endogenous Lifetime and Economic Growth,” *Journal of Economic Theory*, 116, 119–137.

Chakraborty, S., C. Papageorgiou, and F. P. Sebastian (2010): “Diseases, Infection Dynamics and Development,” *Journal of Monetary Economics*, 57(7), 859–872.

Chari, V., and H. Hopenhayn (1991): "Vintage Human Capital, Growth, and the Diffusion of New Technology," *Journal of Political Economy*, 99, 1142–1165.

Chen, H.-J. (2010): "Life Expectancy, Fertility, and Educational Investment," *Journal of Population Economics*, 23(1), 37–56.

Chesnais, J.-C. (1992): *The Demographic Transition: Stages, Patterns and Economic Implications. A Longitudinal Study of Sixty-Seven Countries Covering the Period 1720-1984*. Clarendon Press, Oxford.

Clark, G., and Cummins (2009): "Urbanization, Mortality, and Fertility in Malthusian England," *American Economic Review*, 99(2).

Cutler, D., A. Deaton, and A. Lleras-Muney (2006): "The Determinants of Mortality," *Journal of Economic Perspectives*, 20(3), 97–120.

de la Croix, D., and M. Doepke (2003): "Inequality and Growth: Why Differential Fertility Matters," *American Economic Review*, 93(4), 1091–1113.

——— (2004): "Public versus Private Education when Differential Fertility Matters," *Journal of Development Economics*, 73(2), 607–629.

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

——— (2007): "The Father of Child is Father of Man: Implications for the Demographic Transition," *mimeo, EUI Florence.*

de la Croix, D., T. Lindh, and B. Malmberg (2008): "Swedish Economic Growth and Education Since 1800," *Canadian Journal of Economics*, 41(1), 166–185.

Doepke, M. (2004): "Accounting for Fertility Decline During the Transition to Growth," *Journal of Economic Growth*, 9(3), 347–383.

——— (2005): "Child Mortality and Fertility Decline: Does the Barro-Becker Model Fit the Facts?," *Journal of Population Economics*, 17(2), 337–366.

——— (2007): "Growth Takeoffs," in *New Palgrave Dictionary of Economics*. 2 edn.

Doepke, M., M. Hazan, and Y. Maoz (2007): "The Baby Boom and World War II: A Macroeconomic Analysis," *Northwestern University, mimeo.*

Domeij, D., and M. Floden (2010): "Inequality Trends in Sweden 1978-2004," *Review of Economic Dynamics*, 13(2), 179–208.

Dribe, M. (2004): "Long-Term Effects of Childbearing on Mortality: Evidence from Pre-Industrial Sweden," *Population Studies*, 58(3), 297–310.

Eckstein, Z., P. Mira, and K. I. Wolpin (1999): "A Quantitative Analysis of Swedish Fertility Dynamics: 1751-1990," *Review of Economic Dynamics*, 2, 137–165.

Erosa, A., T. Koreshkova, and D. Restuccia (2011): "How Important Is Human Capital? A Quantitative Theory Assessment of World Income Inequality," *Review of Economic Studies*, 77, 1421–1449.

Falcão, B. L., and R. R. Soares (2008): "The Demographic Transition and the Sexual Division of Labor," *Journal of Political Economy*, 116(6), 1058–1104.

Galor, O. (2005): "From Stagnation to Growth: Unified Growth Theory," in *Handbook of Economic Growth*, ed. by P. Aghion, and S. Durlauf, chap. 4. Elsevier Science B.V., Amsterdam.

Galor, O., and O. Moav (2000): "Ability-Biased Technological Transition and Economic Growth," *Quarterly Journal of Economics*, 115(2), 469–497.

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

Galor, O., O. Moav, and D. Vollrath (2009): "Inequality in Landownership, the Emergence of Human-Capital Promoting Institutions, and the Great Divergence," *Review of Economic Studies*, 76(1), 143–179.

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

Greasley, D., and J. B. Madsen (2010): "Curse and Boon: Natural Resources and Long-Run Growth in Currently Rich Economies," *Economic Record*, 86(274), 311–328.

Greenwood, J., A. Sheshadri, and G. Vandenbroucke (2005): "The Baby Boom and Baby Bust," *American Economic Review*, 95(1), 183–207.

Haveman, R., and B. Wolfe (1995): ""The Determinants of Children's Attainments: A Review of Methods and Findings," *Journal of Economic Literature*, 33(4), 1829–1878.

Hazan, M. (2009): "Longevity and Lifetime Labour Supply: Evidence and Implications," *Econometrica*, 77(6), 1829–1863.

Hazan, M., and H. Zoabi (2006): "Does Longevity Cause Growth? A Theoretical Critique," *Journal of Economic Growth*, 11(4), 363–376.

Jones, L. E., and A. Schoonbroodt (2010): "Complements versus Substitutes and Trends in Fertility Choice in Dynastic Models," *International Economic Review*, 51(3), 671–699.

——— (2011): "Baby Busts and Baby Booms: The Fertility Response to Shocks in Dynastic Models," *mimeo, University of Minnesota.*

Jones, L. E., A. Schoonbroodt, and M. Tertilt (2011): "Fertility Theories: Can They Explain the Negative Fertility-Income Relationship?," in *Demography and the Economy*, ed. by J. B. Shoven, pp. 43–100. NBER.

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

——— (2003): "A Stochastic Model of Mortality, Fertility, and Human Capital Investment," *Journal of Development Economics*, 70(1), 103–118.

——— (2009): "The Uncertain Lifetime and the Timing of Human Capital Investment," *Journal of Population Economics*, 21, 557–572.

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.

Keyfitz, N., and W. Flieger (1968): *World Population: An Analysis of Vital Data.* University of Chicago Press, Chicago.

Kirk, D. (1996): "Demographic Transition Theory," *Population Studies*, 50, 361–387.

Kögel, T., and A. Prskawetz (2001): "Agricultural Productivity Growth and the Escape from Malthusian Trap," *Journal of Economic Growth*, 6, 337–357.

Krantz, O., and L. Schön (2007): *Swedisch Historical National Accounts, 1800-2000.* Almqvist and Wiksell International, Lund.

Lagerlöf, N.-P. (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(1), 116–142.

Lee, R. (2003): "The Demographic Transition: Three Centuries of Fundamental Change," *Journal of Economic Perspectives*, 17(4), 167–190.

Lewis, M. E., and R. Gowland (2007): "Brief and Precarious Lives: Infant Mortality in Contrasting Sites from Medieval and Post-medieval England (AD 850-1859)," *American Journal of Physical Anthropology*, 134(1), 117–129.

Ljungberg, J., and A. Nilsson (2009): "Human Capital and Economic Growth: Sweden 1870-2000," *Cliometrica*, 3, 71–95.

Lutz, W., A. Goujon, and W. Sanderson (2007): "Reconstruction of population by age, sex, and level of education attainment of 120 countries for 1970-2000," in *Vienna Yearbook of Population Research*, pp. 193–235. IAASA, Vienna.

Mitnitski, A., A. Mogilner, and K. Rockwood (2001): "Accumulation of Deficits as a Proxy Measure of Aging," *Scientific World*, 1, 323–336.

Moav, O. (2005): "Cheap Children and the Persistence of Poverty," *Economic Journal*, 115(88–110).

Mturi, A., and A. Hinde (2007): "Fertility Levels and Differentials in the United Republic of Tanzania," *Population Bulletin of the United Nations*, 46, 195.

Murphy, K. M., C. Simon, and R. Tamura (2008): "Fertility Decline, Baby Boom, and Economic Growth," *Journal of Human Capital*, 2(3), 262–302.

Murray, D. R., and M. Schaller (2010): "Historical Prevalence of Infectious Diseases Within 230 Geopolitical Regions: A Tool for Investigating Origins of Culture," *Journal of Cross-Cultural Psychology*, 41(1), 99–108.

Neisser, U., G. Boodoo, T. Bouchard, W. Boykin, N. Brody, S. Ceci, D. Halpern, J. Loehlin, R. Perloff, R. Sternbert, and S. Urbina (1996): "Intelligence: Knowns and Unknowns," *American Psychologist*, 81(2), 77–101.

Nelson, R., and E. Phelps (1966): "Investment in Humans, Technological Diffusion, and Economic Growth," *American Economic Review*, 56(2), 69–75.

Preston, S. H. (1975): "The Changing Relation Between Mortality and Level of Economic Development," *Population Studies*, 29(2), 231–248.

Preston, S. H., and I. T. Elo (1995): "Are Educational Differentials in Adult Mortality Increasing in the United States?," *Journal of Aging and Health*, 7(4), 476–497.

Schön, L. (2008): "Sweden – Economic Growth and Structural Change, 1800-2000," in *EH.Net Encyclopedia*, ed. by R. Whaples.

Searle, S. D., A. Mitnitski, E. A. Gahbauer, T. M. Gill, and K. Rockwood (2007): "A Standard Procedure for Creating a Frailty Index," *BMC Geriatrics*, 8(24), doi:10.1186/1471–2318–8–24.

Skirbekk, V. (2008): "Fertility Trends by Social Status," *Demographic Research*, 18, 145–180.

Smith, C. F., D. F. Sax, S. D. Gaines, V. Guernier, and J.-F. Guegan (2007): "Globalization of Human Infectious Disease," *Ecology*, 88(8), 1903–1910.

Soares, R. (2005): "Mortality Reductions, Educational Attainment, and Fertility Choice," *American Economic Review*, 95(3), 580–601.

——— (2006): "The effect of longevity on schooling and fertility: evidence from the Brazilian Demographic and Health Survey," *Journal of Population Economics*, 19(1), 71–97.

Tamura, R. (2006): "Human Capital and Economic Development," *Journal of Development Economics*, 79(1), 26–72.

Tertilt, M. (2005): "Polygyny, Fertility, and Savings," *Journal of Political Economy*, 113(6), 1341–1371.

Voigtlaender, N., and J. Voth (2010): "How the West 'Invented' Fertility Restriction," *mimeo.*

Wang, L. (2003): "Determinants of Child Mortality in LDCs - Empirical Findings from Demographic and Health Surveys," *Health Policy*, 65(3), 277–299.

Weil, D., and J. Wilde (2009): "How Relevant is Malthus for Economic Development Today?," *American Economic Review*, 99(2), 255–260.

Woods, R. (1997): "Infant Mortality in Britain: A Survey of the Current Knowledge on Historical Trends and Variations," in *Infant and Child Mortality in the Past*, ed. by A. Bideau, B. Desjardins, and H. Perez-Brignoli, chap. 5. Clarendon Press, Oxford.

7 Appendix

Proof of Lemma 1. Consider an individual acquiring human capital of type $j = u, s$. Taking the first order condition of (12) with respect to n_t^i and restricting to an interior solution gives (13), while taking the first order condition with respect to r_t^i gives,

$$-T_t \pi_t n_t^i r_t^{ij} + \gamma \left(\overline{T}_t - \pi_t n_t^i r_t^{ij} - \underline{e}^j \right) \left(q_r(\cdot) r_t^{ij} \right) / q(\cdot) \geq 0. \quad (25)$$

Using (13) to simplify (25) implies $\left[q_r(r_t^{ij}, g_{t+1}) r_t^{ij} \right] / q(r_t^{ij}, g_{t+1}) \geq 1$. Given the functional form (4) this implies (14).

Proof of Lemma 2. The optimal type of human capital maximizes the indirect utility obtained from $j = u, s$. Evaluating the indirect utility substituting for n_t^{ij} with $j = u, s$ from (13) and noting that $r_t^{iu} = r_t^{is} = r_t^*$ from (14) implies that the optimal type of skill depends on,

$$\left(\overline{T}_t - \underline{e}^u\right)^{(T_t+\gamma)} \left(w_t^u h_t^u\right)^{T_t} \gtrless \left(\overline{T}_t - \underline{e}^s\right)^{(T_t+\gamma)} \left(w_t^s h_t^s(a)\right)^{T_t} . \tag{26}$$

Since the indirect utility obtained by acquiring skilled human capital increases with ability, there exists a unique $\widetilde{a}_t$ such that all individuals with $a < \widetilde{a}_t$ optimally choose to acquire u, while those with $a > \widetilde{a}$ optimally choose to obtain s. Solving (26) as equality gives (15).

Proof of Proposition 1. The wage ratio is given by,

$$\frac{w_t^u}{w_t^s} = \frac{1-x_t}{x_t} \left(\frac{\int_{\widetilde{a}_t}^1 h^s(a) f(a) da}{\int_0^{\widetilde{a}_t} h^u f(a) da} \right)^{1-\eta} . \tag{27}$$

Substituting (27) into (15) gives the general equilibrium ability threshold (20). Rearrange (20) to get the equilibrium relationship between $\widetilde{a}_t$ and T_t expressed as

$$\mathcal{G}\left(\widetilde{a}_t\right)^{1-\eta} F\left(x_t\right) - \left(\frac{\overline{T}_t - \underline{e}^s}{\overline{T}_t - \underline{e}^u} \right)^{\frac{T_t+\gamma}{T_t}} = 0 \tag{28}$$

where $\overline{T}_t := min\{T_t, R\}$, $\mathcal{F}(x) := ((1-x_t)/x_t)$ and

$$\mathcal{G}(\widetilde{a}_t) = \frac{(h^u)^{\frac{1}{1-\eta}} \int_{\widetilde{a}_t}^1 h^s(a) f(a) da}{h^s(\widetilde{a}_t)^{\frac{1}{1-\eta}} \int_0^{\widetilde{a}_t} h^u f(a) da} \tag{29}$$

with $\mathcal{G}'(\widetilde{a}_t) < 0$. Notice that $\left[\left(\overline{T}_t - \underline{e}^s\right) / \left(\overline{T}_t - \underline{e}^u\right)\right] \in (0, 1)$ for $T_t \in (\underline{e}^s, \infty)$. For any x_t, the function (28) is therefore defined over the range $\widetilde{a} \in (\underline{a}(x_t), 1]$ where[69]

$$\underline{a}(x_t) : \mathcal{G}\left(\underline{a}(x_t)\right)^{1-\eta} \mathcal{F}(x_t) = 1 \tag{30}$$

Applying calculus it follows that $\partial \underline{a}(x_t)/\partial x_t < 0$ with $\lim_{x\to 0} \underline{a}(x) = 1$ and $\lim_{x\to 1} \underline{a}(x) = 0$. Accordingly for any x_t there exists a level $\overline{\lambda}(x_t) < 1$ which represents the maximum share of the population that for each generation t would acquire skilled human capital in the case in which $T_t \to \infty$. By totally differentiating (28) we have,

$$\frac{d\widetilde{a}_t}{dT_t} = \frac{d\left(\left(\frac{\overline{T}_t - \underline{e}^s}{\overline{T}_t - \underline{e}^u} \right)^{\frac{T_t+\gamma}{T_t}} \right) / dT_t}{\left[(1-\eta)\,\mathcal{G}\left(\widetilde{a}_t\right)^{-\eta} \mathcal{G}'\left(\widetilde{a}_t\right) F\left(x_t\right)\right]} < 0 \tag{31}$$

which is negative since $\mathcal{G}'(\widetilde{a}_t) < 0$ and for $T_t < R$, the numerator is,[70]

$$-\frac{\gamma}{T_t^2} \ln\left(\frac{T_t - \underline{e}^s}{T_t - \underline{e}^u}\right) e^{\ln\left(\frac{T_t - \underline{e}^s}{T_t - \underline{e}^u}\right)} e^{\frac{T_t+\gamma}{T_t}} + \frac{T_t+\gamma}{T_t} \left(\frac{T_t - \underline{e}^s}{T_t - \underline{e}^u}\right)^{\frac{\gamma}{T_t}} \frac{\underline{e}^s - \underline{e}^u}{\left(T_t - \underline{e}^u\right)^2} > 0 \tag{32}$$

For $T_t = \underline{e}^s$ we have $\widetilde{a}_t = 1$ which implies $\mathcal{G}(\widetilde{a}_t) = 0$ so that $\mathcal{G}(\widetilde{a}_t)^{-\eta} = \infty$. Since $\mathcal{G}'(1)$ is a finite number we have that the denominator of (31) goes to infinity as $T_t \to \underline{e}^s$. In turns the numerator has a limit at zero. For $T_t \to \infty$ we have $\widetilde{a}_t \to \underline{a} < 1$ so that the denominator of (31) is a finite number while the numerator has a limit at zero.[71] Hence $\lim_{T_t \to \underline{e}\frac{d\widetilde{a}_t}{dT_t}}{}^s = \lim_{T_t\to\infty} \frac{d\widetilde{a}_t}{dT_t} = 0$ which also implies that the equilibrium locus (21) is convex for $T_t \to \underline{e}^s$ and concave for $T_t \to \infty$.

[69] Since the denominator of (28) has a discontinuity at $\underline{a}$ and the function takes negative values for any $a \leq \underline{a}(x_t)$.

[70] If $\overline{T}_t = R$ then equation (32) reads as $-\frac{\gamma}{T_t^2} \ln\left(\frac{R-\underline{e}^s}{R-\underline{e}^u}\right) e^{\ln\left(\frac{R-\underline{e}^s}{R-\underline{e}^u}\right)} e^{\frac{T_t+\gamma}{T_t}} > 0$.

[71] The same is true if $T_t > R$ since $\lim_{T_t\to\infty} \left[-\frac{\gamma}{T_t^2} \ln\left(\frac{R-\underline{e}^s}{R-\underline{e}^u}\right) e^{\ln\left(\frac{R-\underline{e}^s}{R-\underline{e}^u}\right)} e^{\frac{T_t+\gamma}{T_t}}\right] = 0$.

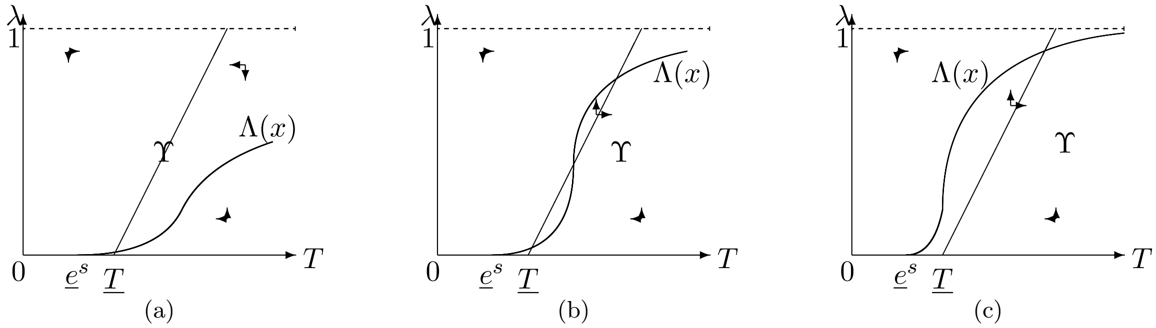


Figure 12: The Process of Development

Proof of Lemma 3. From Proposition 1 for any $T_t > \underline{e}^s$ and any $x_t > 0$, we have $\lambda_t > 0$. From (7) this implies $x_t > x_{t-1}$ for all t with $\lim_{t\to\infty} x_t = 1$; from (8), $g_t > 0$ and $\lim_{t\to\infty} A_t = \infty$ for any $A_0 > 0$. In the limit as $\lambda_t \to 1$, $g_t = \phi$ from (8).

Proof of Proposition 2. The equilibrium relationship linking $\widetilde{a}_t$ and T_t is given in (28). For any T_t, $\widetilde{a}_t$ is an implicit function of x_t. Recall that by implicit differentiation of (20) $\partial\widetilde{a}_t/\partial x_t < 0$ which implies that the equilibrium share of skilled individuals is increasing in $x_t : \partial\lambda_t/\partial x_t > 0$ for any T_t. Consider part *(i)*. If $x_0 \simeq 0$ and $A_0 \simeq 0$ then $\underline{a}\,(0) \simeq 1$; for all $T \in (\underline{e}^s, \infty)$ which implies $\widetilde{a} \simeq 1$ and $\lambda \simeq 0$. In this case the two loci Λ and Υ cross only once for $\lambda \simeq 0$ and $T \simeq \underline{T}$ and the average fertility is given by n^u as implied by (13) evaluated at $T = \underline{T}$. Under these conditions, from (2) the level of income per capita is (arbitrarily) low which, from (6) and (23) implies $\pi_0 \simeq \underline{\pi}$. Phase *(ii)* follows directly from Lemma 3, where $A_\infty \to \infty$, $x_\infty \to 1$, $\lambda_\infty \simeq 1$, $T = \overline{T}$. From (8) this also implies that $g_\infty = \phi$. Finally, since $A_\infty \to \infty$, it follows that $y_\infty \to \infty$ and from (6), $\pi_\infty \simeq 1$ so that fertility is given as in (24).

Figure 12 depicts the evolution of the conditional system given by equations (5) and (21) for the case in which the latter function has a unique inflection point. From *(i)* and *(ii)* the conditional system has a unique steady state for x_0 and x_∞ as illustrated in Figure 12 Panels (a) and (c).

Table 1: Calibration of Parameters

Parameter		Value	Matched Moment (Information Source)
Benchmark Calibration			
Parameters Set Exogenously			
Year of convergence to balanced growth path		2000	First generation with $\lambda > 0.999$
Length of one generation	m	20 years	Average age at first birth (Dribe, 2004, Mturi and Hinde, 2007)
Years before retirement (at age 5)	R	59	Average effective age of retirement in Sweden (OECD)
Production function	η	0.2857	Elasticity of Substitution between skilled and unskilled labor (Acemoglu, 2002)
Parameters Set Endogenously			
TFP growth	ϕ	0.61	Average growth GDP per capita 1995-2010 (ERS Dataset, Historical Statistics Sweden)
Time cost for unskilled/skilled education	$\{\underline{e}^u, \underline{e}^s\}$	{0,12}	Average years of schooling in 1820 and 2000 (Lutz et al., 2007 and Ljungberg and Nilsson, 2009)
Productivity of ability for Human Capital	α	6.1	Spread of log income distribution 2000 (ECHP)
Mean/standard deviation of ability distribution	$\{\mu, \sigma\}$	{0.49,0.066}	Mean and variance of log income in 2000 (ECHP)
Baseline adult longevity/scope for improvement	$\{\underline{T}, \rho\}$	{45,31}	Average life expectancy at age 5 in 1760-1800 and 2000 (Human Mortality Data Base)
Minimum child survival and elasticity parameter	$\{\underline{\pi}, \kappa\}$	{0.5, 0.005}	Child survival probability in 1800 and 2000 (Human Mortality Data Base)
Preferences	γ	9	Gross (total) fertility around 2000 (World Development Indicators)
Function quality of children	$\{\beta, \underline{r}, \delta\}$	$\{0.23, 4.7, 3.54\}$	Pre- and Post-transitional Fertility rates and growth rate of TFP around 1900 (Keyfitz and Flieger (1968), World Development Indicators, Historical Statistics Sweden)
Initial Conditions			
Initial importance of skilled human capital	x_0	0.04	Initial year of calibration, generations before balanced growth is reached
Initial TFP	A_0	15	Level of log GDP per capita Sweden 2000 (Historical Statistics Sweden)
Cross-Country Analysis			
Parameters Set Endogenously			
Baseline adult longevity/scope for improvement	$\{\underline{\underline{T}}, \rho'\}$	{40,36}	Minimum observed life expectancy at age 5 across country in 2000 (UN)
Function quality of children (high fertility)	$\{\beta', \underline{r}', \delta'\}$	$\{0.75, 3, 1.06\}$	Highest fertility rates 1960 (World Development Indicators)
Distribution of baseline adult longevity			Worldwide distribution of Infectious Human Pathogens (Murray and Schaller, 2010)