Educational inequality in Europe, 1870-2000

Peter Foldvari
University of Debrecen and Utrecht University

Bas van Leeuwen
Utrecht University and Free University Amsterdam

Jieli van Leeuwen-Li
Independent researcher

Abstract

In this paper, we develop an annual dataset of average years of education and educational inequality that is comparable accross countries and over time. We find that within country inequality has the expected inverse U-curve before 1950, however, it turns out to be a normal U-curve afterwards. After the effect of changing trends in skill premium though, we find in both periods an inverse U-curve.

As for the effect of education on mean per capita incomes we find a significant, positive coefficient only after 1950 but even then there is no effect in socialist countries. If we include within country educational inequality, we find a positive effect of educational attainment on per capita income in the 19th century.

JEL code: J24, O47, N10
Keywords: average years of education, economic development, inequality, Kuznets-curve

1. Introduction

The relationship between economic development, income inequality and education is one of the main focuses of several branches of modern economies. Inequality of education may affect the accumulation of production factors and the society's capability to adopt new technologies, which is a main concern for development economics (see Barro 1991,Benhabib and Spiegel 1994, Krueger and Lindahl 2001, Pritchett 2001, Cohen and Soto 2007, Zeira 2009), while inequality may also affect the individuals' ability to participate and control decision-making, thereby affecting stability and the degree of rent-seeking, a primary concern for political economy (Perotti 1996, Alesina and Perotti 1996). Recently inequality of education and human capital has also been given increased attention due to the significant expansion in historical data (Morrisson and Murtin 2007, van Zanden *et al.* 2010). Since Knight and Sabot's (1983) theoretical proposition it has been expected that educational inequality may also exhibit a non-linear relationship with development similarly to the celebrated hypothesis of Kuznets's inverted U-curve (Kuznets 1955).[1]

In this paper we use new data on income and educational inequality to address the relation between income inequality and educational inequality in Europe ca. 1820-2008. The paper has the following structure. In section 2 we present the data; this is followed by a discussion of the decomposition of the European income and educational Ginis in a within, between, and an overlap component. Sections 4 and 5 than discuss the relationship between educational and income inequalities and the effect of mean educational attainment with GDP per capita respectively. Our main findings are summarized in Section 6.

2. Data

2.1 Average years of education

The main underlying data of this paper is a dataset on average years of education and educational inequality. The data on population by age class is obtained from Mitchell (2003), and for Iceland from Jonsson and Magnusson (1997). Further, we need data for educational enrolments by age class. For the recent years we start with data from the UNESCO statistical yearbook (various issues). This was brought back in time using Mitchell (2003) and, in case of Iceland Jonsson and Magnusson (1997) and in case of Switzerland Ritzmann-Blickenstorfer (1996).

[1] While several theoretical models were developed (see Banerjee and Newman 1993, Perotti 1993, Galor and Tsiddon 1996) that produce the inverted U-curve, the empirical results are inconclusive. Bruno *et al* (1996), Deininger and Square (1998), Matyas *et al* (1998) and Field (2001) for example cannot find empirical support for it.

For the period after 1960 we follow the method as proposed by Foldvari and Van Leeuwen (2009).They start with the dataset from Cohen and Soto (2007) who provide decadal benchmarks of average years of education and attainment per education level between 1960 and 2010. We use the Barro and Lee (1993; 2001) perpetual inventory method to predict this back and forward. Depending on the data available, this has as a disadvantage that they do not account for differential mortality among people with lower or higher education and dropouts. Hence, when predicting attainment forward, one will underestimate average years of education, while estimation backwards will produce an overestimate. If we assume that in a period of a decade the percentage dropouts and the differential mortality do not change significantly, we can remove the bias by estimating the same year using both forward and backward method and taking the average.

Hence, Foldvari and Van Leeuwen (2009) start with the forward looking perpetual inventory method from Barro and Lee (1993; 2001), but make it more general to allow for different year lengths:

$$h_{0,t} = H_{0,t}/L_t = h_{0,t-i}[1-(L15_t * i/5 * L_t)] + (L15_t * i/5 * L_t) * (1-PRI_{t-i})$$

$$h_{1,t} = H_{1,t}/L_t = h_{1,t-i}[1-(L15_t * i/5 * L_t)] + (L15_t * i/5 * L_t) * (PRI_{t-i} - SEC_t)$$

$$h_{2,t} = H_{2,t}/L_t = h_{2,t-i}[1-(L15_t * i/5 * L_t)] + (L15_t * i/5 * L_t) * SEC_t - (L20_t * i/5 * L_t) * HIGH_t$$

$$h_{3,t} = H_{3,t}/L_t = h_{3,t-i}[1-(L15_t * i/5 * L_t)] + (L20_t * i/5 * L_t)$$

, where h is attainment per level of education (0= no education, 1=prim, 2=sec, 3=high). H is the total number of persons in the population of 15 years and older per level of education, i is the number of years of forward estimation, L is the total population aged 15 years and older, L15 and L20 are the age classes 15-19 and 20-24 year, and PRI, SEC, HIGH are the enrolment ratios in primary, secondary, and higher education.

Rewriting above equations result in a backward estimates:

$$h_{0,t-i} = \left(h_{0,t} - (L15_t * i/5 * L_t) * (1-PRI_{t-i})\right)/(1 - L15_t * i/5 * L_t)$$

$$h_{1,t-i} = \left(h_{1,t} - (L15_t * i/5 * L_t) * (PRI_{t-i} - SEC_t)\right)/(1 - L15_t * i/5 * L_t)$$

$$h_{2,t-i} = \left(h_{2,t} - (L15_t * i/5 * L_t) * SEC_t + (L20_t * i/5 * L_t) * HIGH_t\right)/(1 - L15_t * i/5 * L_t)$$

$$h_{3,t-i} = \left(h_{3,t} - (L20_t * i/5 * L_t) * HIGH_t\right)/(1 - L15_t * i/5 * L_t)$$

Under the assumption that relative mortality and dropouts remain constant in a decade, taking the average of the prediction form both methods results in unbiased estimates of annual average years of education.

For the period prior to 1960 we use a different method. We start by calculating an enrolment ratio by diving enrolment per level of education by the share of the population eligible to follow that level of education. This results in the percentage people in an age class that followed either none, primary, secondary, or higher education. For each year t, this percentage is multiplied by the number of persons in that age class. That is, in order to calculate total attainment in 1950, we multiply age class of people 20 years old with enrolment in primary education 15 years ago, of people 25 years old with enrolment ratio in primary education 20 years ago, etc. The same principle applies to secondary and higher education. This results in the total attainment per age class. Summing the attainment per level of education of the age classes between age 15 and 65, multiplying this with the years of education per education level, and diving this result by the total population aged 15-65 results in average years of education for people aged 15-65.

Two further remarks have to be made. First, this method assumes that the mortality of those with education is equal to those without education. Yet, it allows for different mortality rates among cohorts. This assumption is often seen as acceptable (Cohen and Soto 2007, 55-56). Indeed, as long as the relative mortality of the schooled versus the unschooled does not change, it has no effect on the trend of average years of education. Since we only calculate this variable in the period before ca. 1960, we can make this assumption. Second, in order to calculate average years of education for people aged 15-65, one needs enrolment 50 years back. In order not to miss too many observations at the beginning of the series, we calculated 4 series of average years of education. One for age 15-65, and for age 15-60, 15-55, 15-50 and from age 15-45. The trend in either of these series is completely comparable. Hence, starting with the series covering the population 15-65, we linked the other series with a smaller coverage when going further back in time.

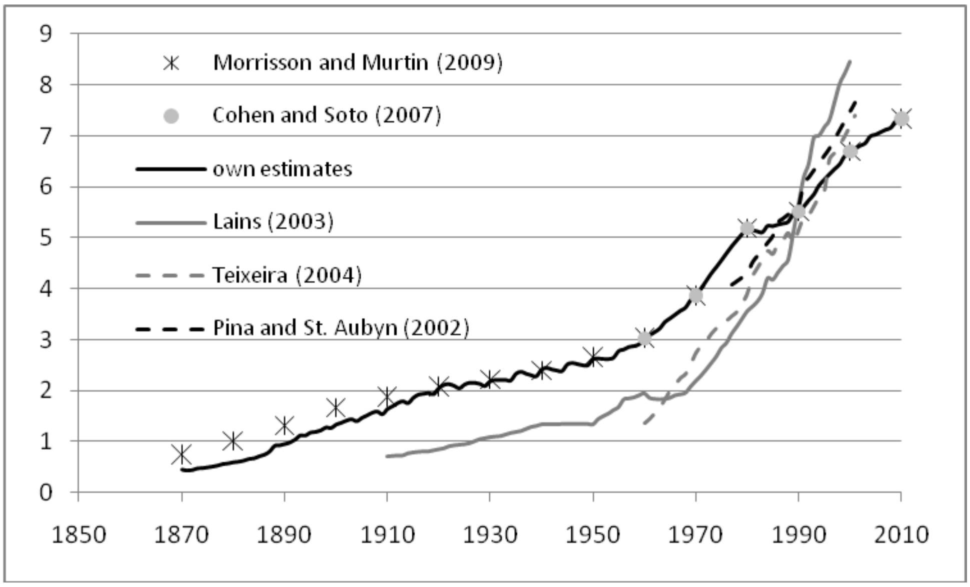
Both methods described above were used for all countries, except for Iceland where we only had Barro and Lee (2001) as benchmark and used method 2 further. Data on Hungary, Poland, Czechoslovakia, Germany, Austria, Romania, Bulgaria were taken from Foldvari and Van Leeuwen (2009b) and, where necessary extended using method 2.

Our dataset has three big advantages over most existing series. First, it includes historical data stretching back to the 1870s. Second, our data are comparable over time and across countries. Third, virtually all other datasets that cover multiply countries present data every 10th or every 5th year whereas our data are available on an annual basis.

Indeed, that comparability is a major issue may become clear when looking at, for example, the Portuguese estimates (see Figure 1). As Figure 1 shows, there are major differences in the estimates of average years of education in Portugal. Whereas Cohen and Soto (2007), Morrisson and Murtin (2009), and our data are based on the population aged 15+, Lains (2003), Texeira (2004), and Pina and St. Aubyn (2002) use the population aged 25+. Clearly, just this difference is coverage is not enough to explain the major difference between the series, hence, part of the explanation may be in the estimation method

Figure 1

Average years of education in Portugal



as well. Indeed, Teixeira (2004) uses a combination of attainment, PIM,and a Kyriacou (1991) regression to project his data back to 1960. The same method, where the regressions were replaced by simple interpolations, is used by Pina and St. Aubyn (2002). However, Cohen and Soto (2007), on which our method and those of Morrisson and Murtin (2007) are partially based, use census data to project average years of education back and forward. This method is generally believed to be the more reliable. Indeed, even though it is well know that Portuguese education expanded massively in the last half of the twentieth century, a quadrupling of average years of education as found in the earlier estimates seems large since this boils down to a compound annual growth rate of 3.5% per annum. Since every year about 2.5% of new persons are added to the human capital stock, this implies that the average level of education of the new entrants must be about 3.4 times higher than the average of the persons already in the stock in order to achieve this growth rate.

A similar problem emerges in Sweden where Ljungberg and Nilsson (2009) estimated average years of education for Sweden between 1870 and 2000. However, since they corrected for school duration (length of school years), they end up with a far steeper growth in average years of education than either Morrisson and Murtin (2007) or this paper. The most comparable dataset is those of Prados de la Escosura and Rosés (2010) from Spain. They use the data largely from Nunez (2005) and Cohen and Soto (2007). Even though they use data in the population aged 15-65 while mostly series refer to the population aged 15+, their data matches ours closely.

A second important point is that our data are comparable across a broad series of European countries. The same applies for Morrisson and Murtin (2009) who report data for every 10th year between 1870 and 2010. Their data are therefore comparable to ours, even though we provide some additional countries (but ultimately has a more limited geographical coverage) and provide annual series. A second broadly comparable dataset is the new dataset of Barro and Lee (2010) which covers every 5th year between 1950 and 2000. Whereas their previous (1993; 2001) estimates were heavily criticized (i.e De la Fuente and Doménech, 2000; Krueger and Lindahl, 2001; Portela, Alessie and Teulings, 2004; Cohen and Soto, 2007), their new estimates use broadly a similar census-based methodology as Cohen and Soto. It is therefore important to compare our estimates to these two datasets.

For the comparison we rely on the method suggested by Krueger and Lindahl (2001, 1115) to acquire a reliability indicator for estimated average years of education. First they assume that there are two alternative estimates (S_1 and S_2) of the real but unobserved series (S). Both estimates have measurement errors (e_1 and e_2) which are assumed to be independent of each other and have zero mean. The OLS estimator for the coefficient if S_1 in a two variable regression between the two estimates, $Cov(S_1,S_2)/Var(S_1)$ has the probability limit $Var(S)/(Var(S)+e_1)$, which can be seen as a indicator of reliability. The higher the coefficient is, the lower the share of measurement error is relative to the signal. Of course, this exercise should not be carried out on non-stationary variables (hence we use the ten or sometime five year differences) and also the uncorrelatedness of the two measurement errors are often an unrealistic assumption, especially since the compared works rely on Mitchell's historical statistics for enrolment numbers. Nevertheless, they sometimes arrive at very different results. If the measurement errors are positively correlated, we can expect the reliability indicator biased upward, for both series.

We used a pooled OLS and also a fixed-effect estimator to estimate the reliability of the two series. The results suggest that for the whole sample, our reliability measure is higher

Table 1

Reliability ratio with the Morrisson and Murtin (2009) data, 1870-2010

	dependent variable			
	ten year difference of av. years of education this text (pooled OLS)	ten year difference of av. years of education Morrison and Murtin (pooled OLS)	ten year difference of av. years of education this text (FE)	ten year difference of av. years of education Morrison and Murtin (FE)
Constant	0.195 (6.06)	0.103 (3.76)	0.232 (5.22)	0.083 (1.91)
ten year difference of av. years of education Morrisson and Murtin	0.680 (13.2)	-	0.614 (7.98)	-
ten year difference of av. years of education this text	-	0.809 (14.9)	-	0.842 (11.4)
R^2	0.55	0.55	0.52	0.52
N	291	291	291	291

Note: robust t-statistics are reported in parentheses

Table 2

Reliability comparison with the Barro-Lee (2010) data, 1960-2010

	every tenth years		every fifth year	
	ten year difference of av. years of education this text (pooled OLS)	ten years difference of av. years of education Barro and Lee (pooled OLS)	five years difference of av. years of education this text (pooled OLS)	five years difference of av. years of education Barro and Lee (pooled OLS)
Constant	0.686 (16.0)	0.570 (4.26)	0.364 (23.8)	0.337 (5.69)
ten year difference of av. years of education Barro and Lee	0.089 (2.04)	-	-	-
ten year difference of av. years of education this text	-	0.331 (2.00)	-	-
five year difference of av. years of education Barro and Lee	-	-	0.048 (1.67)	-
five year difference of av. years of education this text	-	-	-	0.225 (1.70)
R^2	0.03	0.03	0.01	0.01
N	150	150	300	300

Note: robust t-statistics are reported in parentheses than that of Morrisson and Murtin (2009). This also applies to the fixed effects specification.

The second dataset to compare our estimates with is that of Barro and Lee (2010). Their dataset has been significantly updated. As in their previous studies (1993; 2001) they estimate average years for every 5th year, but now they exclusively follow the lead by Cohen and Soto and use attainment by age from the censuses. They claim that their data outperforms those of Cohen and Soto for they use better comparable census data for the OECD. Since our data after 1960 is largely based on Cohen and Soto, it is interesting to test if our data outperform those of Barro and Lee (2010). Second, Barro and Lee use a different method to interpolate the inbetween census years. It is therefore also interesting to test our data every 5th year has less measurement error than the B&L estimates. The results of both tests are given in Table 2. We find that for every tenth year our data (which is effectively the Cohen and Soto (2007) data) clearly outperform the Barro and Lee data. Clearly the reliability ratios for our data are lower than in Table 1 where we compared with Morrisson and Murtin (2009), but this may be caused by the low signal in the Barro and Lee data which downwards biases the reliability ratio of both variables. If we look at every fifth year, we find again that our data has a higher reliability ratio than those of Barro and Lee. Hence, our interpolation technique between the census data seems plausible.

2.2 Educational Gini

Since the data on average years of education as discussed in Section 2.1 also allows us to construct average years of education and attainment per education level, it is now possible to calculate a education Gini per country per year using the method as suggested by Thomas, Wang, and Fan (2000), Checchi (2004) and Castelló and Doménech (2000, 4) to convert these data into educational Ginis. They start with

$$G^h = \frac{1}{2\overline{H}} \sum_{i=0}^{3} \sum_{j=0}^{3} \left| \hat{x}_i - \hat{x}_j \right| n_i n_j$$

Where $\overline{H}$ is average years of schooling in the population aged 15 years and over, i and j are different levels of education, n_i and n_j are the attainment per level of education, and $\hat{x}_i$ and $\hat{x}_j$ are the cumulative average years of schooling at each educational level. This equation can be rewritten as follows:

$$G^h = n_0 \frac{n_1 x_2 (n_2 + n_3) + n_3 x_3 (n_1 + n_2)}{n_1 x_1 + n_2 (x_1 + x_2) + n_3 (x_1 + x_2 + x_3)}$$

Where x stands for the average years of schooling per level of education (0= no education, 1= primary education, 2 = secondary education, and 3 is higher education) divided by the percentage population with at least that level of education attained. n_0, n_1, n_2, and n_3 are the percentages of the population with no-, primary-, secondary-, and higher education respectively.

Using this formula it is possible to calculate for each country in our sample the educational Gini per annum.

2.3 Other data

Besides average years of education and educational inequality, we also require data on income inequality, per capita GDP, and polity variables. Income inequality was taken from a recent study from Van Zanden *et al.* (2010). They estimate income inequality for most countries in the world for benchmark years between 1820 and 2000. Per capita GDP is taken from Maddison (2007). Finally, we use polity related variables from the GUR Polity IV dataset (Marshall and Jaggers 2009).

3. Decomposing income and educational inequality

The use of Gini coefficients for income- and educational inequality is convenient, but there is a price to pay, namely, unlike sum of squares or variance it cannot be simply decomposed into a sum of a within-group and a between-group Gini. Lambert and Aronson (1993) suggest a decomposition according to the following formula:

$$G = G_B + \sum_{k=1}^{n} a_k G_k + R$$

, where G is the total Gini, G_B ius the between group Gini (caused by differences in country averages), G_k is the Gini coefficient of country k, and a_k is the product of the income share and population share of country k. R is the overlap factor, which captures the effect that the income/educational ranges do not overlap. For example, an individual with just 4 years of education may be in the highest quintile of the population in a poor country, but once all country-specific data is added up to calculate the total Gini, the same individual may easily end up in the lowest 20%. As a result, the factor R tends to zero if the distribution of income or educational attainment becomes similar in different countries.

Since with the distribution of education/income the country means are also likely to converge, we can expect a negative relation between the between-country Gini and the overlap factor (this is exactly what find for the educational Ginis in Figure 2). Growing within country inequality, on the other hand, should not necessarily have any monotonous relationship with the overlap factor. It is possible to imagine a situation when the within-country component is rising only because inequality is rising in some of the countries while remains constant in others and distributions diverge. Then the relationship can be negative. If however the within country component is rising because country specific inequalities are rising proportionally everywhere, the overlap factor should not necessarily be affected at all. In certain applications, like the decomposition of income Ginis according to age-groups (Needleman 1979) or pollution inequality across locations (Millimet and Slottje 2002), it may be useful to use the sum of the within-group component and the overlap factor (that is total Gini minus the between-group Gini), a.k.a Paglin Gini (Paglin 1975). This measure is based on an alternative definition of perfect equality: while in the traditional sense perfect equality means that every individual has the same income/educational attainment (total Gini equals zero), in the sense suggested by Paglin, perfect equality means that every individual has the mean income/educational attainment of his/her group, but differences among groups may exist (total Gini can be above zero).

The results from the decomposition are reported in Table 3 for both education and income for the years when both data were available. Since we have annual estimates of education Gini, the decomposition results are reported in Figure 2. For the income Ginis, we followed the aggregation method used by Van Zanden *et al* (2010) and assumed that all within Ginis were lognormally distributed.

Table 3

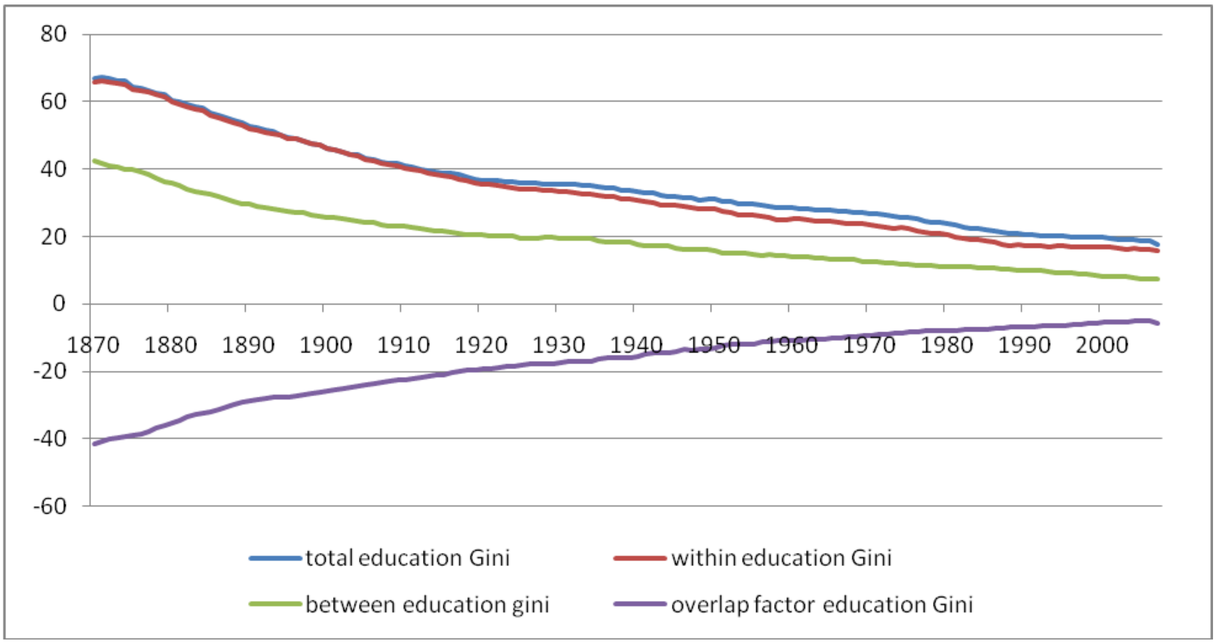
Decomposition of income and educational Ginis in Europe, 1820-2000

	Income Gini				Educational Gini			
	within	between	overlap	Total Gini	within	between	overlap	Total Gini
1820	52.3	13.4	-11.4	54.3				
1850	48.1	17.6	-14.7	51.0				
1870	48.0	21.7	-17.2	52.5	65.9	42.5	-41.3	67.1
1890	44.6	23.4	-18.4	49.6	51.9	29.5	-28.7	52.7
1910	45.6	21.4	-17.0	49.9	40.3	23.0	-22.4	40.8
1929	43.7	25.7	-18.6	50.8	33.6	19.9	-17.9	35.7
1950	38.7	19.9	-14.6	44.0	28.1	15.8	-12.9	31.0
1960	35.2	20.5	-14.7	41.0	25.5	14.0	-11.0	28.5
1970	32.2	19.0	-13.5	37.7	23.6	12.6	-9.3	26.9
1975	37.3	17.0	-13.7	40.6	22.2	11.7	-8.4	25.5

1980	32.3	20.8	-15.2	37.9	20.5	11.1	-7.8	23.8
1985	32.5	21.5	-15.7	38.3	18.7	10.7	-7.4	21.9
1990	32.5	23.6	-14.5	41.7	17.4	10.1	-6.8	20.7
1995	35.1	33.8	-19.2	49.7	17.1	9.3	-6.3	20.1
2000	39.1	34.1	-21.4	51.8	17.0	8.3	-5.4	19.9

The results are as expected. The within country income inequality decreases while the between component grows, which reflects the effect of divergence within our sample. Indeed, he overlap factor decreases, which can be interpreted as a sign of this divergence. Or, if we follow Milanovic's (2002, 70) interpretation that "the more important the overlapping component... the less one's income depends on where she lives", we can conclude that location was more important in 2000 than in 1820 for the individual living standard. This finding also implies a negative relationship between the between-group and the overlap components. At the same time, the increasing divergence in our sample suggest a negative relationship between the within and between Gini when the between educational inequality index decreases.

Figure 2



For educational inequality we find clear decreases in both between and within inequality. This is not surprising since in all countries there was a large scale expansion of public education. The same applies for the within inequality: in all societies education became much more equally distributed. In terms of the overlap factor, the reduction in between country inequality strongly increased the overlap factor converging toward zero. This is a clear sign that in education there was a steady process of convergence in the last 190 years. However, again, the effect of the within educational Gini on the overlap is not monotonous. Where initially an increase of the within Gini might have caused an increase of the overlap factor, the reduction of the between Gini changed this. If the within Gini changed monotonous in all counties, the within Gini would not have an effect on the obverlap Gini. However, it is more likely that the rapid educational expansion in the mid-twentieth century happened unevenly. This implies that the relationship between the within Gini and the overlap Gini turned negative.

In the next section we will discuss within country income-and educational inequality. Since we focus on within country inequality, the overlap factor is of no importance. Both the between and overlap factor are discussed in Section 5.

4. Within country inequality: a composition effect

Since education is expected to affect individual earnings (Becker 1975, Mincer 1974), the distribution of education should affect income distribution/inequality. The relationship is not linear though. Knight and Sabot (1983) use a model with two groups (schooled, unschooled) to derive the nature of this non-linearity. The net effect of educational inequality on income inequality is a sum of two interrelated processes: the first one is the composition effect which is an effect of the expansion of education; the second one is what they call wage compression, i.e. a change in the wage differentials due to changing supply of skilled/unskilled workers. Under the empirically well-established assumption that the standard deviation of the skilled workers is bigger than that of unskilled workers, they show that while the share of skilled workers remains less than 50%, the expansion of education will cause an increase in income inequality, but once the 50% is exceeded it should cause income inequality to fall. This would result in an inverted U-curve like relationship, similarly to the Kuznets-curve. As for the wage compression effect, Knight and Sabot argue that its effect is ambiguous; nevertheless in their derivations they assume that an increase in the supply should rather reduce wage differentials (skill premium).[2] If an increase in the supply of educated labour reduces wage differentials, however, then this component of its effect should necessarily reduce income inequality. The two effects obviously influence the slope of the curve but finally, once the 50% is reached, inequality should drop.

While there are many studies on the relationship between per capita economic growth and education (i.e. Krueger and Lindahl 2001), or inequality of education (Lopez et al 1998, Castello and Doménech 2002) there are relatively few empirical studies in the relationship between income and education inequalities. Checci (2001) uses the Barro-Lee (1996) educational attainment data to estimate educational inequality and explore its relation with income inequality. Even though these Barro-Lee data are known for their bias, their paper has the undeniable merit of having a wide geographical scope with 117 countries. Data availability however only allowed an analysis for the period 1960-95. Just as expected, Checci finds a quadratic relationship between income inequality and educational inequality, but what he finds is a U-curve, not an inverted U-shape as one would expect based on the findings of Knight and Sabot. He also discovers major regional differences; he finds a U-curve for the OECD sub-sample in a fixed-effect panel specification as well as for the whole sample, but the coefficients are insignificant for other regions. Morrison and Murtin (2007) extend their earlier work on income inequality toward human capital inequality, estimated from educational attainment data and assumptions on the rate of returns to education for the period 1870-2010. Even though their primary focus is on testing the Kuznets curve in income, they find a similar inverted U shape in their human capital inequality measure.

Obviously, the existing contradictions in the empirics of the relationship between income and educational inequalities is certainly largely attributable to the poor availability of good quality data, nevertheless, an important part of the story, namely the changes in skill premium/rate of returns to education have somewhat been neglected.

As we saw earlier Knight and Sabot assumed that with growing supply of educated workers, wage differentials should decrease. This is expected if one assumes that wages reflect the marginal product of a given labour variety, and the marginal product is decreasing.

[2] To be more precise they use the following identity from Robinson (1976): $\sigma^2 = p_s\sigma_s^2 + (1-p_s)\sigma_u^2 + p_s(1-p_s)(\bar{y}_s - \bar{y}_u)^2$, where σ, σ_s, and σ_u denote the total income inequality, the income inequality of schooled and unschooled worker respectively, p_s is the share of schooled workers in total population, and finally $\bar{y}_s, \bar{y}_u$ are the mean incomes of the two groups. When they difference this expression with respect to p_s they assume that $\frac{\partial \bar{y}_s}{\partial p_s} < 0, \frac{\partial \bar{y}_u}{\partial p_s} > 0$, that is they assume that the wage of educated individuals is monotonous function of their share in the population i.e. the expansion in education necessarily compresses wages.

In other words, skill premium/rate of returns to education is a decreasing function of educational expansion. Empirics do not always support this assumption however. The wage differentials between college and high school graduates in the United States during the 1960s-1990s exhibited remarkable variation, while the share of college graduates increased steadily. Katz and Murphy (1992) observe that skill premium grew during the second half of the 1960s, dropped in the 70s, and started to grow sharply in the 80s. Mitchell (2005) further explores the movement of skill premium in the US for the 20th century and finds that skill premium fell until the 1940s, grew during the 50s ad 60s, and after drop during the 70s grew steadily from the 80s on. This is all in a period when educational attainment increased and educational inequality decreased. Obviously, the assumption that skill premium is solely affected by the supply of skilled labour must be relaxed.

As for what does affect the relative demand for skilled labour, and hence the skill premium, the literature is abundant. One stream of the literature explains the observed movement of skill premium by Skill Biased Technological Change (SBTC). Katz and Murphy (1992) found that the demand for college graduates increased strongly during the 80s, roughly when the first personal computers entered the market. If computers increased the productivity of college graduates more than the rest, this biased technological development may be a possible answer (Autor et al 1998). Acmeoglu (2002) develops a model in which the skill biasedness of technological change is endogenous: under certain assumptions on the elasticity of substitution among factors of production, technological development may be directed at the more abundant factor. Card and DiNardo (2002) argue that even though this explanation is very popular, the empirical evidence is poor, and other factors, like a reduction in real minimum wages may also offer explanation. This conclusion is also reached by Goldin and Katz (2009) who show that skill biased technological change took place in the entire twentieth century without strongly increasing the skill premium. Only after ca. 1980 they observe an increase in skill biased technical change combined with decreasing growth of college graduates, the skill premium started to grow. In addition, Mitchell (2005) finds a connection between firm size and skill premium in the USA. His argument is based on the flexibility of capital (or alternatively the ratio of fixed to variable costs). When this flexibility reduces, capital can be easier applied to more tasks and the firm size reduces, paired by increasing skill premium. Growing international specialization might also offer an explanation; Wood (1995) sees a connection between the falling share of manufacturing and the growth of the import of low-skill intensive manufactures. In his view this reduced the demand for unskilled workers in developed countries. His argument has been criticized by several authors, nevertheless it cannot be excluded that changes in the composition of trade may affect skill premium in open countries.

Let us see what patterns are to be found in our data. In Table 4 we attempt to capture the relationship between income and educational inequalities with different degrees of

Table 4

Income Gini explained by educational Gini

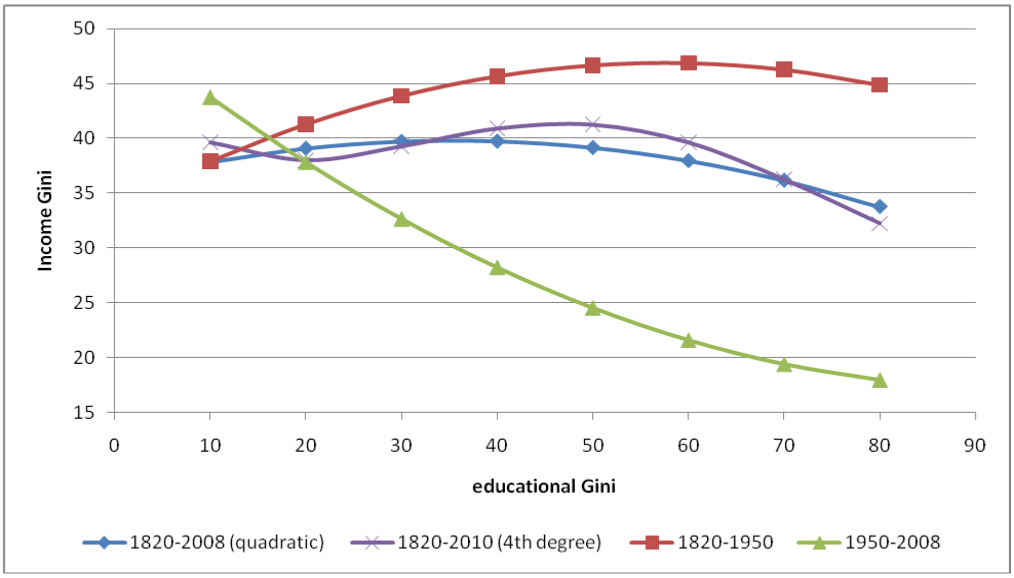
	1820-2008	1820-2008	1820-1950	1950-2008
constant	36.0 (17.5)	47.7 (11.1)	33.68 (9.89)	50.40 (16.78)
GED	0.214 (1.93)	-1.261 (-2.61)	0.459 (3.02)	-0.704 (-3.95)
GED^2	-0.003 (-3.06)	0.054 (2.93)	-0.004 (-2.97)	0.004 (2.17)
GED^3	-	-0.0008 (-2.99)	-	-
GED^4	-	-0.00000385 (2.84)	-	-
R^2	0.620	0.635	0.761	0.738
N	332	332	100	256

Note: Fixed effect specification with country and year specific effects. Robust t-statistics are reported in parentheses.

polynomial functions. Our results suggest that the relationship between the two inequalities changed during the last 180 years. If one wishes to enforce a quadratic relationship on the data, it yields significant coefficients and leads to the expected inverted U curve. In this respect our finding supports that of Morrison and Murtin (2007). But once one allows for a more flexible polynomial form (fourth degree), we find some evidence for a changing relationship (even if we include the mean income and average years of schooling in the regression). In order to have a better picture of what happened to this relationship we split our data in two, a period before and one after 1950. What we find is puzzling, but not completely unexpected: prior to 1950 we find an inverted U-curve as expected but after 1950 the coefficients change sign and we find a U curve, similarly to Checci (2001). The estimated relationships are plotted in Figure 3.

Figure 3

The estimated relationship between educational and income inequalities



We offer an explanation for the observed change in the relationship: skill premium remained relatively stable or even reduced prior to 1950. This was possible because of the lack of skill bias in technological development in combination with fast rising secondary and higher enrolments, but could be also because of the relatively low degree of international commerce. After 1950, however, as we saw for the USA (Mitchell 2005), the demand for educated workers grew so rapidly, together with a reduction in the growth of college graduates, that it finally caused their rate of returns to increase.

Indeed, the generally accepted view that the rate of returns to education should decrease with educational level might not be always true. Morrison and Murtin (2007) estimate the rate of returns to education for the USA using IPUMS Census Data for every tenth year between 1940 and 2000. While the relationship between the log of earning and year spent in education seems to have been concave until 1980 for 1990 and 2000 they rather find it convex (both coefficients of the quadratic function being positive). Morrisson and Murtin do not incorporate this finding in their paper and rather decide to dismiss these results due to the poorer measurement of education for these two years. Nevertheless, it is not surprising that observed growing skill premium translates in the convexity of rate of returns. We have unfortunately much less information on the skill premium in the rest of the world. The widely used parameters for the rate of returns in empirical exercises to convert educational data, usually based on Psacharopulos (1994) or Psacharopoulos and Patrinos (2004), assumes that the rate of returns reduces with the level of education.[3] Hall and Jones (1999) assume 13.4 % for the first 4 years of education, 10.1% for the second 4 years and 6.8% for any additional years, while Morrisson and Murtin (2007) also find diminishing returns to average years of education. Yet, these findings are not incompatabile with rising skill premia over time.

In the followings, we demonstrate by some simple simulations that if skill premium remains constant or decreases (wage compression), while there is an expansion in education (causing more equal distribution of education), we should obtain an inverse U-curve as the relationship between educational and income inequalities. As we pointed out, this seems to be the situation up to the mid-twentieth century. Once we introduce a growth in skill premium in favor of college graduates, however, and slowly eliminate the earning bonus workers with only primary education over non-educated people, we can change the curve so that it becomes an U-shape what seems to coincide with the findings in the second half of the century.

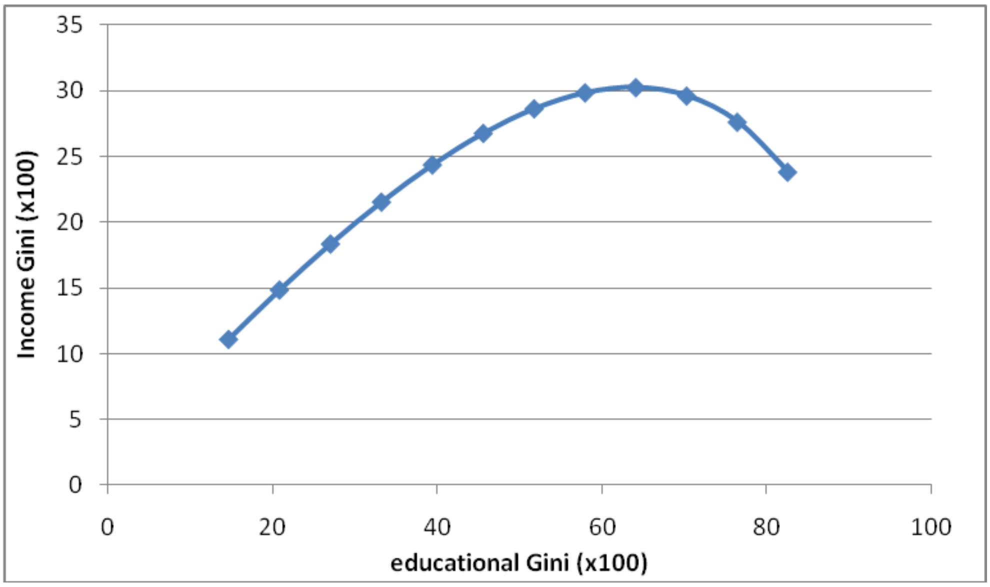
In our simulation we assume that there are four groups of people who differ in terms of their educational attainment only: without education (0 year), primary education (8 years), secondary education (12 years) and higher education (16 years). We assume that in the beginning 80% has no education, 10% has only primary education, while 8% and 2% have completed secondary and tertiary education, which is roughly the situation a European country had to face at the dawn of industrialization with an educational Gini 0.83, We assume a steady increase in the share of educated workers: we assume that the share of tertiary, secondary and primary educated workers grows by 1, 4 and 2 percentage point annually

[3] Hall and Jones (1999) assume 13.4 % for the first 4 years of education, 10.1% for the second 4 years and 6.8% for any additional years, while Morrisson and Murtin (2007) estimate the rate of returns as a linear function of average years of schooling: $\hat{r}_i = 0.125 - 0.004 S_i$.

respectively. So in the 12th period we have a composition of population reminiscent of a modern society: 13% tertiary educated, 52% secondary educated, and 32% primary educated, with only 3% without education. This leads to a relatively equal distribution of education with a Gini 0.15. We keep the rate of returns stable, which is tantamount with constant skill premium, and assume 12%, 8% and 6% rate of returns per additional school year. We further assume that all income differences arise from differences in education and for simplicity we also treat the different groups homogenous (null within group variance). The wage of non-educated workers is set to one. The resulting relationship is the expected inverted U-curve suggested by Knight and Sabot (1983) and found by Morrison and Murtin (2007) and the present paper for the pre-1950 period (Figure 4).

Figure 4

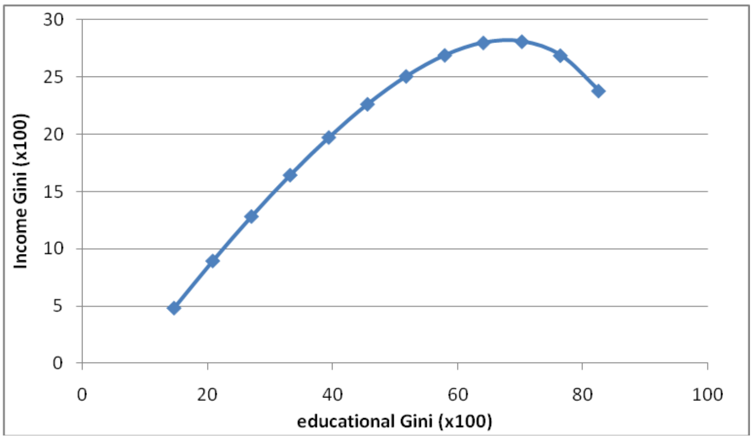
Simulated relationship between income and educational inequalities, with constant skill premium and increasing educational equality



In the second simulation we allow for the wage compression, that is, with increased supply of educated workers the skill premium and rate of returns will decrease, which is what is argued for the first half of the twentieth century. For the first period we use the same skill premium as in the previous simulation, but we reduce the skill premium of tertiary, secondary and primary educated workers by 4, 3 and 1% in every period respectively. These assumptions simulate the wage compression: while the skill premium of college graduates over high school graduates in the first period is 26%, it reduces to 12.6%. The wage bonus of secondary educated over primary educated workers drops even more, from 36% to 8.6%. The effect of wage compression is the expected: the inverted U-curve remains. (Figure 5)

Figure 5

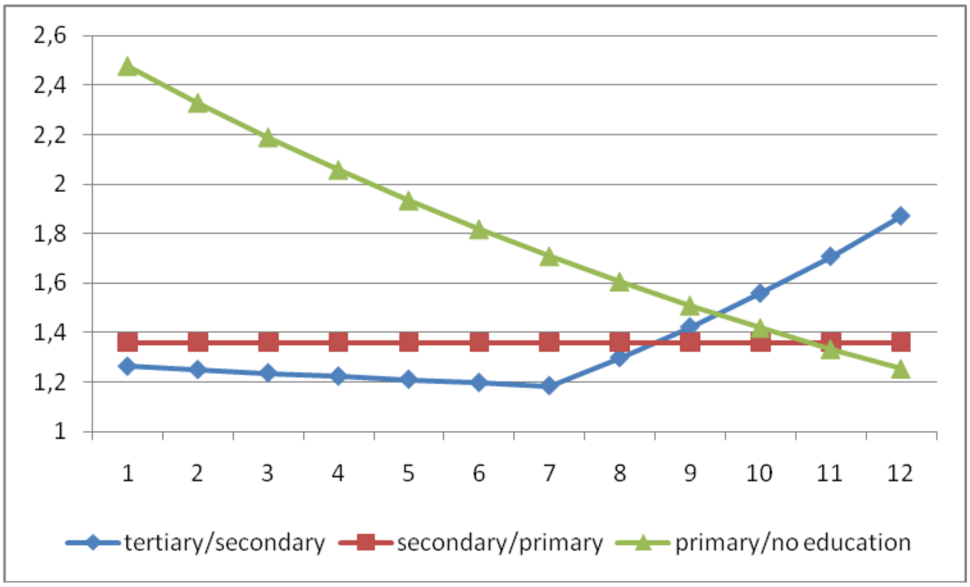
Simulated relationship between income and educational inequalities, with decreasing skill premium and increasing educational equality



However, as we argued, in the first half of the century a wage compression may have taken place, but in the second half there is evidence of increasing skill premia. In the final simulation we therefore assume that the skill premium first decreases but later starts to increase, similarly to what Mitchell (2005) finds for the USA. We assume that the income of primary and secondary educated workers reduces relative to individuals without education by 6% per period, while the earnings of tertiary educated workers experience a reduction in wage

Figure 6

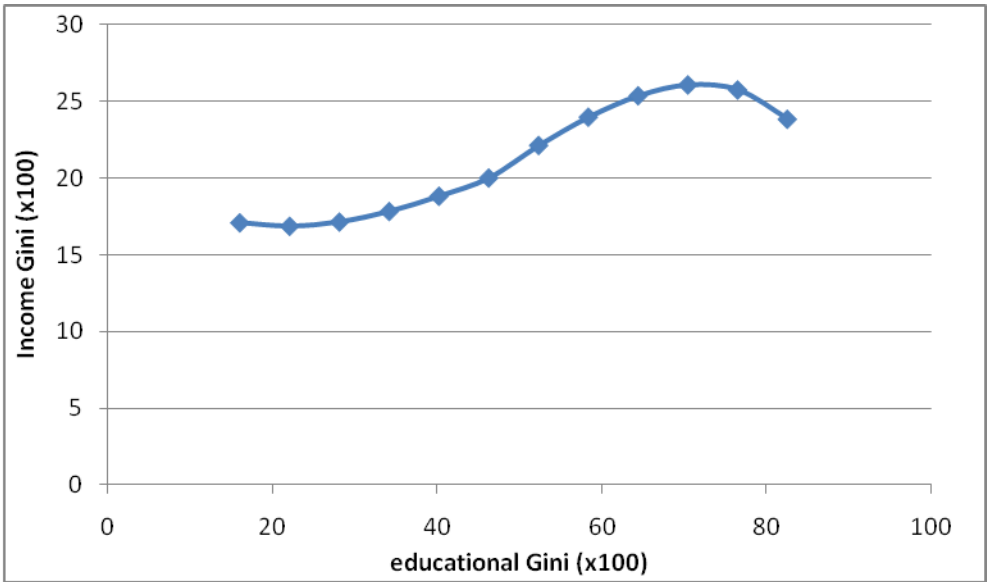
Simulated skill premium for the third round of simulation



by 7% only up to the 6th period, then it begins to rise at 3% a period. This is plotted in Figure 6. The resulting relationship between income and education Ginis is now reminiscent of our estimate with a fourth degree polynomial in Figure 7.

Figure 7

Simulated relationship between income and educational inequalities, with first decreasing then increasing skill premium and increasing educational equality



Even though the parameters of this rudimentary simulation are arbitrary and the results cannot be seen as evidence, nevertheless they demonstrate that changes in the skill premium and especially the convexity of the rate of returns can be an explanation for the changing relationship between the two inequalities.

Unfortunately, a direct test of above hypothesis is prevented by the lack of data on skill premia among detailed occupations.[4] Historical wage data is available for some countries, especially for England and the US, but these usually include wages of far too few occupations that they could serve reliable proxies of skill premium for all levels of education. For this reason, we rather choose an indirect way; instead of trying to include some indicator of skill premium into our regression, we rather try to filter its effect on the coefficient of educational Gini.

Indeed, what we faced in Table 4 is an example of omitted variable bias: skill premium is correlated with the education Gini. If we can find instruments that are correlated with the inequality of education, while not correlated with the skill premium, we could apply a two stage least squares estimation method. We suggest using enrolment data, and two dummy variables (west – OECD country and soc - state socialist country) as instruments in the first stage. We expect enrolment to be strongly related to the equality of education, since they reflect the current state of the distribution of education since more educated parents are more likely to send their children to school. The total enrolment (sum of primary, secondary and tertiary levels) divided by population is a proxy of the general expansion of public education, while the ratio of tertiary to primary enrolment reflects the composition of population.

The idea behind this instrument is again that parents with tertiary level of education will try to send their children to universities/colleges. We believe that current enrolment numbers are not related to skill premium. This might sound counterintuitive, but we can mention a number of explanations. Firstly, one should bear in mind that enrolment to public schools has been since the 19th century regulated by the government and is rarely a question of personal preferences. Indeed, the importance of the State is initiating mass education is well described in the literature (i.e. Ramirez and Ventresca 1992; Cummings 2003). In Europe, several countries countries saw earlier proclamations from King and Church to parents to educate their children. It has been argued that these played an important role in the spread of literacy in Europe (Graff 1987; Mitch 1992; Vincent 2000). This led gradually to the spread of compulsory education. Compulsory education laws were first enacted in Prussia

[4] The data from the ILO October Inquiry initiated in 1924 might offer a solution in the future.

(1763) and Denmark (1814), followed by several South European and Scandinavian countries (Soysal and Strange 1989, 278). In the second half of the nineteenth century compulsory education laws were made in several Northwest European countries, followed by Eastern Europe (Benavot and Resnik. 2006, 11). Yet, in all cases, the State played an important role by enacting compulsory mass eduction be it either because "the establishment of compulsory education addressed narrowly defined educational problems; in others, it was employed as a strategy to "solve" or defer solving long-standing economic, cultural, or social problems."Benavot and Resnik 2006, 13-14). Indeed, in a quantitative analysis Soysal and Strange (1989, 285) find that only the effect of the state on the enactment of compulsory education and increasing enrolments seems to matter.

Secondly, higher income is just one of the possible gains from education; it may also give better social status for the individual and can be a tool of social mobility. These additional benefits are well described. For example, the Mincerian literature often uses parents' education as an instrument for children's education (Card 1999; 2001). Indeed, it has frequently been showed that education is part of social class (i.e. Thomas 1987). For example Chisick (1991) and Van Leeuwen (2000) show that education in the nineteenth century was partly intended to retain the social classes. These findings have even led to several social status indictors being developed partly based on education (i.e. Duncan 1961).

Thirdly, enrolment today can affect the composition of labour force in the future, but since students usually do not enter the labour market in large numbers, and certainly not in the same positions as after their graduation, it should not affect relative wages. To sum up, we have reason to believe that enrolment is at least a predetermined variable.

In the first stage of the regression we intentionally use only a pooled OLS specification, without any country or year specific effects assumed. The reason is that both year and country dummies would be correlated with the skill premium and so our instrumented educational Gini would not be free of these effects. In the second stage we do assume both country and years specific effects, however. We use more instruments than the number of endogenous variables so that the equation is over-identified and we can test the exogeneity of the instruments. We report the results of the first stage in Table 5.

Table5

First stage regression: dependent variable is GED

	1820-2008	1820-1950	1950-2008

Constant	102.1 (72.3)	109.7 (70.5)	79.0 (20.4)
(prim+sec+high)/population	-217.7 (-34.8)	-618.3 (-18.9)	-376.3 (-9.05)
high/prim	-63.3 (-16.4)	-806.4 (-20.0)	-51.1 (-12.2)
$(\text{prim+sec+high})^2/\text{population}^2$	1596.6 (26.5)	1517.9 (10.7)	890.8 (8.20)
$(\text{high/prim})^2$	65.7 (13.0)	4646.6 (13.1)	48.1 (9.37)
D^{west}	-13.7 (-20.1)	-10.0 (-9.16)	-14.6 (-19.5)
D^{soc}	-6.61 (-7.43)	-	-8.15 (-9.17)
R^2	0.600	0.645	0.448
N	3268	1698	1598

First stage regression: dependent variable is GED^2

	1820-2008	1820-1950	1950-2008
constant	9225.9 (70.4)	10416.9 (75.0)	5804.6 (18.1)
(prim+sec+high)/population	-75742.2 (-42.9)	-84208 (-28.9)	-36874.3 (-10.8)
high/prim	-4078.4 (-12.4)	-77532 (-21.2)	-3122.8 (-9.90)
$(\text{prim+sec+high})^2/\text{population}^2$	194209 (35.4)	240116 (19.1)	85683.8 (9.66)
$(\text{high/prim})^2$	4732.3 (11.3)	446780.2 (13.8)	3161.5 (8.45)
D^{west}	-1302.9 (-21.6)	-1018.3 (-10.5)	-1163 (-20.1)
D^{soc}	-677.6 (-8.99)	-	-765.6 (-10.9)
R^2	0.662	0.731	0.439
N	3268	1698	1598

Since we are going to use a quadratic specification, we also instrument the squared educational Gini and use squared instruments as well. All the instruments yield significant coefficients, and the R-squared is always above 0.43, indicating that the instruments are not weak.

In the second stage we use the instrumented educational Gini to find out if we find the expected inverted U-curve for the whole sample (see Table 6). Now we find for the whole

Table 6

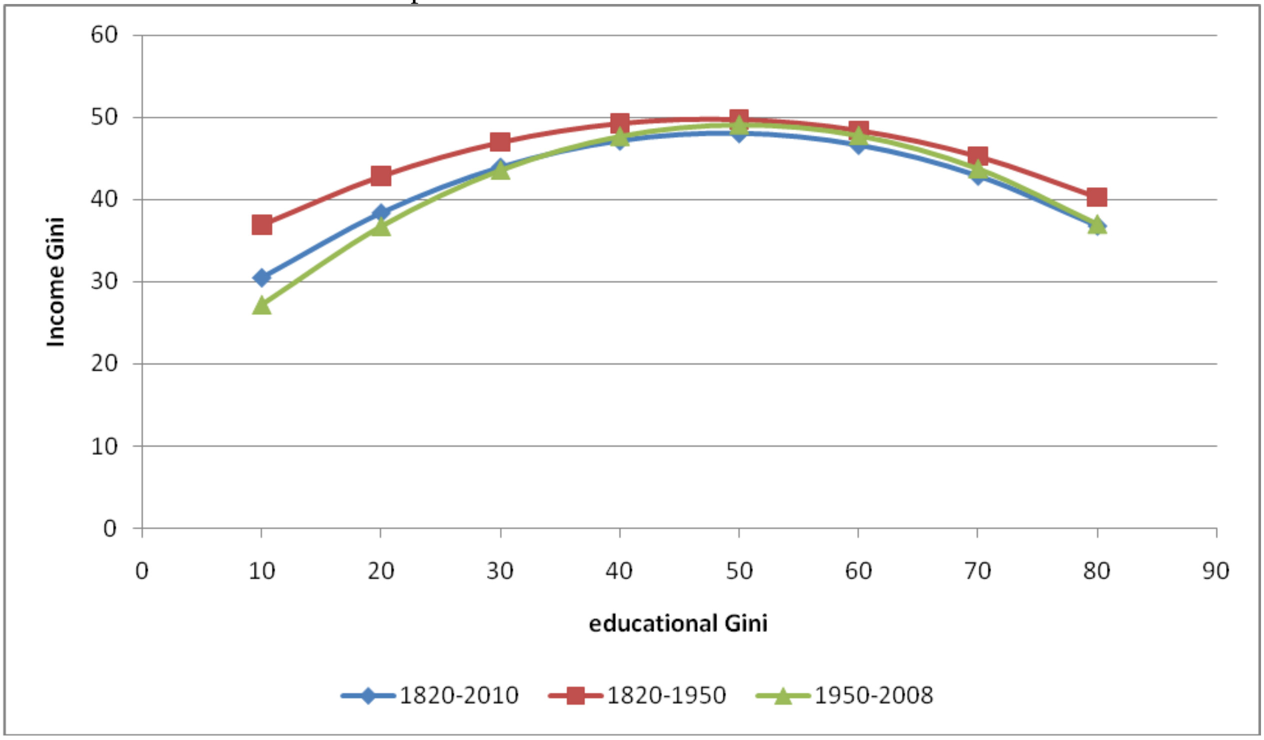
	1820-2008	1820-1950	1950-2008
Constant	20.3 (3.84)	29.2 (4.47)	14.9 (2.26)
GEDIV	1.137 (3.95)	0.866 (2.92)	1.364 (3.22)
$GEDIV^2$	-0.012 (-4.79)	-0.009 (-3.593)	-0.014 (-2.91)
R^2	0.662	0.768	0.703
N	338	107	255
Sargan-test	6.514 (p=0.164)	0.632 (p=0.959)	4.794 (p=0.309)

period the expected coefficients, an inverted U-curve. The estimated curves are more alike than in the non-instrumented specification since the period specific effects have mostly been removed. The Sargan-test cannot reject the null-hypothesis that the instruments are exogenous.

The estimated relationship is plotted in Figure 8. It clearly shows that an inverse U-curve exist for all periods if only the composition effect of education is taken, without the

Figure 8

The estimated relationship between income and educational Ginis from the 2SLS



changing skill premia. Indeed, whereas initially the post-1950 period was initially characterized by an U-curve, filtering out the skill price effect creates the standard inverted U-curve that is characteristic for the composition effect of education.

5. Per capita income and education

The literature on the possible connection between the income differences and educational inequality is abundant. There are several channels through which education may affect economic performance. First, the human capital theory, pioneered by Becker, argues that better educated individuals are more productive which is reflected in their earnings (private returns) but also by a higher productivity of the total labour force through externalities creating social returns. Human capital has been incorporated into the Solow model by Mankiw *et al.* (1992) whose empirical results are consistent with human capital having roughly one-third share in aggregate income, the same magnitude as labour and capital incomes. Human capital plays important role in New Growth Theories. The Uzawa-Lucas model (Lucas 1988, partly based on Uzawa 1965) assumes a two-sector economy, where the first sector employs both physical and human capital to produce income, while the second sector uses human capital to produce additional human capital. Human capital plays important role in Romer's (1990) two-sector model where it is employed in the R&D sector. In both model, under certain assumptions, human capital is not only related to the level of output but also its rate of growth.

Human capital may also affect institutions and therefore indirectly output. Glaeser *et al.* (2007) argues that education promotes civic participation and therefore facilitates democratization. Eicher *et al.* (2009) propose a model in which more educated electorates can more efficiently monitor politicians and limit corruption. The relationship between education and corruption is not monotonous, however, societies with intermediate level of educational attainment may be trapped in a low income/high corruption situation, since the returns to corruption are high, but the voters' ability to monitor the officials' actions are still limited.

The empirical evidence for a relationship between income and education first seemed to be conclusive. Barro (1991) found a significant relationship between primary and secondary enrolment and economic growth between 1960 and 1985. His seminal study has been followed by several attempts, this time utilizing average years of education, an indicator generally believed to be a better proxy of human capital. The results were however less than convincing this time. Even though on micro-level there is plenty of evidence in favour of positive private returns, on macro level sometimes insignificant or even negative coefficients are found (Benhabib and Spiegel 1994, Krueger and Lindahl 2001, Pritchett 2001).

There are several different explanations for the contradictory findings: Krueger and Lindahl (2001) claim that measurement errors in educational variables are mainly responsible and once this is accounted for, a positive and significant relationship is found. Their argument seems to be confirmed by Cohen and Soto (2007) who, using an improved set of estimates of average years of education compared to Barro and Lee (1996), find a positive relationship between income per capita and education. Pritchett (2001) on the other hand suggests that institutional differences may cause education to have different effect in different societies. In countries with bad institutions more human capital may give rise to privately lucrative but socially counterproductive activities, and sometimes the quality of education can be so low that it finally fails to create productive skills. Zeira (2009) points out that the relationship between economic growth and education should ambiguous since the nature of demand for education changed. Industrialization created a mass demand for educated workers, and the availability of abundant human capital is a precondition of adopting new technologies. As such, the costs and barriers to education matters a lot for economic growth and may eventually create a poverty trap.

In the followings we use our data on historical average years of education estimates to find out if education affects economic performance in the short and long-run. Since per capita GDP en average years of education turn out to be non-stationary, I(1) variables, it is possible that there is a long-run relationship between them, that is, they are cointegrated. For this reason we estimate an error-correction representation of this relationship for different sub-periods in a fixed effect panel specification, where both country and year specific effects are assumed for. The results are reported in Table 7.

Table 7
Error-correction models (dependent variable is the first difference of the log of per capita GDP)

variable	FE panel (1820-2008)	FE panel (1820-1900)	FE panel (1900-1950)	FE panel (1950-2008)
constant	0.346 (5.32)	2.051 7.09)	1.043 (4.70)	0.256 (4.89)
$\Delta S(t)$	0.005 (0.28)	0.016 (0.52)	0.007 (0.18)	-0.010 (-0.70)
lny(t-1)	-0.040 (-5.18)	-0.271 (-7.08)	-0.132 (-4.78)	-0.033 (-5.67)
S(t-1)	0.002 (1.29)	0.006 (0.62)	0.004 (0.70)	0.009 (3.58)
D^{soc}	0.058 (4.09)	-	-	0.058 (4.10)
$D^{soc.}\cdot \Delta S(t)$	0.033 (1.12)	-	-	0.054 (1.99)
$D^{soc.}\cdot S(t-1)$	-0.007 (-4.06)	-	-	-0.008 (-4.51)
Polity2(t)	0.0005 (1.19)	-0.002 (-1.68)	0.0005 (0.86)	-0.0005 (-1.98)
R^2	0.218	0.289	0.240	0.302
N	2864	410	966	1525
long run effect of education on output (Wald test)	0.038 (p=0.213)	0.02 (p=0.532)	0.03 (p=0.490)	0.263 (p=0.000)
long run effect of education on output in socialist countries (Wald test)	-0.147 (p=0.011)	-	-	0.014 (p=0.859)

Note: robust t-statistics are reported in parentheses, country and year specific effects included but not reported.

The most important finding is that only for the 1950-2008 sub-sample we find valid error-correction representation, which is an indication that education and per capita income are cointegrated according to the Granger representation theorem (Engle and Granger 1987). Since there are countries in our sample that were under state-socialist regime for a longer period we also tested if this had an effect on the relationship between education and economic performance. The reason behind this hypothesis is similar to Pritchett's (2001) line of reasoning: in these countries institutions were probably less efficient in resource allocation, since government tried to replace the market's role in the allocation of resources, and this might have led to inefficient allocation of human capital. For a similar reason we also use the polity2 variable from the polityIV database as a proxy of political institutions. Unfortunately most institutional datasets have observations only from the 1970s, so we need to assume that the effect of institutions that are less likely to change (here we think of informal institutions in the first place) are captured by the country specific fixed-effects.

As for the effect of education on per capita income, our results suggest education having no immediate impact on output, which is not unexpected. As for the long-run relationship, we find that education had a significant positive relationship with per capita income only in non-socialist countries and only during the second half of the 20th century. Before that, even though the coefficients are positive, they remain insignificant statistically. Our expectation regarding the socialist countries seems to be confirmed: for the whole sample, they are the only group where we find a significant relationship between per capita GDP and the level of educational attainment, but it is negative. When we look at the post-1950 sample, when educational attainment seems to have a positive effect on income in the long-run, in socialist countries we find no significant effect. Our conclusion is that state-led educational expansion[5] and the central distribution of human capital failed to achieve the efficiency of market economies.

It is possible however that our results are affected by a bias due to the inclusion of the lagged per capita income, since our specification is a slightly modified form of a fixed effect ARX(1) specification which is subject to a bias due to a correlation between the error-term and the lagged income variable (Nickell 1981), even though because of the relatively high number of observations this bias is not likely to be very high. Another possible bias can be

[5] That state plays an important role in the expansion of education is not exclusively typical of state-socialist economies, but, using a Gerschenkronian (Gerschenkorn 1962) line of reasoning, it should be true for most late-comers of industrialization.

caused by simultaneity of democracy and income as can be derived from Lipset (1960) and Glaeser et al. (2004). For this reason we use Two-Stage Least Squares where we use the second and third lag of per capita income, and the first lag of the polity2 variables as additional instruments.[6] These variables are predetermined and can serve as instruments, but since the equation is now over-identified we can use a Sargan-test to directly test if the set of instruments we use are exogenous. We assume that the average years of education and its growth rate is not correlated with the present level of output. The results are reported in Table 8.

Table 8

Error-correction models (dependent variable is the first difference of the log of per capita GDP), 2SLS estimates

variable	FE panel (1820-2008)	FE panel (1820-1900)	FE panel (1900-1950)	FE panel (1950-2008)
constant	0.426 (7.89)	1.466 (3.71)	1.344 (7.04)	0.351 (6.55)
$\Delta S(t)$	-0.013 (-0.74)	0.044 (1.24)	-0.035 (-0.93)	-0.011 (-0.77)
lny(t-1)	-0.049 (-7.76)	-0.194 (-3.68)	-0.170 (-7.29)	-0.045 (-7.44)
S(t-1)	0.001 (0.96)	0.004 (0.45)	0.004 (0.73)	0.009 (3.88)
D^{soc}	0.058 (4.09)	-	-	0.067 (4.70)
$D^{soc}\cdot \Delta S(t)$	0.033 (1.12)	-	-	0.037 (1.40)
$D^{soc}\cdot S(t-1)$	-0.007 (-4.06)	-	-	-0.008 (-4.48)
Polity2(t)	0.0005 (1.19)	-0.001 (-0.58)	0.001 (1.88)	-0.0002 (-0.85)
R^2	0.297	0.279	0.277	0.297
N	2824	408	937	1513
long run effect of education on output (p-value)	0.023 (p=0.345)	0.02 (p=0.650)	0.024 (p=0.471)	0.210 (p=0.000)
long run effect of education on output in socialist countries (p-value)	-0.102 (p=0.018)	-	-	0.0004 (p=0.655)
Sargan-test (p-value)	0.050 (p=1.000)	6.143 (p=1.000)	1.012 (p=1.000)	74.93 (P=0.797)
instruments	lny(t-2), lny(t-3), polity2(-1), $\Delta S(t)$, S(t-1), D^{soc}, $D^{soc}\cdot \Delta S(t)$, $D^{soc}\cdot S(t-1)$	lny(t-2), lny(t-3), polity2(-1), $\Delta S(t)$, S(t-1), D^{soc}, $D^{soc}\cdot \Delta S(t)$	lny(t-2), lny(t-3), polity2(-1), $\Delta S(t)$, S(t-1), D^{soc}, $D^{soc}\cdot \Delta S(t)$	lny(t-2), lny(t-3), polity2(-1), $\Delta S(t)$, S(t-1), D^{soc}, $D^{soc}\cdot \Delta S(t)$, $D^{soc}\cdot S(t-1)$

Note: robust t-statistics are reported in parentheses, country and year specific effects included but not reported.

[6] Using the second and third lag of the per capita income as instruments is basically what is suggested by Anderson and Hsio (1981). Their approach had been seen obsolete until recently, but now an increasing number of studies find that the usually preferred GMM and GMM-SYS estimators lead to significantly biased results. See: Bun and Klaassen (2002) or Hauk and Wacziarg (2009).

The Sargan tests in all cases fail to reject the null-hypothesis that the instruments are exogenous. The results do not change significantly, however. We still find that the effect of education is insignificant before 1950; it only turns out to be significant after 1950 in non-socialist countries.

Yet, these results may also been driven by omitted variables. A significant amount of empirical evidence seems to indicate that the within-country component of income inequality also plays an important role in average income levels, and therefore between-country inequality. Torsten and Tabillini (1994) develop a model in which more inequality leads to policies that do not respect property rights and therefore hinder factor accumulation and finally economic growth. Their theoretical proposition is confirmed by an empirical analysis. Perotti (1996) finds that unequal societies have higher fertility but also lower investment in education, paired with less stability, both harmful to growth. A similar argument can be found in Alesina and Perotti (1996) who find that on as sample of 71 countries 1960-1985 that more inequality is usually associated with more political instability and less investments. Not all empirical evidence is supportive of a monotonous relationship, however. Deininger and Squire (1998) find that once initial income, investment and continent specific intercepts are directly included in the regression, income inequality yields negative and insignificant coefficients. Castello and Doménech (2002) focus on measuring the effect of human capital inequality on growth, which they find significant and negative, but once this effect is captured, income inequality yield a positive coefficient. This we can consider as an indication that income inequality may primarily affect growth performance through factor accumulation, which is in line with the theories of political economy.

In the following we therefore rerun our regression with the educational Gini included as regressor. We also experimented with including income Ginis as well, but it reduced the number of useable observations to such an extent that it rendered the results incomparable with our previous results, and also it yielded insignificant coefficients. A possible explanation might be that while measurement error of income inequality does not cause biased estimation if the income Gini is the dependent variable, it can lead to biased results once it is used as a regressor.

Table 9
Error-correction models (dependent variable is the first difference of the log of per capita GDP)

Variable	FE panel (1820-2008)	FE panel (1820-1900)	FE panel (1900-1950)	FE panel (1950-2008)
Constant	0.348 (5.34)	2.139 (7.50)	1.011 (4.48)	0.291 (5.46)
$\Delta S(t)$	0.005 (0.25)	0.083 (2.31)	0.006 (0.15)	-0.017 (-1.14)
lny(t-1)	-0.040 (-5.15)	-0.327 (-7.97)	-0.136 (-4.90)	-0.032 (-5.43)
S(t-1)	0.001 (0.81)	0.053 (3.02)	0.010 (1.26)	0.005 (1.76)
D^{soc}	0.057 (4.03)	-	-	0.061 (4.34)
$D^{soc}\cdot \Delta S(t)$	0.034 (1.12)	-	-	0.056 (2.08)
$D^{soc}\cdot S(t-1)$	-0.007 (-4.04)	-	-	-0.009 (-4.73)
Polity2(t)	0.0004 (1.75)	-0.001 (-1.23)	0.0005 (0.86)	-0.0005 (-2.07)
GED	-0.00005 (-0.49)	0.003 (3.08)	0.0009 (1.53)	-0.0006 (-3.54)
R^2	0.219	0.312	0.243	0.309
N	2864	410	966	1525
long run effect of education on output (Wald test)	0.029 (p=0.423)	0.162 (p=0.001)	0.073 (p=0.218)	0.147 (p=0.082)
long run effect of education on output in socialist countries (Wald test)	-0.155 (p=0.014)	-	-	-0.126 (p=0.186)

Note: robust t-statistics are reported in parentheses, country and year specific effects included but not reported.

After the inclusion of the within Gini the results change somewhat. First of all, now we find a significant long-run relationship between per capita income and education in the 1820-1900 period, which is more in line with the expectations, i.e., skills were demanded increasingly during the Industrial Revolution. The effect of within country educational inequality is not monotonous, however. In the 19th century we have a positive effect, and after 1950 we find a negative coefficient. A possible explanation may be related to Zeira's (2009) idea on the changing nature of the demand for human capital. The second half of the 20th century saw such a widespread use of technologies that could not be operated without relatively high level of education, the 19th century technology was probably much less demanding in this respect and educational inequality remained less detrimental. But this is still far from explaining the positive coefficient for the 19th century. Another explanation lies in the relationship between educational inequality and income inequality. In section 4 we found that the relationship between income and educational inequalities was different in the two periods: prior to 1950 the relationship was predominantly positive, and it became negative subsequently. Since data problems prevent us from including income Ginis in the specification, the coefficient now reflect the effect of income inequality as well. Castello and Domenech (2002) find that once educational inequality is directly included in the regression, the effect of income inequality is positive. The resulting bias is therefore positive, and this may finally cause the coefficient to become positive. On the other hand, the turn of the relationship after 1950 should cause a downward bias in the GED coefficient.

6. Conclusion

This paper discussed the relation between educational and income inequality. Using the method proposed by Foldvari and Van Leeuwen (2009) we construct a dataset with average years of education and educational inequality for Europe stretching back to 1870. We combine these data with estimates of GDP per capita and income inequality from Maddison (2007) and Van Zanden *et al* (2010).

For the within analysis, we find the standard inverse U-curve for the period up to 1950. After 1950, we find a U-curve. This can be explained by increasing prices of skilled labour. After removing the effect of the skill premium, we find again the standard inverse U-curve.

We estimate a standard error correction model of GDP/cap on average years of education. We find that only after 1950 there is a positive and significant coefficient, except for the socialist countries. Theoretically, it may be argued that we have to include the within country educational Gini as well. Indeed, now the coefficients turn significant also before 1950. Interestingly, whereas the within country Gini is positive before 1950, it turns negative after 1950.

References

Acemoglu, D (2002) Directed Technological Change. *The Review of Economic Studies*, Vol. 69, No. 4., pp. 781-809.

Alesina, A. and Perotti, R. (1996), Income Distribution, Political Instability, and Investment. European Economic Review, Vol. 40(6), 1203-1228.

Anderson, T. W. and Hsiao, C. (1981) Estimation of Dynamic Models with Error Components, Journal of the American Statistical Association 76, 497-517.

Autor, D. H. , Katz, L. F. and Krueger, A. B. (1998) Computing Inequality: Have Computers Changed the Labor Market? *The Quarterly Journal of Economics*, Vol. 113, No. 4., pp. 1169-1213

Banerjee, A. and Newman, A. F. (1993). Occupational choice and the process of development. Journal of Political Economy, 101, 274-98.

Barro, Robert and Jong-Wha Lee, April 2010, "A New Data Set of Educational Attainment in the World, 1950-2010." NBER Working Paper No. 15902.

Barro, R. J. (1991) Economic Growth in a Cross Section of Countries. The Quarterly Journal of Economics, Vol. 106, No. 2., pp. 407-443

Barro R. and J.-W. Jong-Wha Lee (1993) International comparisons of educational attainment,' *Journal of Monetary Economics* 32, 363-394.

Barro, R. J., Lee J. W. (2001) International Data on Educational Attainment Updates and Implications. *Oxford Economic Papers* 53 (3); 541-63.

Benavot, Aaron and Julia Resnik, 'Lessons from the Past: A Comparative Socio-Historical Analysis of Primary and Secondary Education,'in: Benavot, Aaron, Julia Resnik, and Javier Corrales, Global Educational Expansion: Historical Legacies and Political Obstacles, American Academy of Arts and Sciences, 2006, 1-90.

Benhabib, J. and Spiegel, M. M. (1994) The role of human capital in economic development evidence from aggregate cross-country data," Journal of Monetary Economics, vol. 34(2), 143-173.

Bruno, M., Ravallion, M. and Squire, L. (1996) Equity and growth in developing countries : old and new perspectives on the policy issues World Bank Policy Research Working Paper Series no. 1563

Bun M. and Klaassen F. (2002) The importance of dynamics in panel gravity models of trade. Tinbergen Institute. Discussion Paper, No. 02–108/2

Card, D. (2001). Estimating the returns to schooling: Progress on some persistent econometric problems. Econometrica,69(5), 1127–1160.

Card, E. and DiNArdo, J. E. (2002) Skill-Biased Technological Change and Rising Wage Inequality: Some Problems and Puzzles, Journal of Labor Economics, 2002, vol. 20, no. 4, 733-783.

Card, D. (1999). The causal effect of education on earnings. In O. Ashenfelter, & D. Card (Eds.). Handbook of Labour Economics, Vol. 3B. Amsterdam: Elsevier.

Castelló, A. and Doménech, R (2002) Human capital inequality and economic growth: some new evidence, The Economic Journal, 112 (478), 187-200

Checci, D. (2001) Education Inequality and Income Inequality, Distributional Analysis Research Programme, Discussion Paper No. 52

Chisick, H., The Limits of Reform in the Enlightenment: Attitudes Towards the Education of the Lower Classes in the Eighteenth Century (London, 1991).

Cohen, D. and Soto, M. (2007) "Growth and human capital: good data, good results," Journal of Economic Growth, vol. 12(1), 51-76.

Cummings, William, 2003. The Institutions of Education: A Comparative Study of Educational Development in Six Core Nations. Oxford: Symposium Books.

Deininger, K. and Squire, L. (1998). New ways of looking at old issues: inequality and growth, Journal of Development Economics, vol. 57, no. 2, pp. 259±87.

De la Fuente, A. and R. Doménech, 2000, Human Capital in Growth Regressions: How Much Difference Does Data Quality Make? *OECD Economics Department Working Papers*, No. 262, pp. 2-68.

Duncan, O. D., "A Socioeconomic Index for All Occupations," in A. Reiss et al., *Occupations and Social Status* (Free Press, 1961).

Eicher, T., García-Peñalosa , C. and van Ypersele, T. (2009) Education, corruption, and the distribution of income, Journal of Economic Growth, Vol. 14 (3), 205-231.

Engle, R. F. and Granger, C. W. J. (1987) Co-integration and Error Correction: Representation, Estimation, and Testing, Econometrica, Econometric Society, vol. 55(2), pages 251-276.

Fields, G. (2001). "Distribution and Development, A New Look at the Developing World". Russel Sage Foundation, New York, and The MIT Press, Cambridge, Massachusetts, and London, 2001.

Földvári, Péter and Bas van Leeuwen, (2009a) 'Average years of education in Hungary: annual estimates 1920-2006,' *Eastern European Economics,*Vol. 47 (2), pp. 5-20.

Földvári Péter, and Bas van Leeuwen, (2009b) 'Capital accumulation and growth in Central Europe, 1920-2006,' manuscript 17 July 2009.

Galor, O. and Tsiddon, D. (1996) Income Distribution and Growth: The Kuznets Hypothesis Revisited, Economica, Vol. 63, No. 250, S103-S117

Gerschenkron, A. (1962), Economic backwardness in historical perspective, a book of essays, Cambridge, Massachusetts: Belknap Press of Harvard University Press.

Glaeser, E.L., La Porta, R., Lopez-de-Silanes, F. and Shleifer, A. (2004) Do institutions cause growth? NBER Working Paper No. 10568.

Glaeser, E. L. - Ponzetto, G. A. M. - Shleifer, A. (2007) Why does democracy need education? Journal of Economic Growth Vol. 12, 77–99.

Goldin, Claudia, and Lawrence F. Katz (2009), *The race between education and technology*, Harvard: Harvard University Press 2009.

Graff, Harvey J. 1987. The Legacies of Literacy: Continuities and Contradictions in Western Culture and Society. Bloomington: Indiana University Press.

Hall, R. E. and Jones C. I. (1999) Why Do Some Countries Produce So Much More Output per Worker than Others? Quarterly Journal of Economics, February 1999, Vol. 114, pp. 83-116.

Hauk, W. R. Jr. and Wacziarg, R. (2009) A Monte Carlo study of growth regressions Journal of Economic Growth (2009) 14, 103–147.

Jonsson, Gudmundur, Magnus S. Magnusson (eds.) (1997), Sogulegar *hagtolur um Island (Iceleandic Historical Statistics)*, Hagstofa Islands: Reykjavik 1997.

Katz, L. F. and Murphy, K. M. Changes in Relative Wages, 1963-1987: Supply and Demand Factors (1992), The Quarterly Journal of Economics, Vol. 107, No. 1, pp. 35-78

Krueger, A. B. And M. Lindahl, 2001, Education for Growth: Why and For Whom? *Journal of Economic Literature*, 39, pp. 1101-36.

Kuznets, S. (1955). Economic growth and income inequality. American Economic Review, 45, 1- 28.

Kyriacou, G. (1991) Level and Growth Effects of Human capital: A Cross-Country Study of the Convergence Hypothesis, Economic Research Reports 19-26, C.V. Starr Center for Applied Economics, New York University.

Lains, Pedro (2003), ‘Catching up to the European core: Portuguese economic growth, 1910–1990,’ *Explorations in Economic history*, Vol. 40, 369-386.
Lambert. P. J. and Aronson, J.R. (1993) Inequality Decomposition Analysis and the Gini Coefficient Revisited, T*he Economic Journal*, Vol. 103 (420), pp. 1221-1227
Lipset, S.M. (1960) Political Man: The Social Basis of Modern Politics. New York: Doubleday.
Ljungberg, Jonas and Anders Nilsson (2009), ‘Human capital and economic growth: Sweden 1870–2000,’ *Cliometrica*, Vol. 3, 71-95.
Lopez, R., Thomas, V. and Yan W. (1998), Addressing the education puzzle: the distribution of education and economic reform," Policy Research Working Paper Series 2031, The World Bank
Lucas, R. (1988) On the mechanics of economic development, Journal of Monetary Economics Vol 22 (1), 3-42.
Maddison, Angus, *Contours of the World Economy, 1-2030 AD. Essays in Macro-Economic History*, Oxford University Press, Oxford, 2007.
Mankiw, N. G., Romer, D. and Weil, D. N. (1992) A Contribution to the Empirics of Economic Growth. The Quarterly Journal of Economics, Vol. 107, No. 2., pp. 407-437.
Marshall, Monty G. and Keith Jaggers (2009), Polity IV Project: Political Regime Characteristics and Transitions, 1800-2009, downloaded from: http://www.systemicpeace.org/polity/polity4.htm.
Matyas, L. , Konya, L. and Macquarie, L. (1998) The Kuznets U-Curve hypothesis: some panel data evidence. Applied Economic Letters 5 , pp. 693-697.
Milanovic, B. (2002). True world income distribution, 1988 and 1993: first calculation based on household surveys alone. *Economic Journal*, vol. 112, pp. 51-92.
Millimet, D. L. and Slottje, D. (2002) An environmental Paglin-Gini', Applied Economics Letters, Vol. 9 (4), 271-274.
Mincer, Jacob. Education, Experience, and Earnings. New York: Columbia University Press, 1974.
Mitch, David. 1992. *The Rise of Popular Vernacular Literacy in Victorian England*. Philadelphia: University of Pennsylvania Press.
Mitchell, B.R, (2003). International Historical Statistics: Europe, 1750-1993. M. Stockton Press, New-York.
Mitchell, M. F. (2005), Specialization and the skill premium in the 20th century International Economic Review Vol. 46, No. 3, 935-955.
Morrisson, C. and Murtin, F. (2007) Education inequalities and the Kuznets curves: a global perspective since 1870. Paris School of Economics, Working Paper No. 2007-12.
Morrisson, C. and Murtin, F (2009) The Century of Education, Journal of Human Capital. Volume 3, Issue 1, Page 1–42, Mar 2009
Needleman, L. (1979) Income Distribution in Canada and the Disaggregation of the Gini Coefficient of Concentration, *Canadian Public Policy* Vol. 5 (4), 497-505
Nickel, Stephen J. (1981): Biases in Dynamic Models with Fixed Effects, Econometrica 49, 1417-1426.
Núñez, C.E., 2005. Educación. In: Carreras, A., Tafunell, X. (Eds.), Estadísticas Históricas de España, siglos XIX y XX. Fundación BBBV, Bilbao, pp. 155–244.
Paglin, M. (1975) The measurement and trend of inequality: a basic revision, American Economic Review, 65, 598-609.
Perotti, R. (1993). Political equilibrium, income distribution and growth: theory and evidence. Review of Economic Studies, 60, 755-66.

Perotti, R. (1996), Growth, Income Distribution, and Democracy: What the Data Say, Journa of Eonomic Growth vol 1., 149-187.

Pina, A., and St. Aubyn, M. (2002), "Public capital, human capital and economic growth: Portugal 1977-2001", *Departamento de Prospectiva e Planeamento*, Ministério das Finanças, October 2002.

Portela, M., R. Alessie, and C. Teulings, 2004, Measurement Error in Education and Growth Regressions. *Tinbergen Institute Discussion Paper, TI 2004-040/3*.

Prados de la Escosura, Leandro, and Joan Rosés (2010 forthcoming), 'Human capital and economic growth in Spain, 1850–2000,' *Explorations in Economic History*.

Pritchett, L (2001) Where has all the education gone? The World Bank Economic Review, Vol. 15 (3), 367-391.

Psacharopoulos, G. (1994), Returns to Investment in Education: A Global Update, World Development, XXII , 1325–1343.

Psacharopoulos, G. and Patrinos, H. A. (2002) Returns to investment in education: a further update. World Bank Policy Research Working Paper 2881

Ramirez, Francisco, and Marc Ventresca. 1992. "Building the Institution of Mass Schooling: Isomorphism in the Modern World." Pp. 47–60 in The Political Construction of Education, ed. Bruce Fuller and Richard Rubinson. New York: Praeger.

Ritzmann-Blickenstorfer, Heiner (ed.) 1996, Historische Statistik der Schweiz =: Statistique historique de la Suisse = Historical statistics of Switzerland, Chronos Verlag, Zurich 1996.

Robinson, S. (1976) A Note on the U Hypothesis Relating Income Inequality and Economic Development, The American Economic Review, Vol. 66, No. 3, pp. 437-440

Romer, P. M. (1990) Endogenous Technical Change. The Journal of Political Economy, Vol. 98, No. 5, pp. 71-102.

Soysal, Yasemin, and David Strange. 1989. "Construction of the First Mass Education Systems in 19th Century Europe." Sociology of Education 62: 277–88.

Teixeira A. C (2004) Measuring aggregate human capital in Portugal. An update up to 2001. FEP Working Paper 152, Faculdade de Economia da Universidade do Porto

Thomas, K. J., 'Numeracy in Early Modern England: The Prothero Lecture', *Transactions of the Royal Historical Society*, 37 (1987), pp. 103-32.

Thomas, V., Wang Y. and Fan X. (2000) Measuring Education Inequality: Gini Coefficients of Education. Mimeo. The World Bank, http://econ.worldbank.org/files/1341_wps2525.pdf.

UNESCO, *UNESCO statistical yearbook*, Paris: UNESCO, various issues.

Uzawa, H. (1965) Optimum Technical Change in An Aggregative Model of Economic Growth, International Economic Review, Vol. 6, No. 1., pp. 18-31.

Van Leeuwen, M.H.D., Logic of Charity: A Simple Model Applied to Amsterdam 1800–1850 (London, 2000).

Van Zanden, J.L., Baten, J., Földvári , P., and Van Leeuwen, B. (Third Draft 2010) 'World Income Inequality 1820-2000', Unpublished paper.

Vincent, David. 2000. The Rise of Literacy: Reading and Writing in Modern Europe. Cambridge: Polity Press.

Wood, A. (1995) How Trade Hurt Unskilled Workers, The Journal of Economic Perspectives, Vol. 9 No. 3, pp. 57-80.

Zeira, J. (2009) Why and How Education Affects Economic Growth, Review of International Economics, 17(3), 602–614.