

Measuring Education in the Century of Change

Extended abstract for a paper to be presented at the 2014 European Population Conference (EPC 2014), Budapest, 25-28 June 2014

Markus Springer, Anne Goujon, Samir K.C. and Michaela Potančoková

Wittgenstein Centre for Demography and Global Human Capital (IIASA, VID/ÖAW, WU), Vienna Institute of Demography (Austrian Academy of Sciences)

Address: Wohllebengasse 12-14, A-1040 Vienna, Austria

Email: markus.springer@oeaw.ac.at

Phone: +43 1 51581 7737

Introduction

This paper describes our approach to establish a Historical Human Capital Database based on two major efforts: First, the collection, harmonization and validation of datasets on educational attainment by age and sex in the 20th century and second, the reconstruction of the gaps in the data time series based on the multi-dimensional cohort-component projection method, that can be used for back-projecting populations. The resulting database shall help in assessing the role of education in the major transformations witnessed during the 20th century whether they were political, social, economic, or demographic.

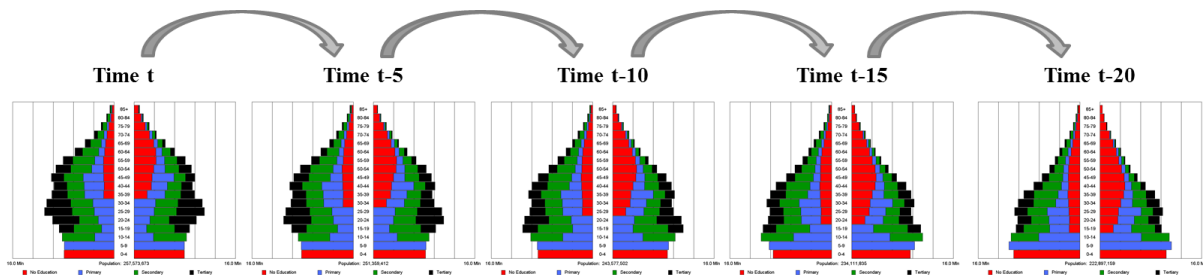
The main base for the backprojection model is the data on educational attainment gathered for the Wittgenstein Centre new projection round (Bauer et al. 2012). However more data points are needed. So far most endeavours aiming at collecting data plus systematically reconstructing and modelling detailed information on past levels of education and literacy (by age and sex) of the population have stopped at 1950 or after (Lutz et al. 2007, Barro and Lee 2010; De la Fuente and Doménech, 2006). A longer perspective can advance our knowledge of human capital formation and its consequences on demographic, socio-economic and technological changes as human capital advancement is highly path dependent (Lutz et al. 2007, Goldin and Katz 2009). The resulting Historical Human Capital Database will enable to study these interdependencies more precisely and in longer time-span.

Data & Methodology

We start from a comprehensive dataset including detailed and accurate data on educational attainment comparable across 171 countries for 2010 (Bauer et.al 2012) and use a multi-dimensional cohort component method to project in five year steps the historical educational composition back to 1970

taking into consideration models of education transitions, mortality and migration differentials by education¹. (K.C. et.al 2013) (see figure 1)

Figure 1 – Schematic Illustration of the Historical Reconstruction



In a second step, the outcome of this reconstruction model is validated against historical datasets on educational stocks from national censuses and surveys to control for the accuracy of the model. Comparing educational compositions of the population 25 years and older of different countries we are able to detect discrepancies between the model and the data to investigate the roots for the deviations, like the over- or underestimation of the pace of educational transition from lower to higher education, to identify inconsistencies caused by external factors like country specific migration patterns like in Israel or Canada, or corrupted census datasets like in Georgia or Indonesia, where historical census over count higher educated people. Whereby most censuses and surveys, especially in the late 20th century provide information on the educational composition of the adult population by age and sex, the theoretical availability of attainment data stands in stark contrast to the actual availability of useful internationally comparable data based on common definitions. The further we go back in time the lower the data availability and the comparability of these datasets becomes.

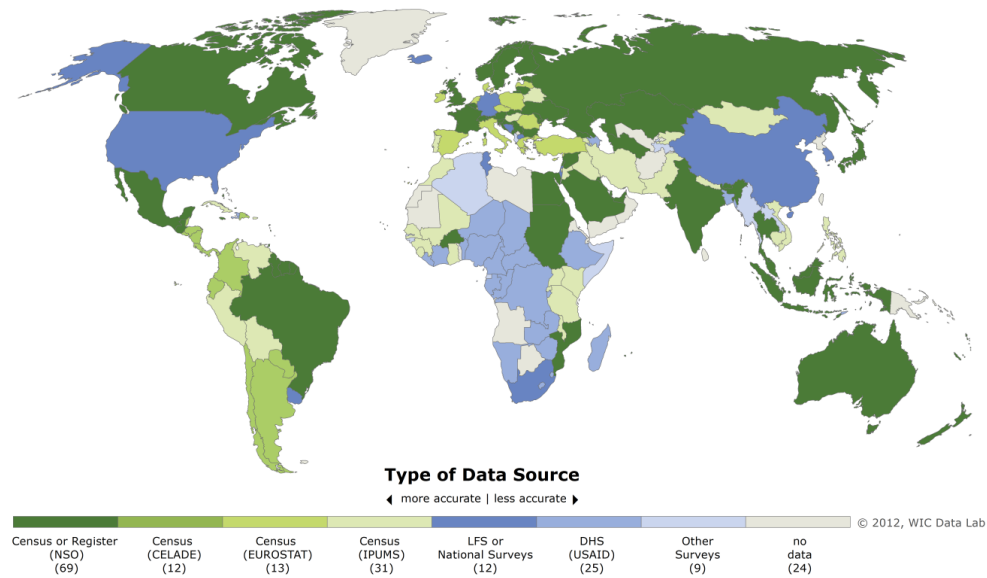
Admittedly the United Nations collected for a number of countries data on educational attainment and literacy from 1945 onwards, reaching back to the 1930s for data on literacy, but there are no comprehensive and harmonized datasets. Other sources are for instance the IPUMS International Database, which covers micro data census samples for a couple of countries since the 1960s. Whenever the validation of the 1970-2010 reconstructed data leads to a dataset that is conform to our quality standards (e.g. Greece 1971 or Pakistan 1973 census), it is used to further back project in time. Since the data available on educational attainment shows a variety of quality and comparability due to the nationally distinct educational systems and has to be harmonized to ensure high quality of the database. Therefore the empirical data has to be transposed from country-specific educational categorisations into the international comparable ISCED 1997 classification.

For the earlier decades of the 20th century the data on educational attainment are unavailable and in general their availability remains scarce until the 1970s. Therefore, the reconstruction will be more

¹ The back projection only reconstructs the share by education which is then applied to estimates of the population by age and sex (from the United Nations Population Division), hence fertility data are not needed.

difficult and will rely on scattered data and will need to be complemented with other sources such as enrolment data, which is more readily available.

Figure 2 – WIC Data Sources on Educational Attainment for the baseyear 2010 by data source



Challenges

The construction of the Historical Human Capital Database will be a step-wise exercise where countries will be added one by one to the database after validation, since the process of modelling and validation poses several challenges, which in particular are:

- **Recovering/validating the empirical data and modelling the missing data:** The lack of comprehensive datasets and timeseries, inconsistent educational categorisations and the variety of data quality, we had to create a multi-dimensional cohort-component projection method to fill the missing gaps to provide a comparable global Historical Human Capital Database.
- **Accounting for structural changes:** These include changes of education systems over time, as well the changes in country borders, entailing changes in the exposed population to education within a country like the extension of compulsory schooling or the change of school systems and types. The control over these changes will assure higher reliability of the data. This is a quite difficult endeavour as despite of the documentation of the United Nation about the school systems since the 1970s no further global documentations exists for the period before.
- **Assessing the role of migration and mortality:** These two demographic determinants influence greatly the educational profile of any population. Mortality differentials by education affect particularly the older population but migration differentials can affect the

whole age distribution. Here again, we have to validate against existing data to account for the differentials. This procedure will advance our knowledge on mortality and migration differentials by levels of education.

The data should be strictly consistent meaning comparable across countries and time, meaning that the Historical Human Capital Database can be used by a broad spectrum of users ranging from researchers in demography and economics to education policy specialists, particularly in advancing our understanding of the types of educational transitions across world countries, and analysing their interactions with the other demographic and socio-economic advancements. The analysis of the database will contribute to more evidence based and therefore more accurate forecasts of the future developments of human capital.

Expected Results

Beside the expected output of the Historical Human Capital Database we will present our methodological approaches, like the multi-dimensional cohort component model, the validation and harmonization procedure for selected countries: Austria, France, Greece, Hungary, Italy, Mexico, Norway, Portugal, Spain, Sweden, United Kingdom & USA

With the establishment of this database we aim to provide a more comprehensive picture of the global educational composition of the past to give a better picture about the future developments.

References:

Barro, R.J. and J.W. Lee. 2010. A New Data Set of Educational Attainment in the World, 1950–2010. Working Paper 15902, National Bureau of Economic Research. Available at: <http://www.nber.org/papers/w15902>

Bauer, R., Potančoková, M., K.C., S. and Goujon, A. 2012. Populations for 171 Countries by Age, Sex, and Level of Education around 2010: Harmonized Estimates of the Baseline Data for the Wittgenstein Centre Projections. IIASA Internal Report (IR-12-016), Laxenburg.

De la Fuente, A., and Doménech, R. 2006. "Human Capital in Growth Regressions: How Much Difference Does Data Quality Make?" *Journal of the European Economic Association* 4(1):1–36.

Flora, Peter. 1983. *State, economy, and society in Western Europe: 1815 – 1975: The growth of mass democracies and welfare states*, I part. Frankfurt am Main: Campus-Verlag.

Flora, Peter. 1987. *State, economy, and society in Western Europe: 1815 - 1975 ; The growth of industrial societies and capitalist economies*, II part. Frankfurt am Main: Campus-Verlag.

Goldin, Claudia, Katz, Lawrence F. 2009. *The Race between Education and Technology*. Harvard University Press.

K.C., S, Potančoková, M., Bauer, R., Goujon, A., and Striessing E. 2013. Summary of Data, Assumptions and Methods for New Wittgenstein Centre for Demography and Global Human Capital (WIC) Population Projections by Age, Sex and Level of Education for 195 Countries to 2100. IIASA Internal Report (IR-13-018), Laxenburg.

Lutz, W., Goujon, A., K.C., S., and Sanderson W. 2007. Reconstruction of Populations by Age, Sex and Level of Educational Attainment for 120 Countries for 1970-2000. IIASA Internal Report (IR-07-002), Laxenburg.

Mitchell, Brian R. 2003. International historical statistics - Europe: 1750 – 2000, 5th edition. Basingstoke: Palgrave Macmillan.

Mitchell, Brian R. 2003. International historical statistics - The Americas: 1750 – 2000, 5th edition. Basingstoke: Palgrave Macmillan.

Mitchell, Brian R. 2003. International historical statistics - Africa, Asia, Oceania: 1750 - 2000. 4th edition. Basingstoke: Palgrave Macmillan.

Morrison Christian, and Murtin, Fabrice. 2009. The Century of Education. *Journal of Human Capital* 3 (1): 1-42.