Journal of Development Economics 122 (2016) 147–169

Contents lists available at ScienceDirect

Journal of Development Economics

journal homepage: www.elsevier.com/locate/devec

Human capital in the long run

Jong-Wha Lee *, Hanol Lee

Economics Department, Korea University, Sungbuk-Ku, Anam-Ro 145, Seoul 02841, Republic of Korea

ARTICLE INFO

Article history:
Received 25 February 2015
Received in revised form 15 January 2016
Accepted 29 May 2016
Available online 1 June 2016

JEL classification:
120
J24

Keywords:
Education
Enrollment
Human capital

ABSTRACT

This study presents new data sets on long-run enrollment ratios, educational attainment, and human capital stock measures for numerous countries. We construct a complete data set of historical enrollment ratios, subdivided by education level and gender, for 111 countries from 1820 to 1945 (at five-year intervals) by using newly compiled census observations and information on the year of establishment of the oldest school in individual countries. Then, by utilizing these enrollment ratios, as well as available census data from 1945 onward on different age groups' educational attainment, we construct a data set of estimated educational attainment, disaggregated by gender and age group, and aggregate human capital stock that spans from 1870 to 2010. The data show that over the past two centuries, there has been remarkable growth in average educational attainment and human capital stock as well as a narrowing of the gap in average educational attainment between nations.

1. Introduction

Human capital has been emphasized as a critical determinant of economic development. In addition, it exerts a significant influence on social and political issues, such as fertility, the education of children, and democracy (Becker, 2009; Barro, 2012).

Since human capital is multifaceted and includes a complex set of human attributes, it is difficult to precisely and quantitatively measure; as such, many have used educational attainment as a proxy and attempted to measure it across various populations.[1]

This study goes beyond previous ones by presenting new long-run data sets on estimated school enrollment and educational attainment, classified by education level and gender, for a sample of 111 countries. These data sets include measures of enrollment ratios among school-age populations from 1820 onward and educational attainment among those aged 15–64 from 1870 onward.

School enrollment provides valuable information about a society's expansion and investment in education. Enrollment data have been widely used by scholars studying comparative education and politics to investigate the development of national education systems (Benavot and Riddle, 1988; Meyer et al., 1992). It has also been used for cross-national comparative analyses of the relationship between education and economic and social development (Psacharopoulos and Woodhall, 1985; Barro, 1991; Benavot, 1996). Recent studies have also used enrollment ratios to construct estimates of a population's educational attainment (Morrisson and Murtin, 2009; Barro and Lee, 2013, 2015).

UNESCO has collected and published enrollment data for every country in the world; however, this information is limited to the post-World War II era. Before that time, the League of Nations and other international agencies made similar efforts to collect such data. Several compendia (Benavot and Riddle, 1988; Mitchell, 2003a, b, c; Lindert, 2004; Banks and Wilson, 2013) also provide historical enrollment statistics for many independent states and colonies.

Yet, available studies do not provide comprehensive data on historical enrollment rates for each education level (primary, secondary, and tertiary). Moreover, no study has attempted to collect extensive data on female enrollment. Our new data set fills this gap by providing more detailed information regarding the enrollment ratios of 111 countries from 1820 to 1945.

Using a variety of sources, we compiled figures for primary, secondary, and tertiary enrollment among females and the population as a whole. Our data set expands the existing ones by using a significant number of new census observations. We also utilize the information on the years of establishment of the oldest schools at different education levels in individual countries. We then used these figures to calculate enrollment ratio estimates. We present the resulting data set on enrollment, disaggregated by education level (primary, secondary, or tertiary) and gender, for 111 nations and former colonies at five-year intervals from 1820 to 1945. By combing historical enrollment ratios with information from post-1950 UNESCO publications, we are able to construct a complete time series of

* Corresponding author.
E-mail addresses: jongwha@korea.ac.kr (J.-W. Lee), hanollee@korea.ac.kr (H. Lee).

[1] See Barro and Lee (2001) and Barro and Lee (2015) for detailed discussions.

http://dx.doi.org/10.1016/j.jdeveco.2016.05.006
0304-3878/

enrollment ratios up to the year 2010. We then analyze this data set in order to study the trends of school enrollment ratios.

The adult population's aggregate stock of education, which has a significant influence on production and social and political life, cannot be gauged by enrollment ratios, though. Attainment of schooling, as reflected in the enrollment of the school-aged population, is indicative of past inflows to schooling. These flows accumulate to create future stocks of educational attainment, but because the education process takes many years, there is a long lag between the flows and stocks.

Many researchers have attempted to construct direct measures of the stock of educational attainment (De La Fuente and Doménech, 2006; Cohen and Soto, 2007). The most widely used of these measures is the number of years that citizens spend in school. For instance, Barro and Lee (1993, 2001, 2013) constructed a data set on educational attainment for a large number of countries by using census and survey observations on attainment as benchmark stocks, and filled in missing observations by using enrollment-ratio and population structure data. The most recent data set provides estimates of educational attainment for the populations of 146 countries, disaggregated by gender and five-year age group, from 1950 to 2010 at five-year intervals. The data set distinguishes between seven different levels of education: no formal education, incomplete primary, complete primary, lower secondary, upper secondary, incomplete tertiary, and complete tertiary. These data were used to calculate the average years of schooling among the adult population both as a whole and at the primary, secondary, and tertiary levels.

The present study constructs estimates of educational attainment before 1950 by extrapolating later data on various age groups' educational attainment and utilizing estimated historical enrollment data. The newly constructed data set presents estimates of educational attainment, classified by age group (15–24, 25–64, and 15–64) and by gender, for 111 countries from 1870 to 1945 at five-year intervals. Combining this data set with existing data from Barro and Lee (2013), we highlight the major trends in educational attainment from 1870 to 2010 both throughout the world and in major regions.[2]

Our new estimates of historical educational attainment improve upon those of previous studies, such as Morrisson and Murtin (2009), Baier et al. (2006), Tamura et al. (2012), and Van Leeuwen and Van Leeuwen-Li (2014). Morrisson and Murtin (2009) presented a historical database on educational attainment in 74 countries from 1870 to 1960 at 10-year intervals. In order to do so, they used historical enrollment data to generate educational stock measures. Baier et al. (2006) utilized a similar method but only extended their study to 34 countries. Tamura et al. (2012) constructed data set on real output, physical capital, and human capital for 168 countries, 69 of which have human capital stock data from 1890 or from an earlier year onward.

Van Leeuwen and Van Leeuwen-Li (2014) presented a data set on average years of education that covers a wider range of countries since 1870.[3] They combined the estimates of Morrisson and Murtin (2009) with the estimates, for the additional countries, from the researches by Van Leeuwen, Van Leeuwen-Li, and Foldvari (e.g., Földvári and Van Leeuwen, 2014). Our estimates make use of a greater number of original sources on enrollment. More importantly, our estimation method uses actual census data on educational attainment by age group as a benchmark. Thus, this new data set contributes to the existing literature by improving the accuracy of educational attainment estimates for the 1870–1945 time periods. Moreover, our data set subdivides education into seven different levels and presents the distribution of educational attainment among the female, male, and total populations.

We also analyze the evolution and distribution of human capital stock worldwide and across nations based on a human capital measure (as defined by Mincer (1974)) and average years of schooling.

2. Estimates of enrollment ratios by education level and gender from 1820 to 1945

2.1. Data and estimation method

Earlier primary data sources present information on total enrollment. In order to make meaningful comparisons between different countries and time periods, we construct enrollment ratios for each education level. Enrollment ratios are calculated by taking the number of students at each education level and dividing it by the number of persons in the relevant school-age population.

There are several means by which to measure enrollment ratios. An "unadjusted enrollment" ratio" is defined as the ratio of all persons enrolled in a given level of schooling to the population within a specified age bracket, which is consistent across all countries. For example, the earlier UNESCO statistics in the 1960s use constant age categories for all countries: 5–14 for the primary level, 15–19 for the secondary level, and 20–24 for the tertiary level. This measure can be biased, though, because it does not consider the national variations in the duration of schooling. For example, if a country requires all children aged 6 to 11 to attend a six-year cycle of primary school, then the unadjusted enrollment ratio, which uses the 5–14 year old population as a denominator, underestimates the true enrollment ratio.

A "gross enrollment ratio" takes into account the differences in school-age populations that stem from disparate national educational systems. It is constructed by dividing the number of all persons enrolled in a given level of schooling by the population of the age group that should be enrolled at that level according to national regulations or customs. For example, in Korea, where primary education begins at the age of 6 and lasts for 6 years, the total number of registered students in primary school would be compared with the population of children aged 6 to 11.

The gross enrollment ratio is widely used in many official publications by UNESCO and national statistical agencies for cross-national comparisons. Though this statistic is more often available for developing countries, it tends to be skewed when a significant number of students repeat grades, which is a typical occurrence in the developing world.[4] In fact, for many advanced countries the gross enrollment ratio often exceeds 100%.

A "net enrollment ratio" is defined as the ratio of students in a designated age group, at a given level of schooling to the total population of that age group. For example, registered primary school students aged 6–11 are compared with the total population of children between the ages of 6 and 11. As such, this statistic does not count the students in primary school who are younger than 6 or older than 11. If there are a great number of such under- or over-aged children at each level of education, then the net enrollment ratio will suffer from measurement errors. In addition, data on net enrollment ratios are not widely available.

We construct "adjusted enrollment ratios," which are appropriate for comparing expansions in primary and secondary school education across nations (Barro and Lee, 2001). These adjusted ratios are a modification of the gross enrollment ratios in that they account for the repetition of grades in primary and secondary schools. The gross enrollment ratio, whether adjusted or unadjusted, is a standard flow measurement of educational investment in a society's new entrants, and it is widely

[2] Barro and Lee (2015) use a similar method to determine educational attainment from 1870 to 2010 for 89 countries. The study does not use information on the establishment years of the oldest schools to construct historical estimates of enrollment ratios.

[3] The data are available at the Clio-Infra (2014) website (https://www.clio-infra.eu).

[4] Fredriksen (1983) estimated that the average gross enrollment ratio at the primary level was 86% for developing countries in 1980; however, when adjusted for repeaters, the figure fell to 73%. The data on repetition ratios are discussed later in this section.