**Working Paper CLIO-INFRA: Average Years of Education**

1. Title

 - Average years of education (Average, total Population 15 years and older), 1850-2010.

2. Author(s)

 - Bas van Leeuwen, Jieli van Leeuwen-Li, and Peter Foldvari

3. Production date

 - 15-3-2013

4. Version

 - 2nd version.

5. Variable group(s)

 - Human capital.

6. Variable(s)

 -Average years of education per country – Average per total population of 15 years and

older.

7. Unit of analysis

 - number of years of formal education.

8. Keywords (5)

 - human capital, education, world.

9. Abstract (200 words)

- The Average years of education in the total population aged 15 years and older is given for the period 1870-2010.

10. Time period

 - 1850-2000.

11. Geographical coverage

 - Entire World.

12. Methodologies used for data collection and processing

- Bibliographical research, research of published and Online Databases, and cross-analysis of various datasets

13. Data quality

 i. Central statistical agencies

 ii. Historical reconstructions

 iii. Estimates

 iv. Conjectures

Virtually all benchmark data after 1960 (i.e. 1960, 1970, 1980, 1990, 2000, and 2010) are derived indirectly from Central Statistical Agencies and therefore probably better classify as historical reconstructions. The remaining data are estimates.

14. Period of collection

 - October/2010 and July/2012.

15. Data collectors

 - Bas van Leeuwen and Jieli van Leeuwen-Li.

16. Sources

The data were derived by combining two main datasets. First, Morrisson and Murtin (2009) supply information of around 78 countries for 10-year intervals. Second,Van Leeuwen, Van Leeuwen-Li and Foldvari, has been published in a range of papers spread over the past years (*e.g.* Van Leeuwen and Foldvari, 2008a; Foldvari *et al.,* forthcoming) and has been made available in a comprehensive format by Clio-Infra in 2012. The last two datasets, from Van Leeuwen *et al.* and Morrisson and Murtin, have been shown to be the ones with the widest coverage, the longest time span and the highest reliability (*e.g.* Foldvari *et al.,* forthcoming). The methodology for the two is also roughly comparable, making it possible to combine the two sets of estimates into a single comprehensive dataset on educational development.

The main sources, with the exception of some country-specific studies, are:

1. Mitchell, B.R. (2007), *International Historical Statistics: Africa, Asia & Oceania, 1750-2005,* Basingstroke [etc.]: Palgrave Macmillan (5th edition).
2. Mitchell, B.R. (2007), *International Historical Statistics: The Americas, 1750-2005,* Basingstroke [etc.]: Palgrave Macmillan (5th edition).
3. Mitchell, B.R. (2007), *International Historical Statistics: Europe, 1750-2005,* Basingstroke [etc.]: Palgrave Macmillan (5th edition).
4. US Department of Commerce, US Census Bureau, *International database*, downloaded from: <http://www.census.gov/population/international/data/idb/informationGateway.php>
5. Unesco, Statistical yearbook, Paris: Unesco, 1964-1999, 1963-1999
6. Blue books of the British colonies (various issues)

Some of the data are also taken from papers:

1. Morrisson, Ch, and F. Murtin (2009), “The Century of Education,” *Journal of Human Capital,* Vol. 3/1, pp. 1-42.
2. Földvári, P., and B. Van Leeuwen, (2009), “Average years of education in Hungary: annual estimates 1920-2006,” *Eastern European Economics*, Vol. 47/2, pp. 5-20.
3. Földvári, P., B. Van Leeuwen, and J. Van Leeuwen-Li (forthcoming), “Educational and income inequality in Europe, 1870-2000,” *Cliometrica,* Forthcoming.
4. Leeuwen, Bas van, Jieli van Leeuwen-Li, and Péter Földvári, 'Regional human capital in Republican and New China: Its spread, quality and effects on economic growth,' 26 July 2011.
5. Leeuwen, Bas van, Jieli van Leeuwen-Li, and Péter Földvári, 'Was education a driver of economic development in Africa? Inequality and income in the twentieth century,' 23 April 2012.
6. Dmitry Didenko, Péter Földvári and Bas van Leeuwen, 'The spread of human capital in the former Soviet Union area in a comparative perspective: Exploring a new dataset,' Journal of Eurasian studies, Vol. 4 (2) 2013, pp. 123-135.

17. Text

As pointed out, most of the data are derived from census data or from existing studies. Yet. before the 1960s, when no benchmark estimates are available anymore, we have to rely solely on a perpetual inventory method. This method is partially based on Schultze and Fernandes (2009). This method is difficult to write down in equations, but perhaps easier to describe in words. If a person is, for example, going to school in 1900, a enrollment rate is calculate for that year (i.e. how many people enter school, divided by the relevant school age children). 10 years later, he/she enters 15. Hence, in 1915 we use the 1900 enrollment ratio and multiply it with the people age 5-10 in 1910. In this way we correct for any mortality that may occur in the meantime. This same procedure is done for all education levels, age classes, and years and leads to an estimate of average years of education with age specific mortality rates.

This data collection was carried out within the framework of the CLIO-INFRA project financed by the Netherlands Organisation for Scientific Research (NWO).

References:

Schulze, Max-Stephan and Fernandes, F. T. (2009), “**Human capital formation in Austria-Hungary and Germany: time series estimates of educational attainment, 1860-1910.**” In: Halmos, Károly and Klement, Judit and Pogány, Ágnes and Tomka, Béla, (eds.) A felhalmozás míve: történeti tanulmányok Kövér György tiszteletére. Századvég Kiadó, Budapest, Hungary.