



HAL
open science

The Kuznets curve of human capital inequality: 1870–2010

Christian Morrisson, Fabrice Murtin

► **To cite this version:**

Christian Morrisson, Fabrice Murtin. The Kuznets curve of human capital inequality: 1870–2010. *Journal of Economic Inequality*, 2013, 11 (3), pp.283 - 301. 10.1007/s10888-012-9227-2. hal-03473868

HAL Id: hal-03473868

<https://sciencespo.hal.science/hal-03473868>

Submitted on 10 Dec 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

The Kuznets curve of human capital inequality: 1870–2010

Christian Morrisson · Fabrice Murtin

Published online: 29 June 2012

Abstract This paper describes the long-term global trends in education inequality since 1870. Inequality in years of schooling is shown to have mechanically decreased along with the decline in the share of illiterate people. In search of a monetary equivalent of years of schooling, we turn to Mincer (1974) human capital inequality. Within countries, we find evidence of an inverted U-shape curve for human capital inequality over time, namely a Kuznets curve for human capital. At the global level, the world inequality in human capital has followed a similar trajectory, first increasing from 1870 to 1970, then decreasing.

Keywords Inequality · Human capital · Education · Kuznets curve

JEL Classifications D31 · E27 · F02 · N00 · O40

1 Introduction

While several studies have described the world distribution of income (e.g. Bourguignon and Morrisson [5], Milanovic [16], Sala-I-Martin [26]), less attention has been paid to the global patterns of the underlying factors of production. In particular, the world distribution of education has never been described over the long run. A better understanding of global inequality in education is important as education convergence is a prerequisite for income convergence. Indeed, education stands as a key factor of production that determines individual workers' productivity, drives the pace of technological progress [1] and triggers key externalities such as the demographic transition [23] or democratization [24]. Besides, education is recognized to be a crucial dimension of welfare alongside income, health and political

C. Morrisson
Université Paris I, 36 Chemin Desvallières, 92410 Ville d'Avray, France

F. Murtin (✉)
OECD, Paris, France
e-mail: fabrice_murtin@hotmail.com

rights. How the world distribution of education has evolved over time is therefore informative on global welfare trends.

Against this background, the paper makes mainly three contributions. First, this study provides estimates of the world distribution of education and of global inequality in years of schooling from 1870 to 2010. Second, it shows that the large decrease in years of schooling inequality is the mechanical consequence of the decline in illiteracy. The third contribution is linked to the search of a monetary equivalent to education. Using the Mincerian definition of human capital, we estimate the distribution of human capital within each country and at the global level.

Whereas existing studies such as Castello and Domenech [7] or Thomas et al. [27] provide a descriptive analysis of inequality in years of schooling both within and between countries, they do not consider the world distribution of years of schooling regardless of country borders. After summarizing the distribution of education by four quantiles (no schooling, only primary schooling, primary and secondary schooling, higher education) in each country, we obtain a distribution of 128 groups (4 quantiles \times 32 large countries) forming the world distribution of years of schooling, and allowing the estimation of schooling global inequality.

Moreover, the cross-country negative correlation between the average of and the inequality in years of schooling has been underlined in several articles (e.g. Castello and Domenech [7], Berthelemy [3]). We complement the latter studies by showing that the observed decline in years of schooling inequality since 1870 is almost entirely explained by the decline in illiteracy, and we provide the theoretical explanation of this relationship.

Last, in line with recent macroeconomic literature (e.g. Bils and Klenow [4], Hall and Jones [11]), we turn to human capital as defined by Mincer [17] to confer a monetary dimension to education. We propose the first estimation of global inequality in human capital, while testing several definitions of human capital. At the national level, we find that human capital inequality has increased within countries, before stabilizing or even decreasing in most regions of the world. When plotted against average human capital, human capital within-country inequality has followed an inverted U-shape curve, namely a “Kuznets curve of human capital inequality”. However, the decrease in human capital inequality is less pronounced when returns to education are assumed to increase with schooling attainment. At the global level, we also find that human capital inequality has increased from 1870 to approximately 1970, and has decreased afterwards.

Section 2 introduces the methodology and the data. Section 3 examines the world distribution of education since 1870. Section 4 focuses on global inequality in education, while Section 5 looks at global inequality in human capital. In Section 6, we describe human capital inequality within countries, while last section concludes.

2 Methodology and data

This section describes the geographical coverage of the data, the construction of national distributions of education, and the computation of global inequality in education.

The definition of countries or geographical areas follows Bourguignon and Morrisson [5], who have selected a sample of large countries and merged countries

of smaller size. Interestingly, having the same sample of ‘macro-countries’ as the latter authors enables a comparison of education and income inequalities. For GDP per capita and population, we updated the data from these authors using the last estimates from Maddison [14]. Country-level data have been averaged to obtain a final sample of 32 macro-countries, which correspond to at least 90% of world population at any period.¹ Each country or country group represents at least 1 per cent of world population or world GDP in 1950. In order to facilitate the presentation of the results, these macro-countries have been aggregated into eight blocks defined geographically, historically or economically: Africa; Latin America excluding Argentine and Chile; Eastern Europe; Western Europe (including Austria, Czechoslovakia and Hungary) and its offshoots in America (Argentina, Chile, Canada and the US) and in the Pacific (Australia and New-Zealand); China; India (including Bangladesh and Pakistan); Japan and Korea; other Asian countries.

The construction of national distributions of education proceeds in two steps. First, we have improved the education database published by Morriison and Murtin [19].² We took into account differential mortality across educational groups to correct the distributions of schooling after 1960, and we refined the US series by taking into account immigration in this country over the 19th and early 20th centuries, while drawing from Murtin and Viarengo [21].

On a second step, we calculated the distribution of education in each country. It is summarized by four groups, namely the percentage p^P of the population displaying only primary schooling, the percentage p^S of the population displaying primary and secondary schooling, the percentage p^H of the population displaying primary, secondary and tertiary schooling, and the complementary part, the percentage p^I of the population that has not attended school. These percentages are inferred from the knowledge of average years $H^{P,S,H}$ of primary, secondary and tertiary schooling, as well as from assumptions on the average durations at school $h^{P,S,H}$ in each grade.³ The national distributions of schooling are given by the following formulas:

$$\begin{cases} H^P = h^P p^P + 6(p^S + p^H) \\ H^S = h^S p^S + 6 p^H \\ H^H = h^H p^H \\ p^P + p^S + p^H + p^I = 1 \end{cases} \quad (1)$$

¹Bourguignon and Morriison [5] have 33 macro-countries as Poland is included within their sample but not in ours.

²It can be downloaded from the website www.fabricemurtin.com together with a companion annex describing data construction in greater details. Morriison and Murtin [19] data set is based on Cohen and Soto [8] for the period 1960–2010, while estimates of average years of schooling have been calculated via perpetual inventory for the period 1870–1960, using data on total enrollments in primary, secondary and higher education since the early 19th century as well as age pyramids. The latter data is available from the work of Mitchell, an historian who published thousands of historical time series covering the 19th and 20th centuries in many countries.

³We consider different repeating and dropout rates across countries and time as in Morriison and Murtin [19], which reflect, at least to some extent, international differences in education quality. Obviously, this is only a partial account of the existing heterogeneity in education quality as measured for instance by PISA studies. Nonetheless, the latter studies started only in 1997. In practice, we have assumed that completed primary was lasting a maximum of six years and tertiary a maximum of four years. These assumptions ensure comparability across time and countries of education distributions, in spite of the many international reforms of schooling systems over that period.

Moreover, we have cross-checked the above estimates of the percentage of non-educated population p^I with a historical database on illiteracy rates that contains 179 observations mainly taken from UNESCO [28] and spanning over the 1870–1950 period. Although we do not expect rates of illiteracy and of school non-attendance to match perfectly as pupils who have attended school for a few years could still be technically illiterate, while people who have never attended school could have received some literacy within their household or at some unreported church-based school, we expect a high level of correlation between the two variables. As a result, this correlation is equal to 0.98 and there is no significant outlier, which suggests a satisfactory degree of data reliability.⁴

The world distribution of education is then obtained as follows. The unit of analysis is a given group of education (people with no schooling, with only primary, with primary and secondary, with primary, secondary and tertiary schooling) in a given country, implying a world distribution of education composed of 4 groups \times 32 macro-countries = 128 groups. Each group is characterized by its share $p_i^{P.S.H}$ of national population and its average number of years of education $h^{P.S.H}$. Accordingly, all groups are pooled across countries and are ranked according to their level of schooling, and then the Lorenz curve is computed while weighting each group by national population. We assumed no heterogeneity in years of schooling within each group. As reported in the companion annex (cf. n. 2 with reference to our website www.fabricemurtin.com), introducing heterogeneity within groups raises slightly the level of world education inequality, but does not modify the observed trends and leaves the main conclusions of the paper largely unchanged.

3 Global trends in educational attainment since 1870

Table 1 presents the distribution of years of schooling at the world level since 1870. In the mid-twentieth century, the world was divided into two classes: Those who have attended school, and those who have not. Over the whole period, there has been a huge reversal, as illiterates and educated individuals are roughly in inverse proportions in 1870 and in 2010. What explains this result is clearly the development of primary schooling, whose attendance involved 24% of the world population in 1870 and 82% in 2010. Moreover, 45% of the world population attended secondary education in 2010, but this development is quite recent since this proportion was about 20% in 1960. In a sense, higher education is the contemporary equivalent of secondary schooling in 1930: 11% of the world population attained higher education in 2010, whereas the population with at least secondary schooling represented 10% of world population in 1930. Overall, the average level of schooling has been multiplied by 6, this increase being unequally spread over the period. Indeed, the absolute increase was less than 3 average years of schooling between 1870 and 1960, but has been equal to 3.5 years over the last fifty years.

The global rise in schooling attainment has been unequally distributed across countries. Table 2 provides a geographical overview of educational attainment broken down in its underlying components (illiteracy, primary, secondary and tertiary schooling). There were three distinct groups in 1870. In Western Europe and

⁴This database is now routinely used by the OECD.

Table 1 The world distribution of schooling

	1870	1880	1890	1900	1910	1920	1930	1940	1950	1960	1970	1980	1990	2000	2010
Illiteracy rate	75.92	73.14	70.13	66.45	64.06	60.06	56.79	50.96	48.73	44.08	36.98	35.03	27.52	20.50	18.22
Share having only primary	20.13	22.15	24.28	27.20	28.41	31.40	33.33	37.85	37.59	37.88	39.57	35.17	37.14	37.99	36.92
Share having only secondary	3.77	4.48	5.38	6.17	7.08	8.01	9.25	10.63	12.84	16.75	20.84	24.83	28.84	32.78	34.61
Share having tertiary	0.19	0.25	0.25	0.40	0.55	0.73	0.91	1.13	1.75	2.45	3.74	5.21	7.00	9.23	10.89
Mean years of schooling	1.24	1.41	1.62	1.86	2.08	2.34	2.62	2.98	3.37	4.02	4.77	5.52	6.33	7.13	7.51
Mean years of primary schooling	1.01	1.14	1.30	1.48	1.64	1.82	2.02	2.29	2.49	2.87	3.27	3.67	4.10	4.49	4.61
Mean years of secondary schooling	0.22	0.26	0.31	0.36	0.42	0.48	0.56	0.65	0.81	1.06	1.36	1.66	1.97	2.31	2.50
Mean years of tertiary schooling	0.01	0.01	0.01	0.01	0.02	0.03	0.03	0.04	0.07	0.09	0.14	0.19	0.26	0.34	0.40

offshoots, average schooling exceeded 3 years. In Latin America, Eastern Europe, Japan, Korea and China, it was comprised between 0.6 and 1 year. In Africa, South Asia (India, Bangladesh and Pakistan) and other Asian countries, average schooling was less than 0.15 years. The illiteracy rate was about 36% in the first group, 80% in the second one, and above 95% in the third one. These figures highlight the huge gap between Western Europe and the rest of the world.⁵

In 2010, the group of less advanced countries is only composed of Africa and South Asia, because average schooling in other Asian countries has increased much more than in India, Bangladesh and Pakistan. In contrast, Japan and Korea, as well as Eastern Europe to a lesser extent, have caught up with Western Europe. In the intermediate group, we find Latin America, China, and other Asian countries with average schooling around 8 years. The difference between Western Europe and this group are about five average years of schooling, which can be roughly decomposed into one year of primary schooling, three years of secondary schooling, and one year of higher education. Figure 1 illustrates clearly the process at work across geographical areas. It is striking that no global convergence in average educational levels has been observed in the postwar period. Illiteracy, which was a common rule in 1870 with rates exceeding 80% everywhere except in Western Europe, is now a regional problem. It remains substantial only in Sub-Saharan Africa and in South Asia with rates around 36% in 2010.

Changes in the world distribution of education since 1870 are described by Table 3, which displays the regional composition of three world quantiles, namely the bottom 60%, an intermediate group composed of the 7th, 8th and 9th deciles, as well as the 10th decile (the first line indicating the share in world population). The main factors explaining the time variations are the differences in the growth rates of average education and of population.⁶ Most of the time, the shares of each group of countries in the world distributions of education and income are often similar in a given year, which is not surprising as schooling and income are highly correlated across countries. However, there are a few interesting exceptions.

Between 1870 and 1910, the respective shares of Africa, South Asia and other Asian countries in the upper deciles (7th–10th) of the world distribution of education have been very low compared with their respective shares in income. On the contrary, these shares have been much higher for schooling than for income in China, and to some extent in Japan and Korea. From that perspective, the current gap in economic development between Africa and China is not new. In 1870, nearly all African population was illiterate, except the Arab population in Northern Africa, whereas in Asia nearly 40% of Chinese and Japanese men could read and write.

⁵China and Japan were ahead of other Asian countries, the Indian empire and Africa, with about one year of schooling (education was higher in Japan than in Korea). This means that around 40% of men and 10% of women could read and write 3000 graphic signs, which requires about 4 years of schooling. A small minority knew several thousand signs after 6 or 8 years of schooling. As the average educational attainment in China and Japan was approximatively the same at the beginning of the eighteenth century, these countries were the only ones in the world that had the same average schooling than Western Europe three centuries ago.

⁶For instance, the shares in world population of Latin America and Africa respectively, have been multiplied by 3 and 2 between 1870 and 2000, whereas the shares of Western Europe and Eastern Europe have decreased.

Table 2 Mean years of schooling and illiteracy rate by geographical area 1870–2010

	Africa	Latin America	Eastern Europe	Europe and offshoots	China	Japan– Korea	South Asia	Other Asian countries
1870								
Total MYS	0.12	0.63	0.94	3.81	1.01	1.00	0.08	0.14
Primary	0.11	0.53	0.84	2.89	0.91	0.98	0.08	0.11
Secondary	0.01	0.09	0.10	0.90	0.09	0.02	0.00	0.02
Illiteracy rate	97.52	86.47	77.62	36.36	75.16	79.77	98.08	97.17
Population (in millions)	66.0	32.4	94.9	229.5	358.0	44.2	257.6	72.9
1910								
Total MYS	0.24	1.31	1.40	6.04	1.21	2.61	0.37	0.28
Primary	0.21	1.06	1.24	4.41	1.10	2.52	0.30	0.24
Secondary	0.02	0.24	0.16	1.56	0.10	0.08	0.07	0.04
Illiteracy rate	95.13	73.37	67.26	13.11	69.86	47.95	93.42	94.01
Population (in millions)	96.2	64.0	148.9	349.3	423.0	59.6	305.0	109.5
1950								
Total MYS	0.82	2.99	5.04	7.93	1.69	6.02	1.16	1.25
Primary	0.68	2.36	4.07	5.35	1.45	4.66	0.87	1.01
Secondary	0.14	0.59	0.86	2.41	0.21	1.25	0.26	0.22
Illiteracy rate	86.46	45.72	2.44	5.20	61.19	15.88	82.70	76.29
Population (in millions)	175.9	140.2	194.4	482.0	546.8	104.7	444.1	192.3
1980								
Total MYS	2.30	5.37	9.02	10.88	4.63	10.45	2.61	4.07
Primary	1.72	3.99	5.69	5.89	3.82	5.77	1.73	3.23
Secondary	0.54	1.20	3.01	4.35	0.78	4.31	0.79	0.76
Illiteracy rate	67.25	28.42	1.15	0.88	31.44	2.96	68.08	39.86
Population (in millions)	366.9	309.4	284.5	636.6	981.2	154.9	852.2	399.2
2010								
Total MYS	4.81	8.17	9.76	12.71	7.69	13.11	4.96	7.51
Primary	3.43	5.31	5.81	5.97	5.07	5.99	3.35	5.14
Secondary	1.26	2.48	3.50	5.41	2.43	5.91	1.38	2.10
Illiteracy rate	34.21	6.35	0.06	0.10	11.58	0.00	37.67	8.14
Population (in millions)	772.1	504.4	297.9	769.2	1339.0	175.8	1509.8	676.3

Furthermore, the share of Western Europe in the education top decile was equal to 60% in 1870, nearly the same amount as its share in the income top decile. If we consider that secondary schooling was the condition of access to technology, in 1910, Western Europe had in some respect the quasi-monopoly of knowledge and technology. Today this monopoly has disappeared. The share of Western Europe in the education top decile is only 40% in 2010 instead of 61% for income.⁷ In

⁷The share of Western Europe among world's top income is the same as in 1870, whereas the share in world's top education has sharply receded since 1910. So Western Europe has kept an important advantage in world income distribution in spite of losing its leadership on top education.

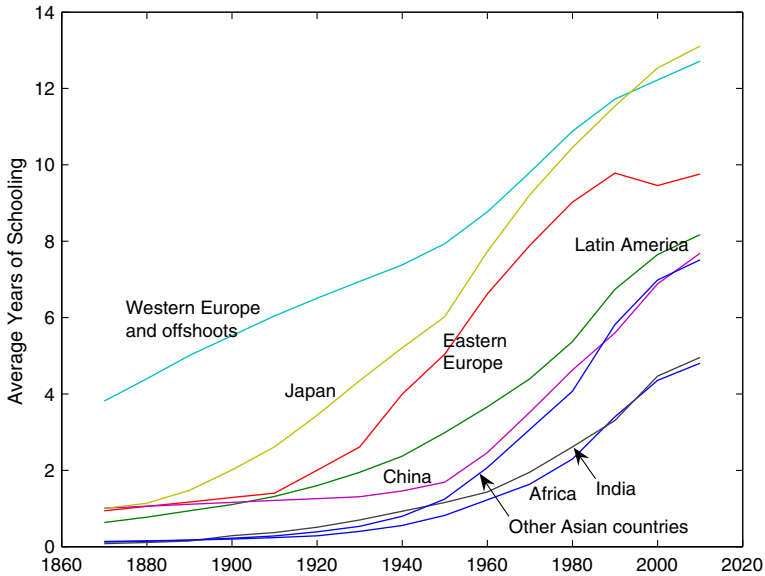


Fig. 1 Average years of schooling by region 1870–2010

comparison, the share of Asia (including China, Japan, Korea, South Asia) in the top decile has increased from 26% in 1910 to 41% in 2010.

4 Inequality in years of schooling

In this section we analyse global inequality in years of schooling as measured by the coefficient of variation, the Gini and the Theil indices,⁸ and we compare it with global income inequality. Table 4 shows an exceptionally high level of inequality in 1870 with a Gini coefficient reaching 0.82 and a Theil index of 1.56, a much higher inequality level than for the income distribution (0.56 for the Gini and 0.67 for the Theil). However, income and education inequality have displayed opposite trends since 1870. Inequality in years of schooling has decreased continuously and rapidly as the Gini coefficient has been divided by more than two and the Theil index amounts today to less than one fifth of its original level. This massive decline in inequality can largely be explained by the long-run decline in illiteracy (see below). On the contrary, income inequality has increased until 1990, reaching 0.65 (Gini) and 0.82 (Theil) at that date. After 1990, global inequality in schooling and global income inequality have both receded.

It is meaningful to draw a comparison between illiteracy and extreme poverty (less than 1 dollar a day). Between 1870 and 1990, the world illiteracy rate has decreased from 76% to 27% and extreme poverty from 75% to 24%. Therefore, the downward

⁸The mean logarithmic deviation was not reported since it is only defined over strictly positive outcomes.

Table 3 Regional distribution of world quantiles of education and income

	Africa		Latin America		Eastern Europe		Europe and offshoots		China		Japan-Korea		South Asia		Other Asian countries		
	S	Y	S	Y	S	Y	S	Y	S	Y	S	Y	S	Y	S	Y	
1870																	
Total	6.7		2.7		9.2		19.5		28.7		3.7		20.8		8.7		
Bottom 60	8.6	8.0	3.1	3.0	9.4	6.2	9.2	4.9	28.2	38.6	3.8	3.4	26.7	26.1	11.1	11.1	9.9
P60-P90	0.7	5.7	1.7	2.5	10.1	16.2	45.7	37.2	35.7	13.0	3.0	4.6	1.9	13.4	1.1	7.5	
Decile 10	0.7	1.5	1.5	1.8	7.8	6.7	57.9	60.0	26.3	13.2	3.3	2.4	1.5	9.8	1.0	4.7	
1910																	
Total	6.3		3.9		11.6		22.0		25.0		3.7		18.6		8.9		
Bottom 60	9.3	8.2	4.4	4.1	12.1	8.4	4.4	3.4	27.1	35.4	2.8	2.9	26.9	26.2	13.0	11.5	
P60-P90	1.0	4.0	3.4	4.0	13.0	19.1	44.4	46.7	27.0	8.5	6.1	5.9	3.5	6.8	1.7	5.0	
Decile 10	0.8	0.9	2.6	2.3	8.8	7.7	61.4	72.5	16.8	6.7	4.9	2.3	3.4	4.1	1.3	3.6	
1960																	
Total	9.3		6.2		9.9		19.4		22.3		4.3		18.6		10.1		
Bottom 60	13.4	13.3	6.3	4.9	6.0	2.4	6.7	0.8	26.6	34.4	1.9	1.6	26.2	29.5	12.8	13.1	
P60-P90	3.4	4.2	6.5	9.3	16.6	25.5	33.3	34.4	18.8	6.1	7.5	10.3	7.0	3.5	7.0	6.7	
Decile 10	2.2	0.9	4.5	4.1	13.5	5.9	53.6	84.3	7.3	0.0	9.4	2.5	6.5	0.0	2.9	2.3	
2010																	
Total	14.6		7.8		6.3		12.5		19.9		3.0		22.9		13.0		
Bottom 60	17.9	22.6	7.8	7.9	5.4	3.9	5.8	0.3	20.4	16.2	1.3	0.0	27.4	33.7	14.0	15.3	
P60-P90	7.9	3.1	8.1	9.0	8.8	12.2	24.3	20.7	20.5	30.3	6.1	5.1	12.9	8.3	11.4	11.4	
Decile 10	3.6	0.4	7.9	4.2	7.5	3.8	39.6	61.4	12.2	11.9	9.0	14.3	11.5	0.0	8.8	4.0	

Note: S stands for years of schooling, Y for income

Table 4 Global inequality in income and years of schooling

	1870	1890	1910	1930	1950	1970	1990	2010
World population—income								
Gini coefficient	0.560	0.589	0.612	0.618	0.643	0.652	0.653	0.619
Theil coefficient	0.672	0.749	0.802	0.781	0.812	0.815	0.816	0.722
Theil coefficient between countries	0.189	0.256	0.305	0.372	0.489	0.499	0.544	0.396
Average income	890	1111	1462	1811	2141	3767	5062	7519
World population—years of schooling								
Coefficient of variation	2.134	1.850	1.624	1.422	1.246	1.005	0.793	0.664
Gini coefficient	0.818	0.776	0.732	0.683	0.631	0.540	0.439	0.370
Theil coefficient	1.557	1.339	1.151	0.970	0.806	0.588	0.411	0.292
Theil coefficient between countries	0.613	0.597	0.526	0.433	0.335	0.205	0.113	0.067
Average years of schooling	1.24	1.62	2.08	2.60	3.31	4.68	6.29	7.47
Illiteracy rate	75.9	70.1	64.0	56.8	48.7	37.0	27.5	18.3
Educated population								
Coefficient of variation	0.582	0.569	0.557	0.554	0.555	0.516	0.424	0.422
Gini coefficient	0.244	0.253	0.257	0.267	0.281	0.270	0.225	0.229
Theil coefficient	0.135	0.133	0.131	0.132	0.138	0.126	0.089	0.090
Average years of schooling	5.13	5.43	5.77	6.02	6.45	7.42	8.68	9.14

trend of these two benchmark indicators, namely the percentages of people who do not have access to education or to a minimum income, is similar and illustrates an unprecedented improvement.

Table 4 describes the decomposition of the Theil index into its two sub-components, namely inequality within and between countries. Regarding schooling, we observe a rapid decrease in within and between-country inequality over the period 1870–2010. Concerning income, the between-country Theil index has largely increased between 1870 and 1990, but it has decreased sharply afterward. This suggests that the catch-up process between advanced and poor countries has started in 1870 for schooling but only in 1990 for income.

Similarly, the structure of global inequality has evolved differently for schooling and income. While the between-country contribution has decreased from 46% in 1910 to only 23% for schooling in 2010,⁹ it has increased continuously from 28% in 1870 to 66% in 1990 for income (Table 4). Likewise, the gap between the poorest region—Africa—and Western Europe is only 1–3 for average schooling instead of 1 to 12 for average income.

To explain the observed trend in schooling inequality, we look more closely at the share of individuals endowed with zero years of schooling.¹⁰ As reported in Table 4, if we exclude the illiteracy group and compute a Gini index on educated individuals only, we find a Gini equal to 0.24 in 1870, 0.28 in 1950, and 0.23 in 2010. So a sizeable share of inequality in years of schooling is explained by illiteracy at a given date, and time variations in schooling inequality reflect mainly illiteracy's decrease. Some studies (Castello and Domenech [7], or Berthelemy [3]) have already pointed at

⁹The latter figure is in agreement with the World Bank [29] estimate (less than 20%).

¹⁰Intuitively, if we remember that the Gini index is twice the area situated below the Lorenz curve, then illiteracy should have a huge impact on this index by shifting away the origin of the curve from zero to the percentage of illiterates in the population.

the negative correlation between years of schooling inequality and average years of schooling, offering various explanations. The following proposition shows that there is a mechanical link between illiteracy and years of schooling inequality (proof in companion Appendix).

Proposition 1 *Let f be the distribution of a random variable X taking values over a domain $[m, M]$ with $0 \leq m < +\infty$ and $M \leq +\infty$. Assume that this distribution can be decomposed as the mixture*

$$f(x) = p\delta_{x=m} + (1 - p)g(x) \quad (2)$$

where $\delta_{x=m}$ is a mass point in the minimum value and g the distribution of the population for which $X > m$, and $p < 1$. Let $\mu(f)$ be the mean outcome for a distribution f , $G(f)$ the corresponding Gini index, and $I_{GE}^\alpha(f)$ the Generalized-Entropy index. The Gini index can be decomposed as follows

$$G(f) = p \frac{\mu(f) - m}{\mu(f)} + (1 - p) \frac{\mu(f) - pm}{\mu(f)} G(g)$$

and similarly for the Generalized-entropy indices $I_{GE}^\alpha(f)$, with $\alpha \neq 0$ ¹¹

$$\begin{aligned} I_{GE}^\alpha(f) &= (1 - p)^{1-\alpha} \left(\frac{\mu(f) - pm}{\mu(f)} \right)^\alpha I_{GE}^\alpha(g) \\ &+ \frac{1}{\alpha^2 - \alpha} \left((1 - p)^{1-\alpha} \left(\frac{\mu(f) - pm}{\mu(f)} \right)^\alpha + pm^\alpha \mu(f)^{-\alpha} - 1 \right) \end{aligned}$$

An application to years of schooling follows immediately. Taking $m = 0$, the proposition shows that the Gini index computed over the whole population is a linear combination of the illiteracy rate and the Gini index computed over the educated population. Formally, one obtains respectively for the Gini and Theil indices (the Theil index decomposition is obtained when $\alpha \rightarrow 1$):

$$\begin{aligned} G(f) &= p + (1 - p)G(g) \\ Theil(f) &= Theil(g) - \ln(1 - p) \end{aligned}$$

The above proposition shows that the decline in illiteracy explains *almost all* of the decrease in years of schooling inequality over the period. Indeed, let us assume that inequality among the educated population remains constant over time and equal to 0.25, its grand average. According to the latter formula, an illiteracy level of 76% should set the Gini index calculated over the whole population at a value of 0.82, while an illiteracy level of 18% would bring it at 0.39. These figures match almost perfectly the actual values of the Gini index calculated over the whole population (0.82 in 1870 and 0.37 in 2010), showing that virtually *all* of the decrease in the Gini index between 1870 and 2010 is explained by illiteracy's decline. Consequently, the cross-country negative correlation between average schooling and schooling inequality depicted in the literature simply reflects the negative correlation between average schooling and illiteracy, which is mechanical.

¹¹The proposition is still valid for the Mean Logarithmic Index, i.e. when $\alpha = 0$, if $m > 0$.

Last and most importantly, inequality indices have been originally constructed for monetary outcomes, but years of schooling do not have such a monetary dimension. A crucial issue in the measurement of education inequality is therefore the search of an equivalence scale for years of schooling. Focusing on human capital is one solution to that problem, which is described in the next section.

5 Inequality in human capital

5.1 Defining human capital

The macroeconomic literature has gradually moved away from considering average years of schooling as a factor of production as in Mankiw et al. [15], to focus on the Mincerian definition of human capital as proposed by Bils and Klenow [4] and Hall and Jones [11]. For an educational group j in a country i at date t let us define human capital $h_{i,j,t}$ as:

$$h_{i,j,t} = e^{r_{i,j,t} S_{i,j,t}}$$

where $S_{i,j,t}$ is average years of schooling of group j and $r_{i,j,t}$ the return to schooling.

As a convenient starting point, we rule out heterogeneity in the return to schooling across time, countries and groups,¹² assuming $r_{i,t} = r = 0.1 \forall i, j, t$. The 10% value is the average world return to schooling arising from the study by Psacharopoulos and Patrinos [25].

As a second step, we argue that the return to schooling declines with the level of educational attainment. In other words, schooling has diminishing returns. As argued by Psacharopoulos and Patrinos [25], the returns to schooling may be higher for primary schooling than for secondary or higher education, although other studies do not fully share the same conclusion (e.g. Duflo [10]). Interestingly, there is evidence of diminishing returns to education in one country over a long period of time (the US between 1940 and 1980, see Murtin and Viarengo [22]). At the cross-country level, the negative correlation between Mincer returns to education and average years of schooling that is reported in Psacharopoulos and Patrinos [25] can be seen as a direct consequence of diminishing returns to schooling.¹³ Using the coefficients estimated from the cross-country regression of average Mincer returns on average years of schooling, we obtain the following benchmark definition of human capital for the educational group j in country i at time t :

$$h_{i,j,t} = e^{0.1254 S_{i,j,t} - 0.002 S_{i,j,t}^2} \quad (3)$$

from which average human capital in country i can be calculated by averaging over all educational groups.

¹²We also rule out any externality of education.

¹³As explained by Mincer [17] and Card [6], diminishing returns to schooling can be captured by a quadratic function of schooling in the regression $\log y_{i,j,t} = a + \rho S_{i,j,t} - \frac{1}{2} k S_{i,j,t}^2 + u_{i,t}$, where y is income and (i, j, t) denote respectively country, educational group and time. Once the above equation is derived with respect to schooling, one obtains the following return to schooling $r_{i,j,t} = \rho - k S_{i,j,t}$, which leads to a negative linear relationship between the average Mincer return to schooling and average years of schooling at the cross-country level.

However, recent studies (e.g. Colclough et al. [9]) have contradicted the above evidence as they have found increasing rather than decreasing returns to education in the recent past. The choice of a particular functional form for the returns to education is important because the rest of the paper relies on it. So we have examined another definition of human capital displaying increasing returns to education:

$$h_{i,j,t} = e^{0.0088 \cdot S_{i,j,t} + 0.0055 \cdot S_{i,j,t}^2} \quad (4)$$

This functional form implies returns to schooling that are close from those described by Colclough et al. [9]. Indeed, the return to primary (respectively secondary and tertiary) schooling is equal to 7.5% (resp. 14.1% and 18.5%), while the latter authors mention 7.4% (resp. 11.0% and 19.8%). With the two latter functional forms, increasing and decreasing returns to education are both equal to 9.5% when average years of education amounts to 7.8 years.

5.2 Results

Table 5 provides estimates of human capital inequality for these three specifications ($r = 10\%$, Mincerian diminishing then increasing returns). The contrast between schooling and human capital inequality is striking, as their trends appear to be opposite until the second half of the 20th century. Indeed, whereas schooling inequality has always decreased, the Gini index of human capital inequality has increased by respectively 0.14 points ($r = 0.10$), 0.08 points (diminishing returns) and 0.30 points (increasing returns) between 1870 and 1970. After 1970, we find that human capital inequality has fallen over time in the simulations with constant and decreasing returns. Inequality keeps on increasing until 1990 in the simulation with increasing returns, then stabilizes (Gini index) or decreases (all other inequality indices). It is striking that increasing returns to education generate much more human capital inequality.

Overall, global human capital inequality has followed an inverted U-shape curve that has peaked somewhere in the second half of the 20th century, then has started falling. This closely mirrors the result brought forward by Morrisson and Murtin [20], who estimate that income inequality between countries has started decreasing in the 1990s and may keep on falling in the forthcoming years due to conditional convergence effects.

The initial increase in human capital inequality cannot be reconciled so easily with declining schooling inequality and constant or decreasing returns to education. To explain this apparent paradox, let us assume that schooling S has a normal distribution with mean m and coefficient of variation s . Laplace transformation of a normal variable simply provides the coefficient of variation of human capital $s(h)$ and a first-order approximation yields

$$s(h) = \frac{\text{Var}(e^{rS})^{1/2}}{E(e^{rS})} = \sqrt{e^{r^2 m^2 s^2} - 1} \simeq rms \quad (5)$$

where r stands for the return to schooling. Now it becomes clear that human capital inequality depends positively on inequality in years of schooling (s), positively on the return to education (r), but also positively on the average level of schooling (m): Due to the convexity of the exponential function, a given level of schooling inequality (as measured by the coefficient of variation) is trans-

Table 5 Global inequality in human capital

	1870	1890	1910	1930	1950	1970	1990	2010
World population								
Constant return equal to 10%								
Coefficient of variation	0.39	0.43	0.47	0.50	0.53	0.54	0.52	0.49
Gini coefficient	0.131	0.159	0.188	0.212	0.241	0.272	0.273	0.264
Theil coefficient	0.057	0.070	0.085	0.097	0.113	0.129	0.124	0.115
Theil coefficient between countries	0.019	0.029	0.040	0.046	0.049	0.052	0.044	0.034
Average human capital	1.18	1.25	1.32	1.41	1.54	1.80	2.13	2.38
Human capital with diminishing returns								
Coefficient of variation	0.30	0.32	0.34	0.35	0.36	0.36	0.33	0.30
Gini coefficient	0.116	0.138	0.157	0.173	0.188	0.197	0.184	0.164
Theil coefficient	0.037	0.044	0.051	0.056	0.061	0.064	0.058	0.047
Theil coefficient between countries	0.014	0.021	0.026	0.029	0.030	0.027	0.020	0.013
Average human capital	1.17	1.22	1.28	1.35	1.44	1.61	1.81	1.95
Human capital with increasing returns								
Coefficient of variation	0.85	0.92	1.09	1.19	1.31	1.33	1.25	1.18
Gini coefficient	0.168	0.212	0.269	0.321	0.390	0.469	0.501	0.509
Theil coefficient	0.158	0.197	0.268	0.327	0.414	0.493	0.504	0.493
Theil coefficient between countries	0.035	0.056	0.089	0.110	0.122	0.150	0.159	0.136
Average human capital	1.22	1.31	1.44	1.59	1.87	2.50	3.37	4.20
Educated population								
Constant return equal to 10%								
Coefficient of variation	0.39	0.39	0.40	0.41	0.42	0.42	0.39	0.39
Gini coefficient	0.154	0.164	0.174	0.185	0.202	0.214	0.202	0.209
Theil coefficient	0.062	0.064	0.067	0.071	0.078	0.082	0.072	0.075
Theil coefficient between countries	0.007	0.009	0.014	0.018	0.018	0.023	0.019	0.018
Average human capital	1.76	1.82	1.89	1.95	2.05	2.27	2.56	2.69
Human capital with diminishing returns								
Coefficient of variation	0.22	0.22	0.22	0.22	0.23	0.22	0.19	0.19
Gini coefficient	0.094	0.100	0.104	0.110	0.119	0.118	0.099	0.101
Theil coefficient	0.022	0.022	0.022	0.023	0.025	0.025	0.018	0.018
Theil coefficient between countries	0.003	0.004	0.006	0.008	0.008	0.008	0.005	0.004
Average human capital	1.69	1.72	1.77	1.80	1.85	1.97	2.12	2.16
Human capital with increasing returns								
Coefficient of variation	1.01	1.00	1.10	1.14	1.18	1.16	1.09	1.06
Gini coefficient	0.325	0.339	0.370	0.393	0.430	0.460	0.463	0.474
Theil coefficient	0.292	0.295	0.341	0.368	0.413	0.435	0.420	0.425
Theil coefficient between countries	0.025	0.032	0.053	0.068	0.062	0.090	0.100	0.097
Average human capital	1.92	2.03	2.22	2.37	2.70	3.39	4.26	4.89

formed into larger human capital inequality as the average level of schooling is rising.¹⁴ Initially, this convexity effect has overcome the equalitarian effect induced by a more equal distribution of years of schooling. On a second step,

¹⁴For instance, with $r = 0.1$, a $s = 10\%$ variation in schooling represents a $s(h) = 0.1 \times 0.1 = 1\%$ variation in human capital when $m = 1$, and a $0.1 \times 0.1 \times 10 = 10\%$ variation in human capital when $m = 10$.

the equalitarian effect has dominated and inequality in human capital has started decreasing.¹⁵

Some limitations of the above framework deserve a quick discussion. First, we have ruled out country-specific returns.¹⁶ Indeed, the goal of the paper is to propose a general framework where education is rescaled along a monetary dimension over 140 years. Second, empirical studies generally use Mincer regressions in the context of wage-earner income. However, wage-earners do not necessarily constitute the bulk of the active population in developing countries. In a sense, it is not certain that our simplistic Mincerian framework would be appropriate to depict the labour market in rural areas. As underlined by Banerjee and Duflo [2], many market failures affect very poor populations, so that marginal productivity and wages do not necessarily match over such imperfect markets (such as those having prevailed among socialist countries). From that perspective, our scale of equivalence in years of schooling does not reflect *actual* human capital (income), but rather *potential* human capital in a counter-factual, well-functioning, labour market.

So far, the findings of the paper are as follows: (i) inequality in years of schooling has declined dramatically because of illiteracy's decline; and (ii) human capital global inequality has increased during approximatively one century, namely between 1870 and 1970–1990, then has started decreasing. In the next section, we analyse the trends of human capital inequality *within countries*.

6 The Kuznets curve of human capital

Figure 2 displays the regional average of human capital inequality within countries, using the eight geographical areas described earlier and constant returns to education.¹⁷ Several facts emerge. First, Western Europe and offshoots are the only region in the world where human capital inequality within countries has been continuously falling since 1870. In all other regions, inequality has increased sharply at least until the mid-20th century. Second, in all regions of the world, human capital within-country inequality has stagnated or has started decreasing in the second half of the 20th century. The timing of that reversal in trends varies across regions. In Japan and Korea, it took place around 1950, around 1970 in Eastern Europe and after 1980 in Latin America, Africa and Other Asian Countries. Human capital inequality has stagnated within China, India, Bangladesh and Pakistan since 1980, and one can expect such a decrease to take place in a close future.

These patterns deeply echo with the well-known Kuznets [12] hypothesis of an inverted U-shaped curve for income inequality within countries. Kuznets argued that, among other causes, sectoral shifts of workers would entail rising income inequality in first place, but a declining trend afterward. This hypothesis has received

¹⁵Decreasing returns to education reinforce the equalitarian effect while increasing returns to education diminish it.

¹⁶Actually, we have tested the effect of introducing country-specific, autocorrelated, shocks on the return to schooling, while running a bootstrap experiment. All results were qualitatively unchanged.

¹⁷Although national distributions of education are composed of only 4 groups, the resulting level of national human capital inequality is robust to mis-measurement in educational attainment as shown in the companion Appendix.

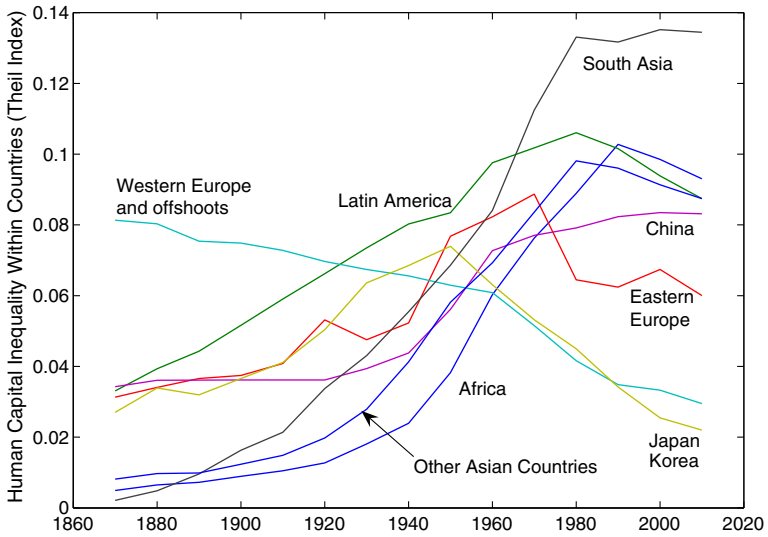


Fig. 2 Inequality in human capital within countries by geographical zone 1870–2010—Theil index and constant returns to education

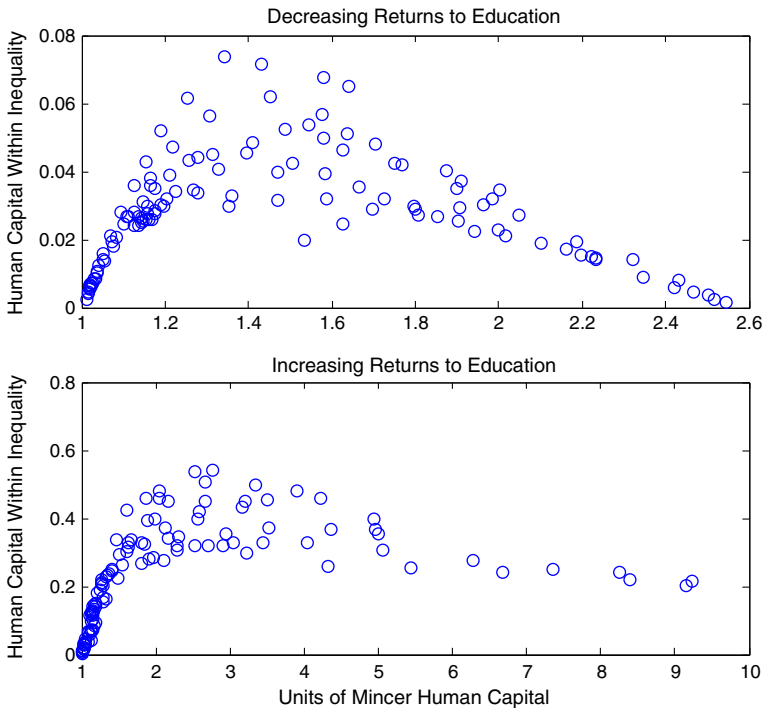


Fig. 3 The Kuznets curve of human capital 1870–2010—Theil index

much attention in the empirical literature and is discussed by Lindert [13] and Morrisson [18] from a historical perspective. In our setting, a similar hypothesis would state that human capital inequality within countries has followed an inverted U-shape curve along educational development.

Figure 3 plots human capital within-country inequality with respect to average human capital for the 32 macro-countries over the period 1870–2010, while selecting decreasing or increasing returns to education, years of schooling. With decreasing returns to education, we find strong evidence of a “Kuznets curve of human capital inequality” over the period 1870–2010. Actually, human capital inequality culminates when countries reach a level of about 3–4 average years of schooling, which roughly corresponds to half of the population being illiterate, 40% receiving primary schooling and 10% secondary schooling (e.g. the world around 1950). It is fairly intuitive that human capital inequality is maximal when the transition from illiteracy to literacy is exactly at mid-course. This is actually what we observe.

With increasing returns to education, the inverted U-shape pattern is still there, but it is logically less pronounced as larger returns to higher education generate more inequality in advanced economies situated on the right tail of the curve, which is rather flat than declining above 5 units of human capital. Still, we do observe a higher inequality level (about 0.4 points of Theil) in the range of 2–3 units of human capital compared to that observed beyond 3 units (about 0.25 points of Theil). This suggests that the Kuznets curve of human capital inequality is robust to the existence of increasing returns to education, although the associated effect has a much lower magnitude.

7 Conclusion

This paper presents the first estimates of the world distribution of years of schooling and of human capital over the last 140 years. An original database on average years of schooling and the distributions of schooling has been built for that purpose, extending past work by Morrisson and Murtin [19]. We have shown that the educational comparative advantage of Western Europe has decreased rapidly since the beginning of the twentieth century. As a consequence the contexts of the two globalization processes, the first in 1860–1914, the second starting in the late 1970s, are very different. In world economic competition, education is a crucial advantage at least because it enables access to technological progress. Today, there is a discrepancy between the lead of Western Europe in the world income distribution and its weight in world education, whereas Asia has increased its educational share substantially over the 20th century.

Furthermore, we have shown theoretically that inequality in years of schooling is a linear function of the illiteracy rate, whose time variation largely dominates the time variation of schooling inequality among educated people. To solve that problem, we have introduced human capital inequality, while using alternatively constant, decreasing or increasing returns to education. As a result, we find that world human capital inequality has increased, peaked in the second half of the 20th century, then started decreasing.

One of the major empirical finding of this paper takes place at the country level. We have exhibited an inverted-U shape curve of human capital inequality

within countries along the process of educational development, namely “a Kuznets curve of human capital inequality”. It happens that human capital within inequality is maximal when the share of illiterate population is close to 50% of national population, an observation that fits well with Kuznets’ original motivation of his hypothesis. However, the potential existence of increasing returns to education may mitigate the equalitarian effect associated with the homogeneization of schooling attainment within countries.

Acknowledgements We thank François Bourguignon for fruitful discussions and significant contributions to this paper. We would like to acknowledge Daniel Cohen and Marcelo Soto for their data and insightful comments, as well as Ximena del Carpio and Francisco Ferreira for their help with World Bank data at an earlier stage of the paper. We are grateful for discussions with Anthony Atkinson, Jean-Claude Berthélemy, Marc Gurgand, Stephan Klasen, Francis Kramarz, Steve Pischke, John Van Reenen, as well as seminar participants at Stanford University, UCLA, EUDN Paris conference, CREST, CSAE Oxford conference, LSE, and PSE. Murtin acknowledges financial support from the Mellon Foundation when he was hosted by the Stanford Centre for the Study of Poverty and Inequality, as well as EU Marie Curie RTN when he was hosted by the Centre for the Economics of Education, London School of Economics. This paper expresses the opinions of the authors and does not necessarily represent the official views of the OECD.

References

1. Aghion, P., Howitt, P.: *The Economics of Growth*. The MIT Press, Cambridge (2009)
2. Banerjee, A., Duflo, E.: The economic lives of the poor. *J. Econ. Perspect.* **21**(1), 141–167 (2007)
3. Berthélemy, J.-C.: To what extent are African education policies pro-poor? *J. Afr. Econ.* **15**(3), 434–469 (2006)
4. Bills, M., Klenow, P.J.: Does schooling cause growth? *Am. Econ. Rev.* **90**(5), 1160–1183 (2000)
5. Bourguignon, F., Morrisson, C.: Inequality among world citizens: 1820–1992. *Am. Econ. Rev.* **92**(4), 727–744 (2002)
6. Card, D.: Estimating the return to schooling: progress on some persistent econometric problems. *Econometrica* **69**(5), 1127–1160 (2001)
7. Castello, A., Domenech, R.: Human capital inequality and economic growth: some new evidence. *Econ. J.* **112**, C187–C200 (2002)
8. Cohen, D., Soto, M.: Growth and human capital: good data, good results. *J. Econ. Growth* **XII**, 51–76 (2007)
9. Colclough, C., Kingdon, G., Patrinos, H.: The changing pattern of wage returns to education and its implications. *Dev. Policy Rev.* **28**(6), 733–747 (2010)
10. Duflo, E.: Schooling and labor market consequences of school construction in Indonesia: evidence from an unusual policy experiment. *Am. Econ. Rev.* **91**(4), 795–813 (2001)
11. Hall, R.E., Jones, C.I.: Why do some countries produce so much more output per worker than others? *Q. J. Econ.* **114**(1), 83–116 (1999)
12. Kuznets, S.: Economic growth and income inequality. *Am. Econ. Rev.* **65**(4), 1–28 (1955)
13. Lindert, P.H.: Three centuries of inequality in Britain and America. In: Atkinson, A.B., Bourguignon, F. (eds.) *Handbook of Income Distribution* (2000)
14. Maddison, A.: *The World Economy. Historical Statistics*. OECD Development Centre, Paris (2008)
15. Mankiw, N.G., Romer, D., Weil, D.N.: A contribution to the empirics of economic growth. *Q. J. Econ.* **107**(2), 407–437 (1992)
16. Milanovic, B.: True world income distribution, 1988 and 1993: first calculation based on household surveys alone. *Econ. J.* **112**(476), 51–92 (2002)
17. Mincer, J.: *Schooling, Experience and Earnings*. NBER, New York (1974)
18. Morrisson, C.: Historical perspectives on income distribution: the case of Europe. In: Atkinson, A.B., Bourguignon, F. (eds.) *Handbook of Income Distribution* (2000)
19. Morrisson, C., Murtin, F.: The century of education. *J.H.C.* **3**(1), 1–42 (2009)
20. Morrisson, C., Murtin, F.: *Internal Income Inequality and Global Inequality 1700–2030*. Working paper, p. 26. FERDI (2011)

21. Murtin, F., Viarengo, M.: American education in the age of mass migrations. *Clometrica* **4**(2), 113–139 (2010a)
22. Murtin, F., Viarengo, M.: The expansion and convergence of compulsory schooling in Western Europe, 1950–2000, *Economica* **78**(311), 501–522 (2011)
23. Murtin, F.: Long-term determinants of the demographic transition: 1870–2000. *Rev. Econ. Stat.* (2012, forthcoming)
24. Murtin, F., Wacziarg, R.: The Democratic Transition. NBER wp. 17432 (2011)
25. Psacharopoulos, G., Patrinos, H.A.: Returns to investment in education: a further update. World Bank working paper no. 2881 (2002)
26. Sala-i-Martin, X.: The world distribution of income: falling poverty and convergence, period. *Q. J. Econ.* **121**(2), 351–397 (2006)
27. Thomas, V., Wang, Y., Fan, X.: Measuring education inequality: Gini coefficients of education, 1960–2000. World Bank research paper no. 2525 (2001)
28. UNESCO: World Illiteracy at Mid-Century. UNESCO, Paris (1957)
29. World Bank: Equity and Development. World Development Report (2005)