The Economic and Demographic Transition, Mortality, and Comparative Development Online Appendix

By Matteo Cervellati and Uwe Sunde*

This Online Appendix contains supplementary material for the paper "The Economic and Demographic Transition, Mortality, and Comparative Development" by Matteo Cervellati and Uwe Sunde (American Economic Journal: Macroeconomics).

* Cervellati: University of Bologna, Piazza Scaravilli 2, 40125 Bologna, m.cervellati@unibo.it. Sunde: University of Munich (LMU), Geschwister-Scholl-Platz 1, D-80539 Muenchen, uwe.sunde@lmu.de.

A. Illustration of the Dynamic System

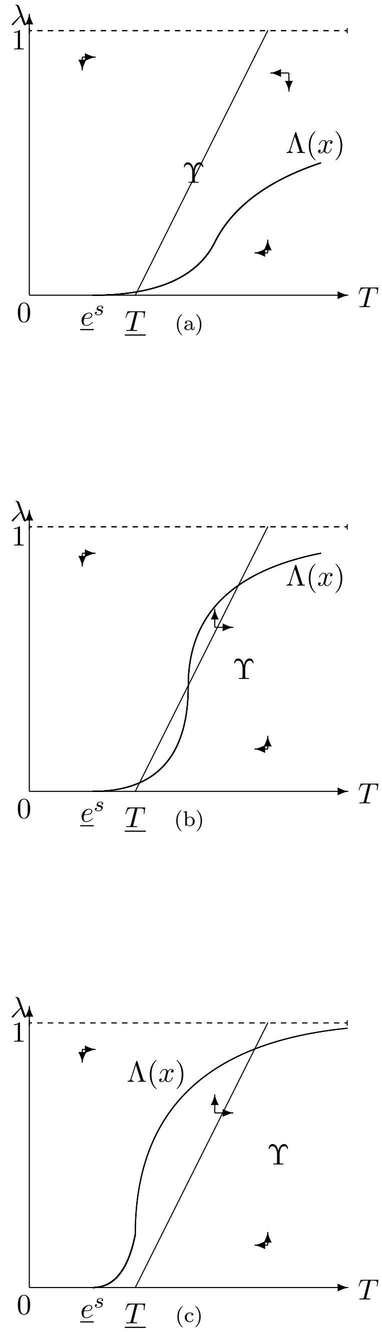


Figure 1. The Process of Development

B. Calibration: Data Sources and Details

Length of a generation. The mean age at first birth for the average country is set to 20 years. in Sweden around 1800 was slightly higher, see Dribe (2004), while age at first birth is still below 20 in pre-transitional countries in Africa nowadays, see Mturi and Hinde (2007).

Age of retirement. Data from `http://www.oecd.org/dataoecd/3/1/39371913.xls`.

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), see main text. Data sources: ERS Dataset (`www.ers.usda.gov`) or historical statistics from the Bank of Sweden (`www.historicalstatistics.org`). Targeting TFP (multi-factor productivity), labor productivity, or the Solow Residual around 2000 (see, e.g., OECD Statistics (`https://stats.oecd.org/Index.aspx?DataSetCode=MFP`) or Donghyun et al., 2012, Table 7) instead of income would deliver very similar values for ϕ.

Production Function. The elasticity of substitution between skilled and unskilled workers is set following the literature. See for instance Acemoglu (2002).

Human Capital. To calibrate some parameters we target a 10 percent pre-transitional share of skilled individuals. The alternative available data sources provide slightly heterogeneous information on enrolment rates in the early 19th Century Sweden, with estimates ranging from about 5 to about 15 percent, see de la Croix, Lindh, and Malmberg (2008) and Ljungberg and Nilsson (2009). The target of the precise value of λ before the transition used for the computation is of little importance for the the obtained parameters, however. The results for alternative parameters obtained by targeting levels of λ up to 0.3 are essentially the same. The average years of schooling in Sweden was 12 years in 2000 (for the cohort age 25-35). Data from Lutz, Goujon, and Sanderson (2007). The earliest available data suggest around 1 year of schooling on average before or around the onset of the transition. The estimates are slightly lower when referring to the entire population alive in Sweden in 2000 since older cohorts are included (for instance 11.4 in the data of Barro and Lee, 2001, and 11.5 years in Ljungberg and Nilsson, 2009). Regarding pre-transitional education levels, the estimates differ somewhat more. Ljungberg and Nilsson (2009) report 1.03 years of schooling in the total Swedish population aged 15-65 in 1870, and 0.1 average standard school years of the population aged 7-14 around 1810-1820, considering absenteeism and length of school years.

Ability Distribution. We estimate the income distribution for Sweden in 2000 using micro data from the ECHP dataset for individual incomes of full-time employees aged 25 to 45, which corresponds to the two last cohorts in the dynamic simulation, and equivalently to the two first generations with $\lambda = 1$ in the data. The income used to estimate the parameters of the ability distribution are converted in US-$ using an average exchange rate of 9 Kroner for one US-$ in 2000. The income distribution is approximately log-normal between the 5th and 95th percentile of the data, with slightly thicker tails. The distribution of log incomes has mean 9.7, standard deviation 0.4, and the lowest and highest observed log-incomes are 6.7 and 12.8, respectively, which implies a maximum spread of 6.1. The moments of the income distribution for the age cohort 25-65 are essentially the same, with the lowest, mean, and highest levels of log income being 6.7, 9.7, and 12.8, respectively, and with a standard deviation of 0.41. The ECHP data are based on surveys and refer to total net income from work, which might explain the small differences between the log income per capita from macro data, which is approximately 10 in 2000, and the mean log income from the micro data that is about 9.7. The relevant data moments extracted from this data set are broadly consistent with other data sources based on register data and alternative surveys for gross earnings, see, Domeij and Floden (2010). The data moments are also close to the ones typically used for the calibration of dispersion in permanent incomes in other OECD countries. For instance, Erosa, Koreshkova, and Restuccia (2011) match a variance of log permanent earnings in the US of 0.36. Robustness checks show that the results are fairly insensitive to varying the dispersion. It is worth noting that the distribution of cognitive ability (or IQ), which is generally measured in the literature as a truncated normal with mean 100 and standard deviation 15, see, e.g., Neisser et al. (1996), would imply a very similar parametrization when normalized for a support $a \in [0,1]$, with $\mu = 0.5$ and $\sigma = 0.075$.

Adult longevity. 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. Data from the Human Mortality Data Base available at `http://www.mortality.org/`. Similar figures are documented for England, France and Italy, see Woods (1997) and Bideau, Desjardins, and Perez-Brignoli (1997) and Lewis and Gowland (2007). In 2000 child mortality in Sweden was around 0.004, which explains the convergence of life expectancy at 5 plus five years of 80.74, and of life expectancy at birth of 80.45.

Child survival probability. Data from `http://www.mortality.org`. The levels of income per capita needed for the computation of the parameters of the function of child survival are extracted from the database of historical statistics of the Bank of Sweden that is freely available online at `www.historicalstatistics.org`. The data are converted to US-$ using an average exchange rate in 2000 of 9 Kroner for one US-$. The income levels used for the calibration of condition (7) are 22,717 and 884 US-$, which correspond to the GDP per capita of Sweden in 2000 and 1800, respectively, in US-$ per 2000.

Preferences. 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). 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. Concerning the cost of raising children, the target $r = 5$ in 2000 is set 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 is about 15 percent which is in line with Doepke (2004) and de la Croix and Doepke (2003). The weight of children relative to own lifetime consumption changes with T_t, as in Soares (2005). For $\gamma = 9$ the relative weight of children compared to per period consumption, γ/T_t, drops from around 0.18 before the transition to around 0.12 in the steady state.

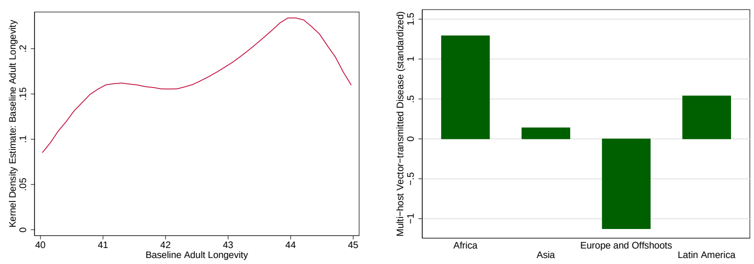
Production function of children's quality. Gross fertility in Sweden in 1800 and 2000 was $n = 2.3$ and $n = 1$. A clear drop in gross fertility occurs around 1900. The data are from Keyfitz and Flieger (1968) and World Development Indicators. The level of TFP and income per capita growth around 1900 vary between 0.7 and 1.7 percent per year. The largest estimates are based on indexed data and include land, see Krantz and Schön (2007), Schön (2008) and Greasley and Madsen (2010). Estimates of TFP and income per capita growth around 1900 vary between 0.7 and 1.7 percent per year. For the calibration we consider the average, 1.2. As an alternative calibration that does not rely on information about the growth rate of technology during the transition, one can also use information on the share of skilled around 1900 and compute the growth rate that is implied by (9). According to estimates by Ljungberg and Nilsson (2009) average years of schooling for the cohort aged 7-14 was around 4 in 1900. Given $\{\phi = 0.61, \underline{e}^u = 0, \underline{e}^s = 12\}$ this implies targeting a level of $\underline{g} = 0.2745$, which delivers essentially the same parametrization.

Initial Conditions. The time axis is set with reference to the convergence to the post-transitional balanced growth path (in terms of λ converging to 1) in 2000. This implies that the choice of $x_0 = 0.04$ determines the beginning of time in the calibration in the stagnation period. This parametrization also implies that the income share of unskilled human capital in total production is larger than 99.9% at the beginning of the simulation, and still above 95% in 1800 just before the transition. The initial level of technology is set targeting the level of GDP per capita in Sweden in 2000 equal to 10.03. Data are from `www.historicalstatistics.org`.

Cross-country differences in life expectancy. For background evidence on the role of a higher exposure to diseases in leading to a faster deficit accumulation and earlier death see, e.g., Mitnitski, Mogilner, and Rockwood (2001) and Searle et al. (2007). Research based on the investigation of skeletons documents that adult longevity during the Mesolithic period was lower in more difficult mortality environments, see Boldson and Paine (2000). As alternative scenario, we target a life expectancy at age five at 45 years (compared to 48 years reflecting Sweden around 1800 just before the transition). The data source is UN Population Statistics available at `www.unstats.un.org`. Data on life expectancy at five for earlier periods are missing for many countries, including most Sub-Saharan Africa countries in 1960. Alternatively, the available information on child mortality and life expectancy at birth in 1960 can be used to derive an estimate of life expectancy at age five. This delivers a very similar target for the highest mortality countries. In 1960 life expectancy at birth was as low as 33 years in some countries like Afghanistan, and child mortality one third. Assuming a constant death rate below the age of 5, these numbers imply a life expectancy at age five between 44 and 45 years. In some countries, like Swaziland life expectancy at birth is just above 30 years still today (data from the CIA World Factbook). This suggests that 45 is possibly a conservative estimate of baseline adult longevity in the worst conceivable mortality environment. Retaining a target of 76 years for life expectancy at age five on the balanced growth path, this implies setting a $\underline{\underline{\mathrm{T}}}$=40 and $\overline{\rho} = 36$ (rather than $\underline{\mathrm{T}}$=45 and $\rho = 31$ as in the benchmark calibration).

Cross-country differences in disease environment. The data in the historical disease prevalence across 113 countries is taken from Murray and Schaller (2010). For each pathogen we construct a binary indicator of whether or not a disease has been present at severe or epidemic levels at least once in the history up to the early 20th century. The diseases include leishmanias, schistosomes, trypanosomes, leprosy, malaria, typhus, filariae, dengue, and tuberculosis. Six of these diseases fall into the class of multi-host vector-transmitted diseases, which are particularly difficult to prevent or eradicate even today because the pathogens survive in multiple hosts (both humans and animals), and which are bound to specific transmission vectors, like mosquitos, which require a particular geographical habitat. The endemicity of the class of multi-host vector-transmitted diseases is fairly insensitive to economic development and globalization, and thus an informative measure of cross-country differences in the extrinsic mortality environment, see Smith et al. (2007). Cervellati, Sunde and Valmori (2012) document the health relevance of the number these pathogens in terms of predicting life expectancy and the likelihood of outbreaks of epidemics. The frequency distribution of the counts of pathogens for all countries of the world is used as distribution of baseline adult longevity within the support $[40, 45]$. The resulting distribution, depicted in Figure 2 in terms of a kernel density plot, is modestly skewed. The frequency of simulated countries with baseline longevity $\underline{T} = 45$ corresponds to the frequency of countries with the lowest observed number of multi-host vector-transmitted diseases ever diagnosed (which includes Sweden). Conversely, the frequency of simulated countries with baseline longevity $\underline{\underline{T}} = 40$ corresponds to the frequency of countries with the highest number of multi-host vector-transmitted pathogens (which include several Sub-Saharan African countries). The distribution on the full support $(40, 45)$ is created by a linear intrapolation of the frequency distribution of the counts of multi-host vector-transmitted diseases on a grid of 0.25 diseases. Figure 2(a) plots the resulting distribution of baseline longevity for the 113 countries of the data by Murray Schaller (2010) data. Figure 2(b) plots the number of diseases in different continents relative to the world average (standardized).

Figure 2. The World-Wide Distribution of Multi-host Vector-transmitted Diseases



(a) Synthetic Distribution of $\underline{T}$: Kernel Density Estimate

(b) Data: Diseases Relative to World Average

C. Data Sources for Time Series and Cross-Section

Time Series for Sweden. Life expectancy and fertility data are taken from the Human Mortality Database (`http://www.mortality.org`), Keyfitz and Flieger (1968) (up to 1960) and World Development Indicators (after 1960), respectively. The Data for GDP, population and GDP per capita is provided by the internet portal for historical Swedish statistics, `www.historia.se` and the Swedish Central Statistical Office, `www.scb.se`. The data on schooling are from de la Croix, Lindh and Malmberg (2008) while the data on average years of schooling are from Ljungberg and Nilsson (2009).

Cross Country Panel Data. We use data from Barro and Lee (2001) as benchmark since they are used more frequently and go back to 1960. The other data sources are Human Mortality Database (`www.mortality.org`), the UN Population Statistics (different historical volumes of the UN Demographic Yearbook, `www.unstats.un.org`), the World Development Indicators at:
(`http://data.worldbank.org/data-catalog/world-development-indicators`).

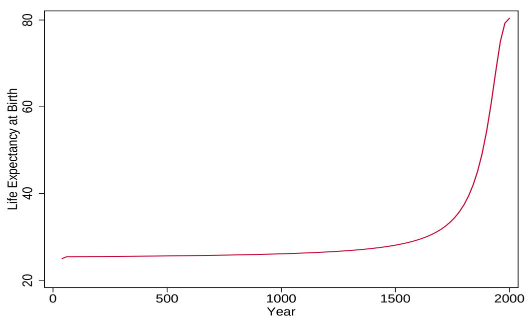
All results are qualitatively and quantitatively very similar using alternative measures like the fraction of total population with at least completed lower secondary education, or the fraction restricted to different age cohorts such as, e.g., age 20-24 years from Lutz, Goujon and Sanderson (2007).

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

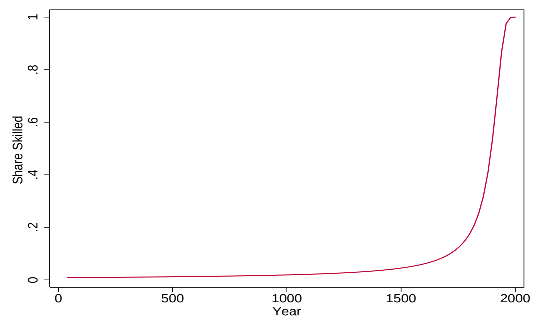
D. Illustration of the Simulated Development Path

Figure 3 depicts the simulated data for the equilibrium share of individuals acquiring skilled human capital and of life expectancy at birth that is obtained from the benchmark calibration. The figure plots the evolution of these variables over the entire simulation period and illustrates the lengthy phase of slow development followed by the endogenous take-off.

Figure 3. Long-Run Development: Simulation of benchmark calibration



(a) Life Expectancy at Birth



(b) Share Skilled λ

E. Bibliography for Online Appendix

The following references refer exclusively to the data sources and articles cited in the Online Appendix.

REFERENCES

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

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

Bideau, Alain, Bertrand Desjardins, and Hector Perez-Brignoli. (1997. *Infant and Child Mortality in the Past*. Clarendon Press, Oxford.

Boldsen, Jesper L., and Richard R. Paine. 2000. "The Evolution of Human Longevity from the Mesolithic to the Middle Ages: An Analysis Based on Skeletal Data," in *Exceptional Longevity: From Prehistory to the Present*, chap. 3. Odense University Press, Odense.

Cervellati, M., U. Sunde, and S. Valmori. 2012. "The Distribution of Infectious Diseases and Extrinsic Mortality Across Countries," *Mathematical Population Studies*, 19(2), 73–93.

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, Thomas Lindh, and Bo Malmberg. 2008. "Swedish Economic Growth and Education Since 1800," *Canadian Journal of Economics*, 41(1), 166–185.

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

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

Donghyun Oh, Almas Heshmati and Hans Loof. 2012. "Technical change and total factor productivity growth for Swedish manufacturing and service Industries," *Applied Economics*, 44(18), 2373–2391.

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

Erosa, Andres, Tamara Koreshkova, and Diego Restuccia. 2011. "How Important Is Human Capital? A Quantitative Theory Assessment of World Income Inequality," *Review of Economic Studies*, 77, 1421–1449.

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

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), 1829–1878.

Keyfitz, Nathan, and Wilhelm Flieger 1968. *World Population: An Analysis of Vital Data*. University of Chicago Press, Chicago.

Krantz, Olle, and Lennart Schön. 2007. *Swedisch Historical National Accounts, 1800-2000.* Almqvist and Wiksell International, Lund.

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), 117–129.

Ljungberg, Jonas, and Anders Nilsson. 2009. "Human Capital and Economic Growth: Sweden 1870-2000," *Cliometrica*, 3, 71–95.

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*, pp. 193–235. IAASA, Vienna.

Mitnitski, Arnold, Alex Mogilner, and Ken Rockwood. 2001. "Accumulation of Deficits as a Proxy Measure of Aging," *Scientific World*, 1, 323–336.

Mturi, Akim, and Andrew Hinde. 2007. "Fertility Levels and Differentials in the United Republic of Tanzania," *Population Bulletin of the United Nations*, 46, 195.

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

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

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

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

Smith, Katherine F., Dov F. Sax, Steve D. Gaines, Vanina Guernier, and Jean-Francois Guegan. 2007. "Globalization of Human Infectious Disease," *Ecology*, 88(8), 1903–1910.

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*, ed. by A. Bideau, B. Desjardins, and H. Perez-Brignoli, chap. 5. Clarendon Press, Oxford.