Human Capital and Income Inequality: New Facts and Some Explanations*

Amparo Castelló-Climent[a] and Rafael Doménech[a,b]

a University of Valencia, Spain
b BBVA Research, Spain

April, 2017

Abstract

This paper computes an updated data set on human capital inequality for 146 countries from 1950 to 2010. Using the new indicators, we document several facts regarding the evolution of income and human capital inequality. The main findings reveal that, in spite of a marked reduction in human capital inequality around the world, the inequality in the distribution of income has hardly changed. In many regions, the income Gini coefficient was almost the same in 2005 as it had been in 1960. In order to find explanations for this puzzle, we first estimate the distribution of wages using recent estimates of rates of return to years of schooling. We find an inverted U-shape relationship between these two Gini coefficients that can be explained by a composition effect resulting from the fall in the share of population with no schooling. When we analyze the sensitivity of our results to different assumptions on the returns to years of schooling, we find the returns do not affect the inverted U-shape in developing countries, where the share of illiterates is high. We also show that exogenous forces such as skill-biased technological progress have offset the effects of the fall in education inequality, which therefore represents an alternative explanation for the low correlation between the changes in labor income and education inequality across countries.

Keywords: education inequality, wage inequality, income inequality.
JEL Classification: I24, O11, O15, O5.

* We are grateful for the comments received at seminars, at the 2014 Royal Economic Society Conference, and at the 31st Annual Meeting of the European Economic Association. A. Castelló would like to thank the Spanish Ministry of Economy and Competitiveness for financial support through ECO2015-65263-P project. R. Doménech thanks the Ministry of Economy and Competitiveness, ECO2014-53150-R for financial support. Contact: amparo.castello@uv.es, rafael.domenech@uv.es.

1. Introduction

In the last few decades, most developing countries have made a great effort to eradicate illiteracy, reducing the number of illiterates by several hundreds of millions. As a result, the inequality in the distribution of education has been reduced by more than half: the average human capital Gini coefficient dropped from 0.55 in 1960 to 0.28 in 2005. However, in spite of the equalizing process in the distribution of education, inequality in the distribution of income has hardly changed. The value of the average income Gini coefficient for the same group of countries was almost the same in 2005 (0.41) as it had been in 1960 (0.42). This trend is not restricted to developing countries: in 1960, the human capital Gini coefficient in the high-income OECD countries was 0.22 and decreased to 0.15 in 2005, whereas the income Gini coefficient has remained unchanged at 0.30.

This paper analyzes the above evidence in detail. In doing so, we compute a very comprehensive data set on human capital inequality variables, covering 146 countries from 1950 to 2010, extending the previous data set elaborated by Castelló and Doménech (2002). In this new version, we use the educational attainment data set by Barro and Lee (2013), which includes more countries and years, reduces some measurement errors, and solves some of the shortcomings pointed out by De la Fuente and Doménech (2006 and 2015) and Cohen and Soto (2007). We also compute a more precise human capital Gini coefficient using seven levels of schooling. The new categories distinguish between those individuals that have completed a level of education and those who are enrolled in a level of schooling but have not necessarily completed that level. This distinction improves the accuracy of the human capital Gini coefficient with regard to previous versions, particularly in developing countries where the dropout rates are extremely high. Using this new data set, the paper reports some interesting new stylized facts regarding the evolution of human capital and income inequality.[1]

We first observe that from 1950 to 2010, there was a significant reduction in human capital inequality around the world. In most countries, the marked reduction in education inequality has mainly been due to the sizeable decline in the share of illiterates (34 pp on average). In most advanced countries, however, there is no clear pattern in the evolution of education inequality, and the human capital Gini coefficient has primarily been determined by the distribution of education among the literate population. Income inequality has remained relatively stable over a period of 45 years , while over that same time, human capital inequality has fallen significantly. Thus, the evidence shows that the reduction in

[1] Previous contributions to the literature have found that income inequality is positively correlated to education inequality and negatively related to education (e.g., Becker and Chiswick, 1966, Ahluwalia, 1976). Others, however, found that schooling inequality has a marginal negative, rather than positive, effect on income inequality (Ram, 1984). De Gregorio and Lee (2002) show that, although countries with higher educational attainments and a more equal distribution of education have a more equal distribution of income, a significant proportion of the variation in income inequality remains unexplained.

human capital inequality has not been accompanied by a decline in the income Gini coefficient. The large expansion in literacy in most developing countries has clearly reached the lowest income groups. In principle, assuming that other drivers of income inequality (such as capital incomes, taxes or public transfers) only partially offset the effects on gross labor income, as more productive workers will be rewarded with higher wages, we would expect an equalizing process in the distribution of income stemming from the increase in the income at the bottom part of the distribution. Thus, prima facie, the stability of the income Gini coefficient is puzzling given the large reduction in human capital inequality.

We provide different complementary explanations for the low correlation observed between the evolution of human capital and income inequality, primarily focusing on labor income inequality. Thus, we first compute the distribution of wages using recent estimates of rates of return to years of schooling for 139 countries provided by Montenegro and Patrinos (2014). When we correlate the human capital Gini and the wage Gini coefficients we find an inverted U-shape relationship between these two inequality indicators.[2] A plausible explanation for the inverted U-shape relationship is the composition effect resulting from the fall in the share of population with no schooling in a dual economy, as suggested by Robinson (1976), Knight and Sabot (1983) and Anand and Kanbur (1993). We find that the maximum Gini coefficient for wages is reached when the share of illiterates is 0.4. Thus, a first explanation for the low correlation between human capital and income inequality could be the non-linearity between the two inequality indicators.

Another explanation for the puzzle is that what matters for labor income is the interaction between returns to education and the level of human capital. The impact of the distribution of education on income inequality will depend not only on the size of the investments in education but also on the rate of return of these investments. Extensive empirical evidence, based on a Mincer (1974) earning function, shows that highly educated workers earn higher wages on average. Nevertheless, whether the returns increase or decrease with the level of education is still subject to debate. The traditional literature suggests the returns decrease with the level of schooling (e.g., Psacharopoulos and Patrinos, 2004), whereas more recent evidence shows that in many countries, the returns to education in the 1990s and 2000s are greater for higher education than for primary or secondary schooling (see, for example, Colclough et al., 2010, or Montenegro and Patrinos, 2014). In any case, as the distribution of wages is not a homothetic transformation of the distribution of years of schooling, the fall in the human capital Gini coefficient does not necessarily imply the fall of wages (and income) inequality, particularly when a large part of the reduction in schooling inequality is explained by the smaller share of illiterates.

[2] Lim and Tang (2008) and Morrison and Murtin (2013) find an inverted U-shape between their measure of human capital income inequality with respect to average years of schooling, which they call the human capital Kuznets curve.

We analyze the sensitivity of the Gini coefficient of the (simulated) labor income to changes in the share of illiterates and the convexity of education returns. We find that the type of returns to years of schooling does not significantly affect the inverted U-shape, especially in less developed countries where there has been a marked reduction in the share of the population with no schooling. Nevertheless, returns to education are relevant in more advanced economies. Thus, in countries with no illiterate population, the simulated wage Gini coefficient increases from 0.309 to 0.363 when returns change from decreasing to increasing returns.

It could also be the case that improvements in literacy, which increase the labor income of the population at the bottom end of income distribution, have coincided with an increase of wages in cohorts with higher education, such that all of them maintain their income shares. This could be the result of the interaction between exogenous forces such as skill-biased technological progress (e.g., Katz and Murphy, 1992) that have increased wages at the top. Using data on skill premia in a relatively large sample of countries from 2000 to 2014, we find evidence confirming that despite the increase in the relative supply of skilled workers, wages at the top are increasing due to skill-biased technological change. We find the earning gap between high-skill and low-skill labor has increased on average 1.7 percent each year.

Finally, we estimate the average contribution of wage inequality to income inequality, finding it to be statistically significant, relatively stable and economically relevant. The high correlation holds when we control for per capita income and its square value, to account for a Kuznets curve. Results suggest that approximately each point of change in the Gini coefficient of wages contributes to a half-point change in the Gini coefficient of income.

The structure of the paper is as follows. Section 2 computes the improved measures of human capital inequality and documents some stylized facts about the evolution of human capital inequality. Section 3 analyzes the distribution of income inequality from 1960 to 2005 and shows some disparities when compared with the evolution of human capital inequality. Section 4 analyzes the empirical support for the explanations we offer for the lack of correlation between the changes in income and education inequality. Section 5 estimates the contribution of wage inequality to total income inequality. Finally, section 6 contains the main conclusions.

2. Evolution of human capital inequality over time

Castelló and Doménech (2002) were the first to provide a comprehensive data set on human capital inequality, taking the educational attainment levels from Barro and Lee (2001) and calculating the Gini coefficient and the distribution of education by quintiles for a large number of countries and periods. However, some studies have shown that the Barro and Lee (2001) data set suffers from several problems. Cohen and Soto (2007) and de la Fuente and Doménech (2006) illustrate that the data show implausible time series profiles for some countries. More recently, Barro and Lee (2013) have addressed most of these concerns in an improved data set that reduces measurement error and improves the accuracy of the estimates by using more information from census data and a new methodology that makes use of disaggregated data by age group.[3] The old and the new measures of the average years of schooling are highly correlated in levels but there is little relationship when the variables are measured in differences. This suggests a lower measurement error in the new indicators derived from a smoother trend in the attainment levels. The attenuation bias in this context is relevant since measurement error is particularly prevalent in the econometric analysis that exploits the within-country variation of the educational data (see Krueger and Lindahl, 2001).

Using the new Barro and Lee (2013) data set, we have updated and expanded the inequality indicators to 146 countries from 1950 to 2010 for five-year time spans, thus obtaining 1898 observations. The data set covers most of the countries in the world, including data for 24 advanced economies, 19 countries in East Asia and the Pacific region, 20 countries in Eastern Europe and Central Asia, 25 countries in Latin America and the Caribbean, 18 countries in the Middle East and North Africa, 7 countries in South Asia, and 33 countries in Sub-Saharan Africa.[4] The large number of observations represents an advantage over other measures of inequality, such as the income Gini coefficient, which has limited coverage across countries and over the years.

We have also improved the methodology used to compute the new inequality indicators. Castelló and Doménech (2002) compute the human capital Gini coefficient with four educational levels. The limitation of that measure is that it does not account for the population that has begun but not completed a given level of schooling. This is particularly relevant in less developed economies with high student dropout rates. We have extended the methodology to include a broader set of educational levels that distinguish between complete and incomplete education. The Gini coefficient has been calculated as follows:

$$Gini^h = \frac{1}{2\overline{H}} \sum_{j=0}^{6} \sum_{j=0}^{6} \left| \widehat{x}_i - \widehat{x}_j \right| n_i n_j \tag{1}$$

where $\overline{H}$ is the average years of schooling in the population 15 years and over, i and j stand

[3] The new methodology in Barro and Lee (2013) fills in the missing observations by backward and forward extrapolation of the census data on attainment levels by age group with an appropriate lag. It also constructs new estimates of mortality rates and completion ratios by education and age group.

[4] We have followed the same classification of countries as in Barro and Lee (2013).

for different levels of education, $\hat{x}$ refers to the cumulative average years of schooling of each level of education, and n is the share of the population with a given level of education: no schooling (0), incomplete primary (1), complete primary (2), lower secondary (3), upper secondary (4), incomplete tertiary (5), and complete tertiary education (7).[5]

Table 1 shows the summary statistics for the average human capital Gini coefficient for some regions. The data show that the group of countries with the largest human capital inequality is South Asia, with an average human capital Gini coefficient equal to 0.676, followed by Sub-Saharan African (SSA) countries (average $Gini^h$ equal to 0.663), and the Middle East and the North African (MENA) region (average $Gini^h$ equal to 0.615). At the other end, the Eastern European and Central Asian countries (EECA) and the advanced economies are the regions where the average years of schooling are more evenly distributed. Lying in between these extremes, the Latin American and Caribbean countries (LAC) and the East Asian and the Pacific region (EAP) have average Gini coefficients of 0.421 and 0.452, respectively.

In spite of the large differences in the distribution of education across regions, there has been a general reduction in human capital inequality worldwide, as shown clearly in Figure 1. In most regions, the decline has been remarkable, with the Gini coefficients reduced by more than half. The general increase in the share of education going to the third quintile and the increase in the ratio of the bottom to the top quintile suggests the improvement in equality has mainly benefited the lowest part of the distribution.[6]

A further examination of the data reveals the large reduction in education inequality has mainly been due to a sizeable decline in illiteracy. Without exception, all the countries in the world that have experienced a great reduction in the share of illiterates also show a similar decline in the human capital Gini coefficient over time, suggesting the reduction in the Gini coefficient over time has been determined, to a great extent, by the decline in the share of illiterates. This fact can be explained by the weight of the share of illiterates in the computation of the human capital Gini coefficient.[7] To illustrate this point, we reorganize

[5] x_i is the duration in years of schooling of each education level, and the cumulative average years of schooling are computed as: $\hat{x}_0 \equiv x_0 = 0$, $\hat{x}_1 \equiv x_1$, $\hat{x}_2 \equiv x_1 + x_2$, $\hat{x}_3 \equiv x_1 + x_2 + x_3$, $\hat{x}_4 \equiv x_1 + x_2 + x_3 + x_4$, $\hat{x}_5 \equiv x_1 + x_2 + x_3 + x_4 + x_5$, $\hat{x}_6 \equiv x_1 + x_2 + x_3 + x_4 + x_5 + x_6$. Appendix A1 describes the procedure followed to compute duration in years of schooling of each education level from Barro and Lee˙s (2013) data set, and shows how the additional information provided by the larger number of educational levels increases the precision of the new Gini coefficient.

[6] We compute the ratio of the bottom to the top quintile as a measure of equality, instead of the top to the bottom quintile as a measure of inequality, since in many countries more than 60 percent of the population were illiterate and therefore the value of the bottom quintile in that case is equal to zero.

[7] Morrison and Murtin (2013) compute global trends in education inequality and show that inequality in years of schooling has mechanically decreased with the decline in the share of illiterates.

Table 1
Summary Statistics

	$Gini^h$										
	Countries	*Mean*	*Std. Dev.*	*Min*	*Max*	$Gini^h_{1950}$	$Gini^h_{2010}$	$3^{rd}Q^h_{1950}$	$3^{rd}Q^h_{2010}$	$1^{st}/5^{th}Q^h_{1950}$	$1^{st}/5^{th}Q^h_{2010}$
World	146	0.412	0.251	0.026	0.997	0.557	0.257	0.202	0.420	0.096	0.278
Advanced	24	0.212	0.116	0.049	0.827	0.242	0.156	0.425	0.499	0.371	0.421
EAP	19	0.385	0.193	0.097	0.923	0.588	0.230	0.159	0.448	0.009	0.262
EECA	20	0.207	0.124	0.026	0.611	0.331	0.099	0.370	0.541	0.173	0.579
LAC	25	0.338	0.169	0.048	0.915	0.456	0.217	0.269	0.457	0.059	0.309
MENA	18	0.575	0.241	0.145	0.997	0.808	0.313	0.046	0.385	0.001	0.133
South Asia	7	0.641	0.227	0.150	0.988	0.780	0.414	0.061	0.279	0.000	0.049
SSA	33	0.614	0.219	0.128	0.963	0.798	0.407	0.027	0.296	0.000	0.105

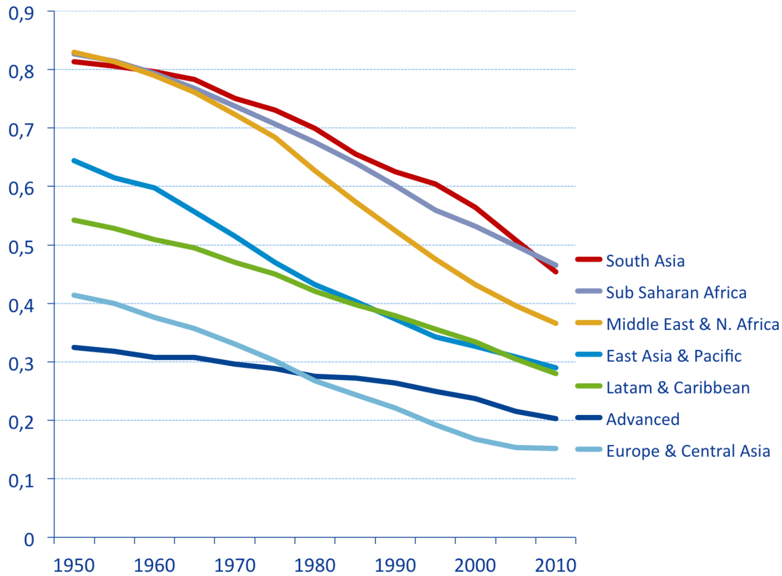

Figure 1: *Human Capital Gini Coefficient of population 15+.*

equation (1) as follows:

$$Gini^h = n_0 + \frac{A}{\overline{\overline{H}}} \tag{2}$$

where:

$$A = \sum_{i=1}^{6}\sum_{j=1}^{6} \left|\widehat{x}_i - \widehat{x}_j\right| n_i n_j$$

The Gini coefficient of education is, therefore, a proportional measure of the share of illiterates in the society. A great reduction in the share of illiterates will translate into a similar reduction in the Gini coefficient. Whether the reduction in the Gini coefficient is greater or smaller than that in the share of illiterates will depend on the changes in the distribution of education among the literates. Given that:

$$Gini^{LIT} = \frac{1}{2\overline{H}^{LIT}} \sum_{i=1}^{6}\sum_{j=1}^{6} \left|\widehat{x}_i - \widehat{x}_j\right| n_i^{LIT} n_j^{LIT} \tag{3}$$

where $Gini^{LIT}$ is the human capital Gini coefficient among the literates, $n_i^{LIT} = n_i/(1 - n_0)$ and n_0 is the share of population with no education, equation (3) can be rewritten as follows:

$$Gini^{LIT} = \frac{1}{(1-n_0)} * \frac{A}{\overline{\overline{H}}} \tag{4}$$

Then, the human capital Gini coefficient can be formally decomposed as a combination of

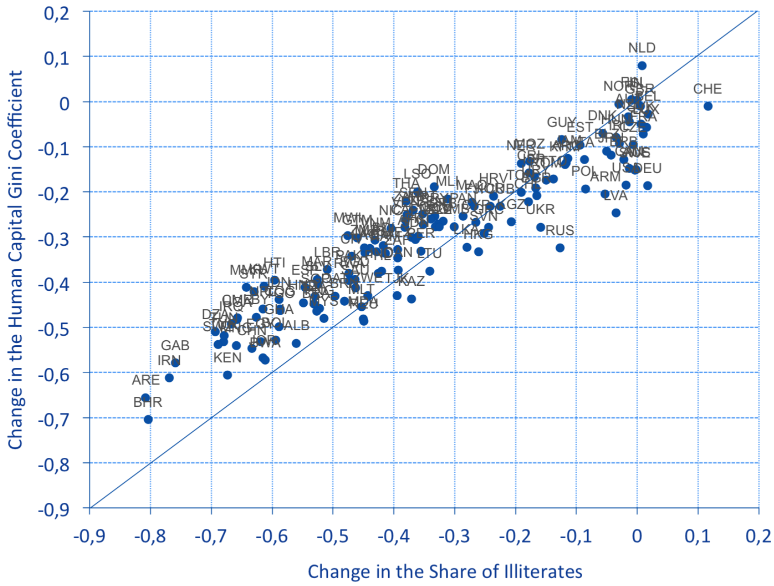

Figure 2: *Change in the human capital Gini coefficient and in the share of illiterates, 1950-2010.*

the share of illiterates and the Gini coefficient among the literates in the following way:

$$Gini^h = n_0 + (1 - n_0)Gini^{LIT} \tag{5}$$

When the share of illiterates is very high, the evolution of the human capital Gini coefficient is mainly determined by the share of illiterates. This fact is illustrated in Figure 2, which plots the relationship between the changes in the distribution of education and the changes in the share of illiterates for the whole sample of countries from 1950 to 2010. In less developed countries, where the share of illiterates in 1950 was high, the reduction in the Gini coefficient has been very similar to the reduction in the share of illiterates, with most of those countries located close to the diagonal line. In the advanced economies, where the share of illiterates is almost zero, the distribution of primary, secondary and tertiary education is what determines the evolution of education inequality.

We can use the previous expression to analyze the contribution of the share of illiterates to the changes in $Gini^h$ from 1950 to 2010:

$$Gini^h_{2010} - Gini^h_{1950} = (n_{0,2010} - n_{0,1950}) + (1 - n_{0,n_{0,2010}})Gini^{LIT}_{2010} - (1 - n_{0,n_{0,1950}})Gini^{LIT}_{1950}$$

In our sample of 146 countries, the average reduction in $Gini^h$ is 0.3 (from 0.557 in 1950 to 0.257 in 2010), whereas the average reduction in the share of illiterates is 0.34. Therefore, the change in n_0 explains on average 114 per cent of the change in $Gini^h$.

Taken together, we can summarize all the preceding evidence in the following stylized facts. First, from 1950 to 2010 there was a significant reduction in human capital inequality around the world. Second, in less developed countries the marked reduction in education inequality has mainly been due to the sizeable decline in the share of illiterates. Third, given their low initial shares of illiterates, in most advanced countries there is no clear pattern in the evolution of education inequality, and the human capital Gini coefficient has been mainly determined by the distribution of education among the literates.

3. Human Capital and Income Inequality

In this section, we analyze the extent to which the reduction in human capital inequality, explained in most countries by the increase in literacy, has been accompanied by a similar change in the distribution of income inequality. We start by comparing the mean values of the human capital and income Gini coefficients for those countries with available income inequality data. We measure income inequality through the net income Gini coefficient taken from the Standardized World Income Inequality Database (SWIID), version SWIID v3.0, which uses a custom missing-data algorithm to standardize WIID from the LIS data set.[8] The data include 75 countries with observations from 1960 to 2005. Table 2 displays the mean values of the income and human capital Gini coefficients for several regions in the world. The table illustrates two main features.

Firstly, if we compare the average value of the income ($Gini^y$) and the human capital ($Gini^h$) Gini coefficients, we observe that the countries with the highest and the lowest inequality in the distribution of income and education do not coincide. The most remarkable example is that of Latin America and the Caribbean, which is one of the regions with the highest income inequality but only moderate inequality in the human capital distribution. At the other extreme, countries in South Asia display high inequality in the distribution of education but relatively low inequality in the distribution of income. We get a similar picture when we look at a large cross-section of countries, as in Figure 3, which shows the correlation between the income and the human capital Gini coefficients in 2005. Several key points emerge from this figure. First, in most countries, income inequality was greater than human capital inequality. Thus, only 35 of 124 countries are below the diag-

8 Most of the studies that have analyzed the determinants and the effects of income inequality have used the UNU/WIDER-UNDP World Income Inequality Database (WIID), which is an updated version of Deininger and Squire's (1996) data set and reports income inequality measures for developed as well as developing economies. However, there are concerns about the poor quality of income inequality data covering multiple countries due to problems of cross-country comparability and the incompleteness of coverage across countries and over time (e.g., Atkinson and Brandolini, 2001). The most reliable data set on income inequality is the Luxemburg Income Study (LIS), which provides improved data for income inequality measures in terms of their quality and comparability across countries. Nevertheless, the main drawback of the LIS data set is that it only contains data for a reduced sample of advanced economies, mostly starting in 1980, which reduces the sample size considerably.

Table 2
Summary Statistics

	Countries	$Gini^h$	$Gini^h_{1960}$	$Gini^h_{2005}$	$Gini^y$	$Gini^y_{1960}$	$Gini^y_{2005}$
World	75	0.352	0.454	0.243	0.373	0.381	0.379
Advanced	22	0.212	0.233	0.160	0.294	0.307	0.308
East Asia and the Pacific	9	0.341	0.524	0.225	0.377	0.397	0.382
Europe and Central Asia	6	0.139	0.181	0.085	0.247	0.241	0.293
Latin America and the Caribbean	15	0.300	0.388	0.214	0.470	0.481	0.463
Middle East and North Africa	7	0.582	0.794	0.358	0.395	0.397	0.384
South Asia	4	0.578	0.698	0.411	0.330	0.321	0.371
Sub-Saharan Africa	12	0.577	0.744	0.399	0.478	0.461	0.448

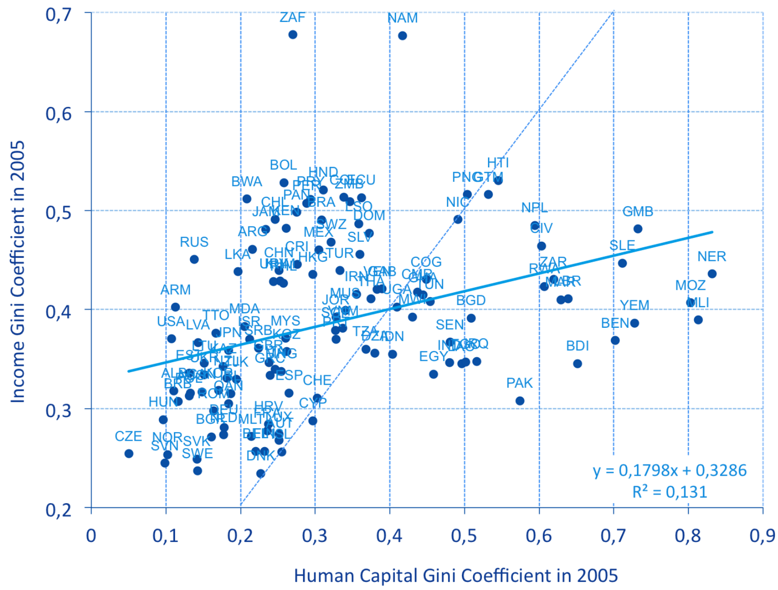

Figure 3: *Human capital and income Gini coefficients across countries in 2005.*

onal representing equal values for human capital and income Gini coefficients. Second, the variance in education inequality is higher than in income inequality. The income Gini coefficient in 2005 takes values between 0.235 and 0.678, whereas the minimum and maximum values of the human capital Gini coefficient are 0.050 and 0.832, respectively. Most of the countries with the highest income inequality are located in Latin America, and those with the highest education inequality are in Africa and South Asia. On average, the advanced economies have the lowest inequality in the distribution of education and income. However, there are some advanced economies, such as USA or Japan, with income Gini coefficients higher than those in India, Pakistan and Egypt. Third, the correlation between the income and the human capital Gini coefficients is not very high (0.362).

The second key feature to emerge from Table 2 is that education and income inequality have evolved in a different manner over last decades. The last two columns of Table 2 display the values of the income Gini coefficient at the beginning and at the end of the sample period. The data indicate that the income Gini coefficient has remained quite stable over a period of 45 years. This evidence is illustrated in Figure 4, which plots the evolution of the income Gini coefficient for all the regions and available time periods. An interesting feature is that, in spite of some variations over short periods of time, in most of the regions the income Gini coefficient in 2005 was very similar to what it had been in

1960, reflecting the long-term stability of the income Gini coefficient despite the significant reduction in human capital inequality. While Figure 1 shows a notable reduction in education inequality over time, mainly due to a reduction in the illiterate population, Figure 4 indicates the inequality in the distribution of income has scarcely changed.

This evidence is corroborated in Figure 5, which highlights the absence of correlation between the change in income and human capital Gini coefficients in a sample of 75 countries from 1960 to 2005. Even though there are some countries in which both income and education inequality have increased (e.g., Norway, Great Britain and the Netherlands) and others where both variables have decreased (e.g., Kenya, Taiwan, Senegal and Colombia, among others), in a large number of countries changes in income and education inequality display a non-significant correlation. For example, in countries such as China, India, Singapore, USA, Argentina, Australia and many others, there has been a reduction in the inequality of education distribution and an increase in the inequality of income distribution. Additionally, as Appendix 2 shows, we have also checked that the reduction in the number of illiterates and, therefore, in education inequality, has not increased the share of income accruing to the bottom quintile.

In sum, the past decades have witnessed a marked decline in education inequality, driven by the decrease in the illiterate population, which has not been accompanied by similar reductions in income inequality. Both across world regions and a large sample of countries, income inequality has remained relatively stable over a period of 45 years. The evidence presented in this section suggests the correlation between the evolution of income and education inequality across countries and over time is weak, since other factors may have offset the effects of the reduction in human capital inequality on income inequality.

4. Explaining the low correlation between changes in income and education inequality

4.1 Factors other than human capital

Before explaining in more detail the effects of changes in human capital inequality on income inequality, it is necessary to account for the fact that income is also affected by factors other than human capital. As an illustration, we decompose total income into its labor (where wages, w, depend on human capital, H) and physical capital (i.e., the product of capital, K, and its return, r) components, as well as taxes (T) and public transfers (TR):

$$Y = w(H) + rK + TR - T \tag{6}$$

Given the previous expression it is clear that components other than human capital may

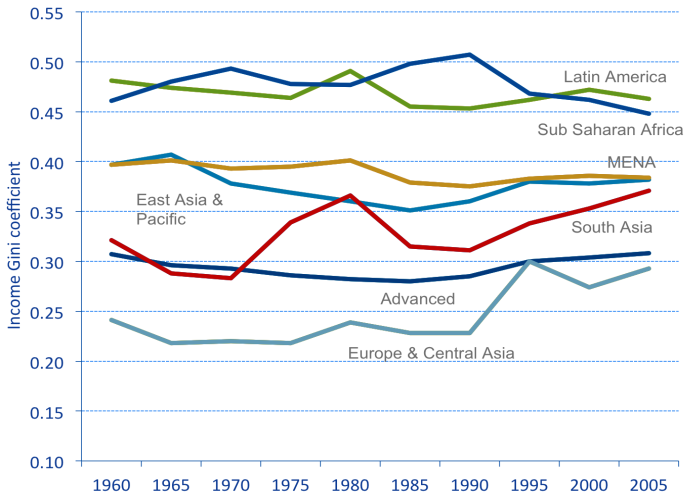

Figure 4: *Evolution of the income Gini coefficient across regions, 1960-2005.*

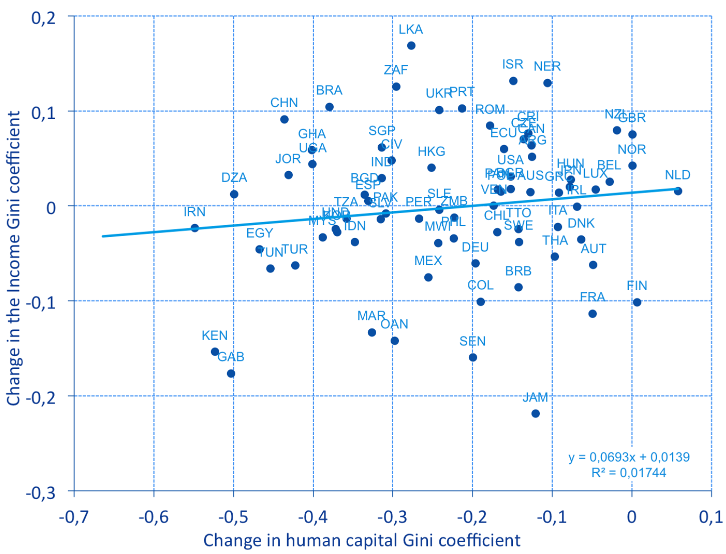

Figure 5: *Change in income and human capital Gini coefficients across 75 countries, 1960-2005*

explain changes in income inequality.

We focus first on the relevance of the last two components, comparing the evolution of the income Gini coefficient before and after taxes and transfers. The correlation between the levels of the two variables is very high, but falls to 0.69 in terms of their changes between 1960 and 2005. Whereas the gross income Gini coefficient has fallen by 4.7 pp, the corresponding net value has remained almost constant (-0.02 pp). As a result, the correlation between the changes in human capital and income Gini coefficients is slightly higher for the gross (0.25) than for the net income (0.13) definition.

Changes in the income Gini coefficient are also affected by composition effects of changes in the distribution of labor and capital income. Until recently, the stability of the labor share of income was considered a stylized fact in macroeconomics. However, Karabarbounis and Neiman (2013) have challenged this fact. They show evidence of a decline in the labor share in a large majority of countries and industries since 1975. At the same time, although the evidence is very scarce due to the lack of data on asset distribution for a broad set of countries, Piketty and Zucman (2014) have documented an increase in wealth-to-income ratios since the 1970s in the top eight developed economies. And for other countries for which we have only recent data and no evidence about the changes over time, the distribution of wealth is more unequal than the distribution of income . Additionally, Checchi and García-Peñalosa (2010) have shown that the labor share is negatively correlated with the income Gini coefficient. Although this potential explanation is beyond the scope of this paper given the lack of data, the increasing relevance of capital income in many countries suggests that the composition effects of income shares may be a potential explanation for the lack of correlation between the changes in human capital and income inequality.

In the rest of the paper we focus on how the changes in the distribution of human capital have affected the distribution of the labor income component, $w(H)$, taking into account the changes over time in the returns to education.

4.2 Wage inequality in dual economies and returns to education

To analyze the correlation between human capital and labor income inequality we need data on wages for different levels of schooling. Since these data are not available for many countries, in this section we estimate wages for seven levels of schooling and compute a Gini coefficient for wage inequality. Following the pioneering research by Mincer (1974), there has been a large number of contributions showing the close relationship between wages (w) and years of schooling (H). Among this literature, the recent paper by Montenegro and Patrinos (2014) reports estimated rates of return to years of schooling for 139 countries since the late 1950s. These authors first estimate the following wage equation for

139 countries:

$$\ln w_i = \alpha + \beta_P D_{P,i} + \beta_S D_{S,i} + \beta_T D_{T,i} + \beta_1 X_i + \beta_2 X_i^2 + \mu_i$$

where D_P, D_S and D_T are dummy variables for primary, secondary and tertiary schooling, respectively, and X is labor market age experience. Given the estimated values of β_P, β_S and β_T, the rates of returns are derived from the following expressions:

$$\begin{aligned} r_P &= \beta_P / S_P \\ r_S &= (\beta_S - \beta_p)/(S_S - S_P) \\ r_T &= (\beta_T - \beta_S)/(S_T - S_S) \end{aligned}$$

where S_P, S_S and S_T are years of primary, secondary and tertiary schooling, respectively. For the 139 countries in the sample the average returns are $\overline{r_P} = 0.106$, $\overline{r_S} = 0.072$, and $\overline{r_T} = 0.152$.

Since Barro and Lee (2013) provide data for completed (C) and incomplete (I) years of schooling, assuming that the average years of schooling for each level of education is $S_{IP} = 3$, $S_{CP} = 6$, $S_{IS} = 3$, $S_{CS} = 6$, $S_{IT} = 2$ and $S_{CT} = 4$, we compute wages for each level of human capital in each country as:

$$\begin{aligned} \ln w_{NS,i} &= \alpha_i \\ \ln w_{IP,i} &= \alpha_i + 3r_{P,i} \\ \ln w_{CP,i} &= \alpha_i + 6r_{P,i} \\ \ln w_{IS,i} &= \ln w_{CP,i} + 3r_{S,i} \\ \ln w_{CS,i} &= \ln w_{CP,i} + 6r_{S,i} \\ \ln w_{IT,i} &= \ln w_{CS,i} + 2r_{T,i} \\ \ln w_{CT,i} &= \ln w_{CS,i} + 4r_{T,i} \end{aligned}$$

Note that we consider heterogeneity across countries in their rates of returns, although we assume that there is no time variability and use the specific time average for each country. Once we have simulated the wage for each education level in every country, we can compute the Gini coefficient of wages and compare it with the Gini coefficient of human capital. Bear in mind that this is just a crude approximation to the real distribution of wages, since we are considering only between-group and not within-group differences in wages in the seven groups of schooling levels.

Figure 6 plots the human capital and the wage Gini coefficients, that is, $Gini^h$ and $Gini(W^s)$ respectively. The figure shows a clear inverted U-shape relationship between the two Gini coefficients for 139 countries from 1950 to 2010, in five-year time spans. When

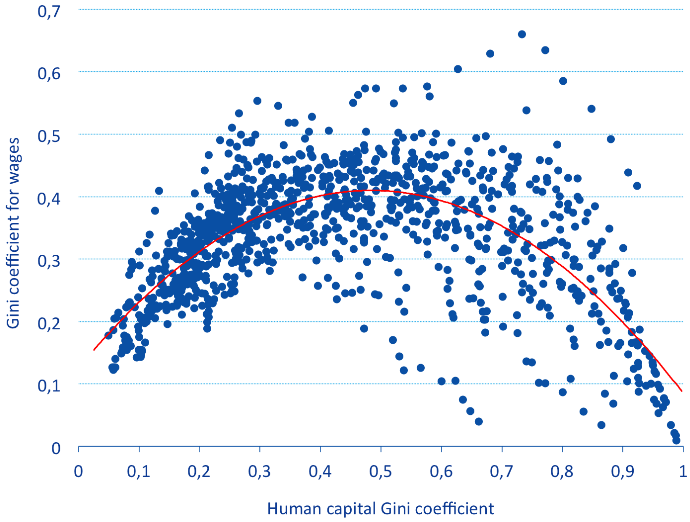

Figure 6: *Gini coefficients for human capital and simulated wages, 139 countries, 1950-2010.*

education inequality is high, labor income and education inequality correlate negatively. When education inequality is low, the two inequality indicators are positively related

The main explanation for the inverted U-shape relationship shown in Figure 6 is the composition effect of the share of population with no schooling. As we have shown in the previous section, in most countries the fall in the human capital Gini coefficient is explained by the fall in the share of population with no schooling (n_0), whereas the inequality within the other six levels of education (from incomplete primary to completed tertiary) only plays a role in the group of more advanced economies. However, the effect of the fall in n_0 on the Gini coefficient for wages is non-linear for the same reason that Robinson (1976), Knight (1976), Knight and Sabot (1983), Anand and Kanbur (1993) or Fields (1993) have demonstrated: in an economy with two population groups, a transfer of workers from the low- to the high-education group raises the inequality of wages until the high-education group reaches a certain share. This share depends on the difference in the means and the variances of wages between the two groups.

Anand and Kanbur (1993) also showed that, when there is no within-sector inequality in the two population groups, commonly-used inequality measures start at zero, then increase continuously to an interior maximum, before finally dropping to zero. Remember that, by construction, the group with no schooling (n_0) has no within-sector wage inequality, whereas the second group with at least incomplete primary education ($1 - n_0$) exhibits within-group inequality, as a result of the differences in wages between the six groups of population with some education. Therefore, in our case the Gini coefficient for wages

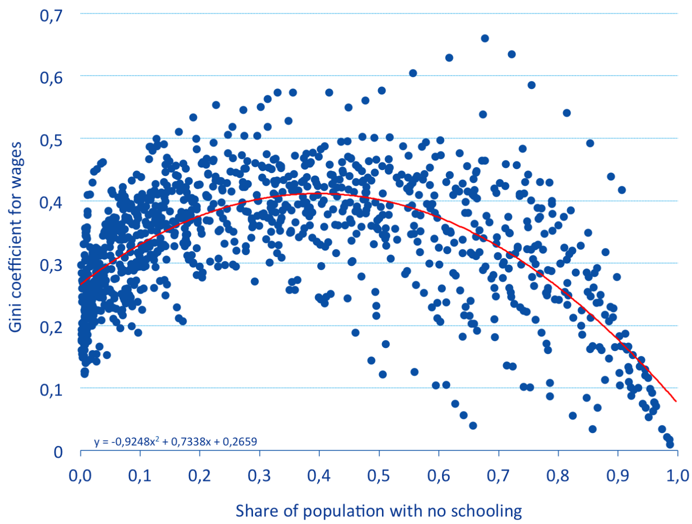

Figure 7: *Share of population with no schooling and Gini coefficient for simulated wages, 139 countries, 1950-2010.*

would start at zero when $n_0 = 1$, would reach an interior maximum and would fall to the Gini coefficient for the six groups with some education (with an average equal to 0.266) when $n_0 = 0$. As Figure 7 shows, when we fit a second-order equation for the sample of countries, the maximum Gini coefficient for wages is reached when $n_0 = 0.4$.

Lim and Tang (2008) also estimate a Mincerian measure of human capital income from 1960 to 2000, assuming the same world average returns that decrease with the level of education. They show that the Gini coefficient for their measure of human capital income has an inverted U-shape curve with respect to average years of schooling, which they call the human capital Kuznets curve. We have extended their results to our sample of 139 countries allowing for country heterogeneity of returns and we have corroborated their inverted U-shape curve, but with a significantly higher number of average schooling years: 5.5 instead of the 4.2 years that they find. Morrison and Murtin (2013) also report a human capital Kuznets curve along educational development for 32 macro-countries over the period 1870-2010, imposing homogeneity of returns across countries and using only four levels of education.

How sensitive are the results to changes in the rates of return to years of education? When we assume the returns are the same for all countries and education levels, and equal to 0.101 (the average return to schooling estimated by Montenegro and Patrinos) we obtain a very similar inverted U-shape relationship between the two Gini coefficients as in Figure 6. In this case, the maximum Gini coefficient for the fitted second-order equation is obtained when $n_0 = 0.35$.

In order to provide a more detailed analysis, we have simulated the extent to which the inverted U-shape relationship between the share of population with no schooling and the Gini coefficient for simulated wages is affected by the type of returns. In Figure 8 we present the results of this simulation. The variable in the x-axis is n_0.[9] Education returns vary from decreasing ($\overline{r_P} = 0.10$, $\overline{r_S} = \overline{r_P} - 0.05/2$, and $\overline{r_T} = \overline{r_P} - 0.05$) to increasing ($\overline{r_P} = 0.10$, $\overline{r_S} = r_P' + 0.05/2$, and $\overline{r_T} = \overline{r_P} + 0.05$). The vertical axis represents the different values of $Gini(W^s)$ obtained using the approach described in the previous subsection.

Three main results can be obtained from Figure 8. First, the type of returns to years of schooling does not affect the inverted U-shape, which is dominated by the composition effects driven by no schooling. Second, increasing returns drive up inequality, particularly when n_0 is equal to zero: $Gini(W^s)$ increases from 0.309 to 0.363, when returns change from decreasing to increasing. By definition, when n_0 is equal to one then $Gini(W^s) = 0$ and returns to years of schooling have no effect on inequality. Therefore, the effects of increasing returns to education on inequality will be greater in advanced economies than in developing countries. Third, going from decreasing to increasing returns reduces slightly the value of n_0 at which $Gini(W^s)$ reaches its maximum level.

Note that, while the share of population with no education is still large, the increase in wage inequality as a result of the transfer of workers from the low- to the high-education group is, according to Fields (1979), a statistical artifact and not an economically meaningful worsening of the income distribution. On the contrary, this is a transitory effect of an economic development process that is good in absolute income terms and that reverts when n_0 falls sufficiently and more people are educated, completing at least primary schooling. Eradicating illiteracy and completing primary schooling are, therefore, necessary conditions for the subsequent improvement of per capita income and equality, showing that there is no trade-off between them. It is not a sufficient condition because, as discussed before, other factors such as the increase in the capital income share, wealth inequality or a less redistributive fiscal system may more than offset the fall in wage inequality. As we analyze in the next subsection, skill-biased technological changes may prevent the fall in wage inequality despite the improvement in education, therefore also explaining why the fall in human capital inequality has not been accompanied by a reduction in income inequality. This is a particularly interesting hypothesis when developing countries become advanced economies, where n_0 is close to zero and the composition ef-

[9] The shares of population for the different levels of education with some schooling ($n_i, i = 1, ..., 6$) have been simulated according to the fitted value obtained from regressing n_i on a quadratic function of $(1 - n_0)$:

$$\widehat{n_i} = \widehat{\alpha_i}(1 - n_0) + \widehat{\beta_i}(1 - n_0)^2$$

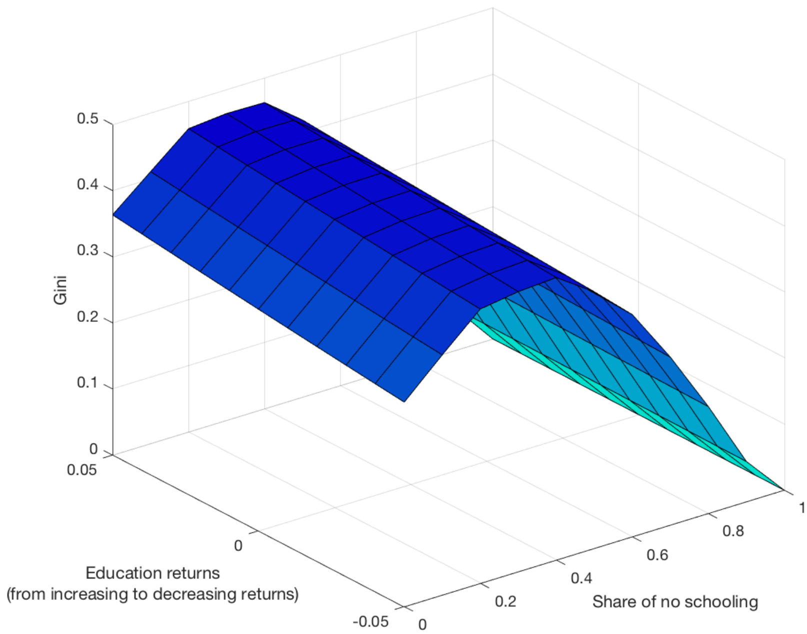

Figure 8: *Sensitivity of* $Gini(W^s)$ *to changes in the share of population with no schooling and the type of returns to years of schooling.*

fect discussed above does not come into play.

4.3 Skill-biased technological change

Despite the increase in the supply of educated workers, the demand for skills could have kept pace with the human capital investment, so that wage dispersion might have remained unchanged. The reasoning behind this potential explanation is represented in Figure 9. In this example, an individual (or quintile) with no education in t becomes educated in $t+1$. However, the increase in their income is the same as for an individual with high education, who benefits from the increase in wages due to skill-biased technological change (SBTC). As a result, although there is a reduction in human capital inequality, income inequality remains unchanged, helping to explain the puzzle documented in Section 3.

The "canonical model" of the race between education and technological change (e.g., Katz and Murphy, 1992, Card and Lemieux, 2001, or Acemoglu and Autor, 2012, among others) provides a well-founded explanation of the effects of skill-biased technological

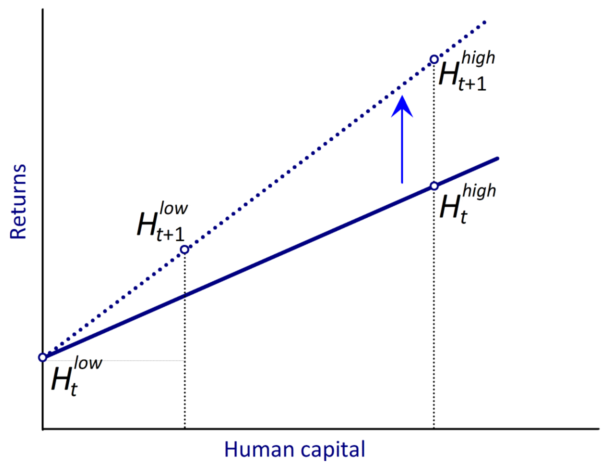

Figure 9: *Skill-biased technological change and human capital*

change. The motivation behind this literature is the observation that in the United States and other developed countries, in spite of the growing supply of college graduate workers, there has been an increase in wage inequality, proxied by the increase in the wage of college graduate workers relative to the wages of high school graduates. This model argues that the returns to skills are determined by a race between the demand for skills, driven by a skill-biased technological change, and the increase in the supply of skills. When the relative demand increases faster than the relative supply, wage dispersion rises. Conversely, when the supply outpaces the demand, wage dispersion decreases.

To test the canonical model of the race between education and technological change, we rely first on the information provided by the OECD in *Education at a Glance* (2015) for a sample of 33 countries from 2000 to 2013, which includes some emerging countries such as Brazil, Mexico and Turkey. The evidence presented in Figure 10 for the sample average is consistent with skill-biased technological change. Despite the massive 107 percent increase in the share of the adult population with tertiary education with respect to those with below upper secondary education, relative earnings have also increased by around 16 percent in 14 years, from 1.85 to 2.15.

Following the seminal work of Katz and Murphy (1992), we relate the earning gap or skill premium (i.e., the wage ratio of skilled and unskilled workers w_{H_t}/w_{L_t}) to the relative supply of skills (H/L) and the relative technology trend (A_H/A_L), proxied by a

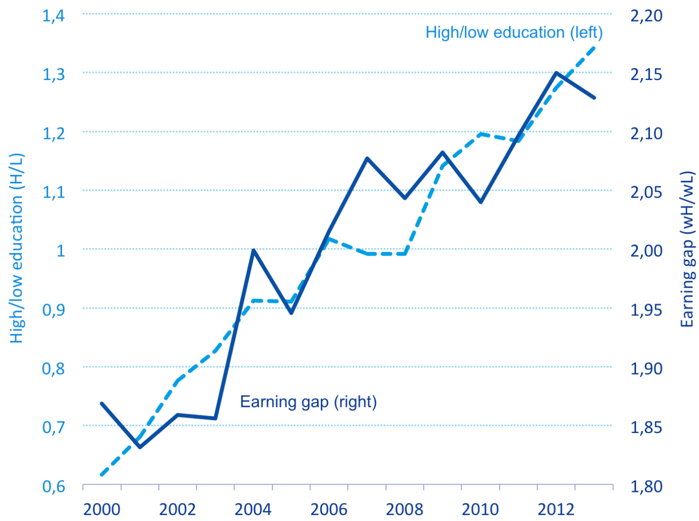

Figure 10: *Relative supply of population with high education with respect to population with below upper secondary education (left axis) and the earning gap (right axis). Average for 33 countries, 2000-2013.*

time trend (t):[10]

$$\ln\left(\frac{w_{H_t}}{w_{L_t}}\right) = \frac{\sigma - 1}{\sigma}\gamma_0 + \frac{\sigma - 1}{\sigma}\gamma_1 t - \frac{1}{\sigma}\ln\left(\frac{H_t}{L_t}\right) \tag{7}$$

where σ is the elasticity of substitution between high-skill and low-skill labor, H refers to the share of population 25-64 years old with tertiary education, and L is the share of population with below upper secondary education.

The estimated coefficients for an unbalanced sample of 33 countries from 2000 to 2013 (293 observations) are displayed in the first column of Table 4. As we can see, the relative supply H/L enters with the correct sign. The elasticity of substitution between the population with tertiary education and those with below upper secondary is about 8.58 (■ $= 1/0.117$). The coefficient of the time trend is 0.017, that is, the earning gap has increased by an average of 1.7 percent each year.

We check the robustness of the results to alternative indicators for the demand for skilled labor. In column (2) we add the log of the share of high-technology exports in total manufactured exports (X^{high}/X), taken from the World Development Indicators. We expect that a greater production of high-technology goods is likely to increase the demand

[10] It is common in the literature to assume that there is a log-linear increase in the demand for skills over time coming from technology, captured as follows:

$$\ln\left(\frac{A_{H,t}}{A_{L,t}}\right) = \gamma_0 + \gamma_1 t$$

Table 4
Dependent Variable: $\ln \frac{w_H}{w_L}$

	Reduced sample		Extended sample		Less developed	
	(1)	(2)	(3)	(4)	(5)	(6)
$\ln \frac{H}{L}$	-0.117^a	-0.155^a	-0.098^a	-0.108^a	-0.132^a	-0.136^a
	(0.024)	(0.026)	(0.017)	(0.018)	(0.020)	(0.020)
t	0.017^a	0.022^a	0.005^c	0.009	-0.008^b	-0.006
	(0.004)	(0.007)	(0.003)	(0.005)	(0.004)	(0.007)
$\ln \frac{X^{high}}{X}$		0.108^a		0.038^a		0.039^a
		(0.025)		(0.010)		(0.011)
R^2	0.155	0.217	0.320	0.339	0.284	0.330
Obs.	293	291	508	504	280	278
N	33	33	55	55	32	32

Notes: OLS Regression. Robust standard errors in parenthesis. a, b and c are 1, 5, and 10 percent significance levels respectively. Regression from 2000 to 2013.

for high-skill labor. The effect is statistically significant: the results show that a higher share of exports in high-technology goods and services is related to a higher skill premium. Additionally, the coefficient and the statistical significance of the time trend are slightly higher and the coefficient of $\ln H/L$ is in line with the estimates of the previous column.

A concern with previous estimates is that most of the countries in columns (1) and (2) are advanced economies, and the only countries in the sample considered middle income are Mexico and Brazil. In the remaining columns, we check whether SBTC could also be an explanation for the lack of correlation between human capital and income inequality in less developed economies. In column (3) we extend the sample with the addition of 22 emerging countries for which the wage gap from 2000 to 2013 has been computed, using the education returns estimated by Montenegro and Patrinos (2014). We also include a dummy variable equal to 1 for the new countries added. As we can see, the coefficient of lnH/L is similar to the one in column (1), but the coefficient of the time trend is now lower (0.005), although still statistically significant at the 10 percent level. In columns (5) and (6) we restrict the sample to less developed economies.[11] The evidence suggests that SBTC could have offset the effect of the fall in human capital inequality in less developed countries as well. On the supply side, as expected, we find the relative wage of high-skilled workers varies positively with their relative supply. On the demand side, whereas the coefficient on the trend is negative, a larger share of high-technology exports in total manufactured exports, which is likely to increase the demand for skilled workers, has a positive and statistically significant effect on the skill premium.

[11] This sample includes all countries in the extended sample minus the advanced economies.

5. Simulated wages and income inequality

We complete the analysis of the relationship between years of schooling and income inequality with an estimation of the contribution of the simulated wages inequality to total income inequality.[12]

The literature on inequality provides different methods to compute the contribution of a particular component of income, factor or subgroup of population to income inequality (see, for example, the review by Cowell and Fiorio, 2011). Most of these decompositions, such as those proposed by Shorrocks (1982) and Fields (2003), compute the absolute contribution (S_j) of a particular income source or factor (Y_j) according to a weight (s_j) such that:

$$S_j = s_j I(Y)$$

and

$$\sum_j s_j = 1$$

where $I(Y)$ in an inequality measure of total income (Y) that, as in the case of the Gini coefficient, is symmetric, continuous and equal to zero if incomes are equal for all individuals in the distribution.[13]

Although these methods are very useful, they do not relate the inequality of a particular factor or component to the inequality of total income. More relevant to our case is the type of decomposition proposed by Fei, Ranis, and Kuo (1978) and Pyatt, Chen, and Fei (1980), who decompose total income inequality in terms of the inequality distributions of its components. In the case of the Gini coefficient of total income, it can be decomposed as:

$$Gini(Y) = \sum_j \phi_j R_j Gini(Y_j) \tag{8}$$

where $Gini(Y_j)$ is the Gini coefficient of income source Y_j, ϕ_j is the share of income from factor j in total income and R_j is the rank correlation ratio:

$$R_j = \frac{Cov(Y_j, F_y)}{Cov(Y_j, F_j)}$$

that is, the correlation coefficient between Y_j and the ranking of Y,where F_j and F_Y are the cumulative distribution of Y_j and Y respectively. The product of the Gini coefficient of Y_j

[12] Both Lim and Tang (2008) and Morrison and Murtin (2013) have analyzed the relationship between years of schooling and the distribution of simulated wages, but not with respect to income inequality. Conversely, Fölvári and van Leeuwen have analyzed the relationship between years of schooling and income inequality, without estimating a distribution of the Mincerian human capital income, obtaining a (non-inverted) U-curve from 1950 to 2000. Only when they instrument the Gini coefficient of years of education, taking into account the effect of the unobserved skill premium, do they find an inverted U-curve from 1950 to 2000.

[13] According to Shorrocks (1982), the weight of a particular factor is given by $s_j^S = \frac{cov(Y_j,Y)}{\sigma^2(Y)}$.

and its rank correlation ratio is usually denominated the pseudo-Gini coefficient of income from factor j or the concentration ratio.

Two problems arise when applying this method to our case. First, we do not have information on individual wages in different countries and periods. As discussed previously, they have been approximated using a Mincerian equation that relates education returns and years of schooling in different groups of populations. Therefore, for each country i and each year t, simulated wages (W^s) are just an approximation of true wages (W), such that

$$W_{it}^s = W_{it} + \varepsilon_{it}$$

where ε is a measurement error, which implies that $Gini(W_{it})$ is not necessarily equal to $Gini(W_{it}^s)$. Second, both ϕ_j and R_j vary across countries (i) and years (t).

According to equation (8), the contribution of the Gini coefficient of wages to income inequality is given by

$$\phi_{wit} R_{wit} Gini(W_{it})$$

The empirical evidence presented by Deutsch and Silber (2004) for 23 countries between 1983 and 1990 is that R_w is close to one, with an average equal to 0.992, ranging from 0.938 in Rwanda to 1.002 in Pakistan. Meanwhile, wages are the most important source of income: the average value of ϕ_w is 0.542, ranging from 0.105 in Rwanda (where the most important income source is entrepreneurial income) to 0.940 in Japan.[14] The evidence also shows that the share of wages has a high correlation with the log of per capita income (equal to 0.654) whereas entrepreneurial income exhibits a negative correlation (−0.776). According to Deutsch and Silber, these correlations and composition effects can explain the Kuznets curve: the rising section of the curve is mainly the consequence of the increasing share of wages, whereas the declining section is explained by the decreasing share of entrepreneurial income and the increasing role played by public transfers, which more than compensate for the rising contribution of property income inequality.[15]

Given the lack of data for ϕ_{wit} and R_{wit} in our large sample of countries and periods, we cannot compute an exact decomposition of $Gini(Y)$. Instead, we approximate the contribution of $Gini(W_{it}^s)$ to $Gini(Y_{it})$ through the estimation of the following equation:

$$Gini(Y_{it}) = \alpha + \beta_t Gini(W_{it}^s) + \lambda_t Gini(W_{it}^s) \ln y_{it} + \delta_t + u_{it} \tag{9}$$

[14] Using the OECD Income Distribution Database, the share of all sources of employee income in mean disposable income in a sample of 36 countries from 2004 to 2013 is 83%.

[15] According to the influential work by Kuznets (1955), inequality in the distribution of income first increases and then decreases as per capita income rises. In the first stages of development, per capita income as well as income inequality increase, since there is a movement of population from the agricultural sector, characterized by low per capita income and low inequality, to an industrial sector in which per capita income and income inequality are higher. Subsequently, when most of the population is in the industrial sector, per capita income increases and income inequality is reduced.

assuming that

$$\phi_{wit} R_{wit} \simeq \beta_t + \lambda_t \ln y_{it} \tag{10}$$

and

$$Gini(W_{it}) = Gini(W_{it}^s) + u_{it}^W$$

where y_{it} is per capita income. Note that the country heterogeneity of ϕ_{wit} is approximated by $\ln y_{it}$, as suggested by the results of Deutsch and Silber (2004).

Although, given the lack of data, equation (9) gives us an indirect approximation of the contribution of inequality in years of schooling to income inequality, it should be note that OLS estimates of β and λ could be biased if the residuals in (9) are correlated with $Gini(W_{it}^s)$. This could be the case if omitted variables (e.g., the Gini coefficients of other sources of incomes) are correlated with $Gini(W_{it}^s)$. Additionally, measurement errors induce a bias towards zero.

We begin by estimating equation (9) assuming that $\lambda_t = 0$. The solid line in Figure 11 shows the estimated values of β_t from 1960 to 2010, with a 95 per cent confidence interval. As we can see, β_t is statistically significant from 1980 onwards and relatively stable, with an average equal to 0.48, slightly below the average of 0.64 in the sample of Deutsch and Silber (2004). In column (1) of Table 5 we allow for time dummies (δ_t) for each period but impose the same value β for the whole sample, estimating a value of 0.402 for the coefficient of $Gini(W_{it}^s)$.

In column (2) we assume that $\lambda_t = 0.159$, which is the coefficient estimated for the log of per capita income in a regression with $\phi_{wi} R_{wi}$ as dependent variable, for the sample of 23 countries used by Deutsch and Silber (2004). As we can see, in this case β is 0.582, much closer to the average of 0.64 obtained for $\phi_{wi} R_{wi}$ by these authors. We have also checked that the estimated values β_t from 1960 to 2010 when $\lambda_t = 0.159$ are above the estimates represented in Figure 11.[16]

To mitigate concerns of omitted variables bias, in column (3) we add the log of per capita income and its square value. Controlling for the square value is important as we found a non-linear relationship between human capital and wage inequality. In line with the Kuznets curve, we find an inverted U-shape relationship between the level of development and income inequality, which does not disappear even when $Gini(W^s)$ is included as an explanatory variable.[17] More importantly, we also find that the association

[16] To facilitate the comparisons of the estimated coefficient of $G(W^s)$ in column (2) of Table 5, we have defined $\ln y$ in deviations with respect to its sample mean. Therefore, the value of $\beta = 0.582$ is the coefficient of $G(W^s)$ when the normalized value of $\ln y$ is equal to zero.

[17] In a cross-section of countries, Ahluwalia (1976) found empirical support for the Kuznets hypothesis but Anand and Kabur (1993) showed that Ahluwalia's results were not robust to the use of alternative functional forms or different data sets. Using panel data, Deininger and Squire (1998) did not find an inverted U-shape relationship between the level of income and the Gini coefficient for the majority of countries in their sample (40 out of 49), whereas Barro (2000) finds some evidence in favor of the Kuznets curve.

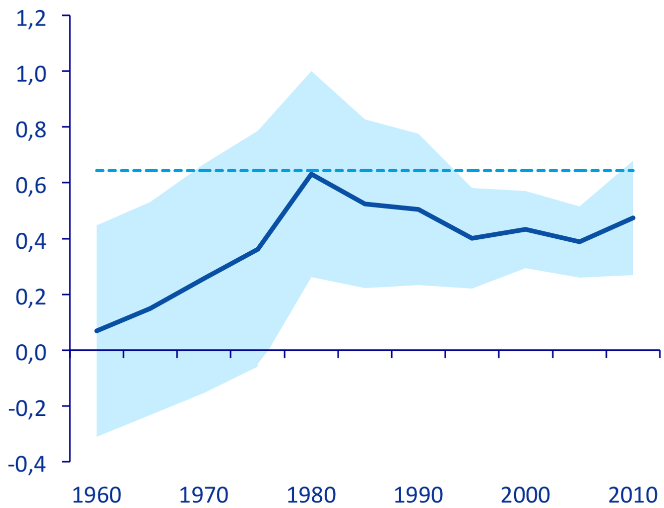

Figure 11: *Estimated contribution of $G(W^s)$ to income inequality.*

between $Gini(W^s)$ and income inequality still holds: the wage Gini coefficient is positive and statistically significant at the 1 percent level, although its value is now lower than in columns (1) and (2).

In columns (4) to (6), we replicate the same estimations as in previous columns but now assuming that the rates of returns are constant over time, homogeneous across countries and the same for all education levels ($\overline{r_P} = \overline{r_S} = \overline{r_T} = 0.1$). This assumption allows us to increase the number of available observations in the regression by 58%. The results are quite similar to the previous ones, regardless of the assumption on the returns, corroborating that the Gini coefficient of simulated wages has a significant and relevant effect on income inequality.

Therefore, the results of this subsection suggest that approximately each point of change in the Gini coefficient of wages contributes to a half-point change in the Gini coefficient of income. The quantitative effect is similar whether we assume increasing or constant returns to scale. Thus, most of the effect stems from differences in the inequality in the distribution of education across levels of schooling.

6. Conclusions

This paper has documented the trends in inequality in years of schooling from 1950 to 2010 using an improved data set on human capital inequality. The evidence shows that most countries have experienced a very intense reduction in human capital inequality, mainly

Table 5
Dependent variable: income inequality $Gini(Y)$

	(1)	(2)	(3)	(4)	(5)	(6)
$Gini(W^s)$	0.402	0.582	0.216	0.420	0.608	0.219
	(9.89)	(11.1)	(4.86)	(10.8)	(11.5)	(5.72)
$\ln yGini(W^s)$		0.159^r	0.159^r		0.159^r	0.159^r
		–	–			
$\ln y$			0.188			0.151
			(5.10)			(4.77)
$(\ln y)^2$			−0.015			−0.014
			(7.01)			(7.64)
R^2	0.122	0.175	0.540	0.097	0.108	0.648
Obs.	652	627	627	1042	990	990
δ_t	Y	Y	Y	Y	Y	Y

Notes: OLS regression with robust standard errors and t-ratios in parenthesis. Regressions from 1960 to 2010 in a 5-year span.

due to an unprecedented decrease in the share of illiterates, which has not been accompanied by a similar reduction in income inequality. Changes in the distribution of capital incomes, in the share of labor incomes and other factors may have offset improvements in the distribution of human capital. In this paper, we explore the relevance of wages in order to explain the lack of correlation between changes in the distributions of human capital and income.

We use a Mincerian approximation for 139 countries from 1950 to 2010 to estimate wage inequality and find a clear inverted U-shape relationship between the Gini coefficient of years of schooling and the Gini coefficient of the simulated wages. The main explanation for this inverted U-shape relationship is the composition effect of the share of population with no schooling. This result is consistent with the hypothesis that in an economy with two population groups, a transfer of workers from the low- to the high-educated group raises the inequality of wages, up to the point where the high-educated group reaches a certain share. The maximum Gini coefficient for wages is reached when the share of illiterates is 40 percent, which corresponds to an average of 5.5 years of schooling. Therefore, these results imply that, if countries wish to reduce income inequality through education, they should ensure that the entire population has completed at least primary schooling (six years on average).

We check the sensitivity of the results to different returns. If returns to primary schooling are low, a large reduction in the share of illiterates may not translate to a sizeable increment in the wages of the population at the bottom end of the income distribution, when a smaller share of the population is also improving the human capital at the top. We find that returns to years of schooling do not affect the inverted U-shape relationship, although increasing returns exacerbates inequality, particularly when the share of illiterates is very low. Therefore, the effects of increasing returns to education on inequality will be greater in advanced economies than in developing countries. We also find that going from decreasing to increasing returns slightly reduces the value of the share of illiterates at which the Gini coefficient of simulated wages reaches its maximum level.

In line with the skill-biased technological change hypothesis, we also provide evidence of an increment in the wage gap between wages at the top and at the bottom of the wage distribution in spite of an increase in the relative supply of skilled workers, particularly in advanced economies. Thus, it could also be possible that improvements in literacy and wages of the population at the bottom end of income distribution have also coincided with increased wages in other cohorts of population with higher education, such that all of them maintain their income shares.

We finish the analysis by showing that the estimated average contribution of wage inequality to income inequality is statistically significant, relatively stable from 1980 onwards and economically relevant: approximately each point of change in the Gini coefficient of wages contributes on average to a half-point change in the Gini coefficient for income, which is also affected by the changes in its composition and other important factors, such as changes in the distribution of other sources of incomes and the impact of fiscal redistribution policies.

The evidence presented in this paper is highly relevant for development policies. Many governments have made a great effort to eradicate illiteracy rates, but these policies have not been accompanied by a more even distribution of income, due to the presence of different offsetting forces. This evidence does not imply that educational policies have failed to reduce poverty or improve the wages and the standards of living of hundreds of millions through better education. On the contrary, eradicating illiteracy and completing primary schooling are necessary conditions for the subsequent improvement in per capita income and inequality, showing that there is no trade-off between them. Better education is crucial in order to increase average earnings per worker, to avoid the negative effects of skill-biased technological progress and to offset other driving forces that may contribute to greater income inequality.

7. References

Acemoglu, D. and D. Autor (2012). "What Does Human Capital Do? A Review of Goldin and Kazt˙s The Race Between Education and Technology." *Journal of Economic Literature*, Vol. 50(2), 426-463.

Aitchison, J. and J.Brown (1957). *The Lognormal Distribution*. Cambridge University Press

Ahluwalia, M. S. (1976). "Income Distribution and Development: Some Stylized Facts." *American Economic Review*, Vol. 66(2), 128-35.

Alesina, A. and D. Rodrik (1994). "Distributive Politics and Economic Growth." *Quarterly Journal of Economics*, Vol. 109, 465-90.

Anand, S. and S.M.R. Kanbur (1993). "The Kuznets Process and the Inequality-Development Relationship." *Journal of Development Economics*, Vol. 40, 25-52.

Atkinson, A. B. and A. Brandolini (2001). "Promise and Pitfalls in the Use of 'Secondary' Data Sets: Income Inequality in OECD Countries as a Case Study." *Journal of Economic Literature*, 39, 771-99.

Atkinson, A.B., T. Piketty and E. Saez (2011). "Top Incomes in the Long Run of History." *Journal of Economic Literature*, Vol. 49(1), 3-71.

Autor, D. H., L. Katz and M. S. Kearney (2006). "The Polarization of the U.S. Labor Market." *American Economic Review*, Vol. 96(2), 189-94.

Banerjee, A. V. and E. Duflo (2003). "Inequality and Growth: What Can the Data Say?" *Journal of Economic Growth*, Vol. 8, 267-99.

Barro, R. J. (2000). "Inequality and Growth in a Panel of Countries," *Journal of Economic Growth*, 5, 5-32.

Barro, R. J. and J. W. Lee (2001): "International Data on Educational Attainment Updates and Implications." *Oxford Economic Papers*, no. 3, 541-63.

Barro, R. and J-W Lee (2013). "A New Dataset of Educational Attainment in the World, 1950-2010." *Journal of Development Economics*, 104, 184–198.

Becker, G. S. (1962). "Investment in Human Capital: A Theoretical Analysis." *Journal of Political Economy*, Vol. 70 (5), 9-49.

Becker, G. S. and B. R. Chiswick (1966). "Education and the Distribution of Earnings," *American Economic Review*, Vol. 56 (1), 358-69.

Card, D. and T. Lemieux (2001). "Can Falling Supply Explain the Rising Return to College for Younger Men? A Cohort-Based Analysis." *Quarterly Journal of Economics*, Vol. 116 (2), 705-46.

Castelló-Climent, A. (2010). "Channels Through Which Human Capital Inequality Influences Economic Growth." *Journal of Human Capital*, Vol. 4(4), 394-450.

Castello, A. and R. Doménech (2002): "Human Capital Inequality and Economic Growth: Some New Evidence." *The Economic Journal*, 112 (March), 187-200.

Checchi, D. and C. García-Peñalosa (2010): "Labour Market Institutions and the Personal Distribution of Income in the OECD." *Economica*, 77(307), 413-450.

Cohen, D. and M. Soto (2007). "Growth and Human Capital: Good Data, Good Results." *Journal of Economic Growth*, 12(1), 51-76.

Colclough, C.; G. Kingdon and H. Patrinos (2010). "The Changing Pattern of Wage Returns to Education and its Implications." *Development Policy Review*, Vol. 28(6), 733-47.

Council of Economic Advisers (2012). *Annual Report*. United States.

Cowell, F. A. and C. V. Fiorio, 2011). "Inequality Decompositions - A Reconciliation." *Journal of Economic Inequality*, 9, 509–528.

Deininger, K. and L. Squire (1996). "A New Data Set Measuring Income Inequality." *World Bank Economic Review*, 10, 565-91.

Deininger, K. and L. Squire (1998). “New Ways of Looking at old Issues: Inequality and Growth.” *Journal of Development Economics*, 57, 259-87.

De Gregorio, J., and J. W. Lee (2002). “Education and Income Distribution: New Evidence from Cross-Country Data.” *Review of Income and Wealth*, Vol. 48(3), 395–416.

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

de la Fuente, A. and R. Doménech (2015). “Educational Attainment in the OECD, 1960–2010. Updated Series and a Comparison with Other Sources.” *Economics of Education Review*, 48, 56-74.

Deutsch, J. and J. Silber (2004). “Measuring the Impact of Various Income Sources on the Link Between Inequality and Development: Implications for the Kuznets curve.” *Review of Development Economics*, 8(1), 110-127.

Epifani, P. and G. Gancia (2008). “The Skill Bias of World Trade.” *The Economic Journal*, vol. 118, 927-60.

Fei, J. C. H., G. Ranis, and S. W. Y. Kuo (1978). “Growth and the Family Distribution of Income by Factor Components.” *The Quarterly Journal of Economics*, 92(1), 17-53.

Fields, G. S. (1979). “A Welfare Economic Approach to Growth and Distribution in the Dual Economy.” *The Quarterly Journal of Economics*, 93(3), 325-353.

Fields, G. S. (1993). “Inequality in Dual Economy Models.” *The Economic Journal*, 103(420), 1228-1235.

Fields, G. S. (2003). “Accounting for Income Inequality and Its Change: a New Method with Application to Distribution of Earnings in the United States.” *Research in Labor Economics*, 22, 1–38

Forbes, K. (2000). “A Reassessment of the Relationship Between Inequality and Growth,” *American Economic Review*, Vol. 90 (4), 869-87.

Goldberg, P.K. and N. Pavcnik (2007). “Distributional Effects of Globalization in Developing Countries.” *Journal of Economic Perspectives*, Vol. XLV, 39-82.

Goldin, C. and L. F. Katz (2007). “Long-Run Changes in the Wage Structure: Narrowing, Widening, Polarizing.” *Brookings Papers on Economic Activity*, Vol. 38(2), 135-68.

Griliches, Z. and J. Hausman (1986). “Errors in Variables in Panel Data.” *Journal of Econometrics*, 31, 93-118.

Halter, D., M. Oechslin and J. Zweimüller (2014). ˇInequality and Growth: the Neglected Time Dimension.” *Journal of Economic Growth*, Vol. 19, 81-104.

Hanushek, E. A. and L. Woessmann (2012). “Do Better Schools Lead to More Growth? Cognitive Skills, Economic Outcomes, and Causation.” *Journal of Economic Growth*, Vol. 17(4), 267-321.

Jaumotte, F., S. Lall and Ch. Papageorgiou (2013). “Rising Income Inequality: Technology, or Trade and Financial Globalization?” *IMF Economic Review*, Vol. 61(2), 271-309.

Karabarbounis, L. and B. Neiman (2013). “The Global Decline of the Labor Share.” *The Quarterly Journal of Economics*, 129(1), 61-103.

Katz, L. F. and K. M. Murphy (1992). “Changes in Relative Wages, 1963-1987: Supply and Demand Factors.” *Quarterly Journal of Economics*, vol. 107 (1), 35-78.

Knight, J. B. (1976). “Explaining Income Distribution in Less Developed Countries: A Framework and an Agenda,” *Oxford Bulletin of Economics and Statistics*, 38(3), 161-177.

Knight, J. B., and R.H. Sabot (1983). "Educational Expansion and the Kuznets Effect." *The American Economic Review*, 73(5), 1132-1136.

Krueger, A. and M. Lindahl (2001). "Education for Growth: Why and For Whom?" *Journal of Economic Literature*, vol. XXXIX, 1101-36.

Kuznets, S. (1955). "Economic Growth and Income Inequality," *American Economic Review*, vol. 45(1), 1-28.

Lane, P.R., and G.M. Milesi-Ferretti (2007). "The External Wealth of Nations Mark II: Revised and Extended Estimates of Foreign Assets and Liabilities, 1970–2004." *Journal of International Economics*, Vol. 73 (2), 223–50.

Lemieux, T. (2006). "Postsecondary Education and Increasing Wage Inequality." *American Economic Review*, Vol. 96(2), 195-99.

Lim, A. S. K., and K. K. Tang (2008). "Human Capital Inequality and the Kuznets Curve." *The Developing Economies*, 46(1), 26-51.

Mincer, J. (1974). *Schooling, Experience and Earnings*. Columbia University Press.

Montenegro, C. E. and H. Patrinos (2014). "Comparable Estimates of Returns to Schooling around the World." Policy Research Working Paper, 7020. The World Bank.

Morrisson, C. and F. Murtin (2013). "The Kuznets curve of human capital inequality: 1870–2010." *Journal of Economic Inequality*, 11, 283–301.

Ning, G. (2010). "Can Educational Expansion Improve Income Inequality in China? Evidences from the CHNS 1997 and 2006 Data." IZA Discussion Paper No. 5148.

OECD (2011). *Divided We Stand: Why Inequality Keeps Rising*, OECD Publishing.

OECD (2015). *Education at a Glance 2015*, OECD Publishing.

Perotti, R. (1996). "Growth, Income Distribution and Democracy: What Data Say." *Journal of Economic Growth*, Vol. 1, 149-87.

Persson, T. and G. Tabellini (1994). "Is Inequality Harmful for Growth? Theory and Evidence." *American Economic Review*, Vol. 84, 600-21.

Psacharopoulos and H. Patrinos (2004). "Returns to Investment in Education: A Further Update." *Education Economics*, Vol. 12 (2), 111-34.

Pyatt, G., C. Chen and J. Fei (1980). "The distribution of income by factor components." *Quarterly Journal of Economics*, 95, 451-474.

Ram, R. (1984). "Population Increase, Economic Growth, Educational Inequality, and Income Distribution: Some Recent Evidence." *Journal of Development Economics*, Vol. 14, 419–28.

Robinson, S. (1976). "A Note on the U Hypothesis Relating Income Inequality and Economic Development." *The American Economic Review*, 66(3), 437-440.

Shorrocks, A.F. (1982). "Inequality Decomposition by Factor Components." *Econometrica*, 50(1), 193–211.

Slonimczyk, F. (2013). "Earnings Inequality and Skill Mismatch in the U.S.: 1973-2002." *Journal of Economic Inequality*, Vol. 11, 163-194.

Solt, F. (2016). "The standardized world income inequality database." *Social Science Quarterly*.

8. Appendix 1. Computation of duration levels

Barro and Lee (2013) provide data on total average years of schooling (TYR). Years of education are available for different levels of schooling, namely primary (PYR), secondary (SYR) and tertiary (HYR) education. The dataset also provides information on the highest level attained, disaggregated in total and complete levels. For example, for the population aged 15 years and oveR, the attainment levels include the share of population with total primary (PRI), primary completed ($PRIC$), total secondary (SEC), secondary completed ($SECC$), tertiary ($HIGH$), and tertiary completed ($HIGHC$). We compute incomplete attainment levels by subtracting the complete value from the total attainment in each educational level.

The calculation of the Gini coefficient requires the duration of each level of education (x_i). We use Barro and Lee˙s (2013) data set to compute duration as follows:

$$HYR = DURH * HIGH \tag{11}$$

$$DURH = \frac{HYR}{HIGH} \tag{12}$$

where $DURH$ stands for the duration in years of tertiary education. Expression (14) can be disaggregated into complete and incomplete education. Thus, the average years of schooling of tertiary education can be rewritten in terms of both levels of education:

$$HYR = DURH^{INC} * HIGH^{INC} + DURH^{C} * HIGH^{C} \tag{13}$$

where the superscripts INC and C account for incomplete and complete education respectively. As we do not have information on the duration of each level, we assume the duration of incomplete levels to be half that of the corresponding complete level of schooling. Rearranging the above expressions gives the duration of completed tertiary education,

$$DURH^{C} = \frac{HYR}{[HIGH^{INC}/2] + HIGH^{C}} \tag{14}$$

A similar procedure is used to compute the duration of secondary education

$$DURS = \frac{SYR}{SEC + HIGH} \tag{15}$$

$$DURS^{C} = \frac{SYR}{[SEC^{INC}/2] + [SEC^{C} + HIGH]} \tag{16}$$

and primary schooling

$$DURP = \frac{PYR}{PRI + SEC + HIGH} \tag{17}$$

$$DURP^{C} = \frac{PYR}{[PRI^{INC}/2] + [PRI^{C} + SEC + HIGH]} \tag{18}$$

The additional information provided by the larger number of educational levels makes the new Gini coefficient more precise than previous versions.

9. Appendix 2. Further evidence

We have checked whether improvements in the income going to the first quintile have been accompanied by improvements in the share of income going to the top quintile, with the result that the income Gini coefficient remains largely unchanged over time. The problem with testing this hypothesis is that the most reliable data on the different parts of the income distribution, comparable across countries, are those of the LIS data set. However, as stated above, those data only cover a few wealthy economies and the greatest reductions in the share of illiterates have taken place in developing countries. As an alternative, we have used data on the distribution of income by deciles from the UNU-WIDER World Income Inequality Database (WIID), which consists of an update of the Deininger & Squire database from the World Bank, new estimates from the Luxembourg Income Study and Transmonee, and other new sources. The version we have used is UNU-WIDER World Income Inequality Database, Version 2.0c, May 2008. Given that there are very few observations with which to create a balanced panel, under the criteria of using only high quality data and the same original source, we have taken the first and the last available observation of income deciles in each country and we have computed the difference in several income quintiles. Then, for each country, we have computed the absolute changes in income for the time period for which the income deciles are available.

Figure 12 plots the within-country increase in the share of income going to the first and fifth quintiles. The figure indicates that there is no systemic positive relationship between the two variables. In contrast, the correlation is clearly negative. Furthermore, we find weak evidence suggesting that a reduction in the share of illiterates has resulted in an increase in the share of income going to the poorest 20 percent of the population. Figure 13 displays the change in the share of illiterates for the population 15 years and over and the change in the first quintile in the distribution of income. As we can observe, there is not a marked negative relationship between the two variables. In countries such as Mexico, Spain, Brazil and Indonesia there has been a significant reduction in the share of illiterates but the income accruing to the bottom quintile has remained almost constant.[18]

[18] This sample of countries does not show a clear positive relationship between the increment in the share of illiterates and the increment in the income going to the poorest 60 percent of the population (cumulative third quintile). Nor do we find that an increase in the share of education going to the top education quintile is related to an increase in the share of income going to the top income quintile.

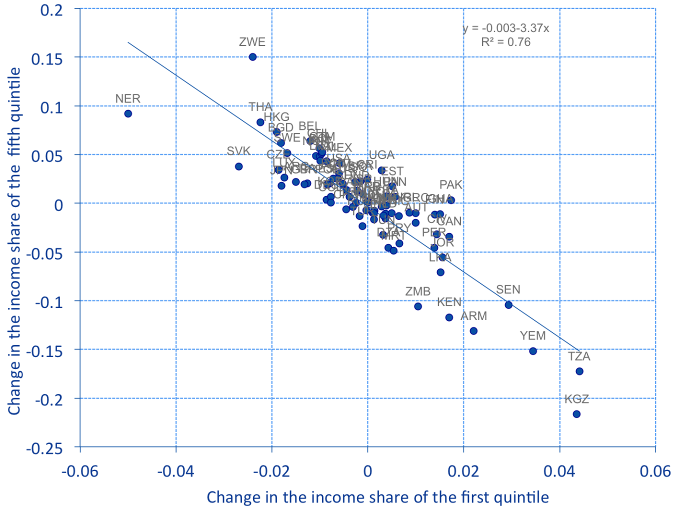

Figure 12: *Change in the income shares of the first and fifth quintiles.*

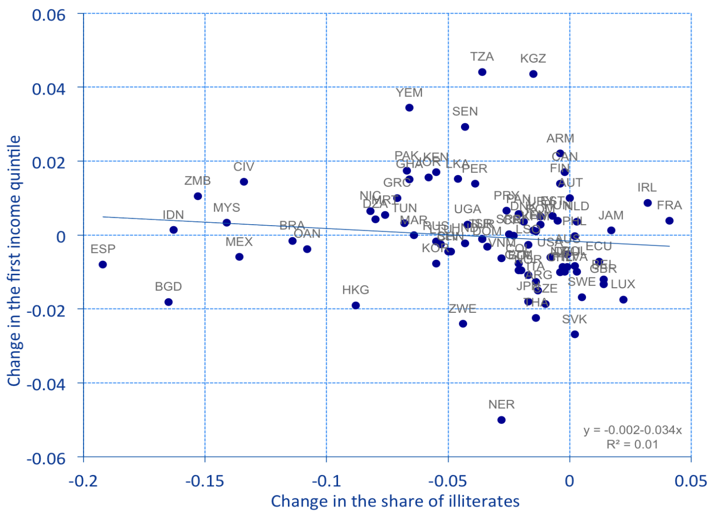

Figure 13: *Changes in the shares of illiterates and the first income quintile.*