American Economic Journal: Macroeconomics 2015, 7(3): 189–225
http://dx.doi.org/10.1257/mac.20130170

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

By Matteo Cervellati and Uwe Sunde*

This paper develops a quantifiable unified growth theory to investigate cross-country comparative development. The calibrated model can replicate the historical development dynamics in forerunner countries like Sweden and the patterns in cross-country panel data. The findings suggest a crucial role of the timing of the onset of the economic and demographic transition for explaining differences in development. Country-specific differences in extrinsic mortality are a candidate explanation for differences in the timing of the take-off across countries and the resulting worldwide comparative development patterns, including the bimodal distribution of the endogenous variables across countries. (*JEL* I12, J11, J13, N33, N34, O41, O47)

Explaining the differences in economic development across the world is a central objective of research in macroeconomics. While there exists a considerable body of investigations of the determinants of long-run development, as discussed in Section I, the empirical literature has only been loosely connected to unified growth theories that investigate nonlinear development dynamics and the mechanics behind the endogenous transition from stagnation to sustained growth. The primary contribution of this paper is to carry out a systematic quantitative analysis of a prototype

* Cervellati: University of Bologna, Piazza Scaravilli 2, 40125 Bologna (e-mail: m.cervellati@unibo.it); Sunde: University of Munich (LMU), Geschwister-Scholl-Platz 1, D-80539 Muenchen (e-mail: uwe.sunde@lmu.de). The authors wish to thank Graziella Bertocchi, Francesco Billari, Javier Birchenall, Michele Boldrin, Raouf Boucekkine, Michael Burda, Nancy Chau, Carl-Johan Dalgaard, Matthias Doepke, David de la Croix, Oded Galor, Nezih Güner, Marcus Hagedorn, Omar Licandro, Omer Moav, Fabrice Murtin, Giovanni Prarolo, Alexia Przkawetz, Rodrigo Soares, Vegard Skirbekk, Chiara Strozzi, Nico Voigtländer, Joachim Voth, Romain Wacziarg, and David Weil, as well as participants at the Workshop on Demographics, Health and Economic Development at Stanford University, the Society for the Advancement of Economic Theory Meeting (SAET) in Vigo, the Society of Economic Dynamics Conference (SED) in Vancouver, the Verein für Socialpolitik (VfS) Councils for Population Economics and Macroeconomics, the VfS annual conference in Graz, the United Nations Development Division in New York, University of California at Los Angeles Anderson, the Norwegian School of Economics Bergen, the University Carlos III, Madrid, the University of Cologne, the University of Copenhagen, the University of Göteborg, the University of Edinburgh, the University of Modena, University Milan Bicocca, University of Munich, the University of Paris I, 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 London School of Economics-University College London development seminar, London, Queen Mary University, London, University of Uppsala, University of Helsinki, the European Society of Population Economics 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 Allied Social Science Association meeting in San Francisco, and the Conference on Long-Run Factors in Comparative Development at Brown University for helpful discussions and comments. We are also particularly grateful to David de la Croix and Vegard Skirbekk for sharing their data.

† Go to http://dx.doi.org/10.1257/mac.20130170 to visit the article page for additional materials and author disclosure statement(s) or to comment in the online discussion forum.

unified growth model, and to study the implications of nonlinear dynamics and of delays in the development process, for cross-country comparative development. The hypothesis motivating this analysis is that different countries follow a similar nonlinear development process, as suggested by the striking similarities in the economic and demographic transition, but crucially differ in the actual timing of take-off to sustained growth, and that comparative development patterns are related to delays in this development process. As a potential determinant of these delays, the analysis investigates the role of differences in the extrinsic mortality environment across countries.

The analysis proceeds in three steps. Section II proposes a simple unified growth theory that can be used to investigate quantitatively the patterns of comparative development in the relevant economic and demographic variables. The framework is based on an occupational choice model with unskilled and skilled human capital and endogenous differential fertility. Bidirectional feedbacks between the education composition of the population, technological progress, and demographic change (in particular mortality) eventually trigger a growth take-off. The paper contributes a simple prototype unified growth model with analytical predictions that are in line with the stylized facts of long-run economic development, in terms of income and education, and of demographic development, in terms of adult longevity, child mortality, and (gross and net) fertility.

Section III presents a quantitative version of the model. In Section IIIA the model is calibrated to the long-run development patterns of Sweden, the textbook case of long-run economic and demographic development. The calibration strategy involves setting the (time-invariant) parameters of the model by targeting data moments on the balanced growth path in 2000 and, in some cases, before the onset of the transition around 1800. The simulated model produces the endogenous evolution of the economy over a long period of time (from year 0 to year 2000) that includes the onset of the transition and the convergence to the balanced growth path. Section IIIB compares the simulated and the actual data for Sweden over the period 1760–2000. The comparison between simulated and actual data along the full transition from stagnation to sustained growth documents the ability of the model to reproduce the development patterns in the different dimensions of the historical data, including those that had been more difficult to replicate quantitatively in the previous literature (like, e.g., the reduction in net fertility).

Section IV explores the ability of the model to account for cross-country comparative development patterns. This analysis effectively constitutes an exploration of the quantitative implications of a prototype unified growth model since no data moments of the cross-country analysis are targeted for the calibration. The results are therefore informative on the model's fit to empirical patterns "out of sample." Section IVA investigates the hypothesis that all countries follow a similar nonlinear development path with a main difference being the timing of the transition. The simulated correlations between education and other equilibrium variables, including adult longevity, fertility, and income per capita are compared to the respective correlations in cross-country data in 1960 and 2000. Despite the underlying nonlinear development dynamics, the simulated data display the same monotonic, and almost linear, cross-country correlations between the equilibrium share of educated individuals and all other central variables as found in the empirical data. The model also reproduces other cross-sectional data patterns, such as a hump-shaped relationship between life expectancy and the subsequent change in the education composition, or the well documented concave relationship between income per capita and life expectancy known as the "Preston Curve."

Section IVB explores the role of cross-country differences in extrinsic mortality ("baseline longevity") for the development process. We simulate a counterfactual economy that differs from the benchmark model calibrated for Sweden only in terms of baseline longevity, which is calibrated by targeting data moments for the countries with the highest mortality in 2000 in the world (instead of Sweden in 1800). The results document that empirically reasonable cross-country differences in baseline longevity can result in substantial delays (of more than a century) of the economic and demographic transition, despite leaving the cross-sectional data patterns, including the Preston Curve, essentially unaffected. The analysis also illustrates that the timing of development for the different continents (with Europe on one extreme and Africa on the other) is broadly consistent with measures of extrinsic mortality, such as the exposure to human pathogens. The results thereby provide a link between the unified growth literature and the empirical literature on the fundamental determinants of long-run development, and contribute to the empirical debate about the role of life expectancy for development.

Finally, Section IVC presents the results from the simulation of an artificial world composed of countries that differ in terms of baseline longevity, but that are otherwise identical to the benchmark calibration for Sweden. The results deliver cross-country distributions of all variables of interest that match quantitatively the actual world distributions, which are bimodal in 1960 and rather unimodal in 2000. The findings show that the unified growth framework provides a natural explanation for the changing bimodality of the distributions in the different central variables due to the changes of all variables during the transition to the balanced growth path. The quantitative results also suggest an acceleration in the development path of today's developing countries, compared to the development of the European forerunners, in terms of demographic conditions (mortality and fertility) and (to a lower degree) economic development.

The analytical derivations and proofs are relegated to the Appendix. The details of the calibration, the data sources, and additional material are made available in the online Appendix.

I. Related Literature

Addressing the research question of this paper requires a model that is able to reproduce the main stylized facts of the economic and demographic transitions, and that is suitable for a quantitative analysis of the long-run development path, including the endogenous transition phase. The general equilibrium framework presented in this paper is based on an occupational choice model with unskilled and skilled human capital. Education and fertility decisions crucially depend on economic and demographic conditions (in particular technology and mortality) whose dynamics are modeled through intergenerational externalities. The resulting prototype unified growth framework builds on Galor and Weil (2000), Cervellati and Sunde (2005), and

Soares (2005).[1] The model features differential fertility across different education groups, similar to de la Croix and Doepke (2003), and produces qualitative predictions that are in line with the stylized facts.[2]

The paper contributes to the existing literature of quantitative studies of long-term development. A quantitative analysis that exploits comparative statics around the balanced growth path as in most of the existing literature (see, e.g., Jones, Schoonbroodt, and Tertilt 2011, for a survey) is not suited for the purpose of this paper, which requires studying the endogenous take-off of the transition from quasi-stagnation to growth. The logic of analysis is closer in spirit to the unified growth studies presented by Lagerlöf (2003); Doepke (2004); Strulik and Weisdorf (2008); and de la Croix and Licandro (2012), which simulate the dynamic long-run evolution of an economy, including the demographic dynamics. One advantage of the prototype model presented in this paper is that it can reproduce the patterns of the economic and demographic transition (in terms of both fertility and mortality) with a parsimonious set of time-invariant parameters that have a clear economic interpretation and that can be calibrated targeting observable data moments.[3]

The calibrated model is not only suitable for the analysis of the evolution of one economy over time, as the literature cited above, but also for the investigation of patterns of comparative development at different moments in time. The solution of the prototype unified growth model presented below is always interior, so that the take-off is not generated by the exit from corner solutions of the dynamic system. This technical feature permits conducting smooth comparative statics on the main variables of interest, such as baseline longevity, and investigating their role for the differential delays in take-off. In this context, the paper contributes to the few quantitative papers in the literature that investigate the timing of the transition, such as Ngai (2004) and Chakraborty, Papageorgiou, and Sebastian (2010). To our knowledge, this paper offers the first application of a unified growth framework for the quantitative analysis of cross-country comparative development patterns.

The quantitative results shed new light on some unsettled empirical questions. The findings suggest that even moderate differences in extrinsic mortality can have relevant implications for comparative development patterns by delaying the

[1] See also Doepke (2004) for an investigation of the interplay between technology and fertility; de la Croix and Licandro (1999); Kalemli-Ozcan, Ryder, and Weil (2000); and Boucekkine, de la Croix, and Licandro (2002, 2003) for investigations of the role of mortality; and Aiyar, Dalgaard, and Moav (2008) for a modeling of technology dynamics in pretransitional economies. Changes in mortality have also been modelled as resulting from rational investments in health, see de la Croix and Licandro (2012) and Dalgaard and Strulik (2014).

[2] In particular, the theory can rationalize the observed drop of net fertility below the pretransitional levels, which is one of the defining elements of the demographic transition as conceptualized by demographers, see, e.g., Chesnais (1992). This fact has been difficult to rationalize with fertility theories based on the quantity-quality trade-off or child mortality, see Kalemli-Ozcan (2003) and Doepke (2005). Building on an occupational choice framework, rather than only a quantity-quality, the proposed unified growth theory can generate changes in net fertility without having to impose restrictive assumptions on the utility function, see also Mookherjee, Prina, and Ray (2012).

[3] Previous quantitative studies that are not based on a unified growth framework have investigated the long-run development dynamics by exogenously calibrating the dynamics of relevant variables (like technology or mortality) to match the empirical time series, see Eckstein, Mira, and Wolpin (1999); Kalemli-Ozcan (2002); de la Croix, Lindh, and Malmberg (2008); Bar and Leukhina (2010); and Cervellati and Sunde (2013). Differently from that, our approach involves calibrating the time-invariant parameters of the functional relationships that produce the (endogenous) dynamics of the relevant variables (like technology and mortality), which are not matched by construction.

economic and demographic transition, thereby complementing the evidence on the deep determinants of long-run development, see, e.g., Spolaore and Wacziarg (2013) and Ashraf and Galor (2013). By demonstrating that such differences in mortality affect the nonlinear development dynamics but leave the cross-sectional patterns between central variables, such as income, life expectancy, education, and fertility essentially unchanged, the results offer a rationale for the mixed empirical findings based on panel data and linear regression frameworks.[4] The results also show that the unified growth framework can generate the concave relationship between life expectancy and income per capita, the so-called Preston Curve, whose underlying mechanisms are still not well understood.[5] Finally, the quantitative analysis provides a natural explanation for the observation of the bimodality in the distribution of income, see Azariadis and Stachurski (2005) and life expectancy, see Bloom and Canning (2007b), across the world, and predicts similar bimodalities in the distributions of education and fertility that are consistent with the data.

II. A Prototype Unified Growth Model

This section presents the theoretical framework with the functional forms that are applied in the calibration in Section III, even though they are not needed for the analytical results in Section IIB. The functional forms are specified in line with the previous literature and the available evidence, and to minimize the number of parameters.

A. *Set-Up*

Population Structure.—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, of length $K_t = k$, and adulthood, with duration 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 represents both life expectancy at age k and the maximum duration of adulthood.[6] In the model, T_t is a summary statistic of the effective time available during adulthood and can be also interpreted as a "health augmented" time endowment of adults.

[4] The evidence from linear estimation frameworks on the effect of life expectancy on growth is mixed, see, e.g., Acemoglu and Johnson (2007) and Lorentzen, McMillan, and Wacziarg (2008). These findings can be reconciled by explicitly accounting for nonlinear dynamics of life expectancy, population, and education that are consistent with the economic and demographic transition, see Cervellati and Sunde (2011 and 2015). The results also complement quantitative studies of the consequences of exogenous variation in mortality for development, see, e.g., Ashraf, Lester, and Weil (2008).

[5] This relationship has been documented for the first time by Preston (1975). See Deaton (2003) and Bloom and Canning (2007a) for a discussion of the debate. To our knowledge, the only other theory that can provide a theoretical rationale for the Preston Curve is by Dalgaard and Strulik (2014).

[6] This modeling of child survival and adult longevity follows Soares (2005). It is formally isomorphic to a "perpetual youth" modeling, where longevity is one over the age of independent adult survival probability. Considering a deterministic longevity simplifies the set-up by abstracting from uncertainty and, as discussed below, by allowing for a direct match between the simulated and empirical data for child mortality and life expectancy at age five. In the quantitative analysis, $1 - \pi_t$ corresponds to child mortality, and T_t corresponds to life expectancy at age five (so that $k = 5$). Assuming a constant death rate before age 5, life expectancy at birth is $\pi_t(5 + T_t)$.

Reproduction is asexual and takes place at age $m \geq k$, which is the length of a generation. A cohort 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 cohort. Every individual of cohort t is endowed with an innate ability $a \in [0, 1]$, which is randomly drawn from a distribution $f(a)$. For the calibration of the model we assume a (truncated) normal distribution of ability with mean μ and standard deviation σ.

Preferences and Production.—During childhood individuals are fed by their parents and make no choices. At the beginning of adulthood, those individuals that survive childhood make decisions about their own education and their fertility to maximize their (remaining) lifetime utility. Individuals derive utility from own consumption, c, and the quality, q, of their (surviving) offspring, πn. The lifetime utility of an individual of generation t is additively separable and given by

$$\int_0^{T_t} \ln \, c_t(\tau) \, d\tau + \gamma \ln \left(\pi_t n_t q_t\right), \tag{1}$$

where $\gamma > 0$ is the weight of the utility that parents derive from their surviving children relative to their own lifetime consumption as adults.[7]

We set the subjective discount rate to zero and assume that individuals perfectly smooth consumption within their adult period of life, $c_t(\tau) = c_t$ for all τ. This allows abstracting from the path of consumption during the life cycle, which is of no primary relevance for the investigation of the long-term evolution of the economy across generations.[8] 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 offspring.[9]

The inputs of production are *skilled* human capital, denoted by s, and *unskilled* human capital, denoted by u. We treat human capital as inherently heterogenous across generations. In line with the literature on vintage human capital, this reflects the view that individuals acquire their skills in environments characterized by the availability of a particular technology. The aggregate stocks of human capital of each type, H_t^u and H_t^s, supplied by generation t are used to produce the unique consumption good with a constant returns to scale technology

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

where $\eta \in (0, 1)$. Generation t only operates the technological vintage t, which is characterized by the relative productivity of skilled human capital, $x_t \in (0, 1)$, and total factor productivity (TFP) A_t. The production function (2) is a specialized (CES) formulation of the vintage production function by Chari and Hopenhayn (1991). As

[7] The utility formulation follows Soares (2005). As in Becker and Lewis (1973) parents derive utility from the quality of their children, which allows studying the change in the quantity-quality trade-off in the simplest way.

[8] See also Rogerson and Wallenius (2009) for a similar assumption, which is equivalent to assuming a small open economy facing a zero discount rate.

[9] The actual formulation of the utility function and the fact that longevity implicitly affects the weight of the utility from consumption and children is irrelevant for the results. As shown in a previous version of the paper, one could equivalently assume that individuals derive utility from average per period lifetime consumption and children as in Galor and Weil (2000).

in Boucekkine, de la Croix, and Licandro (2002), the vintage of technology is linked to generation-specific knowledge in terms of skilled and unskilled human capital.[10]

The returns to human capital are determined in general equilibrium on competitive markets and equal marginal productivity,

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

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 $j = \{u, s\}$. Individual ability is relatively more important in producing skilled human capital. 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 she decides to become skilled, and $h^u(a) = e^{\alpha\mu}$ if she decides to be unskilled. An individual that decides to become skilled, respectively unskilled, pays a fixed cost, measured in terms of adult time, of $\underline{e}^s > \underline{e}^u \geq 0$.[11]

Raising a child involves a time cost $r_t = \tilde{r}_t + \underline{r}$ where $\underline{r} > 0$ is a fixed time cost that needs to be spent and $\tilde{r}_t \geq 0$ is the extra time that can be spent voluntarily in addition.[12] The time spent with a child increases the child's quality according to

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

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 quality (the quality time $\tilde{r}_t$). 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).

The 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), $R > 0$.[13] The effective time available for productive activities during adulthood is therefore bounded from above by $\overline{T}_t = \min\{T_t, R\}$. An individual with education

[10] Vintage models that relax the assumption that human capital is perfectly homogenous across different age cohorts are empirically appealing in the context of long-term development, where cohorts of workers of different age acquire knowledge of different technologies. This vintage structure is not needed for the main mechanism and the analytical results, but it allows for a transparent quantitative analysis as 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.

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

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

[13] The assumption of a limit R, which may be due to compulsory retirement or some other 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 older ages. In the quantitative analysis, the parameter R is calibrated exogenously to match the effective retirement age.

$j = \{u, s\}$ cannot use more than the available time and cannot spend more than the total earnings for total consumption.[14] The budget constraint conditional on being skilled or unskilled, $j = \{u, s\}$, is thus given by

$$T_t c_t = \left(\overline{T}_t - \underline{e}^j - \pi_t n_t r_t\right) w_t^j h_t^j(a). \tag{5}$$

The problem of an individual with ability a born in generation t is to choose the type of human capital to be acquired, $j \in \{u, s\}$, the number of children, n_t, and the time invested in raising each child, r_t, so as to maximize utility (1) subject to (5).

Adult Life Expectancy and Child Survival.—In line with the available evidence, we consider a differential impact of human capital and income on adult longevity and child mortality.[15]

Adult longevity of generation t is assumed to be increasing in the share of skilled individuals in the parent generation,

$$T_t = \Upsilon(\lambda_{t-1}) = \underline{T} + \rho\lambda_{t-1}, \tag{6}$$

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.[16] Since $\lambda \in (0, 1)$, the maximum level of adult longevity is given by $\overline{T} = \underline{T} + \rho$.

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

$$\pi_t = \Pi(\lambda_{t-1}, y_{t-1}) = 1 - \frac{1 - \underline{\pi}}{1 + \kappa\lambda_{t-1} y_{t-1}}, \tag{7}$$

with $\kappa > 0$ and where $1 > \underline{\pi} > 0$ is the baseline child survival that would be observed in an economy with $\lambda_{t-1} y_{t-1} = 0$.[17]

Technology.—Technological progress, which takes place with 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, is skill biased. The relative

[14] See the Appendix for the expressions of the time and resource constraints.

[15] Environmental factors, in particular macroeconomic conditions, are crucial determinants of individual health but appear to affect adult longevity and child mortality in different ways. Cutler, Deaton, and Lleras-Muney (2006) suggest 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 relatively more important for increasing the survival probability of children, see Wang (2003) for a survey.

[16] This reduced form modeling allows going beyond the assumption that changes in mortality are fully exogenous (as in, e.g., Jones and Schoonbroodt 2010) in the simplest and most parsimonious way. The evolution of longevity could be made endogenous to human capital by extending the model to the consideration of optimal investments in health along the lines of de la Croix and Licandro (2012).

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

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, $1 - x_{t-1}$,

$$(8) \qquad \frac{x_t - x_{t-1}}{x_{t-1}} = X(\lambda_{t-1}, x_{t-1}) = \lambda_{t-1}(1 - x_{t-1}).$$

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,[18]

$$(9) \qquad g_{t+1} = \frac{A_{t+1} - A_t}{A_t} = G(\lambda_t) = \phi\lambda_t\,, \ \phi > 0.$$

B. *Analytical Results*

This section derives analytical results, in terms of optimal decisions, intragenerational general equilibrium, and the dynamic evolution of the economy over time, that are useful for the interpretation of the quantitative analysis.

Equilibrium Fertility.—The first order conditions uniquely identify the optimal fertility and the time spent raising children conditional on the type of human capital acquired by each individual.[19] The resulting average fertility in the population is given by

$$(10) \quad n_t^* = N(T_t, \lambda_t, \pi_t) = \frac{\gamma}{(T_t + \gamma)\, r_t^* \pi_t}\left[(1 - \lambda_t)(\overline{T}_t - \underline{e}^u) + \lambda_t(\overline{T}_t - \underline{e}^s)\right],$$

where λ_t denotes the share of individuals of generations t that acquire skilled human capital and r_t^* is the optimal time invested in children.[20]

To facilitate the interpretation of the quantitative results in Section III, let us briefly comment on the role of demographic variables for gross and net fertility. Gross fertility is decreasing in π_t through a *substitution effect*, but net fertility is independent of π_t. The effect of adult longevity on fertility is more complex. Higher adult longevity T_t increases gross fertility as long as $T_t < R$ due to a positive

[18] This can be seen as a reduced form of endogenous growth models such as Aghion and Howitt (1992), where ϕ can be interpreted as the average size of an innovation and the labor involved in research is increasing in λ_t.

[19] See Lemma 1 in the Appendix for the optimal individual choices and their derivation.

[20] As characterized in Lemma 1 in the Appendix, the optimal time investment in children is given by,

$$(11) \qquad r_t^* = \max\left\{\underline{r}, \frac{1 - \left[1/\left(\delta(1 + g_{t+1})\right)\right]}{1 - \beta}\underline{r}\right\}.$$

Along the lines of Galor and Weil (2000), when technical progress g_{t+1} is too low, parents may optimally decide not to invest any extra time in raising their children beyond the minimum level, so that $r_t^* = \underline{r}$. Provided that a positive extra time is invested in raising children, faster technological progress g_{t+1} increases r_t^* and reduces optimal fertility for unskilled and skilled individuals.

income effect, but decreases gross fertility when $T_t \geq R$ as the income effect turns negative.[21] In addition, a higher T_t reduces fertility by a *differential fertility* effect if, as shown below, it increases the share of skilled workers, λ_t, who have fewer children, see Skirbekk (2008) for evidence. The existence of differential fertility implies that increases in adult longevity may materialize in a reduction of both gross and net fertility. An indirect effect arises from the effect of the share of skilled individuals on technological progress.[22]

Intragenerational General Equilibrium.—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. The corresponding unique share λ_t of individuals that find it optimal to acquire skilled human capital is increasing in the relative wage w_t^s / w_t^u, decreasing in e^s, increasing in adult longevity T_t, and is unaffected by child mortality π_t.[23]

The general equilibrium for generation t is given by the share of skilled individuals λ_t^* where individual optimal choices and market wages are jointly determined.

PROPOSITION 1: *For any* $\{T_t \in (\underline{e}^s, \infty), \pi_t \in (0,1), x_t \in (0,1)\}$ *there exists a unique*

$$\lambda_t^* = \Lambda(T_t, x_t) \tag{12}$$

and H_t^{j*}, w_t^{j*} *for* $j = u, s$, *for which individual optimal education decisions are consistent with market wages. 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 proof of Proposition 1, and the explicit characterization of the function $\lambda_t^* = \Lambda(T_t, x_t)$ are reported in the Appendix. The key state variables affecting λ_t^* are adult longevity T_t and the relative importance 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 nonlinear, however. When T_t is low the locus $\Lambda(T_t, x_t)$ is convex, and large increases in T_t are needed to induce a significant fraction of individuals to acquire skilled human capital since the fixed cost $\underline{e}^s > \underline{e}^u$ prevents a large part of the population from receiving sufficient lifetime earnings when becoming skilled. When 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

[21] Increases in longevity above R (so that $\overline{T}_t = R$) reduce fertility. A longer expected time in retirement requires devoting more income to consumption (to keep a constant consumption over the life cycle) thereby lowering fertility.

[22] Fertility is decreasing with the time invested in children, in line with a standard quantity-quality trade-off. The quantity-quality trade-off is not directly affected by adult longevity and child mortality, however, and the optimal time spent raising a child does not depend on the type of human capital acquired by parents. Parents substitute, however, quantity for quality in the face of technological progress, which depends on λ_t. Higher longevity therefore reduces fertility also indirectly by changing the future parental investments in the quality of children.

[23] See Lemma 2 in the Appendix.

returns to human capital of either type, which drive down the relative wage w^s/w^u.[24] To shorten notation in the following we denote by λ_t the equilibrium share of skilled workers.

Development Dynamics.—The dynamic path is given by a sequence $\{T_t, x_t, \lambda_t, A_t, \pi_t, n_t\}$ for $t = [0, 1, \ldots, \infty)$, which results from the evolution of the nonlinear first-order dynamic system,

$$\begin{cases} T_t &= \Upsilon(\lambda_{t-1}) \\ x_t &= X(x_{t-1}, \lambda_{t-1}) \\ \lambda_t &= \Lambda(T_t, x_t) \\ A_t &= A_{t-1}(1 + G(\lambda_{t-1})) \\ \pi_t &= \Pi(T_{t-1}, x_{t-1}, \lambda_{t-1}, A_{t-1}) \\ n_t &= N(T_t, \lambda_t, \pi_t) \end{cases} . \tag{13}$$

Notice that the system is block recursive.[25] 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 the importance of skilled human capital for production. Total factor productivity, A_t, child mortality, π_t, and fertility, n_t, only depend on past levels of T_t, λ_t, and x_t, and do not affect the evolution of the dynamic system (13) in terms of these variables.

The development process involves reinforcing feedbacks between increases in human capital, and increases in adult longevity and technological progress. The different phases of development are illustrated in Figure 1 in the online Appendix.

PROPOSITION 2: [*Economic and Demographic Transition*] *For a sufficiently low* x_0, *the development path is characterized by:*

(*i*) *An initial phase with* $\lambda \simeq 0$, *low longevity*, $T \simeq \underline{T}$, *high child mortality* $\pi \simeq \underline{\pi}$, *slow income growth, and gross fertility given by*

$$n \simeq \gamma \frac{\underline{T} - \underline{e}^u}{(\underline{T} + \gamma)\, \underline{r}\, \underline{\pi}}. \tag{14}$$

[24] Characterizing the second derivative of $\lambda(T_t)$ analytically 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 III can be shown numerically and can be established analytically when imposing assumptions on the shape of the ability distribution (like, e.g., a uniform distribution).

[25] Adult longevity T_t is as in (6), while the evolution of x_t is characterized by (8). The share of skilled, λ_t, in turn is determined by the intragenerational equilibrium implied by Proposition 1. TFP, A_t, evolves as in (9), while child survival probability, π_t, evolves according to (7) and also depends on y_{t-1} and, therefore on $T_{t-1}, x_{t-1}, \lambda_{t-1}$, and A_{t-1}. Fertility is determined in (10). A noteworthy feature of the dynamic system (13) is that all variables are characterized by interior solutions with the speed of their dynamics changes varying over time until the balanced growth path is reached. This is convenient for the quantitative analysis since it allows smooth comparative statics.

(*ii*) *A final phase of balanced growth in income per capita, with* $\lambda \simeq 1, T \simeq \overline{T}$, *low child mortality* $\pi \simeq 1$, *and with gross fertility given by*[26]

$$n \simeq \gamma \frac{\min\{\overline{T}, R\} - \underline{e}^{s}}{(\overline{T} + \gamma)\,\overline{r}}. \tag{15}$$

(*iii*) *An endogenous transition from* (*i*) *to* (*ii*).

Adult longevity and the share of skilled individuals affect the timing of the transition to the balanced growth path, whereas fertility and child mortality do not affect the dynamics of the economy. A lower baseline adult longevity $\underline{T}$ implies a later onset of the economic and demographic transition, since higher levels of technology x_t are required to induce the endogenous disappearance of the initial phase and the take-off to a balanced growth path.

III. Quantitative Analysis of Long-Run Development

A. *Benchmark Calibration*

The calibration of a unified growth framework requires setting the time-invariant parameters of a model that produces a dynamic evolution. This evolution is not limited to the balanced growth path but includes the transition in the different variables. More specifically, calibrating the model proposed in Section II requires setting the values of 15 parameters that characterize the utility and production function $\{\gamma, \eta\}$, technological progress ϕ, adult longevity $\{\underline{T}, \rho\}$, 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, we allow for the possibility that individuals retire at some exogenously given age R. 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 the evolution of all variables of interest is determined endogenously by the model along the development path, for $t = [0, 1, \ldots, \infty)$ and it involves a phase of quasi-stagnant development, which is eventually followed by the endogenous transition and convergence to the balanced growth path.

For the calibration, we use data for Sweden as the prototypical example of the economic and demographic transition. Data of comparably high quality are available for Sweden since the mid-eighteenth century, which makes it a natural benchmark for evaluating the quantitative fit of the model in terms of long-term development patterns.[27] Some parameters are set (exogenously) by matching directly observable counterparts in the data for Sweden or following the parametrization of existing quantitative studies. A second set of parameters is calibrated by solving the equilibrium conditions of the model and matching them to observable data moments

[26] The optimal investment in children, $\overline{r}$, corresponds to the growth rate on the balanced growth path, $g_{t+1} = \phi$. See Lemma 3 in the Appendix.

[27] An earlier version demonstrated that the model equally well captures development dynamics in England.

on the balanced growth path. In the model, the balanced growth path is reached when all individuals get involved in formal education, $\lambda \simeq 1$; for Sweden this corresponds to the year 2000.[28] The calibration of the parameters of some functions requires solving systems of simultaneous equations by exploiting information on data moments at two points in time. In these cases we target data moments both on the balanced growth path (in 2000) and before the onset of the transition (in 1800).

When comparing the simulated data to the actual time series data for Sweden, the targeted moments (most of them referring to 2000 and few to 1800) will be matched by construction. The comparison between simulated and actual data along the transition from (quasi-)stagnation to sustained growth (and in particular in the period 1750–2000 for which we have complete time series data) is not matched by construction and is thus informative for evaluating the fit of the model to the data.

Below we give a brief description of the calibration of the model. Table A1 contains summary information about the data moments used as targets, the data sources, and the calibrated parameters. Due to space limitations, the details of the calibration, the data sources, and the discussion of the sensitivity of the parametrization to alternative calibration strategies are reported in the online Appendix.

Parameters Set Exogenously.—The length of generations, m, the age of retirement R and the fixed (time) cost of education are set to match observable counterparts. The elasticity of substitution in the production function, η is set in line with the literature.

Length of a Generation: The length of a generation is set to $m = 20$ years. Across countries the average age of women at first birth before the demographic transition is approximately 20 years.[29] A twenty-year frequency also allows for a direct match of the simulated data with cross-country panel data without the need for interpolation.

Age of Retirement: The average effective retirement age was around 64 in Sweden in 2000. Since R is the number of years before retirement as perceived at the end of childhood, at age $k = 5$, we set $R = 59$.[30]

Human Capital: To set the fixed (time) cost of education, $\{\underline{e}^u, \underline{e}^s\}$, we target the average years of schooling in Sweden (for the cohort age 25–35), which was 12 years in 2000; see Lutz, Goujon, and Sanderson (2007). The available data

[28] The enrollment shares in Sweden have essentially reached 100 percent in primary and lower secondary education after 1980 and 1995, respectively.

[29] 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 pretransitional countries in Africa nowadays, see Mturi and Hinde (2007).

[30] The data source is Organisation for Economic Co-operation and Development (2015). In spite of substantial changes in the health status at old age (which may facilitate old age labor supply) and the introduction of welfare programs (that anticipated retirement), the average age of retirement was relatively stable across historical cohorts in western countries. See Hazan (2009) and Strulik and Vollmer (2013). In an earlier version, the analysis abstracted from the possibility of retirement before death, with similar results.

suggest approximately one year of schooling around the onset of the transition.[31] This implies setting $\underline{e}^s = 12$ and $\underline{e}^u = 0$.

Production Function: The elasticity of substitution between skilled and unskilled workers is set to $1/(1 - \eta) = 1.4$ in the literature (see Acemoglu 2002) so that $\eta = 0.285$.

Parameters Set by Solving the Model.—The parameters of the function driving the evolution of *TFP*, the ability distribution and the preferences for fertility are set by solving the model moments and matching them to the corresponding data moments for Sweden for 2000 (i.e., on the balanced growth path).

Technological Progress: The parameter of *TFP*, ϕ, is set to match the average annual growth rate of income per capita on the balanced growth path (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. This implies targeting a growth factor of 1.61 over a twenty-year period. Given the function (9), and with $\lambda = 1$ along the balanced growth path, we set $\phi = 0.61$.[32]

Ability Distribution: The parameter α, which determines the importance of ability for the acquisition of individual human capital, and the moments of the ability distribution $\{\mu, \sigma\}$ are calibrated by targeting data for the income distribution in Sweden in 2000. The income distribution on the balanced growth path only involves the production of skilled workers, since $\lambda \simeq 1$. The empirical income distribution is approximately log-normal between the fifth and ninety-fifth percentile, with slightly thicker tails.[33] The individual (per period) income of 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 is matched by setting $\alpha = 6.1$.[34] Matching the data moments implies setting $\alpha\mu = 3$ and $\alpha\sigma = 0.4$,

[31] 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 pretransitional 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.

[32] See, e.g., the historical statistics from the Bank of Sweden (Edvinsson 2014). Targeting estimates of growth in multifactor productivity, labor productivity, or the Solow residual deliver similar magnitudes for ϕ.

[33] We use micro data from the European Community Household Panel (Eurostat 2014) 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. Incomes are converted to US dollars using an average exchange rate of 9 kroner for one US dollar in 2000. The relevant data moments extracted from this dataset are broadly consistent with other data sources based on register data and alternative surveys for gross earnings, see Domeij and Floden (2010).

[34] 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. 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

which for $\alpha = 6.1$, implies $\mu = 0.49$ and $\sigma = 0.066$. A similar calibration would be obtained from a typical distribution of cognitive ability as a proxy for ability in acquiring skilled human capital.[35]

Preferences: The parameter γ is calibrated targeting gross fertility $n = 1$ along the balanced growth path, which is also equivalent to targeting the net reproduction rates approximately at replacement levels, setting child survival to $\pi = 0.996$ consistent with Sweden in 2000.[36] The time spent in raising children is determined endogenously in the model and changes with the growth rate of income and technology. We set a target for the number of years spent raising a child in 2000 of $r = 5$.[37] With $\lambda = 1$, $\pi = 0.996$, $R = 59$, and $r = 5$ this delivers $\gamma = 9$.

The remaining parameters relating to the evolution of adult longevity, child mortality and fertility, are set by solving systems of simultaneous equations and targeting data for 2000 and for 1800, which represents the latest pretransitional period for which reliable data are available for all the variables of interest. As discussed above, baseline longevity affects the timing of the take-off while child mortality and fertility do not affect the timing of the transition as a consequence of the block-recursiveness of the system (13) and the lack of scale effects.

Adult Longevity: Ideally, the levels of $\underline{T}$, which represents baseline longevity, and of ρ, the scope for improvements in longevity, would be calibrated exogenously. Given the lack of reliable historical data, the calibration of these two parameters is done by solving two equations of the type $T_t = \underline{T} + \rho\lambda_{t-1}$ at two points in time using data for 2000 and 1800. This appears a reasonable strategy since the share of educated individuals is still very small until the onset of the transition in 1800. To solve the system we consider a share of skilled workers of $\lambda = 0.1$ before the transition, which roughly corresponds to the enrollment rates in early nineteenth century Sweden.[38] Life expectancy at age 5 in Sweden was approximately 48 around

income being 6.7, 9.7, and 12.8, respectively, and with a standard deviation of 0.41. 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, Koreshkova, and Restuccia (2011) match a variance of log permanent earnings in the United States of 0.36. Robustness checks show that the results are fairly insensitive to the particular parameter of the dispersion.

[35] 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 to a support $a \in [0, 1]$, with $\mu = 0.5$ and $\sigma = 0.075$, with very similar results.

[36] 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 2014). 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.

[37] The target is in line with the estimates by Haveman and Wolfe (1995). This is equivalent to setting a target for the share of work life that is spent in raising a child which, is about 15 percent, as in Doepke (2004) or de la Croix and Doepke (2003).

[38] Even though the available data sources provide slightly heterogeneous information on enrollment rates in the early nineteenth century Sweden, the estimates range from about 5 to about 15 percent, see de la Croix, Lindh, and Malmberg (2008) and Ljungberg and Nilsson (2009). The precise value of λ before the transition is therefore of little importance and the results for alternative parameters obtained by assuming for 1800 levels of λ up to 0.3 are essentially the same.

1800 and 76 in 2000.[39] With these targets, the parameters of the function (6) 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 (7) delivers a baseline child survival probability of $\underline{\pi} = 0.5$ and a $\kappa = 0.005$.[40]

Production Function of Children's Quality: The parameters $\{\beta, \underline{r}, \delta\}$ are calibrated by targeting the levels of gross fertility for Sweden in 1800 and 2000, and the growth rate of technology in 1900, which is taken as the period of the exit from the corner solution of zero investments in children's quality in light of the pronounced drop in fertility in Sweden around this time.[41] To calibrate the parameters of the function, in (4), we use the optimal time investment by parents in children and the minimum growth rate of technology $\underline{g}$ for which parents spend some positive extra time in raising children. Given the targets we obtain $\{\beta = 0.23, \underline{r} = 4.7, \delta = 3.54\}$.

Initial Conditions.—We finally also need to determine the initial conditions in terms of the importance of skilled human capital in the production function, x_0, and the level of total factor productivity A_0. Given these initial productivity parameters, the dynamic system (13) generates the endogenous evolution of all variables of interest along the development path for $t = [0, 1, \ldots, \infty)$. The initial importance of skilled human capital in the production function, x_0, only affects the number of generations before the take-off in the simulation. Choosing x_0 sufficiently low implies simulating the model before the onset of the phase transition that triggers the convergence to the balanced growth path. Setting $x_0 = 0.04$, the simulation converges to the balanced growth path (which is assumed to be reached when λ exceeds 0.999) in 100 generations thereby covering the period from year 0 A.D. until 2000.[42] The initial level of TFP is a scale parameter that does not affect the

[39] 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 Database 2014). Similar figures are documented for England, France, and Italy, see Woods (1997) and Bideau, Desjardins, and Perez-Brignoli (1997) and Lewis and Gowland (2007). Also note that, as discussed below, in 2000 child mortality is around 0.004, which implies the convergence of life expectancy at 5 plus 5 years of 80.74, and of life expectancy at birth of 80.45.

[40] The data on child survival are from the Human Mortality Database (2014). The levels of income per capita needed for the computation of the parameters are taken from the historical statistics from the Bank of Sweden (Edvinsson 2014), converted to US dollars using an average exchange rate in 2000 of 9 kroner for one US dollar. The income levels used for the calibration of condition (7) are \$22,717 and \$884, which correspond to the gross domestic product (GDP) per capita of Sweden in 2000 and 1800, respectively, in US dollars per 2000.

[41] Gross fertility in Sweden in 1800 and 2000 was $n = 2.3$ and $n = 1$, respectively. The data are from Keyfitz and Flieger (1968) and the World Development Indicators (2014). As documented in the demographic literature, and as clearly visible in the time series reported below, a noticeable drop in gross fertility occurs in Sweden around 1900. We take 1900 to be the period of the exit from the corner solution of the quantity-quality trade-off. Estimates of TFP and income per capita growth around 1900 vary between 0.7 and 1.7 percent per year, see Krantz and Schön (2007), Schön (2008) and Greasley and Madsen (2010). We set the level of $\underline{g}$ to 1.2 percent per year with a corresponding growth factor over a 20-year generation of 0.27.

[42] Setting a smaller x_0 only implies increasing the number of generations before the take-off, that is, it only implies starting the simulation further back in time.

endogenous evolution of the system (13), but only affects the level of production. We set $A_0 = 15$ to match the level of log GDP per capita in the year 2000.[43]

B. *Time Series Results*

The dynamic evolution of the model economy is characterized by a long period of slow development followed by a (comparatively) rapid transition to a sustained growth path that takes place over a time horizon of about 200 years.[44]

Figure 1 restricts attention to the period 1750–2000 and compares the simulated data to the corresponding time series of historical data from Sweden. To interpret the results, recall that the different data moments in 2000 are matched by construction as they reflect the balanced growth path to which many parameters are calibrated. Panels A and B of Figure 1 report the evolution of life expectancy at birth (T0) and at age five (plus five years, T5), and child mortality rates, respectively. The calibration targets life expectancy at age five as well as child mortality at two points in time (in 1800 and 2000). During the transition, however, the values of these variables are generated by the simulated model and are not constrained to match data moments by construction. Nevertheless, 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. Figure 1, panels C and D plot the share of skilled individuals, λ, against the primary school enrollment 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 enrollment 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 the emergence of school systems, the model's dynamics fit the data well.

Figure 1, panel 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 investment in child quality around 1900. The simulation matches the initial and terminal levels, as well as the intermediate transition. The eventual reduction in net fertility, which, as discussed in Section I, has been difficult to generate in previous quantitative studies, is matched by the model due to the presence of the differential fertility effect that is absent in models exclusively based on the quantity-quality trade-off.[45] Compared to the historical data, the model does not

[43] The data are from Edvinsson (2014).

[44] When considered over the entire simulation period from year 0 to 2000, the simulated data for the variables of interest exhibit a lengthy phase of slow development, followed by the endogenous take-off around 1800. This is illustrated in Figure 3 in the online Appendix.

[45] The change in the quantity-quality trade-off is small and 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. The endogenous cost of raising children is actually very similar before the onset of the transition and on the balanced growth path, with levels of 4.7, and 5, respectively. Assuming a fixed cost of raising children at post-transition levels leaves the benchmark time series of fertility essentially unchanged. This is not the case for the quantity-quality function, which is calibrated by targeting data moments for the high fertility countries (see also the discussion of Figure 3).

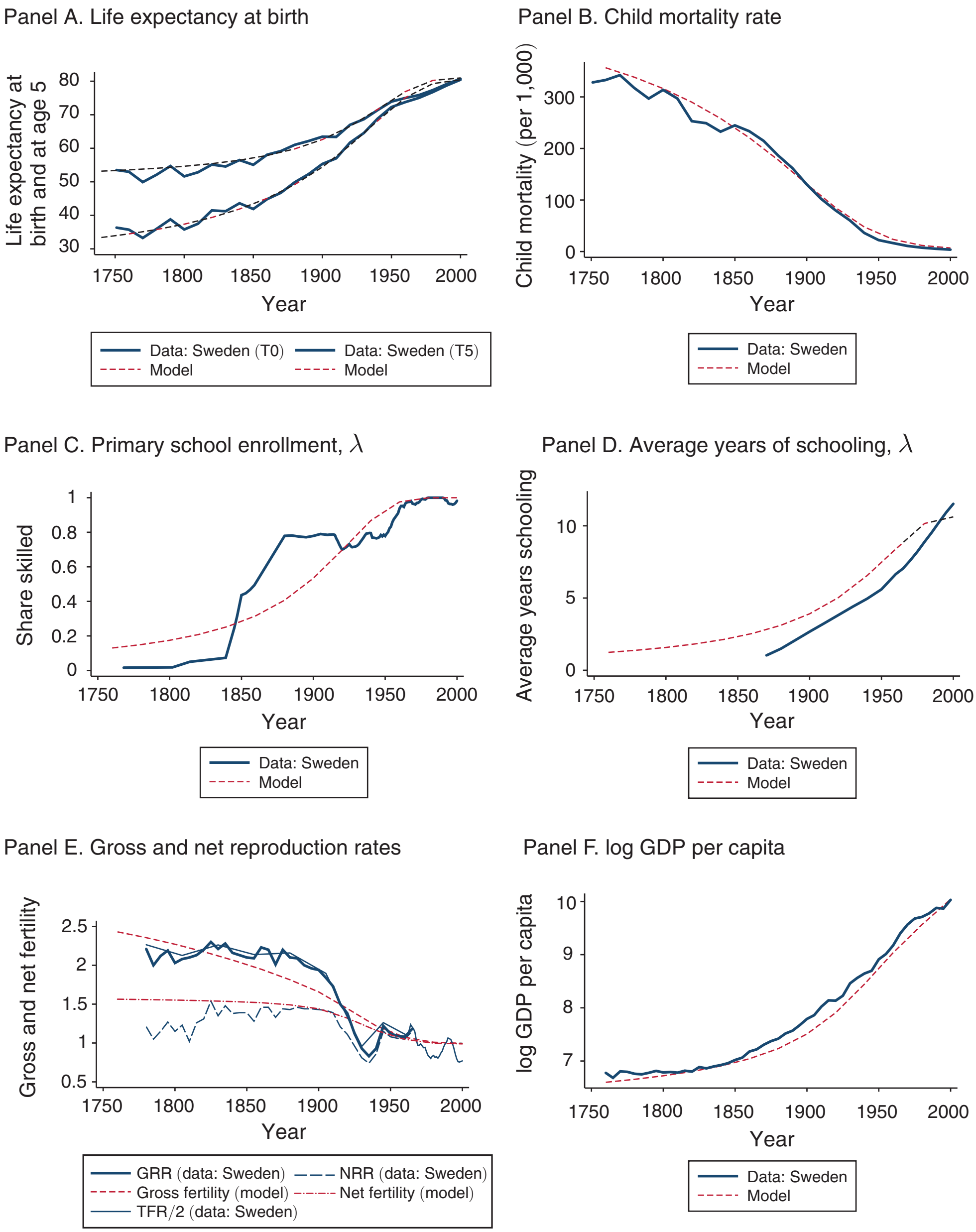

Figure 1. Long-Run Development: Simulation of Benchmark Calibration of the Model and Historical Data for Sweden 1750–2000

Source: Life expectancy and child (under 5) mortality are taken from the Human Mortality Database (2014). Primary school enrollment data are taken from de la Croix, Lindh, and Malmberg (2008), average years of schooling data are from Ljungberg and Nilsson (2009). Fertility and GDP per capita data are from historical statistics from Statistics Sweden (Edvinsson 2014).

match, however, the quantitative increase in net fertility that is observed during the early phase of the demographic transition.

Finally, Figure 1, panel F depicts the evolution of income per capita. The level of initial technology and the elasticity of technological progress were calibrated to match the level and growth rate of income per capita in 2000. The evolution of income per capita matches the data series over the entire period including the acceleration during the transition.

IV. Accounting for Comparative Development

This section investigates to what extent the model can account for comparative development patterns. The analysis is motivated by the observation that the stylized patterns of long-run development dynamics are very similar across countries and times, including forerunners like Sweden and England as well as countries that entered their demographic and economic transition much later than the European forerunners. Demographers such as Kirk (1996) notice that “in non-European countries undergoing the demographic transition in the mid-twentieth century, the regularities are impressive.”

The analysis proceeds in three steps. In Section IVA we compare the data obtained from the benchmark calibration to cross-country panel data for the period 1960–2000. The analysis provides a first investigation of the hypothesis that differences in development across the world might be accounted for by different delays in the economic and demographic transitions, and thus by delays in the take-off from quasi-stagnation to sustained growth. Section IVB explores the possibility that differences in the country-specific extrinsic mortality environment, in terms of baseline longevity $\underline{T}$, account for these delays. We investigate this by performing controlled variations in baseline longevity, thereby quantifying the role of mortality differences for comparative development patterns. Section IVC pushes the analysis one step further by simulating an artificial world composed by countries that are identical in all dimensions except baseline longevity. The simulated data are compared to the empirical world distribution of adult longevity, child mortality, education, fertility and income (in 1960 and 2000).

The analysis in this section can be viewed as an attempt to address the open question whether comparative development patterns can be explained by delays in the time of the take-off by exploring the prototype unified growth model that has been calibrated to the long-run development patterns of Sweden. Since the calibration is not based on data moments from cross-country data, and only uses differences in one parameter, extrinsic mortality, as potential explanation for the delays in development, the results of this section can be used to judge the ability of the model to fit the data “out of sample.” The results therefore represent a joint investigation of the general mechanics generated by the model and of the role of differences in baseline longevity.

A. *Simulated Data and Cross-Country Panel Data*

We begin the analysis by evaluating the ability of the model, calibrated for Sweden, to account for comparative development patterns. 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.

Figure 2 presents the data generated by the simulation of the calibrated model (as depicted in Figure 1), but plotted against the key variable driving the transition, the share of skilled workers λ, at the respective point in time (rather than as time series). These simulated data are plotted together with corresponding cross-country data for 1960 and 2000.[46] Panels A, B, and C, plot the data on life expectancy at birth, child mortality, and income per capita against the share of educated individuals, λ. The cross-sectional interpretation of the calibrated data fits the cross-country data patterns quantitatively well and the relation appears stable over the 40-year horizon. For low levels of λ (which correspond to the less developed countries) the actual data exhibit higher life expectancy and child survival probabilities than predicted by the model especially for 2000.

Figure 2, panel D plots the share of skilled individuals 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. Again the calibration performs well in matching the data, but 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 suggests that, compared to Sweden or other European countries for the same level of initial share of educated individuals, the developing countries have experienced an acceleration in education acquisition over the last 40 years.

Even though the role of baseline longevity will be explored in more detail in the next section, it is useful to illustrate the role of differences in $\underline{T}$ for the cross-sectional patterns already at this point. To this end, Figure 2, also plots simulated data from an alternative ("High Mortality") model that is based on the exact same calibration, but with a baseline longevity that is five years lower than in the benchmark calibration (i.e., $\underline{T} = 40$).[47] The results show that the cross-sectional patterns and the correlations between λ and life expectancy, child mortality, and income per capita are essentially identical to those generated by the benchmark calibration with higher baseline longevity.

Figure 3, panels A and B present the results for gross and net fertility.[48] 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, but it substantially underestimates the fertility levels for pretransitional countries with low levels of λ. Sweden, like other European countries, displays pretransitional fertility levels that are particularly low in a worldwide perspective. The literature has offered

[46] As empirical counterpart of λ across countries we 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. See the online Appendix for further information.

[47] As discussed in Section IVB a five year difference in baseline longevity is in line with the empirical evidence on pretransitional life expectancy in the lowest and highest mortality countries.

[48] 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.

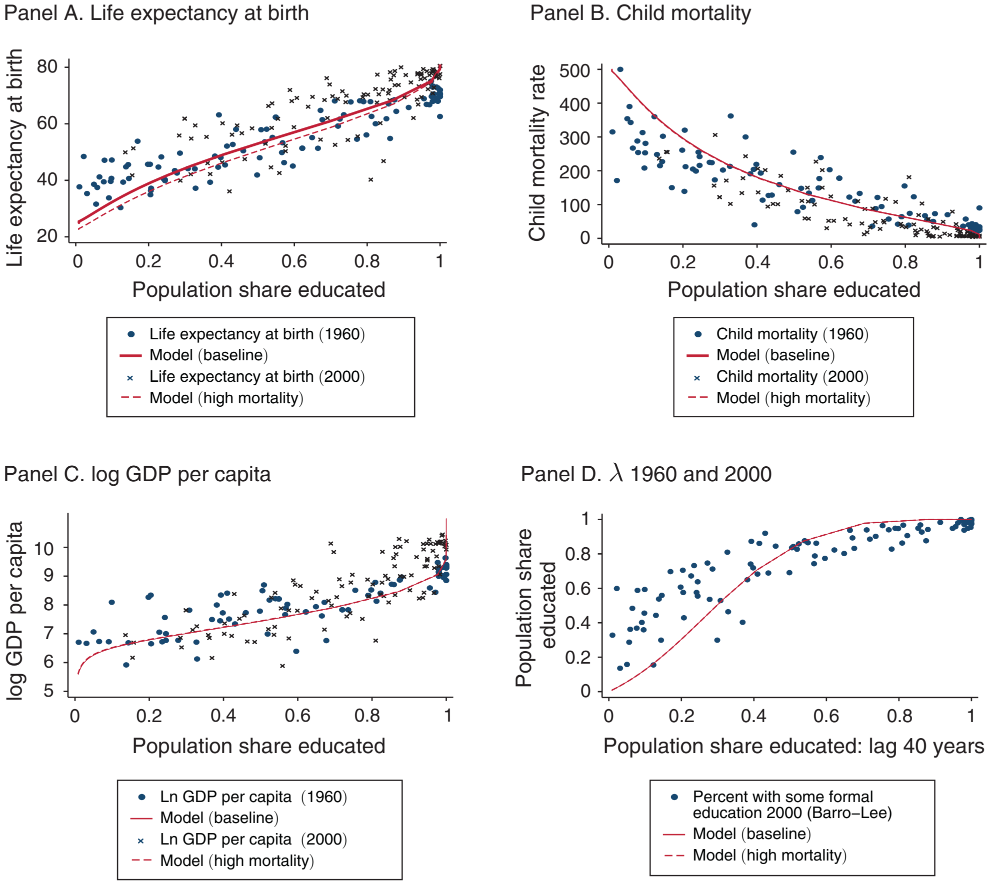

Figure 2. Education, Mortality, and Income [Simulation and Data (1960 and 2000)]

Note: λ is measured as 1-population share with no schooling from Barro and Lee (2001).

Source: Human Mortality Database (2014), United Nations Population Statistics (2014), and World Development Indicators (2014)

several hypotheses that try to explain why historically the cost of raising children was comparatively high in pre-industrial Europe. To explore the quantitative implications of lower costs for children, we consider an alternative calibration of the quantity-quality trade-off by targeting data moments for pretransitional countries with the highest recorded fertility in 2000. This alternative parametrization of the quantity-quality trade-off substantially improves the match between the simulated model and the data.[49]

[49] That fertility levels have been comparatively low in Europe compared to other regions is well documented and the reasons have been investigated recently, see, e.g., Moav (2005), Strulik and Weisdorf (2014), and Voigtländer and Voth (2013). 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 that accounts for the fact that the average total fertility rates of the highest fertility countries was around 7, or above, in 2000, as compared to about 5 for pretransitional Europe. Changing the target for the pretransitional fertility to $n = 3.5$ and recalibrating the parameters accordingly delivers $\{\beta' = 0.75, r' = 3, \delta' = 1.06\}$. The kink in the simulated data in Figure 3, panel A and B corresponds to the exit from the corner solution of the extra time invested in children. Recall that the cross-sectional patterns

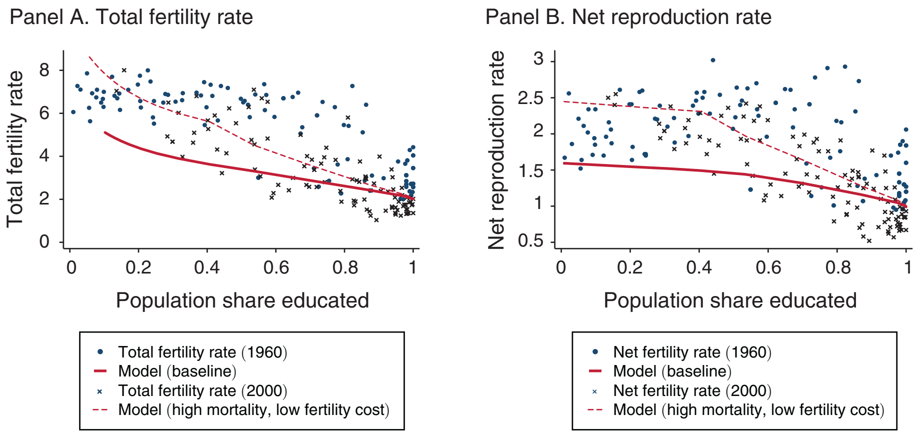

Figure 3. Education and Fertility [Simulation and Data (1960 and 2000)]

Note: λ is measured as (1-population share with no schooling) from Barro and Lee (2001).

Source: United Nations Population Statistics (2014), and World Bank World Development Indicators (2014)

The theory also predicts the existence of nonlinear dynamics that link economic and demographic variables during the process of long-term development. The nonlinearity of the equilibrium locus Λ, characterized in Proposition 1, implies that the changes in λ are largest in the intermediate range, where the locus has its steepest slope. The increase in the share of skilled workers is relatively large in countries with intermediate levels of adult longevity, but relatively small in pretransitional and post-transitional countries. The model therefore predicts a nonmonotonic relationship between life expectancy and subsequent changes in λ. As discussed in Section I the existence of a nonlinear effect of life expectancy on education has relevant implications for empirical investigations. Panels A and B of Figure 4 depict the relationship between life expectancy in 1960 and the change in the share of individuals with no formal education over the following 20 and 40 years in the data (including a quadratic regression line), in comparison to the respective data from the benchmark calibration. The model matches the data well, although it somewhat underestimates the improvements in the change in education in countries with lower initial life expectancy. Compared to the historical experience of Sweden, education improvements in the poorest countries were comparatively large in the period 1960–2000.

Another direct implication of the development dynamics of the model is the existence of a concave relationship between life expectancy and income. During the early phases of development, low longevity induces little human capital accumulation as proxied by the population share skilled, λ, and consequentially income is low. As longevity improves, incentives to become skilled increase, and incomes rise. As development continues, however, further improvements in life expectancy lose momentum as the skill composition, adult longevity, and child survival converge to their natural upper bounds, whereas income remains on a sustained growth

depicted in Figure 2 are unaffected by the actual calibration of the quantity-quality trade-off because the dynamic system (13) is block recursive and does not involve any scale effect.

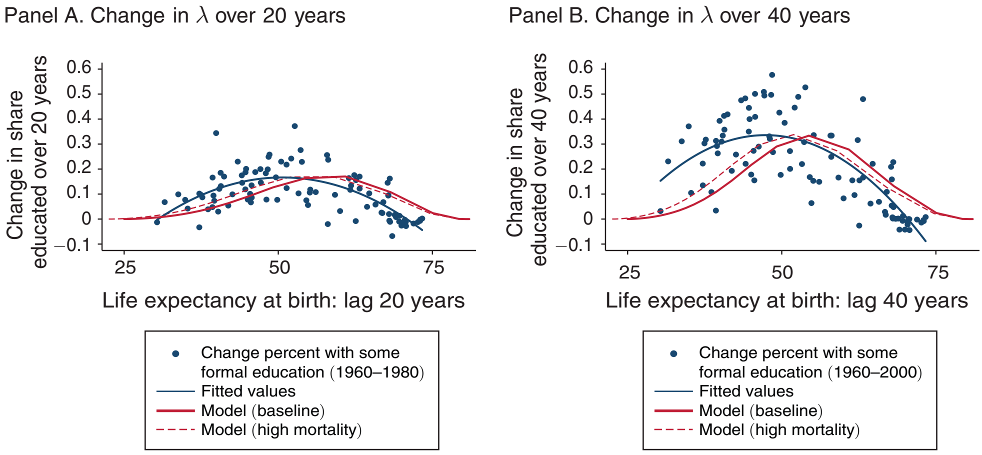

FIGURE 4. LIFE EXPECTANCY AND CHANGES IN EDUCATION [SIMULATION AND DATA]

Note: λ is measured as 1-population share with no schooling.
Source: Barro and Lee (2001)

path. This prediction is in line with the empirical pattern between income and life expectancy that is known as the Preston Curve (see Preston 1975), and that is considered a stylized fact in demography. Figure 5 shows that the patterns implied by the simulation closely match the empirical Preston Curve in the data (in 1960 and 2000). The prototype unified growth model provides a natural explanation for the Preston Curve, whose determinants and mechanics are still debated in the literature, as discussed in Section I.

B. *Different Levels of Baseline Longevity and Comparative Development*

In this section, the calibrated model is used to investigate the quantitative role of differences in 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.

We calibrate an alternative scenario of baseline adult longevity, $\underline{T}$, that reflects the worst mortality environment of all countries. This calibration targets data moments for pretransition countries with the highest observed adult mortality in 2000, which corresponds to targeting a life expectancy at age 5 of 45 years (compared to 48 years reflecting Sweden around 1800 just before the transition). This level is in line with the lowest available measure in 2000 for life expectancy at birth and, accounting for the respective skill composition, implies setting $\underline{T} = 40$ (as compared to the baseline $\underline{T} = 45$).[50]

[50]Retaining a target of 76 years for life expectancy at age five on the balanced growth path, and using condition (6), this implies setting a $\underline{T} = 40$ and $\overline{\rho} = 36$ (rather than $\underline{T} = 45$ and $\rho = 31$ as in the benchmark

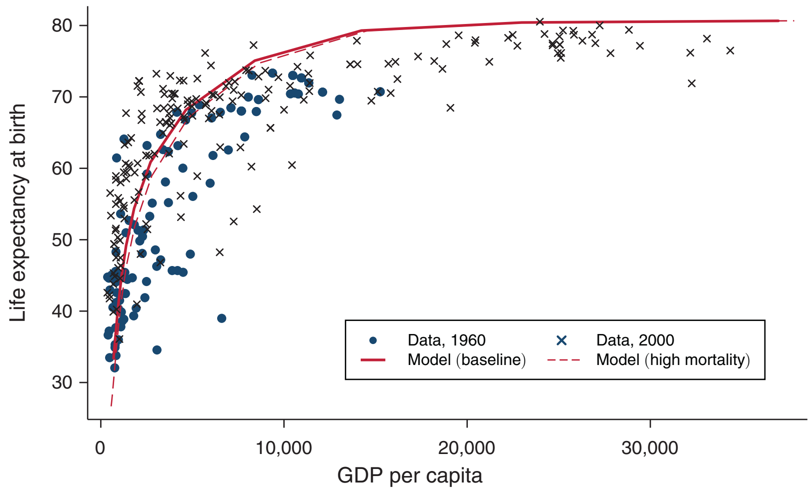

Figure 5. The Preston Curve

Source: Human Mortality Database (2014) and World Development Indicators (2014)

The results reported in Figure 2 have shown that a five year lower baseline longevity leaves the cross-sectional patterns generated by the simulated model essentially unaffected. From the dynamic system (13), a lower baseline adult longevity implies a lower population share of skilled individuals in equilibrium for any given level of technology and education of the previous generation, and therefore a delayed take-off. To investigate the quantitative importance of this prediction, we replicate the analysis with the baseline adult longevity recalibrated to $\underline{T} = 40$ to reflect countries with the highest extrinsic 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.

Figure 6 plots life expectancy at birth, the share of skilled individuals, total fertility rates, and income per capita for the benchmark calibration and for the alternative calibration with low baseline longevity. The delay in the transition, spans about 7 generations or 140 years, as a consequence of imposing $\underline{T} = 40$ rather than $\underline{T} = 45$. The joint consideration of the simulation in Figures 2 and 6 therefore suggests that differences in baseline longevity may be relevant to explain the delay in comparative development, but their effect is hard to detect by estimating linear regressions with

calibration). The data source is United Nations Population Statistics (2014). Data on life expectancy at five for earlier periods are missing for many countries, including most sub-Saharan African countries in 1960. Alternatively, the available information on child mortality and life expectancy at birth in 1960 can be used to derive an estimate 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 5 between 44 and 45 years. In some countries, like Swaziland, life expectancy at birth is just above 40 years still today (United Nations Population Statistics 2014). This suggests that 45 is possibly a conservative estimate of baseline adult longevity in the worst conceivable mortality environment.

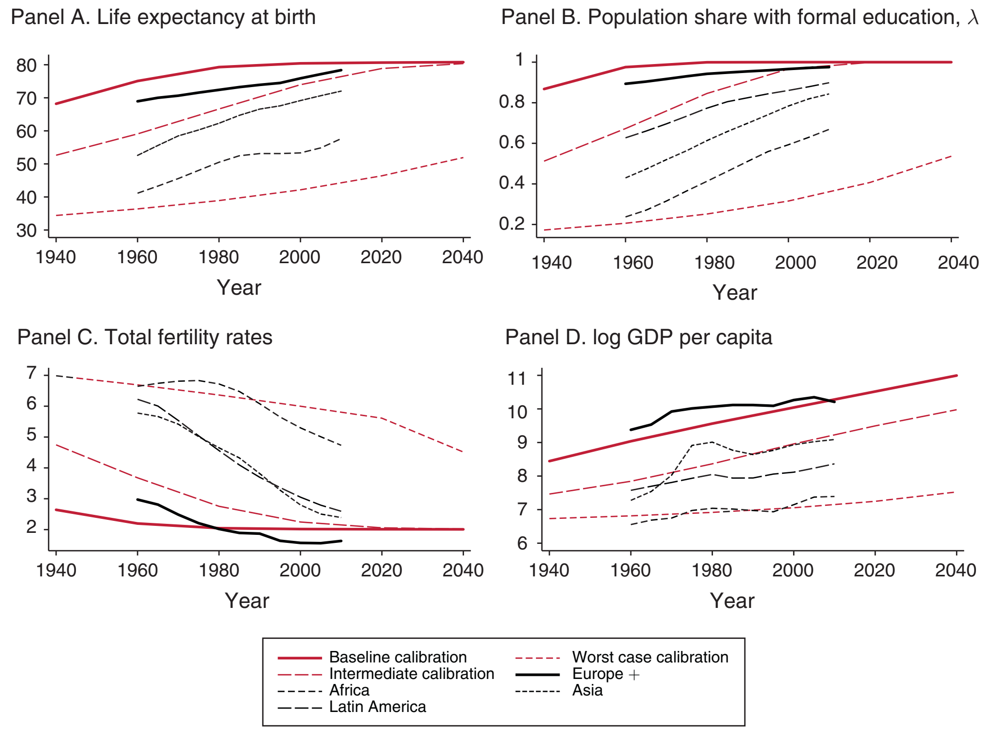

FIGURE 6. THE ROLE OF LOWER BASELINE LONGEVITY FOR COMPARATIVE DEVELOPMENT

Note: λ is measured as 1-population share with no schooling.

Source: Human Mortality Database (2014), Barro and Lee (2001), United Nations Population Statistics (2014), and World Development Indicators (2014)

cross-country panel data. In fact, apart from the timing of the take-off, the different countries experience a very similar development process. This is illustrated in the figure by also plotting the dynamic simulation for a country with intermediate baseline longevity that converges to the balanced growth path in 2040 (rather than 2000).[51]

Figure 6 also plots the corresponding empirical development dynamics over the period 1960–2010 for different continents. The development dynamics of Europe and Western offshoots are captured by the baseline calibration rather well. On the other end of the spectrum, African countries display a substantially delayed development in all four dimensions. By construction, Africa as a whole still displays a somewhat better development performance than the calibration for $\underline{\underline{T}} = 40$ for the worst conceivable scenario with the highest disease burden in the world. The development dynamics of Asia and Latin America lie between Europe and Africa. This pattern is consistent with estimates of the extrinsic mortality environment.

[51] As discussed in more detail in the next section, we estimate a distribution of baseline longevity using information on the number of multi-host vector-transmitted pathogens to simulate the distribution of baseline longevity for a world of artificial countries. The baseline longevity of the intermediate country plotted in Figure 6 corresponds to the median of the estimated distribution (which is 43.25).

Europe and Western offshoots display the lowest, and African countries display the highest disease burden in the world, while Asian and Latin American countries display an intermediate level of disease exposure.[52]

C. *Accounting for the Worldwide Distribution of Comparative Development*

The analysis so far supports the view that the dynamic evolution of the different countries is characterized by a similar process that involves a long period of slow development followed by a rapid transition to a sustained growth path. A main difference across countries appears to be the actual timing of the take-off. A direct implication of this view is that, at a given point in time, relatively few countries are observed during their transition (since for most of its history each country is either pretransitional or post-transitional, while the transition is comparably quick). As a result one should therefore expect the cross-sectional distribution of all variables of interest to display two modes corresponding to the mass of countries that are still pretransitional or on the balanced growth path, respectively, as characterized in Proposition 2.[53] While intuitive, this implication has not been derived and empirically investigated in the existing literature.

To do so, we simulate an artificial world composed of countries that are identical in all parameters except for baseline adult longevity $\underline{T}$. To create a meaningful world-wide distribution of baseline longevity in the range $\left[\underline{\underline{T}} = 40, \underline{T} = 45\right]$, which is needed to simulate a cross-sectional distribution of the variables of interest, we exploit information on cross-country differences in the historical prevalence of particular (multi-host vector-transmitted) infectious diseases. For their particular features, the existence of the respective pathogens in a country is closely connected to local country-specific biological and climatological conditions, and they have not been eradicated in any country. The historical distribution of these pathogens is therefore a good proxy for the country-specific extrinsic mortality, see Cervellati, Sunde, and Valmori (2012). The calibration is based on a count index that exploits information on whether a pathogen had been detected in a country. These data have been collected from sources from the early twentieth century and therefore reflect extrinsic mortality across the world before major health innovations, and their worldwide dissemination, which took place with the so-called epidemiological revolution after World War II. This means that the analysis does not rely on the spread of the disease in terms of the number of fatalities or infected cases, which were potentially endogenous to development already in the nineteenth century. We construct a distribution of baseline longevity parameters for 113 artificial economies in the range $\left[\underline{\underline{T}} = 40, \underline{T} = 45\right]$ and simulate a world of these artificial countries that only differ in terms of their baseline adult longevity and are otherwise identical.[54]

[52] See Figure 2(b) in the online Appendix and the discussion below.

[53] The precise shape of the distribution depends on the actual distributions of the underlying country-specific characteristics that drive the delay in the take-off. Nonetheless, the bimodality should be detectable regardless of the particular distribution as long as sufficiently many countries are still pretransitional.

[54] The distribution of baseline longevity parameters for 113 economies is created using the empirical distribution of multi-host vector-transmitted pathogens across countries. The details, including the number of diseases in different continents relative to the world average and the artificial distribution, are reported in the online Appendix.

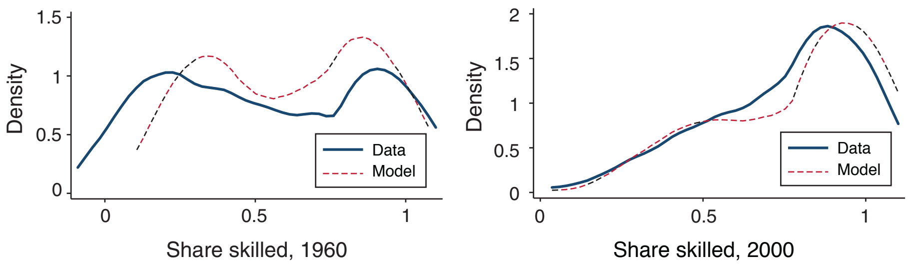

Figure 7. Distributions: Education [Simulation and Data (1960 and 2000)]

Note: λ is measured as 1-population share with no schooling.
Source: Barro and Lee (2001)

The data generated by the simulation of the artificial world are then pooled and used to estimate the simulated cross-country distribution of all variables of interest. These distributions are then compared to the corresponding distributions obtained from cross-county data in 1960 and in 2000. In interpreting the results it is useful to keep in mind that the only assumed difference across countries is baseline longevity.

Figure 7 plots the simulated distributions of education 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. Figure 8 does the same for the distributions of life expectancy and child mortality. For all variables the expected bimodality is clearly apparent in 1960 both in the simulated and the actual data, and the empirical patterns are broadly matched in terms of the support, the shape of the distribution, and location of the modes. By 2000 both the simulated and the empirical distributions tend to be more unimodal since more countries have undergone the transition.[55] An interesting observation that is in line with some of the insights from the analysis in the previous section is that by 2000 the bimodality in the actual distribution is less visible compared to the simulated world. This again suggests that the model underestimates the timing of the take-off in the less developed countries over the last decades. In other words, there is an anticipation in the take-off that accelerates development in these countries compared to the historical path followed by the European forerunners. This effect, which is particularly visible for child mortality and adult longevity, might result from health spillovers from the developed to the developing countries. Such spillovers are not considered in the simulated world.

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 alternative parametrization of the quantity-quality function that was calibrated targeting data moments for these countries as described in Section IVA. Figure 9 presents the results for total fertility rates and net reproduction rates. The simulation fits the data by roughly capturing the modes at low and high levels of fertility, as well as the shape of the distribution and its change over the 40-year

[55] The bimodality of the simulated distribution is not due to the calibrated distribution of baseline longevity. Alternatively, we have performed the exercise considering a uniform distribution of baseline longevity over the range $[\underline{T} = 40, \overline{T} = 45]$, with a very similar pattern of bimodality and of changes in the distributions over time.

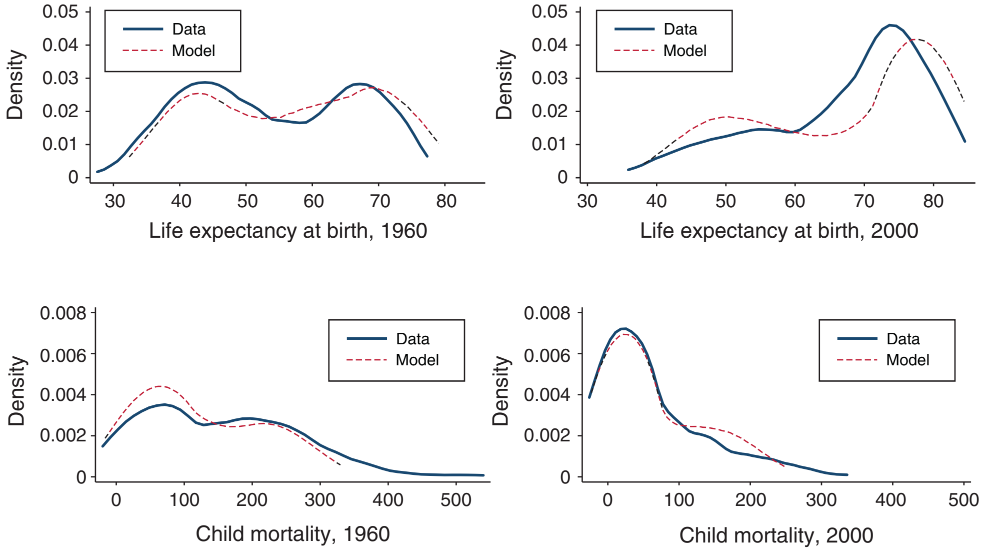

Figure 8. Distributions: Mortality [Simulation and Data (1960 and 2000)]

Source: Human Mortality Database (2014)

horizon.[56] Also in this case the disappearance of the low mode is faster in the real data compared to the artificial world.

Finally, Figure 10 depicts the world-wide distribution of incomes per capita. Compared to the demographic variables the worldwide income distribution is matched less accurately both in terms of the support of the distribution and in terms of change over time. Notice that the model is limited in capturing the world income distribution partly by construction. The model is calibrated to Sweden (which was not among the most developed countries even in an historical perspective). Also, the artificial world does not account for any other relevant country-specific determinants of cross-country comparative development that have been studied in the literature.[57]

V. Concluding Remarks

This paper proposes a simple prototype theory of the economic and demographic transition that generates the endogenous evolution of mortality, education, fertility, and income. The model is calibrated to historical data for Sweden and allows for a

[56] The actual calibration of the production function of children's quality 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 underestimate the location of the mode for high fertility. This suggests that differences in the cost of raising children across countries are potentially more important for the cross-country differences in pretransitional fertility levels than differences in mortality.

[57] The model 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 spillovers or transfers of technology and innovations. Also, while the samples used for the density plots in Figures 8 and 9 are balanced for the observation periods 1960 and 2000, for GDP the sample for 1960 only contains 72 countries due to data availability, but 90 countries in 2000, so that the density plots obtained from data are not perfectly comparable.

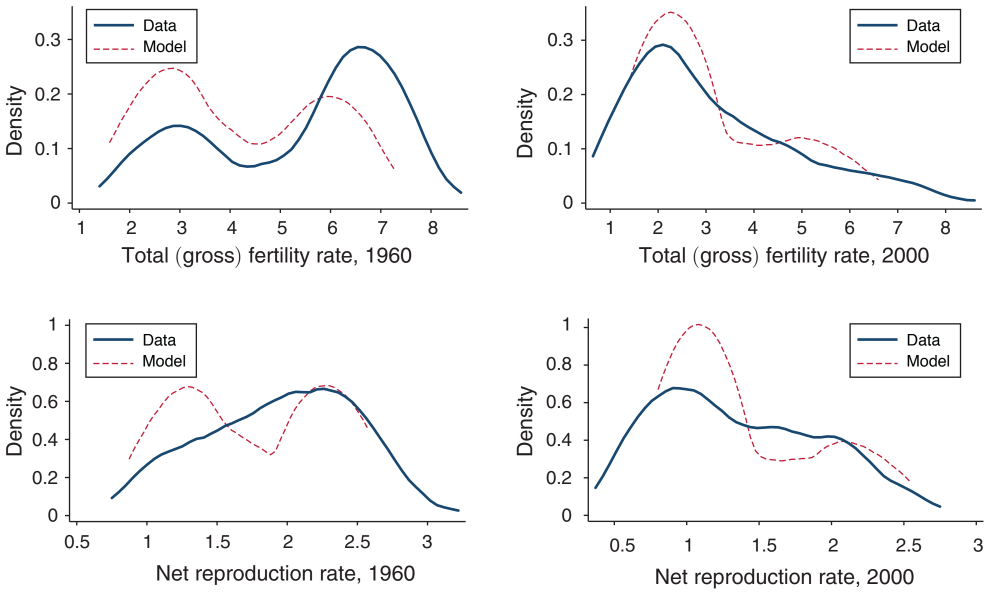

Figure 9. Density Distributions of Fertility [Simulation and Data (1960 and 2000)]

Source: UN Population Statistics (different historical volumes of the UN Demographic Yearbook, www.unstats.un.org)

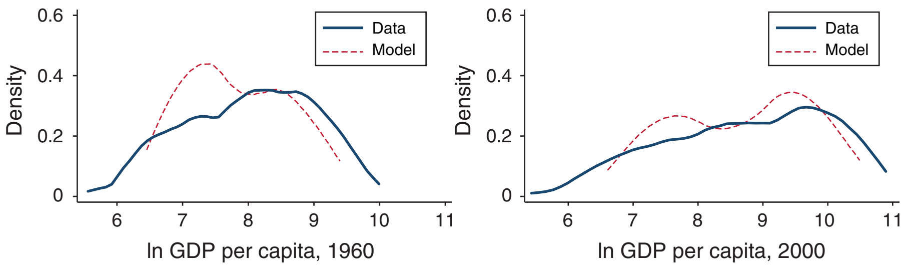

Figure 10. Distribution of Income per Capita [Simulation and Data (1960 and 2000)]

Source: World Development Indicators (2014) and United Nations Population Statistics (2014)

first systematic quantitative analysis of the implications of a unified growth model for long-run development and comparative development patterns. The results document the ability of the unified growth framework to rationalize both historical and cross-country development patterns. The findings provide support for the view that all countries follow similar nonlinear development processes, characterized by a long period of quasi-stagnation followed by rapid economic and demographic transitions. The findings also show the ability of the unified growth framework to account for cross-country development patterns, suggesting that the differential timing of the take-off is a crucial determinant of comparative development differences.

As a candidate explanation for delays in development, the analysis focused on cross-country differences in extrinsic mortality (in terms of the exposure to pathogens). The quantitative role of mortality differences for the timing of the take-off has been isolated by performing counterfactual exercises. The results show that even moderate differences in baseline longevity can be relevant for comparative development differences by inducing sizable delays in the take-off of growth, while leaving the cross-country correlations essentially unaffected. The findings also provide valuable insights for the design of empirical investigations from the perspective of unified growth theory. The results show, in particular, that linear empirical specifications may lead to misleading conclusions about the empirical role of relevant determinants of long-run growth.

The results of this paper suggest some directions for further research. In a next step, the unified growth framework can be applied to compare the quantitative relevance of other country-specific determinants of comparative development beyond the role of extrinsic mortality investigated in this paper. Moreover, while instructive regarding the main mechanics, the analysis in this paper has completely abstracted from cross-country spillovers, for instance in technological and medical knowledge, or other interactions between countries at very different stages of development. The findings suggest that, compared to the historical experience of the European forerunners, the development of some (but not all) developing countries exhibit an acceleration, as documented by the differences between the simulated and the real world after 1960. Extending the unified growth framework to the explicit consideration of cross-country spillovers therefore appears another fruitful direction for future research.

Appendix: Derivations and Proofs

With the assumptions made in Section IIA, the utility can be expressed as,

$$U(c_t, \pi_t n_t q_t) = T_t \ln c_t + \gamma \ln (\pi_t n_t q_t). \tag{A1}$$

The time budget faced by an individual is given by

$$\overline{T}_t \geq l_t + \underline{e}^j + \pi_t n_t r_t. \tag{A2}$$

In addition, the individual faces a resource constraint

$$l_t w_t^j h_t^j(a) \geq T_t c_t, \tag{A3}$$

where l_t is the total time spent working. Given the utility function (A1) both constraints will be binding at the optimum. Combining (A2) and (A3) delivers the budget (5) in the text. Maximizing utility (A1) subject to (5) is equivalent to maximizing

$$T_t \ln \left[(1/T_t)\left(\overline{T}_t - \underline{e}^j - \pi_t n_t r_t\right) w_t^j h_t^j(a)\right] + \gamma \ln (\pi_t n_t q_t). \tag{A4}$$

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

$$n_t^j = \frac{\gamma\left(\overline{T}_t - \underline{e}^j\right)}{(T_t + \gamma)\, r_t^j \pi_t}, \tag{A5}$$

where r_t^j *is given by*

$$r_t^j = r_t^* = \max\left\{\underline{r}, \frac{1 - \left[1/(\delta(1 + g_{t+1}))\right]}{1 - \beta}\underline{r}\right\}. \tag{A6}$$

PROOF OF LEMMA 1: Consider an individual acquiring human capital of type $j = u, s$. Taking the first order condition of (A4) with respect to n_t and restricting to an interior solution gives (A5), while taking the first order condition with respect to r_t^j gives

$$-T_t \pi_t n_t r_t^j + \gamma\left(\overline{T}_t - \pi_t n_t r_t^j - \underline{e}^j\right)\left(q_r(\cdot) r_t^j\right)/q(\cdot) \geq 0. \tag{A7}$$

Using (A5) to simplify (A7) implies $\left[q_r\left(r_t^j, g_{t+1}\right) r_t^j\right]/q\left(r_t^j, g_{t+1}\right) \geq 1$. Given the functional form (4) this implies (A6).[58] ∎

LEMMA 2: *For any* $\{w_t^s, w_t^u, T_t, \pi_t\}$ *there exists a unique* $\tilde{a}_t$ *implicitly defined by*

$$\frac{h_t^s(\tilde{a}_t)}{h_t^u} = \left(\frac{\overline{T}_t - \underline{e}^u}{\overline{T}_t - \underline{e}^s}\right)^{\frac{T_t + \gamma}{T_t}} \frac{w_t^u}{w_t^s}, \tag{A8}$$

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

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^j with $j = u, s$ from (A5) and noting that $r_t^u = r_t^s = r_t^*$ from (A6) implies that the optimal type of skill depends on

$$\left(\overline{T}_t - \underline{e}^u\right)^{(T_t + \gamma)} \left(w_t^u h_t^u\right)^{T_t} \gtrless \left(\overline{T}_t - \underline{e}^s\right)^{(T_t + \gamma)} \left(w_t^s h_t^s(a)\right)^{T_t}. \tag{A9}$$

Since the indirect utility obtained by acquiring skilled human capital increases with ability, there exists a unique $\tilde{a}_t$ such that all individuals with $a < \tilde{a}_t$ optimally choose to acquire u, while those with $a > \tilde{a}$ optimally choose to obtain s. Solving (A9) as equality gives (A8). ∎

[58] From (A6) there is a unique $\underline{g} > 0$ (implicitly given by $r_t^*(\underline{g}) = \underline{r}$) such that for any $g_{t+1} > \underline{g}$ then $r_t^* > \underline{r}$ and $dr_t^*/dg_{t+1} > 0$.

PROOF OF PROPOSITION 1: The aggregate levels of human capital are given by

$$\text{(A10)} \qquad H_t^u = N_t \int_0^{\tilde{a}_t} h_t^u f(a)\, da \;\text{ and }\; H_t^s = N_t \int_{\tilde{a}_t}^1 h_t^s(a) f(a)\, da.$$

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

$$\text{(A11)} \qquad \frac{w_t^u}{w_t^s} = \frac{1-x_t}{x_t}\left(\frac{H_t^s}{H_t^u}\right)^{1-\eta} = \frac{1-x_t}{x_t}\left(\frac{\int_{\tilde{a}_t}^1 h^s(a) f(a)\, da}{\int_0^{\tilde{a}_t} h^u f(a)\, da}\right)^{1-\eta}.$$

Substituting (A11) into (A8) gives the general equilibrium ability threshold

$$\text{(A12)} \qquad \frac{h_t^u\left(\int_{\tilde{a}_t}^1 h^s(a) f(a)\, da\right)^{1-\eta}}{h^s(\tilde{a}_t)\left(\int_0^{\tilde{a}_t} h_t^u f(a)\, da\right)^{1-\eta}} = \frac{x_t}{1-x_t}\left(\frac{\overline{T}_t - \underline{e}^s}{\overline{T}_t - \underline{e}^u}\right)^{\frac{T_t+\gamma}{T_t}}.$$

Since there is a one-to-one relationship between the share of skilled workers λ_t and the threshold ability $\tilde{a}_t$, this also characterizes implicitly the equilibrium share of skilled individuals, λ_t, where H_t^u is decreasing in λ_t and H_t^s is increasing in λ_t. Rearrange (A12) to get the equilibrium relationship between $\tilde{a}_t$ and T_t expressed as

$$\text{(A13)} \qquad \mathcal{G}(\tilde{a}_t)^{1-\eta} F(x_t) - \left(\frac{\overline{T}_t - \underline{e}^s}{\overline{T}_t - \underline{e}^u}\right)^{\frac{T_t+\gamma}{T_t}} = 0,$$

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

$$\text{(A14)} \qquad \mathcal{G}(\tilde{a}_t) = \frac{(h^u)^{\frac{1}{1-\eta}} \int_{\tilde{a}_t}^1 h^s(a) f(a)\, da}{h^s(\tilde{a}_t)^{\frac{1}{1-\eta}} \int_0^{\tilde{a}_t} h^u f(a)\, da},$$

with $\mathcal{G}'(\tilde{a}_t) < 0$. Notice that $\left[\left(\overline{T}_t - \underline{e}^s\right)/\left(\overline{T}_t - \underline{e}^u\right)\right] \in (0,1)$ for $T_t \in \left(\underline{e}^s, \infty\right)$. For any x_t, the function (A13) is therefore defined over the range $\tilde{a} \in \left(\underline{a}(x_t), 1\right]$ where[59]

$$\text{(A15)} \qquad \underline{a}(x_t) : \mathcal{G}\left(\underline{a}(x_t)\right)^{1-\eta} \mathcal{F}(x_t) = 1.$$

Applying calculus, $\partial \underline{a}(x_t)/\partial x_t < 0$ 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

[59] Since the denominator of (A13) has a discontinuity at $\underline{a}$ and the function takes negative values for any $a \leq \underline{a}(x_t)$.

share of the population that for each generation t would acquire skilled human capital in the case in which $T_t \rightarrow \infty$. By totally differentiating (A13) we have

$$\text{(A16)} \qquad \frac{d\tilde{a}_t}{dT_t} = \frac{d\left(\left(\frac{\overline{T}_t - \underline{e}^s}{\overline{T}_t - \underline{e}^u}\right)^{\frac{T_t+\gamma}{T_t}}\right) / dT_t}{\left[(1-\eta)\,\mathcal{G}(\tilde{a}_t)^{-\eta}\mathcal{G}'(\tilde{a}_t)\,\mathcal{F}(x_t)\right]} < 0,$$

which is negative since $\mathcal{G}'(\tilde{a}_t) < 0$. For $T_t = \underline{e}^s$ we have $\tilde{a}_t = 1$ which implies $\mathcal{G}(\tilde{a}_t) = 0$ so that $\mathcal{G}(\tilde{a}_t)^{-\eta} = \infty$. Since $\mathcal{G}'(1)$ is a finite number we have that the denominator of (A16) goes to infinity as $T_t \rightarrow \underline{e}^s$. In turn, the numerator has a limit at zero. For $T_t \rightarrow \infty$ we have $\tilde{a}_t \rightarrow \underline{a} < 1$ so that the denominator of (A16) is a finite number while the numerator has a limit at zero. Hence $\lim_{T_t \rightarrow e} \frac{d\tilde{a}_t^s}{dT_t} = \lim_{T_t \rightarrow \infty} \frac{d\tilde{a}_t}{dT_t} = 0$ which also implies that the equilibrium locus (12) is convex for $T_t \rightarrow \underline{e}^s$ and concave for $T_t \rightarrow \infty$. ■

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

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 (8) this implies $x_t > x_{t-1}$ for all t with $\lim_{t\rightarrow\infty} x_t = 1$; from (9), it follows that $g_t > 0$ and $\lim_{t\rightarrow\infty} A_t = \infty$ for any $A_0 > 0$. In the limit as $\lambda_t \rightarrow 1$, $g_t = \phi$ from (9). ■

PROOF OF PROPOSITION 2:

The equilibrium relationship linking $\tilde{a}_t$ and T_t is given in (A13). For any T_t, $\tilde{a}_t$ is an implicit function of x_t. Recall that by implicit differentiation of (A12) $\partial\tilde{a}_t/\partial x_t < 0$, 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 $\tilde{a} \simeq 1$ 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 (A5) evaluated at $T = \underline{T}$. Under these conditions, from (2) the level of income per capita is (arbitrarily) low which, from (7) and (14) implies $\pi_0 \simeq \underline{\pi}$. Part (ii) follows directly from Lemma 3, where $A_\infty \rightarrow \infty$, $x_\infty \rightarrow 1, \lambda_\infty \simeq 1, T = \overline{T}$. From (9) this also implies that $g_\infty = \phi$. Finally, since $A_\infty \rightarrow \infty$, it follows that $y_\infty \rightarrow \infty$ and from (7), $\pi_\infty \simeq 1$ so that fertility is given as in (15). Part (iii) follows from combining part (i), part (ii), and Lemma 3.

Figure 1 in the online Appendix depicts the evolution of the conditional system given by equations (6) and (12) 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 panels A and C of Figure 1 in the online Appendix. ■

Table A1—Summary Information on Calibration of Parameters

Parameter		Value	Matched moment (information source)
Panel A. 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, Sweden)
Time cost for unskilled/skilled education	$\{\underline{e}^u, \underline{e}^s\}$	{0,12}	Years of schooling in 1820 and 2000 (Lutz, Goujon, and Sanderson 2007; Ljungberg-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 LE at 5 in 1760–1800 and 2000 (Human Mortality Database)
Minimum child survival and elasticity parameter	$\{\underline{\pi}, \kappa\}$	{0.5, 0.005}	Child survival probability in 1800 and 2000 (Human Mortality Database)
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/Post-transitional fertility, TFP growth 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)
Panel B. Cross-country analysis			
Parameters set endogenously			
Baseline adult longevity/scope for improvement	$\{\underline{T}, \rho'\}$	{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			World distribution of human pathogens (Cervellati, Sunde, and Valmori 2012)

REFERENCES

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

Acemoglu, Daron, and Simon Johnson. 2007. "Disease and Development: The Effect of Life Expectancy on Economic Growth." *Journal of Political Economy* 115 (6): 925–85.

Aghion, Philippe, and Peter Howitt. 1992. "A Model of Growth Through Creative Destruction." *Econometrica* 60 (2): 323–51.

Aiyar, Shekhar, Carl-Johan Dalgaard, and Omer Moav. 2008. "Technological Progress and Regress in Pre-Industrial Times." *Journal of Economic Growth* 13 (2): 125–44.

Ashaf, Quamrul, and Oded Galor. 2013. "The 'Out of Africa' Hypothesis, Human Genetic Diversity, and Comparative Economic Development." *American Economic Review* 103 (1): 1–46.

Ashaf, Quamrul H., Ashley Lester, and David N. Weil. 2008. "When Does Improving Health Raise GDP?" In *NBER Macroeconomics Annual 2008*, Vol. 23, 157–204. Chicago: University of Chicago Press.

Azariadis, Costas, and John Stachurski. 2005. "Poverty Traps." In *Handbook of Economic Growth*, Vol. 1, Pt. A, edited by Philippe Aghion and Steven Durlauf, 295–384. Amsterdam: North-Holland.

Bar, Michael, and Oksana Leukhina. 2010. "Demographic Transition and Industrial Revolution: A Macroeconomic Investigation." *Review of Economic Dynamics* 13 (2): 424–51.
Barro, Robert J., and Jong-Wha Lee. 2001. "International Data on Educational Attainment: Updates and Implications." *Oxford Economic Papers* 53 (3): 541–63.
Becker, Gary S., and H. Gregg Lewis. 1973. "On the Interaction between the Quantity and Quality of Children." *Journal of Political Economy* 81 (2, Pt. 2): S279–88.
Bideau, Alain, Bertrand Desjardins, and Hector Perez-Brignoli. 1997. *Infant and Child Mortality in the Past*. Oxford: Clarendon Press.
Bloom, David E., and David Canning. 2007a. "Commentary: The Preston Curve 30 Years On: Still Sparking Fires." *International Journal of Epidemiology* 36 (3): 498–99.
Bloom, David E., and David Canning. 2007b. "Mortality Traps and the Dynamics of Health Transitions." *Proceedings of the National Academy of Sciences* (*PNAS*) 104 (41): 16044–49.
Boucekkine, Raouf, David de la Croix, and Omar Licandro. 2002. "Vintage Human Capital, Demographic Trends, and Endogenous Growth." *Journal of Economic Theory* 104 (2): 340–75.
Boucekkine, Raouf, David de la Croix, and Omar Licandro. 2003. "Early Mortality Declines at the Dawn of Modern Growth." *Scandinavian Journal of Economics* 105 (3): 401–18.
Cervellati, Matteo, and Uwe Sunde. 2005. "Human Capital, Life Expectancy, and the Process of Development." *American Economic Review* 95 (5): 1653–72.
Cervellati, Matteo, and Uwe Sunde. 2011. "Life expectancy and economic growth: The role of the demographic transition." *Journal of Economic Growth* 16 (2): 99–133.
Cervellati, Matteo, and Uwe Sunde. 2013. "Life Expectancy, Schooling, and Lifetime Labor Supply: Theory and Evidence Revisited." *Econometrica* 81 (5): 2055–86.
Cervellati, Matteo, and Uwe Sunde. 2015. "The Economic and Demographic Transition, Mortality, and Comparative Development: Dataset." *American Economic Journal: Macroeconomics*. http://dx.doi.org/10.1257/mac.20130170.
Cervellati, Matteo, and Uwe Sunde. 2015. "The effect of life expectancy on education and population dynamics." *Empirical Economics* 48 (4): 1445–78.
Cervellati, Matteo, Uwe Sunde, and Simona Valmori. 2012. "The Distribution of Infectious Diseases and Extrinsic Mortality Across Countries." *Mathematical Population Studies* 19 (2): 73–93.
Chakraborty, Shanka, Chris Papageorgiou, and Fidel Pérez Sebastian. 2010. "Diseases, infection dynamics and development." *Journal of Monetary Economics* 57 (7): 859–72.
Chari, V. V., and Hugo Hopenhayn. 1991. "Vintage Human Capital, Growth, and the Diffusion of New Technology." *Journal of Political Economy* 99 (6): 1142–65.
Chesnais, Jean-Claude. 1992. *The Demographic Transition: Stages, Patterns and Economic Implications. A Longitudinal Study of Sixty-Seven Countries Covering the Period 1720–1984*. Oxford: Clarendon Press.
Cutler, David, Angus Deaton, and Adriana Lleras-Muney. 2006. "The Determinants of Mortality." *Journal of Economic Perspectives* 20 (3): 97–120.
Dalgaard, Carl-Johan, and Holger Strulik. 2014. "Optimal Aging and Death: Understanding the Preston Curve." *Journal of the European Economic Association* 12 (3): 672–701.
de la Croix, David, and Matthias Doepke. 2003. "Inequality and Growth: Why Differential Fertility Matters." *American Economic Review* 93 (4): 1091–1113.
de la Croix, David, and Omar Licandro. 1999. "Life Expectancy and Endogenous Growth." *Economics Letters* 65: 255–63.
de la Croix, David, and Omar Licandro. 2012. "The Child is Father of Man: Implications for the Demographic Transition." *Economic Journal* 123 (567): 236–61.
de la Croix, David, Thomas Lindh, and Bo Malmberg. 2008. "Swedish economic growth and education since 1800." *Canadian Journal of Economics* 41 (1): 166–85.
Deaton, Angus. 2003. "Health, Inequality, and Economic Development." *Journal of Economic Literature* 41 (1): 113–58.
Doepke, Matthias. 2004. "Accounting for Fertility Decline During the Transition to Growth." *Journal of Economic Growth* 9 (3): 347–83.
Doepke, Matthias. 2005. "Child mortality and fertility decline: Does the Barro-Becker model fit the facts?" *Journal of Population Economics* 18 (2): 337–66.
Domeij, David, and Martin Floden. 2010. "Inequality Trends in Sweden 1978–2004." *Review of Economic Dynamics* 13 (1): 179–208.
Dribe, Martin. 2004. "Long-term effects of childbearing on mortality: Evidence from pre-industrial Sweden." *Population Studies* 58 (3): 297–310.

Eckstein, Zvi, Pedro Mira, and Kenneth I. Wolpin. 1999. “A Quantitative Analysis of Swedish Fertility Dynamics: 1751–1990.” *Review of Economic Dynamics* 2 (1): 137–65.
Edvinsson, Rodney. 2014. “GDP per capita in Sweden 1800–2000.” http://www.historia.se/tableC.xls (accessed August 11, 2014).
Erosa, Andres, Tatyana Koreshkova, and Diego Restuccia. 2011. “How Important Is Human Capital? A Quantitative Theory Assessment of World Income Inequality.” *Review of Economic Studies* 77 (4): 1421–49.
Eurostat. 2014. European Community Household Panel. http://ec.europa.eu/eurostat/web/microdata/european_community_household_panel (accessed August 11, 2014).
Galor, Oded. 2005. “From Stagnation to Growth: Unified Growth Theory.” In *Handbook of Economic Growth*, Vol. 1A, edited by Philippe Aghion and Steven N. Durlauf, 171–293. Amsterdam: North-Holland.
Galor, Oded, and David N. Weil. 2000. “Population, Technology, and Growth: From Malthusian Stagnation to the Demographic Transition and Beyond.” *American Economic Review* 90 (4): 806–28.
Greasley, David, and Jakob B. Madsen. 2010. “Curse and Boon: Natural Resources and Long-Run Growth in Currently Rich Economies.” *Economic Record* 86 (274): 311–28.
Haveman, Robert H., and Barbara Wolfe. 1995. “The Determinants of Children’s Attainments: A Review of Methods and Findings.” *Journal of Economic Literature* 33 (4): 1829–78.
Hazan, Moshe. 2009. “Longevity and Lifetime Labour Supply: Evidence and Implications.” *Econometrica* 77 (6): 1829–63.
Human Mortality Database. 2014. “Sweden.” Max Planck Institute for Demographic Research. http://www.mortality.org/ (accessed August 11, 2014).
Jones, Larry E., and Alice Schoonbroodt. 2010. “Complements Versus Substitutes And Trends In Fertility Choice In Dynastic Models.” *International Economic Review* 51 (3): 671–99.
Jones, Larry E., Alice Schoonbroodt, and Michelle Tertilt. 2011. “Fertility Theories: Can They Explain the Negative Fertility-Income Relationship?” In *Demography and the Economy*, edited by John B. Shoven, 43–100. Chicago: University of Chicago Press.
Kalemli-Ozcan, Sebnem. 2002. “Does Mortality Decline Promote Economic Growth?” *Journal of Economic Growth* 7 (4): 411–39.
Kalemli-Ozcan, Sebnem. 2003. “A stochastic model of mortality, fertility, and human capital investment.” *Journal of Development Economics* 70 (1): 103–18.
Kalemli-Ozcan, Sebnem, Harl E. Ryder, and David N. Weil. 2000. “Mortality decline, human capital investment, and economic growth.” *Journal of Development Economics* 62 (1): 1–23.
Keyfitz, Nathan, and Wilhelm Flieger. 1968. *World Population: An Analysis of Vital Data*. Chicago: University of Chicago Press.
Kirk, Dudley. 1996. “Demographic Transition Theory.” *Population Studies* 50 (3): 361–87.
Krantz, Olle, and Lennart Schön. 2007. *Swedish Historical National Accounts, 1800–2000*. Stockholm: Almqvist and Wiksell International.
Lagerlöf, Nils-Petter. 2003. “From Malthus to Modern Growth: Can Epidemics Explain the Three Regimes?” *International Economic Review* 44 (2): 755–77.
Lewis, Mary E., and Rebecca 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): 117–29.
Ljungberg, Jonas, and Anders Nilsson. 2009. “Human Capital and Economic Growth: Sweden 1870–2000.” *Cliometrica* 3 (1): 71–95.
Lorentzen, Peter, John McMillan, and Romain Wacziarg. 2008. “Death and Development.” *Journal of Economic Growth* 13 (2): 81–124.
Lutz, Wolfgang, Anne Goujon, and Warren 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*, 193–235. Vienna: Vienna Institute of Demography.
Moav, Omer. 2005. “Cheap Children and the Persistence of Poverty.” *Economic Journal* 115 (500): 88–110.
Mookherjee, Dilip, Silvia Prina, and Debraj Ray. 2012. “A Theory of Occupational Choice with Endogenous Fertility.” *American Economic Journal: Microeconomics* 4 (4): 1–34.
Mturi, Akim J., and Andrew Hinde. 2007. “Fertility levels and differentials in the United Republic of Tanzania.” *Population Bulletin of the United Nations* 46: 195–207.
Neisser, Ulrich, Gwyneth Boodoo, Thomas J. Bouchard, Jr., A. Wade Boykin, Nathan Brody, Stephen J. Ceci, Diane F. Halpern, et al. 1996. “Intelligence: Knowns and Unknowns.” *American Psychologist* 51 (2): 77–101.

Ngai, L. Rachel. 2004. "Barriers and the transition to modern growth." *Journal of Monetary Economics* 51 (7): 1353–83.

Organisation for Economic Co-operation and Development (**OECD**). 2015. "Average effective age of retirement in 1970–2012 in OECD countries (time series)." http://www.oecd.org/els/emp/Summary_1970+values.xls (accessed May 13, 2015).

Preston, Samuel H. 1975. "The Changing Relation Between Mortality and Level of Economic Development." *Population Studies* 29 (2): 231–48.

Rogerson, Richard, and Johanna Wallenius. 2009. "Micro and macro elasticities in a life cycle model with taxes." *Journal of Economic Theory* 144 (6): 2277–92.

Schön, Lennart. 2008. "Sweden: Economic Growth and Structural Change, 1800–2000." In *EH. Net Encyclopedia*, edited by Robert Whaples. http://eh.net/encyclopedia/sweden-economic-growth-and-structural-change-1800-2000/.

Skirbekk, Vegard. 2008. "Fertility Trends by Social Status." *Demographic Research* 18 (5): 145–80.

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

Spolaore, Enrico, and Romain Wacziarg. 2013. "How Deep Are the Roots of Economic Development?" *Journal of Economic Literature* 51 (2): 325–69.

Strulik, Holger, and Sebastian Vollmer. 2013. "Long-Run Trends of Human Aging and Longevity." *Journal of Population Economics* 26 (4): 1303–23.

Strulik, Holger, and Jacob Weisdorf. 2008. "Population, food, and knowledge: A simple unified growth theory." *Journal of Economic Growth* 13 (3): 195–216.

Strulik, Holger, and Jacob Weisdorf. 2014. "How Child Costs And Survival Shaped The Industrial Revolution And The Demographic Transition." *Macroeconomic Dynamics* 18 (1): 114–44.

United Nations Population Statistics. 2014. UN Demographic Yearbook, various volumes. http://unstats.un.org/UNSD/demographic/products/dyb/dyb2.htm (accessed August 11, 2014).

Voigtländer, Nico, and Joachim Voth. 2013. "How the West 'Invented' Fertility Restriction." *American Economic Review* 103 (6): 2227–64.

Wang, Limin. 2003. "Determinants of Child Mortality in LDCs: Empirical Findings from Demographic and Health Surveys." *Health Policy* 65 (3): 277–99.

Woods, Robert. 1997. "Infant Mortality in Britain: A Survey of the Current Knowledge on Historical Trends and Variations." In *Infant and Child Mortality in the Past*, edited by Alain Bideau, Bertrand Desjardins, and Héctor Pérez-Brignoli, 74–88. Oxford: Clarendon Press.

World Development Indicators. 2014. World Bank, http://data.worldbank.org/data-catalog/world-development-indicators (accessed August 11, 2014).