Human Capital and Income Inequality: Some Facts and Some Puzzles*
Amparo Castelló-Climent\textsuperscript{a} and Rafael Doménech\textsuperscript{a,b}

October 2012

Abstract
Using a broad number of indicators from an updated data set on human capital inequality for 146 countries from 1950 to 2010, this paper documents several facts regarding the evolution of income and human capital inequality. The main findings reveal that, in spite of a large reduction in human capital inequality around the world driven by a decline in the number of illiterates of several hundreds of millions of people, the inequality in the distribution of income has hardly changed. In many regions, the income Gini coefficient in 1960 was very similar to that in 2005. Therefore, improvements in literacy are not a sufficient condition to reduce income inequality, even though they improve life standards of people at the bottom of the income distribution. Increasing returns to education, external effects on wages of higher literacy rates or the simultaneous concurrence of other exogenous forces (e.g., globalization or skill-biased technological progress) may explain the lack of correlation between the evolution of income and education inequality.

Keywords: Distribution of education, income inequality, human development, panel data.

JEL: I24, I25, O15, O50.

* A. Castelló would like to thank financial support from the Spanish Ministry of Economy and Competitiveness through ECO2011-29283 project and the Ramón y Cajal Programme. R. Doménech thanks the financial support from the Ministry of Economy and Competitiveness, ECO2011-29050. Contact: amparo.castello@uv.es, rafael.domenech@uv.es.
\textsuperscript{a} University of Valencia.
\textsuperscript{b} BBVA Research
1. Introduction

This paper documents important empirical regularities regarding the evolution of human capital and income inequality, using an updated data set on human capital inequality for 146 countries from 1950 to 2010. Fostered by the United Nations Millennium Development Goal to achieve universal primary education by 2015, most developing countries have made a great effort to eliminate illiteracy rates. As a result, the human capital Gini coefficient was 0.549 in 1960 and only 0.280 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 income Gini coefficient for the same group of countries was almost equal in 1960 (0.415) than in 2005 (0.411). This fact is not restricted to developing countries alone, in 1960 the human capital Gini coefficient in the high income OECD countries was 0.224 and reduced to 0.153 in 2005, whereas the income Gini coefficient has remained unchanged taking values of 0.300 and 0.302, respectively. This paper analyzes the above evidence in more detail and tests several hypothesis that can explain the lack of correlation between the evolution of human capital and income inequality.

To analyze the evolution of human capital inequality over the last decades, we have updated the dataset constructed by Castelló and Doménech (2002) using the new version of the Barro and Lee (2010) dataset. This new version includes more countries and years, reduces some measurement errors and solves some of the shortcomings revealed by De la Fuente and Domenech (2006) and Cohen and Soto (2007). The new data set encompasses 146 countries from 1950 to 2010 in a five-year period span.

Using the new data set, we document interesting facts regarding the trends in income and human capital inequality. In the first place, from 1950 to 2010 there has been a significant reduction in human capital inequality around the world. The reduction has been more pronounced in developing countries but it is also present in the developed economies. The evidence also shows that the improvements in equality have mainly benefited the lowest part of the distribution. We examine this evidence further and decompose the Gini coefficient as a combination of the share of illiterates and the Gini coefficient among the literates. Interestingly, the components of the Gini coefficient suggests two stages in the evolution of overall human cap-
ital inequality. In the first stage, when the share of illiterates is very high, the evolution of the human capital Gini coefficient is mainly determined by the share of illiterates. Thus, the great effort in developing countries to increase the literacy rates has translated into a considerable reduction in the human capital Gini coefficient over the years. For example, in Sub-Saharan Africa, almost 80 percent of the population was illiterate in 1950, this number has reduced more than half and the share of illiterates was 32 percent in 2010. As a result, the human capital Gini coefficient for this region reduced from 0.8 in 1950 to 0.41 in 2010. In the later stages, when the share of illiterates is low, the evolution of the human capital Gini coefficient is mainly determined by the evolution of the Gini among the literates. At this stage, there has not been a unique pattern in the evolution of human capital inequality over time. In the advanced economies, some countries display an increase in the human capital Gini coefficient and in the Gini among the literates (e.g. Finland, Iceland, Austria, the Netherlands or Great Britain), whereas in others, there has been a reduction in both inequality indicators (e.g. Germany, Australia, USA, Sweden or Canada, among others).

However, the evolution of income inequality has been quite different to that of human capital. In spite of some variations in the short term, in most of the regions the income Gini coefficient in 2005 was very similar to that in 1960. The analysis of the evolution of income inequality goes back to the work by Kuznets (1955), which maintained the existence of an inverted U shape in the relationship between inequality and development, that is, inequality in the distribution of income increases and later decreases as per capita income rises.\(^2\) The evidence shows that the only region that displays an increase (from 1960 to 1990) followed by a reduction in income inequality is Sub-Saharan Africa. In the rest of the regions, the tendency has been an increase in the income Gini coefficient in the recent years, in spite of the level of development of the region. In fact, in the high income OECD countries, we observe an U-shape in the evolution of income inequality.

\(^2\) Due to the scarcity of available income inequality data for long periods, it is difficult to test empirically this hypothesis and the evidence up to now has been mixed. For example, 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 favour of the Kuznets' curve.
over time, so that the reduction in inequality from 1960 to the middle of the 1970’s and the increase afterwards has made the income Gini coefficient in 1960 (0.3) to be identical to that in 2005 (0.302).

The rather stability of the income Gini coefficient is puzzling given the large reduction of human capital inequality. One should expect that a large decline in human capital inequality would translate into a decline in income inequality as well. We check potential candidates to explain this puzzle. A plausible explanation could be that the massive increment in the literate population attaining primary schooling has reduced income inequality by improving the share of income accruing to the lowest income quintiles but, at the same time, there have been other factors that have increased income inequality by increasing the share of income going to the top income quintiles, so that the income Gini coefficient has slightly varied. In fact, Atkinson et al. (2011) show that, in the second half of the twentieth century, many countries have experienced an increase in the top income shares. Nevertheless, when we test this hypothesis we do not find any evidence indicating that the changes in the share of income of the first quintile are positively correlated with the changes in the share of income accruing to the top quintile.

On the other hand, we do not find either a significant relationship between the changes in the lowest education quintile and the changes in the shares of the lowest income quintile. This result is surprising since it suggests that the large reduction of illiteracy, or the massive increment of population with primary education, has not translated into a significant increase in the share of income accruing to the poorest population. Also striking is the finding of a negative and statistically significant correlation between the changes in the income Gini coefficient and the variation in the Gini among the literates.

A possible explanation of this puzzle is that improvements in literacy, which increase the wage of population at the bottom end of income dis-

---

3 The “canonical model” of the race between education and technological change (e.g. Goldin and Katz, 2008) could be a plausible candidate. However, it is difficult to test this model for a broad number of countries since there are not available data of wages for workers with different levels of education. Moreover, the theory is aimed to explain wage dispersion and earnings dispersion is not the same as income inequality (e.g. Gottschalk and Smeeding, 1997).
tribution, have also coincided with an increase of wages in other cohorts of population with higher education, such that all of them maintain their incomes shares. The latter could be an indirect consequence of the fall in illiteracy, due to the external effects of having a more educated population, or, alternatively, just a spurious correlation if the improvements in education at the bottom have coincided with other exogenous forces (e.g., globalization or skill-biased technological progress) that have increased wages at the top. The problem is that, due to the lack of data for a large sample of countries and years, we cannot test these hypothesis. Nevertheless, casual information for some countries where this information is available reveals that real wages for low levels of education have remained almost constant from 1960 onwards, whereas wages for higher levels of education have exhibited a clear upward trend. If this was also the case for other countries, then increases in literacy rates could rise wages and life standards of people at the bottom of the income distribution with no improvement in their income share.

An alternative explanation for the lack of correlation between the evolution of income and that of education inequality could be that the returns to education are increasing. If the returns to education are convex, an extra year of education at the primary level brings a smaller increase in wages than it does at the higher levels of education. In this scenario, an increment of education at the lower levels might not necessarily increase the share of income at the bottom end of the income distribution. We use aggregate data on real output per worker from the PWT 7.0 and compute the social returns to education for different levels of schooling. The evidence suggests that one year increment in university is almost 3 times more productive than an increase in one year in primary education. In some specifications, the estimated coefficient of primary schooling is even negative or not significant. The increasing returns to education are found when we control for country specific effects with the fixed effect and the first difference estimators, and are robust to the use of instrumental variables. Thus, low returns to primary education could explain why, in spite of a more even distribution of human capital over the years, we do not observe an equalizing effect in the distribution of income as well. 4

4 This result is in line with recent evidence that uses data for the 1990s and 2000s and shows that in many countries the returns to education are greater for higher education than
The evidence found in this paper is in line with the literature that analyses the evolution of world inequality over time (e.g. Milanovic, 2002; Bourguignon and Morrison, 2002; Sala-i-Martin, 2002), which obtains that most of the total income inequality observed around the world is mainly due to differences in per capita income across countries. Instead, we focus on income inequality for individual countries, and show that most of the variation in income inequality is across-countries and, on average, the income Gini coefficient slightly changes within countries over time.

A strand of the literature has focused on the effect of income and human capital inequality on economic growth.\(^5\) In cross-section and panel data models, Castelló and Doménech (2002) and Castelló-Climent (2010) show that a more uneven distribution of human capital has a negative influence on the growth rates of per capita income. However, the evidence on the effect of income inequality on growth has been mixed. Whereas cross-section estimates show that income inequality is harmful for growth (see Alesina and Rodrik, 1994, Persson and Tabellini, 1994; Perotti, 1996), panel data models indicate the relationship is not linear (e.g. Barro, 2000 and Banerjee and Duflo, 2003) or even positive (e.g. Forbes, 2000). The evidence presented in this paper shows that while human capital inequality data might have enough signal in both cross-section and panel models, the rather stability of the income Gini coefficient may be a potential explanation for the mixed evidence found in the literature, since specifications that exploit the within country variation of the income inequality data may have low signal for identification. When the signal in the data is very low, the effect on growth is prone to be influenced by measurement error and the results may be more sensitive to the countries included in the sample and the period of analysis.

The structure of the paper is as follows. Section 2 computes the improved measures of human capital inequality and documents some styl-
ized 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 aims to explain the lack of correlation between income and education inequality. Finally, section 5 contains the main conclusions.

2. Evolution of human capital inequality over time

The most comprehensive data set on human capital inequality measures is that of Castelló and Doménech (2002), which takes the educational attainment levels from Barro and Lee (2001) and calculates the Gini coefficient and the distribution of education by quintiles for a broad number of countries and periods. However, recent studies have shown that the perpetual inventory method in Barro and Lee (2001) suffers from several problems. The main drawback is that it utilizes very few census data points and depends crucially on enrollment rates, which are usually criticized for being overstated in developing countries. Cohen and Soto (2007) and de la Fuente and Doménech (2006) illustrate that Barro and Lee (2001) dataset show implausible time series profiles for some countries. As Castelló and Doménech (2002) utilizes the Barro and Lee’s (2001) data set, the inequality measures are subject to the same criticisms as the average years of schooling. This paper uses the improved Barro and Lee (2010) data set that reduces measurement error by using more information from census data and a new methodology that makes use of disaggregated data by age group. The new inequality indicators are available for 146 countries from 1950 to 2010 in a 5 year span and include a total of 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 East 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.

6 In line with Cohen and Soto (2007), the methodology fills 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.

7 We have followed the same classification of countries as in Barro and Lee (2010).
between the old and the new human capital Gini coefficient shows that the correlation for the overlapping observations is very high in levels (0.975), but it is very low when the variables are measured in 10-year difference (0.278), suggesting a lower measurement error in the new human capital Gini coefficient derived from a smoother trend in the attainment levels.\textsuperscript{8} A better quality in the education data is very important since it might attenuate the bias caused by measurement error, which is particularly prevalent in econometric analysis that exploit the within country variation of the data (see Krueger and Lindahl, 2001).

Table 1 shows the summary statistics for the average human capital Gini coefficient for these regions. Following Catelló and Doménech (2002), the human capital Gini coefficient has been defined as follows:

\[
Gini^h = \frac{1}{2\bar{H}} \sum_{i=0}^{3} \sum_{j=0}^{3} \left| \bar{x}_i - \bar{x}_j \right| n_i n_j
\]  

(1)

where \( \bar{H} \) are the average years of schooling in the population 15 years and above, \( i \) and \( j \) stand for different levels of education, \( \bar{x} \) refers to the cumulative average years of schooling of each level of education and \( n \) are the share of population with a given level of education. We consider 4 levels of schooling: no schooling (0), primary (1), secondary (2) and tertiary (3) education.

According to Table 1, the region with the largest human capital inequality is South Asia, with an average human capital Gini coefficient equal to 0.641, followed by Sub-Saharan African (SSA) countries (average \( Gini^h \) equal to 0.614) and the Middle East and the North African (MENA) region (average \( Gini^h \) equal to 0.575). At the other end, the East European and Central Asian countries (EECA), with an average human capital Gini of 0.207, and the advanced economies, with a \( Gini^h \) equal to 0.212, are the regions where the average years of schooling are more evenly distributed. In the middle of the extremes, the Latin American and the Caribbean countries (LAC) and the East Asian and the Pacific region (EAP) have average Gini coefficients of 0.338 and 0.385, respectively.

\textsuperscript{8} Barro and Lee (2010) also find that the old and the new measures of the average years of schooling are highly correlated in levels and there is a scarcerly relationship when the variables are measured in differences.
<table>
<thead>
<tr>
<th>Countries</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
<th>$Gini_{1950}$</th>
<th>$Gini_{2010}$</th>
<th>3rd Q $Gini_{1950}$</th>
<th>3rd Q $Gini_{2010}$</th>
<th>1st/5th Q $Gini_{1950}$</th>
<th>1st/5th Q $Gini_{2010}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>World</td>
<td>146</td>
<td>0.412</td>
<td>0.251</td>
<td>0.026</td>
<td>0.997</td>
<td>0.557</td>
<td>0.202</td>
<td>0.420</td>
<td>0.096</td>
<td>0.278</td>
</tr>
<tr>
<td>Advanced</td>
<td>24</td>
<td>0.212</td>
<td>0.116</td>
<td>0.049</td>
<td>0.827</td>
<td>0.242</td>
<td>0.156</td>
<td>0.425</td>
<td>0.499</td>
<td>0.371</td>
</tr>
<tr>
<td>EAP</td>
<td>19</td>
<td>0.385</td>
<td>0.193</td>
<td>0.097</td>
<td>0.923</td>
<td>0.588</td>
<td>0.230</td>
<td>0.159</td>
<td>0.448</td>
<td>0.009</td>
</tr>
<tr>
<td>EECA</td>
<td>20</td>
<td>0.207</td>
<td>0.124</td>
<td>0.026</td>
<td>0.611</td>
<td>0.331</td>
<td>0.099</td>
<td>0.370</td>
<td>0.541</td>
<td>0.173</td>
</tr>
<tr>
<td>LAC</td>
<td>25</td>
<td>0.338</td>
<td>0.169</td>
<td>0.048</td>
<td>0.915</td>
<td>0.456</td>
<td>0.217</td>
<td>0.269</td>
<td>0.457</td>
<td>0.059</td>
</tr>
<tr>
<td>MENA</td>
<td>18</td>
<td>0.575</td>
<td>0.241</td>
<td>0.145</td>
<td>0.997</td>
<td>0.808</td>
<td>0.313</td>
<td>0.046</td>
<td>0.385</td>
<td>0.001</td>
</tr>
<tr>
<td>South Asia</td>
<td>7</td>
<td>0.641</td>
<td>0.227</td>
<td>0.150</td>
<td>0.988</td>
<td>0.780</td>
<td>0.414</td>
<td>0.061</td>
<td>0.279</td>
<td>0.000</td>
</tr>
<tr>
<td>SSA</td>
<td>33</td>
<td>0.614</td>
<td>0.219</td>
<td>0.128</td>
<td>0.963</td>
<td>0.798</td>
<td>0.407</td>
<td>0.027</td>
<td>0.296</td>
<td>0.000</td>
</tr>
</tbody>
</table>
The evolution of education inequality over time is displayed in Table 1, which shows the value of the Gini coefficient for each region in 1950 and in 2010. The comparison of the Gini coefficient in both periods indicates that, in spite of the large differences in the distribution of education across regions, there has been a general reduction of human capital inequality worldwide. In most of the regions, the decline has been spectacular and their Gini coefficients have been more than halved. For example, in 1950 the Middle East and North Africa was the region with the greatest inequality in the distribution of education, with a Gini coefficient about 0.808, in 2010 the Gini coefficient of this region was only 0.313.

Fact 1. From 1950 to 2010 there has been a significant reduction in human capital inequality around the world.

Whether the reduction in education inequality has been due to variations at the bottom, middle or upper part of the distribution cannot be disentangle by looking at the Gini coefficient alone, since the Gini index is an absolute measure of inequality and does not give information on the different parts of the distribution. Thus, Table 1 also displays the values of the 3rd quintile and the ratio between the bottom to the top quintile. The numbers show a general increment in the share of education going to the third quintile and a general increase in the ratio of the bottom to the top quintile as well, suggesting that the improvement in equality has mainly benefited the lowest part of the distribution.

A further examination of the data reveals that the large reduction of education inequality has mainly been due to a sizeable decline in the share of illiterates. Without exception, all the regions in the world have experienced a great reduction in the share of illiterates that has implied a decline by more than halve in the population with no education. This fact is illustrated in Figures 1 and 2, which display the evolution over time of the human capital Gini coefficient and the share of the population 15 and above that are illiterates. Interestingly, the comparison of Figures 1 and 2

9 We compute 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 the value of the third quintile in that case is equal to zero.
indicates the evolution of the share of illiterates and the Gini coefficient is almost identical. The strong similarity in both figures suggests the reduction in the Gini coefficient over time has been determined, to a great extent, by the decline in the share of illiterates.

**Fact 2.** In most countries the large reduction of education inequality has mainly been due to the sizeable decline in the share of illiterates.

A closer look at the formula of the human capital Gini coefficient highlights the importance of the share of illiterates in its computation. Defining \( x_i \) as the average years of schooling of each education level and reorganizing we get:\(^{10}\)

\[
Gini^h = n_0 + \frac{n_1 (n_2 x_2 + n_3 (x_2 + x_3)) + n_2 n_3 x_3}{H} \tag{2}
\]

This formula points out that the Gini coefficient of education is a proportional measure of the share of illiterates in the society. Thus, 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 lower than that in the share of illiterates will depend on the changes in the distribution of education among the literates. In fact, given that:

\[
Gini^{LIT} = \frac{1}{2H^{LIT}} \sum_{i=0}^{3} \sum_{j=0}^{3} |\tilde{x}_i - \tilde{x}_j| n_i^{LIT} n_j^{LIT} \tag{3}
\]

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, eq. (2) can be rewritten as follows:

\[
Gini^{LIT} = \frac{1}{(1 - n_0)} \ast \frac{n_1 (n_2 x_2 + n_3 (x_2 + x_3)) + n_2 n_3 x_3}{H} \tag{4}
\]

then, it is straightforward to show that the human capital Gini coefficient can be formally decomposed as a combination of the share of illiterates and the Gini coefficient among the literates in the following way:

\(^{10}\)The cumulative average years of schooling are computed as: \( \tilde{x}_0 \equiv x_0 = 0, \tilde{x}_1 \equiv x_1, \tilde{x}_2 \equiv x_1 + x_2, \tilde{x}_3 \equiv x_1 + x_2 + x_3. \)
Figure 1: Human Capital Gini coefficient of population15+.

Figure 2: Share of illiterates of population15+.
Human Capital and Income Inequality

\[ Gini^h = n_0 + (1 - n_0)Gini^{LIT} \] (5)

The empirical evidence corroborates that countries with the greatest share of illiterate population display the greatest Gini coefficient of education. On the other end, when the share of illiterates is almost zero, the distribution of primary, secondary and tertiary education is what determines the degree of education inequality. The relationship between changes from 1950 to 2010 in the distribution of education and in the share of illiterates for the whole sample of countries is in Figure 3. The figure shows that in countries where in 1950 the share of illiterates was high, the reduction in the Gini coefficient has been very similar to that of the reduction in share of illiterates since most of the countries are located close to the diagonal line. Countries such as Bahrain, Botswana, Kenya or Uganda, among others, have experienced an important increase in the literacy rates and, as a result, in 2010 the share of illiterates was below 30 percent and the Gini coefficient was also below 0.3. On the other extreme, however, there are countries where the increment in the literacy rates has been more moderate. For instance, in Niger, Mozambique, Gambia, Mali or Sierra Leone the share of illiterates in 2010 was still above 60 percent and the Gini coefficient was above 0.6 as well.

In those countries in which in 1950 most of the population were literate, the figure shows a different pattern. With the exception of Spain, Portugal, Greece, Cyprus and Korea, which also had a substantial population that were illiterate in 1950 and have reduced both illiteracy and education inequality over time, most of the wealthy economies display a slightly variation in the share of illiterates and a larger variance in the changes of the Gini coefficient.

This is illustrated in Figure 4, which displays the absolute change in the Gini coefficient and the share of illiterates in the high income OECD economies. In this sample of countries, we observe a small reduction in the share of illiterates and a considerable difference in the evolution of education inequality across countries. For instance, countries such as Finland, Iceland, the Netherlands, Austria or UK show an increase in the human capital Gini coefficient over time, whereas the inequality in the distribution of education decreased in Sweden, Canada, USA, Australia or Germany.
Figure 3: Change in the human capital Gini coefficient and in the share of illiterates, 1950-2010.

Figure 4: Change in the human capital Gini coefficient and the share of illiterates. High income OECD countries, 1950-2010.
Nevertheless, the figure also shows that in some developed economies the change in the share of illiterates is positive (e.g. Belgium, France, Germany, the Netherlands, Switzerland and Great Britain). This fact is surprising since the tendency worldwide has been the elimination of illiteracy and we do not find any developing country in which the share of population with no education has increased.\textsuperscript{11}

In order to analyze further this evidence we have estimated the following equation:

$$\Delta Gini^h = \alpha + \beta (1 + \gamma d_{OECD}) \Delta n_0$$  \hspace{1cm} (6)

where $d_{OECD}$ is a dummy variable for the high income OECD countries. The results, displayed in Table 2, show that, as expected, the coefficient of $n_0$ is statistically significant and below 1. The interaction term with the dummy is negative and statistically significant for the high income OECD countries (column (1)) and negative but not significant when the dummy excludes Spain, Portugal, Greece and Korea (column (2)). Additionally, the statistical adjustment is very high and the changes in the share of illiterates explain more than 90 percent of the variation in the changes of the human capital Gini coefficient. If in eq. (6) we replace $\Delta n_0$ by $\Delta Gini^{LIT}$, the estimated coefficient of the Gini among the literates is lower, whereas the coefficient of the interaction term with the OECD dummy is larger in absolute value.\textsuperscript{12}

\textsuperscript{11} A plausible explanation for the increment in the population with no education in some developed countries could be that part of the new immigrants are illiterate.

\textsuperscript{12} The estimated coefficients for the Gini among the literates are as follows:

$$\Delta Gini^h = -0.302 \pm 0.016 + 0.452 \pm 0.176 \Delta Gini^{LIT} - 0.968 \pm 0.653 d_{OECD} \Delta Gini^{LIT}$$

with $R^2 = 0.032$. 

- 15 -
Human capital and income inequality

Table 2
Dependent variable: $\Delta Gini_{1950−2010}^h$

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta n_{1950−2010}$</td>
<td>0.814$^a$</td>
<td>0.813$^a$</td>
</tr>
<tr>
<td></td>
<td>(0.020)</td>
<td>(0.020)</td>
</tr>
<tr>
<td>$\Delta n_{1950−2010} \ast d_{OECD}$</td>
<td>-0.078$^b$</td>
<td>-0.189$^b$</td>
</tr>
<tr>
<td></td>
<td>(0.035)</td>
<td>(0.145)</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.022$^b$</td>
<td>-0.021b</td>
</tr>
<tr>
<td></td>
<td>(0.009)</td>
<td>(0.009)</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.929</td>
<td>0.928</td>
</tr>
<tr>
<td>Countries</td>
<td>146</td>
<td>146</td>
</tr>
</tbody>
</table>

Note: OLS Regression. Robust standard errors in parenthesis. $a$, $b$ and $c$ are 1, 5 and 10 per cent significance levels respectively. In column (2), the dummy for the high income OECD countries excludes Spain, Portugal, Greece and Korea.

Fact 3. In advanced countries there is not a 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.

In summary, the evidence in this section has shown that the great effort in many developing countries to increase literacy has implied a sizeable reduction in the share of illiterates, which has translated into a similar reduction in the human capital Gini coefficient over time. In advanced economies, with a small share of illiterate population in 1950, the Gini coefficient has been mainly determined by the distribution of education among the literates. In these countries there is not a clear pattern in the evolution of education inequality: whereas some countries display a reduction in the Gini coefficient, other countries have experienced a more uneven distribution of education over time.

3. Human capital and income inequality

In this section we analyze to what extent the reduction in human capital inequality, explained in most countries around the world by the increase in
literacy, has affected the distribution of income inequality.

Data on income inequality have been an issue of concern due to their low quality and related problems of comparability across countries and the scarcity of coverage across countries and over time. 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, Atkinson and Brandolini (2001) have warned about the measurement error in cross-country regressions when using income inequality data. In fact, these authors show that even the high-quality data of Deininger and Squire (1996) for OECD countries contain problems since definitions and data collection methods differ across countries. The most reliable dataset on income inequality is the Luxemburg Income Study (LIS) that 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 starting mainly in 1980, which reduces the sample size considerably.

In this paper we measure income inequality through the net income Gini coefficient taken from the Standardize World Income Inequality Database (SWIID), version SWIID v3.0, which uses a custom missing-data algorithm to standardize WIID from the LIS dataset. Table 3 presents the mean values of the human capital Gini coefficient and the income Gini coefficient for those countries for which income inequality data are available. The data include 75 countries with observations from 1960 to 2005. There are some interesting facts that can be drawn from this table.

In the first place, if we compare the average value of the income (Gini\textsuperscript{i}) and the human capital (Gini\textsuperscript{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 join with a moderate inequality in the human capital distribution. On the other extreme, the countries in South Asia display high inequality in the distribution of education and a relative low inequality in the distribution of income.
## Table 3

**Summary Statistics**

<table>
<thead>
<tr>
<th>Region</th>
<th>Countries</th>
<th>( \text{Gini}^h )</th>
<th>( \text{Gini}^h_{1960} )</th>
<th>( \text{Gini}^p )</th>
<th>( \text{Gini}^p_{1960} )</th>
<th>( \text{Gini}^y )</th>
<th>( \text{Gini}^y_{1960} )</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>World</strong></td>
<td>75</td>
<td>0.352</td>
<td>0.454</td>
<td>0.243</td>
<td>0.373</td>
<td>0.381</td>
<td>0.379</td>
</tr>
<tr>
<td><strong>Advanced</strong></td>
<td>22</td>
<td>0.212</td>
<td>0.233</td>
<td>0.160</td>
<td>0.294</td>
<td>0.307</td>
<td>0.308</td>
</tr>
<tr>
<td><strong>East Asia and the Pacific</strong></td>
<td>9</td>
<td>0.341</td>
<td>0.524</td>
<td>0.225</td>
<td>0.377</td>
<td>0.397</td>
<td>0.382</td>
</tr>
<tr>
<td><strong>Europe and Central Asia</strong></td>
<td>6</td>
<td>0.139</td>
<td>0.181</td>
<td>0.085</td>
<td>0.247</td>
<td>0.241</td>
<td>0.293</td>
</tr>
<tr>
<td><strong>Latin America and the Caribbean</strong></td>
<td>15</td>
<td>0.300</td>
<td>0.388</td>
<td>0.214</td>
<td>0.470</td>
<td>0.481</td>
<td>0.463</td>
</tr>
<tr>
<td><strong>Middle East and North Africa</strong></td>
<td>7</td>
<td>0.582</td>
<td>0.794</td>
<td>0.358</td>
<td>0.395</td>
<td>0.397</td>
<td>0.384</td>
</tr>
<tr>
<td><strong>South Asia</strong></td>
<td>4</td>
<td>0.578</td>
<td>0.698</td>
<td>0.411</td>
<td>0.330</td>
<td>0.321</td>
<td>0.371</td>
</tr>
<tr>
<td><strong>Sub-Saharan Africa</strong></td>
<td>12</td>
<td>0.577</td>
<td>0.744</td>
<td>0.399</td>
<td>0.478</td>
<td>0.461</td>
<td>0.448</td>
</tr>
</tbody>
</table>
Nevertheless, for a large sample of countries Figure 5 shows that the correlation between income and the human capital Gini coefficients in 2005 was positive although not statistically significant. This figure illustrates three main features. First, in most countries income inequality was greater than human capital inequality. Thus, only 21 of 75 countries are below the diagonal 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 0.026 and 0.809, 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, as stated before, the correlation between the income and the human capital Gini coefficients was not very high.

**Fact 4.** For a large sample of countries, the correlation between income and human capital Gini coefficients is low and non-significant. The average of income inequality is larger than for human capital, whereas its variance is lower.

The last two columns of Table 3 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, despite the significant reduction in human capital inequality and in the share of illiterate population from 1950 to 2010.

This fact is also illustrated in Figure 6, which plots the income Gini coefficient for all the regions and available time periods. The figure shows that, in spite of some variations over short periods of time, in most of the regions the income Gini coefficient in 2005 is very similar to that in 1960. Thus, whereas Figure 1 shows a notorious reduction of education inequality over time, mainly due to a reduction of the illiterate population (Figure 3), Figure 6 indicates the inequality in the distribution of income has scarcely changed.
Figure 5: Human capital and income Gini coefficients across countries in 2005.

Figure 7 corroborates this evidence on the absence of correlation between the change in income and human capital Gini coefficients in a sample of 75 countries from 1960 to 2005. In fact, in 36 countries there is a reduction in the human capital Gini coefficient and an increase in the income Gini. Thus, even though there are some countries in which both income and education inequality have increased (e.g., New Zealand, Great Britain and the Netherlands) and others were both variables have reduced (e.g. Kenya, Taiwan, Senegal or Colombia, among others), in a large number of countries changes in income and education inequality display a negative correlation. For example, in countries such as China, India, Singapore, USA, Argentina, Australia and many others, there have been a reduction in the inequality in the distribution of education and an increase in the inequality in the distribution of income. On the contrary, in Austria, Finland and France the inequality in the distribution of education has increased and income inequality has been reduced.
Fact 5. Both across world regions and a large sample of countries, income inequality has remained relatively stable, despite the significant reduction in human capital inequality from 1960 to 2005.

The result that the reduction in human capital inequality has not been accompanied by a improvement in the income Gini coefficient is quite puzzling, since it could lead to the suggestion that a more equalitarian distribution of education has not had any effect on the distribution of income. Therefore it is convenient to analyze further the relationship between both variables and the alternatives explanations to this puzzle.

4. Some explanations for the lack of correlation between income and education inequality

The rather stability of the income Gini coefficient is puzzling given the large reduction of human capital inequality over time. One should expect that a large decline in human capital inequality would translate into a decline in income inequality as well. We check potential candidates to explain the lack of correlation between the evolution of income and education inequality. The first candidate is the “canonical model” of the race between education and technological change (e.g. Katz and Murphy, 1992; Card and Lemieux, 2001; Acemoglu and Autor, 2012). The motivation behind this literature is the observation that in the United States and other developed countries, in spite of the increment in the supply of college graduate workers, there has been a recent 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 of skills, influenced by a skilled bias technological change, and the increase in the supply of skills. When the relative demand increases faster than does the relative supply, wage dispersion rises, when the supply outpaces the demand, wage dispersion reduces. Thus, a plausible explanation for the long term stability of income inequality could be that despite the increase in the supply of educated workers, the demand for skills has kept pace with the human capital investment, so that wage dispersion has remained unchanged in the long term. Although the evidence shows that the ratio of skilled to unskilled workers has increased...
Figure 6: Evolution of the income Gini coefficient across regions, 1960-2005.

Figure 7: Change in income and human capital Gini coefficients across 75 countries, 1960-2005.
not only in the high income OECD countries but also in the less developed economies, it is difficult to perform a proper test of this model for a broad number of countries since there are not available data of wages for workers with different levels of education. Moreover, this theory is aimed to explain wage dispersion and, as stated by Gottschalk and Smeeding (1997), earnings dispersion is not the same as income inequality. The measures of income inequality not only include wages but also income from self-employment, capital rents, transfers and taxes. Next, we test alternative hypothesis.

4.1. Evidence on bottom and top incomes shares

A plausible explanation of Fact 5 is that the reduction in the number of illiterates and, therefore, in education inequality has reduced income inequality by increasing the income at the bottom quintile but, at the same time, other factors might have raised income inequality by increasing the share of income accruing to the top quintile. As a result, the overall income inequality, measured through the income Gini coefficient, could have

13 Using data of the attainment levels by Barro and Lee (2010), in the advanced economies, the ratio of the share of population that has completed tertiary education to the share of population that has completed only primary has increased from 0.069 in 1950 to 2.592 in 2010. In this group of countries, the standard deviation in 2010 was very large (3.842) with a minimum value of 0.081 in Portugal and a maximum value of 15.412 in the U.S. In the less developed countries, the increment has been lower although important in quantitative terms, the ratio increased from 0.108 in 1950 to 0.632 in 2010.

14 See Atkinson and Brandolini (2006) for a broader discussion on the difference between earnings dispersion and income inequality. García-Peñalosa and Orgiazzi (2012) use data from the Luxemburg Income Study for 6 wealthy economies and conclude that the impact of increased dispersion of wages on household income inequality varied across countries. Whereas variation in wage dispersion is an important determinant of the increment in income inequality in the U.S., the share of self-employment income and income capital inequality accounts for a relevant share of the variation in income inequality in the other countries.

15 Atkinson et al. (2011) show that in the second half of the twentieth century, many countries have experienced an increase in the top income shares. The degree of the increment has varied dramatically across countries. Among the 22 countries for which there are available data, Western English speaking countries, China and India, and to a lesser extent, Southern European and Nordic countries, display a substantial increase in the top income shares in the last decades, whereas there is no increase or a modest one in Continental Europe (France, Germany, Netherlands, Switzerland) and Japan.
remained quite stable.

The problem to test this hypothesis is that the most reliable data on the different parts of the income distribution, comparable across countries, is that of the LIS data set. However, as stated above, it includes data for a few wealthy economies and the greatest reductions in the share of illiterates have taken place in the 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 there are very few observations 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 same time period for which the income deciles are available.

Using these data we have checked whether improvements in the income going to the first quintile have also been accompanied by improvements in the share of income going to the top quintile so that both changes have made the income Gini coefficient to remain largely unchanged over time.

Figure 8 plots the increment within a country in the share of income going to the first and fifth quintiles. The figure clearly shows a negative correlation between both variables. In some countries there has been an increase in the share of income going to the bottom quintile join with a reduction in the share of income going to the top quintile. In other countries income inequality has increased by both a reduction in the income going to the poorest 20 percent of the population and by an increase in the share of income going to the top 20 percent of the richest population. Therefore, a positive correlation in the variations at the bottom and the top income quintiles can not be a plausible explanation for the stability of the income Gini coefficient over time.

In fact, we do not find any evidence suggesting that a reduction in the share of illiterates has resulted in an increase in the share of income going
Figure 8: Change in the income shares of the first and fifth quintiles.

Figure 9: Changes in the shares of illiterates and of the first income quintile.
to the poorest 20 percent of the population. Figure 9 displays the change in the share of illiterates for the population 15 years and above and the change in the first quintile in the distribution of income. As we can observe, there is not a marked negative relationship among both variables. In countries such as Mexico, Spain, Brasil or Indonesia there has been a significant reduction in the share of illiterates but the income accruing to the bottom quintile has remained almost constant.\textsuperscript{16}

Overall, there is not robust and significant relationship between the changes in income shares of different quintiles and those in the distribution of education.

4.2. Income inequality and human capital inequality among literates

It is possible that the changes in the distribution of income are not influenced by the changes in the total distribution of education but by the changes in the distribution of education among the literate, that is, by the movements in the distribution among primary, secondary and university education. We test this hypothesis in Figure 10, which shows the relationship between the change in the income Gini coefficient and the change in the Gini among the literate population during the period 1960-2005. The figure illustrates that the correlation among both variables is even negative. There are countries such as Brazil, Hong Kong, USA or Australia where the increment in income inequality is accompanied by a reduction in the inequality of education among the literates. On the other extreme, there are several countries that have experienced the opposite situation; an increase in the inequality in the distribution of education among the literate joint with a reduction in the inequality in the distribution of income. Some examples include Mexico, Colombia, Thailand, Finland or France.

Therefore, this preliminary evidence leads to the suggestion that the changes in the distribution of education among the literate population can-

\textsuperscript{16} This sample of countries does not show either 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). At the other extreme, we do not find either that an increment 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. In fact, the correlation between the increment in the income and in the education of the respective top quintiles is almost zero.
not be the driving force behind the changes in the distribution of income inequality. Next, we explore this evidence further.

4.3. Delayed effects on income inequality

One plausible explanation for the lack of correlation between the change in the distribution of education and that in the distribution of income is that education does not have a contemporaneous effect on income. It takes time for the improvements in the distribution of education to have an effect on the distribution of income. Thus, the reduction in the share of illiterates and the increments in different levels of schooling may not affect the distribution of current income but it may influence the distribution of income some years later.

We test this hypothesis by measuring the change in the human capital Gini coefficient lagged several periods. Specifically, the change in the income Gini coefficient is dated from 1980 to 2005 and the change in the educational variables from 1960 to 1980. The results of an OLS regression for 73 countries are reported in Table 4. Column (1) shows that the lagged change in the human capital Gini coefficient displays a positive coefficient, yet the estimate is not statistically significant at the standard levels. However, given that the changes in income and education inequality may be conditioned by their initial values, the estimate in column (1) may be biased. We augment the specification by controlling for the initial level of income and human capital inequality and the initial level of development, as measured by the real GDP per capita in 1960.\textsuperscript{17} The results, displayed in column (2), indicate that adding these controls changes the coefficient of the increment in the human capital Gini coefficient from positive to negative, though the estimate is not statistically significant either. In columns (3) and (4) we replace the change in the human capital Gini coefficient by the change in the share of illiterates and the change in the human capital Gini coefficient among the literates. Whereas the positive coefficient of the change in the share of illiterates depends on the set of controls, the results indicate that, even controlling for the initial levels of inequality and development, the coefficient of the change in the Gini among the literates is neg-

\textsuperscript{17} For the countries that do not have data for income inequality and per capita income in 1960, we fill the gaps with data on 1965 or 1970.
Human Capital and Income Inequality

HUMAN CAPITAL AND INCOME INEQUALITY

ative and statistically significant. Thus, in line with the previous evidence, there are countries that have experienced an increase in the inequality in the distribution of income, in spite of the reduction in the Gini among the literates, and others in which the increase in the inequality of education among the literates has been followed by a reduction in income inequality. In columns (5-8) we examine whether these correlations differ between rich and poor economies. The results show that, in the high income OECD countries, the changes in the share of illiterates are positive and significantly related to the changes in the income Gini coefficient. However, this positive correlation disappears in the less developed economies. On the other hand, the negative estimates of the lagged changes in the human capital Gini coefficient and the Gini among the literates hold in both samples.

Overall, the result show that the lagged change in the human capital Gini coefficient does not have any significant effect on the subsequent change in the income Gini coefficient. Thus, the delayed effect of human capital inequality cannot be an explanation for the lack of correlation between the evolution of income and human capital inequality. In fact, the regression results suggest that the lagged change in the Gini among the educated population is negative and significantly correlated with the change in the income Gini coefficient not only in the whole sample, but also in the high income and less developed economies.

4.4. Returns to education

Low returns in primary education could explain why, in spite of the observed reduction in the share of illiterates, there is not an increment in the income of the bottom quintiles. In fact, the effect of a more equalitarian distribution of income, derived from an increment at the lower levels of schooling, may be diluted if the returns to education are convex.

We examine this hypothesis by estimating an aggregate production function of the form

\[ Y = AK^\alpha (hL)^{1-\alpha} \]  \hspace{1cm} (7)

where \( Y \) is aggregate income, \( A \) stands for total factor productivity, \( K \) is the aggregate stock of physical capital, \( h \) refers to the human capital per worker and \( L \) is the total number of workers. Rewriting the expression in
Table 4
Dependent Variable: $\Delta Gini_{1980-2005}^h$

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>High Income OECD</th>
<th>Less Developed</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta Gini_{1960-1980}^h$</td>
<td>0.109</td>
<td>-0.148</td>
<td></td>
<td>-0.096</td>
<td>-0.209</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.082)</td>
<td>(0.104)</td>
<td></td>
<td>(0.122)</td>
<td>(0.134)</td>
<td></td>
</tr>
<tr>
<td>$\Delta n_{1960-1980}$</td>
<td></td>
<td></td>
<td>0.171$^b$</td>
<td>-0.091</td>
<td>0.775$^c$</td>
<td>-0.150</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.074)</td>
<td>(0.107)</td>
<td>(0.389)</td>
<td>(0.117)</td>
</tr>
<tr>
<td>$\Delta Gini_{1960-1980}^LIT$</td>
<td>-0.438$^a$</td>
<td>-0.279$^b$</td>
<td></td>
<td>-0.319$^b$</td>
<td>-0.526$^c$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.163)</td>
<td>(0.131)</td>
<td></td>
<td>(0.124)</td>
<td>(0.307)</td>
<td></td>
</tr>
<tr>
<td>$lny_{1960}$</td>
<td></td>
<td></td>
<td></td>
<td>-0.014$^c$</td>
<td>-0.018$^c$</td>
<td>-0.009</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.008)</td>
<td>(0.010)</td>
<td>(0.012)</td>
</tr>
<tr>
<td>$Gini_{1960}^h$</td>
<td>-0.342$^a$</td>
<td>-0.305$^a$</td>
<td></td>
<td>-0.069</td>
<td>-0.059</td>
<td>-0.377$^a$</td>
</tr>
<tr>
<td></td>
<td>(0.091)</td>
<td>(0.097)</td>
<td></td>
<td>(0.157)</td>
<td>(0.134)</td>
<td>(0.097)</td>
</tr>
<tr>
<td>$Gini_{1960}^h$</td>
<td>-0.103$^b$</td>
<td></td>
<td></td>
<td>-0.093</td>
<td>-0.112$^b$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.046)</td>
<td></td>
<td></td>
<td>(0.103)</td>
<td>(0.049)</td>
<td></td>
</tr>
<tr>
<td>$n_{1960}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.099$^b$</td>
<td>0.278$^c$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.042)</td>
<td>(0.150)</td>
</tr>
<tr>
<td>$Gini_{1960}^LIT$</td>
<td>-0.018</td>
<td></td>
<td></td>
<td>-0.133</td>
<td>-0.085</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.150)</td>
<td></td>
<td></td>
<td>(0.180)</td>
<td>(0.218)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.013</td>
<td>0.283$^a$</td>
<td>0.034$^a$</td>
<td>0.299$^a$</td>
<td>0.292$^a$</td>
<td>0.295$^a$</td>
</tr>
<tr>
<td></td>
<td>(0.009)</td>
<td>(0.088)</td>
<td>(0.008)</td>
<td>(0.094)</td>
<td>(0.096)</td>
<td>(0.105)</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.024</td>
<td>0.397</td>
<td>0.147</td>
<td>0.417</td>
<td>0.117</td>
<td>0.289</td>
</tr>
<tr>
<td>Countries</td>
<td>73</td>
<td>73</td>
<td>73</td>
<td>73</td>
<td>22</td>
<td>51</td>
</tr>
</tbody>
</table>

Note: OLS regression. Robust Standard errors in parenthesis. a, b, and c are 1, 5, and 10 per cent significance level.
HUMAN CAPITAL AND INCOME INEQUALITY

per worker terms yields

\[ \frac{Y}{L} = A(K/L)^{\alpha} h^{1-\alpha} \]  

(8)

Following the literature on the private returns to education, the simplest macro specification of a Mincerian human capital production function can be written as follows:\(^{18}\)

\[ h = e^{\theta S} \]  

(9)

where \( S \) are the average years of schooling in the workforce and \( \theta \) is the return to education. Taking logs to eq. (8) yields

\[ \ln y = \ln A + \alpha \ln k + (1 - \alpha) \theta S \]  

(10)

We can decompose \( S \) as the addition of the average years of schooling of different levels of education

\[ S = S^{PRIM} + S^{SEC} + S^{TERT} \]  

(11)

Thus, the aggregate returns to primary, secondary and university can be computed by estimating the following specification:

\[ \log(y_{i,t}) = \beta_0 + \beta_1 \log(k_{i,t}) + \beta_2 S^{PRIM}_{i,t} + \beta_3 S^{SEC}_{i,t} + \beta_4 S^{TERT}_{i,t} + \gamma_i + \delta_t + \mu_{i,t} \]  

(12)

where \( y_{i,t} \) is real GDP per capita in country \( i \) measured at year \( t \), \( k_{i,t} \) is the stock of physical capital per worker, \( S^{PRIM}, S^{SEC}, S^{TERT} \) are the average years of schooling for the population 15 years and above in primary, secondary and tertiary education, \( \delta_t \) is a time-specific effect, \( \gamma_i \) stands for specific characteristics in every country that are constant over time and \( \mu_{i,t} \) collects the error term that varies across countries and across time. The results of the estimation of this specification would essentially establish that

\(^{18}\) Based on the work of Mincer (1974), the private returns to education are computed by estimating a wage equation in which the log of the individual’s wage is linearly related to the individual’s years of schooling, years of experience, years of experience squared and other characteristics such as gender, type or labor contract and sector of economic activity.
more education is good (if $\hat{\beta}_i > 0$) and which education level is more productive. For example, if the returns to education are convex we would expect that $\hat{\beta}_4 > \hat{\beta}_3 > \hat{\beta}_2$.

Table 5 displays the results of estimating eq. (12) by different econometric techniques. Column (1) assumes that $\gamma_i = 0$ and reports the OLS estimates. The results show that the estimated coefficients of all the levels of schooling are positive and statistically significant. The evidence also suggest that the returns to education are increasing with the level of schooling; one year increment in university is almost 5 times more productive than the increase in one year in primary education. Although this preliminary evidence is suggestive, OLS estimates may be biased since they fail to control for specific characteristics of countries that scarcely change over time. Column (2) shows that controlling for country specific effects with the fixed effect (FE) estimator reduces the size of all the educational variables. However, the FE estimates retain the convex effect of the different levels of education, yet the coefficient of the average years of primary education is no longer significant. We also test the robustness of the results by using instrumental variables in order to control for the endogeneity of education. In line with Barro and Lee (2010), we use the parent’s education lagged 10 years as instruments for the current years of schooling. Specifically, we use the average years of schooling in primary, secondary and university of the population 40-74 years old lagged 10 years as instrument for the educational variables, and the capital stock per worker lagged 10 years as instrument for the capital stock. As in the previous estimates, columns (3) and (4) show that the FE estimates are lower that their OLS counterpart and that the coefficients of the educational variables are positive and increasing with the level of education. An alternative way to control for unobservable heterogeneity is to estimate eq. (12) in first differences. This exercise is useful since it is in line with the previous evidence reported in most of the figures, which plotted the correlation between the changes in the income distribution and the changes in the educational variables. Column (5) reports the results for the 10-year increments and column (6) displays the long term difference over a period of 60 years. When data in 1950 or 2010 is missing, we replace the gap with the first and the last available observation. In that case, we compute the differences in the educational variables for the same years.
the coefficient of the average years of university education remains positive and higher than that of the secondary education, whereas that of the primary schooling is not significant at the standard levels.

Overall, the evidence shows that, although the increasing returns to education may still be driven by some other omitted variables, different econometric techniques point to a convex relationship between the aggregate level of output per worker and the years of schooling in primary, secondary and university education. The evidence shows that the estimated coefficient of the average years of primary education is lower than that of any other level of schooling. Moreover, in some regressions the coefficient is not statistically significant and, in others, the estimate is even negative. Thus, in spite of the large increment in the population with primary education, the low returns of an additional year of schooling at the primary level might have prevented low educated individuals to climb in the income ladder. Therefore, increasing returns to education could be a plausible explanation of the lack of correlation between the change in the income Gini coefficient and the global reduction of education inequality, mainly driven by an increment in the share of literates.

4.5. External effects and other alternative explanations

Another possible explanation of this puzzle is that improvements in literacy, which increase the wage of population at the bottom end of income distribution, have also coincided with an increase of wages in other cohorts of population with higher education, such that all of them maintain their incomes shares. The latter could be an indirect consequence of the fall in illiteracy, due to the external effects of having a more educated population (e.g., Lucas, 1988), or, alternatively, just a spurious correlation if the improvements in education at the bottom have coincided with other exogenous forces (e.g., globalization or skill-biased technological progress) that have increased wages at the top.

The problem is that, due to the lack of data for a large sample of countries and years, we cannot test these hypothesis. Nevertheless, casual information for some countries where this information is available reveals that real wages for low levels of education have remained almost constant or even present a negative trend from 1960 onwards, whereas wages for higher levels of education have exhibited a clear upward trend. In Figure
## Table 5

**Dependent Variable: Log of output per worker**

<table>
<thead>
<tr>
<th>Instrumental Variables</th>
<th>10-year</th>
<th>First Dif</th>
<th>First Dif</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1950-2010</td>
<td>1950-2010</td>
<td></td>
</tr>
<tr>
<td>OLS FE</td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \log(K/L) )</td>
<td>0.071(^a)</td>
<td>0.358(^a)</td>
<td>0.055(^a)</td>
</tr>
<tr>
<td>( S^{PRIM} )</td>
<td>0.100(^a)</td>
<td>0.014</td>
<td>0.112(^a)</td>
</tr>
<tr>
<td>( S^{SEC} )</td>
<td>0.165(^a)</td>
<td>0.060(^a)</td>
<td>0.206(^a)</td>
</tr>
<tr>
<td>( S^{TERT} )</td>
<td>0.477(^a)</td>
<td>0.173(^a)</td>
<td>0.197(^b)</td>
</tr>
<tr>
<td>Constant</td>
<td>3.231(^a)</td>
<td>4.273(^a)</td>
<td>3.181(^a)</td>
</tr>
<tr>
<td>( R^2 )</td>
<td>0.593</td>
<td>0.609</td>
<td>0.621</td>
</tr>
<tr>
<td>Obs.</td>
<td>1367</td>
<td>1367</td>
<td>1093</td>
</tr>
<tr>
<td>Countries</td>
<td>137</td>
<td>137</td>
<td>137</td>
</tr>
<tr>
<td>Country dummies</td>
<td>NO</td>
<td>YES</td>
<td>NO</td>
</tr>
<tr>
<td>Time dummies</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
</tbody>
</table>

*Note: Robust Standard errors in parenthesis. \(^a\), \(^b\), and \(^c\) are 1, 5, and 10 per cent significance level. The dependent variable is the log of the real GDP per worker. The explanatory variables are the log of the capital stock per worker and the average years of primary, secondary and university education of the population 15 years and above. The instrument for the log(K/L) is the 10-year lagged variable. The instruments for the average years of schooling in primary, secondary and university are the corresponding average years of schooling of each educational level among the population 40-74 years old lagged 10 years.*
we show the average annual real earnings by worker education level in the USA, from 1963 to 2010, taken from the Annual Report of the Council of Economic Advisers (2012). As we observe, the average earnings of workers with less than high school were in 2010 clearly below the levels in the 60s. On the contrary, workers with bachelor’s degree or higher experienced a significant increase in their earnings. According to the Council of Economic Advisers, a substantial fraction of the overall increase in income inequality in the USA is related to this growing divergence in earnings between workers with more years of education and those with fewer years of education. Using a sample of 12 countries from 1985 to 2005, the OECD (2011) also finds that the majority of countries have experienced an increase in the skill wage gap between high and low-skilled workers, across all sectors. Additionally, the evidence presented by the OECD (2012) for 31 countries from 2000 to 2010 is also consistent with this explanation. For the OECD average, the relative earnings of population with education below upper secondary, with respect to those with upper secondary and post-secondary non-tertiary education, have fallen from 80 to 76 per cent from 2000 to 2010. On the contrary, the relative earnings of population with ter-
Human Capital and Income Inequality

tertiary education have increased from 149 to 159 per cent in the same period, implying that their earnings gap with the population with education below upper secondary has increased in 14 percentage points in just ten years.

If this evidence was also the case for emerging countries, then increases in literacy rates (with the corresponding fall in human capital inequality) could rise wages and life standards of people at the bottom of the income distribution with no improvement in their income shares. In this situation, although the average worker benefits from higher wages, the increase in income of more educated people at the bottom is compensated by higher wages of cohorts with higher education, such that income inequality remains unchanged.

5. Conclusions

This paper documents trends in human capital inequality from 1950 to 2010 using an improved data set on human capital inequality measures. The evidence shows that most countries have experienced a dramatic drop in human capital inequality, due mainly to an unprecedented decrease in the share of illiterates, which has not been accompanied by a similar reduction in income inequality.

A closer examination of the data shows that the evolution of education inequality presents two stages. In early stages of development, as the share of illiterate population is very high, the evolution of the human capital Gini coefficient is determined by the evolution of the illiterate population. This stage is observed in most developing countries, in which the large reduction in the share of illiterates has led a significant decrease in the human capital Gini coefficient over time. In later stages of development, as most of the population are literate, the evolution of the human capital Gini coefficient is determined by the Gini among the literates. In this stage, however, there has not been a clear pattern in the evolution of human capital inequality. In some advanced countries we observe an increase in human capital inequality, whereas other countries have experienced a more even distribution of education.

Has income inequality reduced as much as education inequality? The answer is a clear no. In fact, the evidence shows that in many countries, a decrease in the inequality in the distribution of education has been ac-
Human Capital and Income Inequality

accompanied by an increase in income inequality. Thus, variations in the
distribution of education inequality have not contributed to reduce the in-
equality in the distribution of income. Additionally, the evidence shows
that increasing literacy is not a sufficient condition to reduce income in-
equality.

A plausible explanation for this puzzle could be that returns to school-
ing are increasing with the level of education. Thus, if returns to primary
schooling are low, a large reduction in the share of illiterates may not be
reflected into a sizeable increment in the wages of the population at the
bottom end of the income distribution. Using data for real GDP per worker
for a broad number of countries, we compute aggregate returns for dif-
ferent levels of education. The findings reveal that the returns to tertiary
education are much higher than those of secondary and primary school-
ing. In some specifications, the returns to primary education are even neg-
ative or non statistically significant. These findings are consistent with a
low quality educational system at the lower levels of schooling, which may
lead to higher literacy, but does not contribute to skill accumulation (e.g.,
Hanushek and Woessmann, 2012).

An alternative explanation is that improvements in literacy and wages
of population at the bottom end of income distribution have also coincided
with an increase of wages in other cohorts of population with higher ed-
ucation, such that all of them maintain their incomes shares. The latter
could reflect external effects of having a more educated population or just
a spurious correlation if the improvements in education at the bottom have
coincided with other exogenous forces (e.g., globalization or skill-biased
technological progress) that have increased wages at the top. In the latter
case, improvements in education and human capital inequality observed
in many countries would have avoided significant increases in income in-
equality.

The evidence presented in this paper is relevant because, in many de-
veloping countries, governments have made a great effort in eradicating
illiteracy rates, but these policies have not been accompanied by a more
even distribution of income. However, this evidence does not imply that
educational policies have not reduced poverty and improved wages and
the standards of living of hundreds of millions with better education. On
the contrary, better education is crucial to increase average earnings per
HUMAN CAPITAL AND INCOME INEQUALITY

worker, and the eradication of illiteracy is a necessary condition to ensure access to higher levels of education for all people.

References

Human Capital and Income Inequality

HUMAN CAPITAL AND INCOME INEQUALITY

tary Economics, 22, 3-42.

Working Papers

09/01 K.C. Fung, Alicia García-Herrero and Alan Siu: Production Sharing in Latin America and East Asia.

09/02 Alicia García-Herrero, Jacob Gyntelberg and Andrea Tesei: The Asian crisis: what did local stock markets expect?

09/03 Alicia García-Herrero and Santiago Fernández de Lis: The Spanish Approach: Dynamic Provisioning and other Tools.

09/04 Tatiana Alonso: Potencial futuro de la oferta mundial de petróleo: un análisis de las principales fuentes de incertidumbre.

09/05 Tatiana Alonso: Main sources of uncertainty in formulating potential growth scenarios for oil supply.

09/06 Ángel de la Fuente y Rafael Doménech: Convergencia real y envejecimiento: retos y propuestas.


09/08 Alicia García-Herrero, Philip Woolbridge and Doo Yong Yang: Why don’t Asians invest in Asia? The determinants of cross-border portfolio holdings.

09/09 Alicia García-Herrero, Sergio Gavilá and Daniel Santabárbara: What explains the low profitability of Chinese Banks?

09/10 J.E. Boscá, R. Doménech and J. Ferri: Tax Reforms and Labour-market Performance: An Evaluation for Spain using REMS.


09/12 J.E. Boscá, R. Doménech and J. Ferri: Search, Nash Bargaining and Rule of Thumb Consumers.

09/13 Angel Melguizo, Angel Muñoz, David Tuesta y Joaquín Vial: Reforma de las pensiones y política fiscal: algunas lecciones de Chile.


09/15 Angel Melguizo, Angel Muñoz, David Tuesta and Joaquín Vial: Pension reform and fiscal policy: some lessons from Chile.

09/16 Alicia García-Herrero and Tuuli Koivu: China’s Exchange Rate Policy and Asian Trade.


09/18 Alicia García Herrero y Daniel Santabárbara García: Una valoración de la reforma del sistema bancario de China.


10/01 Carlos Herrera: Rentabilidad de largo plazo y tasas de reemplazo en el Sistema de Pensiones de México.


10/05 Soledad Hormazabal D.: Gobierno Corporativo y Administradoras de Fondos de Pensiones (AFP). El caso chileno.


10/07 Rafael Doménech y Juan Ramón García: ¿Cómo Conseguir que Crezcan la Productividad y el Empleo, y Disminuya el Desequilibrio Exterior?

10/08 Markus Brückner and Antonio Ciccone: International Commodity Prices, Growth, and the Outbreak of Civil War in Sub-Saharan Africa.

10/09 Antonio Ciccone and Marek Jarocinski: Determinants of Economic Growth: Will Data Tell?

10/10 Antonio Ciccone and Markus Brückner: Rain and the Democratic Window of Opportunity.

10/11 Eduardo Fuentes: Incentivando la cotización voluntaria de los trabajadores independientes a los fondos de pensiones: una aproximación a partir del caso de Chile.

10/12 Eduardo Fuentes: Creating incentives for voluntary contributions to pension funds by independent workers: A primer based on the case of Chile.


10/16 Soledad Zignago: Determinantes del comercio internacional en tiempos de crisis.

10/17 Angel de la Fuente and José Emilio Boscá: EU cohesion aid to Spain: a data set Part I: 2000-06 planning period.

10/18 Angel de la Fuente: Infrastructures and productivity: an updated survey.

10/19 Jasmina Bjeletic, Carlos Herrera, David Tuesta and Javier Alonso: Simulaciones de rentabilidades en la industria de pensiones privadas en el Perú.


10/21 Máximo Camacho and Rafael Doménech: MICA-BBVA: A Factor Model of Economic and Financial Indicators for Short-term GDP Forecasting.

10/22 Enestor Dos Santos and Soledad Zignago: The impact of the emergence of China on Brazilian international trade.

10/23 Javier Alonso, Jasmina Bjeletic y David Tuesta: Elementos que justifican una comisión por saldo administrado en la industria de pensiones privadas en el Perú.
10/24 Javier Alonso, Jasmina Bjeletic y David Tuesta: Reasons to justify fees on assets in the Peruvian private pension sector.


10/26 Carlos A. Herrera: Long-term returns and replacement rates in Mexico’s pension system.

10/27 Soledad Hormazábal: Multifondos en el Sistema de Pensiones en Chile.

10/28 Soledad Hormazábal: Multi-funds in the Chilean Pension System.

10/29 Javier Alonso, Carlos Herrera, María Claudia Llanes y David Tuesta: Simulations of long-term returns and replacement rates in the Colombian pension system.

10/30 Javier Alonso, Carlos Herrera, María Claudia Llanes y David Tuesta: Simulaciones de rentabilidades de largo plazo y tasas de reemplazo en el sistema de pensiones de Colombia.

11/01 Alicia García Herrero: Hong Kong as international banking center: present and future.


11/03 Angel de la Fuente: Human capital and productivity

11/04 Adolfo Albo y Juan Luis Ordaz Díaz: Los determinantes de la migración y factores de la expulsión de la migración mexicana hacia el exterior, evidencia municipal.

11/05 Adolfo Albo y Juan Luis Ordaz Díaz: La Migración Mexicana hacia los Estados Unidos: Una breve radiografía.

11/06 Adolfo Albo y Juan Luis Ordaz Díaz: El Impacto de las Redes Sociales en los Ingresos de los Mexicanos en EEUU.

11/07 María Abascal, Luis Carranza, Mayte Ledo y Arnoldo López Marmolejo: Impacto de la Regulación Financiera sobre Países Emergentes.


11/09 Angel de la Fuente y Rafael Doménech: El impacto sobre el gasto de la reforma de las pensiones: una primera estimación.

11/10 Juan Yermo: El papel ineludible de las pensiones privadas en los sistemas de ingresos de jubilación.

11/11 Juan Yermo: The unavoidable role of private pensions in retirement income systems.

11/12 Angel de la Fuente and Rafael Doménech: The impact of Spanish pension reform on expenditure: A quick estimate.


11/14 David Tuesta: Una revisión de los sistemas de pensiones en Latinoamérica.


11/16 Adolfo Albo y Juan Luis Ordaz Díaz: La Migración en Arizona y los efectos de la Nueva Ley “SB-1070”.

11/17 Adolfo Albo y Juan Luis Ordaz Díaz: Los efectos económicos de la Migración en el país de destino. Los beneficios de la migración mexicana para Estados Unidos.


11/19 Angel de la Fuente y José E. Boscá: Gasto educativo por regiones y niveles en 2005.

11/20 Máximo Camacho and Agustín García Serrador: The Euro-Sting revisited: PMI versus ESI to obtain euro area GDP forecasts.
11/21 **Eduardo Fuentes Corripio**: Longevity Risk in Latin America.
11/22 **Eduardo Fuentes Corripio**: El riesgo de longevidad en Latinoamérica.
11/24 **Javier Alonso, Rafael Doménech y David Tuesta**: Public Pension Systems and the Fiscal Crisis in the Euro Zone. Lessons for Latin America.
11/25 **Adolfo Albo y Juan Luis Ordaz Díaz**: Migración mexicana altamente calificada en EEUU y Transferencia de México a Estados Unidos a través del gasto en la educación de los migrantes.
11/26 **Adolfo Albo y Juan Luis Ordaz Díaz**: Highly qualified Mexican immigrants in the U.S. and transfer of resources to the U.S. through the education costs of Mexican migrants.
11/27 **Adolfo Albo y Juan Luis Ordaz Díaz**: Migración y Cambio Climático. El caso mexicano.
11/28 **Adolfo Albo y Juan Luis Ordaz Díaz**: Migration and Climate Change: The Mexican Case.
11/29 **Ángel de la Fuente y María Gundín**: Indicadores de desempeño educativo regional: metodología y resultados para los cursos 2005-06 a 2007-08.
11/30 **Juan Ramón García**: Desempleo juvenil en España: causas y soluciones.
11/31 **Juan Ramón García**: Youth unemployment in Spain: causes and solutions.
11/32 **Mónica Correa-López and Beatriz de Blas**: International transmission of medium-term technology cycles: Evidence from Spain as a recipient country.
11/33 **Javier Alonso, Miguel Angel Caballero, Li Hui, María Claudia Llanes, David Tuesta, Yuwei Hu and Yun Cao**: Potential outcomes of private pension developments in China.
11/34 **Javier Alonso, Miguel Angel Caballero, Li Hui, María Claudia Llanes, David Tuesta, Yuwei Hu and Yun Cao**: Posibles consecuencias de la evolución de las pensiones privadas en China.
11/35 **Enestor Dos Santos**: Brazil on the global finance map: an analysis of the development of the Brazilian capital market
11/36 **Enestor Dos Santos, Diego Torres y David Tuesta**: Una revisión de los avances en la inversión en infraestructura en Latinoamérica y el papel de los fondos de pensiones privados.
11/37 **Enestor Dos Santos, Diego Torres and David Tuesta**: A review of recent infrastructure investment in Latin America and the role of private pension funds.
11/38 **Zhigang Li and Minqin Wu**: Estimating the Incidences of the Recent Pension Reform in China: Evidence from 100,000 Manufacturers'.
12/01 **Marcos Dal Bianco, Máximo Camacho and Gabriel Pérez-Quiros**: Short-run forecasting of the euro-dollar exchange rate with economic fundamentals.
12/03 **George Chouliarakis and Mónica Correa-López**: A Fair Wage Model of Unemployment with Inertia in Fairness Perceptions.
12/05 **Javier Alonso, Miguel Angel Caballero, Li Hui, María Claudia Llanes, David Tuesta, Yuwei Hu and Yun Cao**: Potential outcomes of private pension developments in China (Chinese Version).
12/06 **Alicia García-Herrero, Yingyi Tsai and Xia Le**: RMB Internationalization: What is in for Taiwan?
12/07 **K.C. Fung, Alicia García-Herrero, Mario Nigrinis Ospina**: Latin American Commodity Export Concentration: Is There a China Effect?
12/08 **Matt Ferchen, Alicia García-Herrero and Mario Nigrinis**: Evaluating Latin America's Commodity Dependence on China.

12/10 Maximo Camacho and Jaime Martinez-Martin: Real-time forecasting US GDP from small-scale factor models.

12/11 J.E. Boscá, R. Doménech and J. Ferri: Fiscal Devaluations in EMU.

12/12 Ángel de la Fuente and Rafael Doménech: The financial impact of Spanish pension reform: A quick estimate.

12/13 Biliana Alexandrova-Kabadjova, Sara G. Castellanos Pascacio, Alma L. García-Almanza: The Adoption Process of Payment Cards -An Agent- Based Approach:


12/15 Sara G. Castellanos, F. Javier Morales y Mariana A. Torán: Análisis del Uso de Servicios Financieros por Parte de las Empresas en México: ¿Qué nos dice el Censo Económico 2009?

12/16 Sara G. Castellanos, F. Javier Morales y Mariana A. Torán: Analysis of the Use of Financial Services by Companies in Mexico: What does the 2009 Economic Census tell us?


12/18 Chen Shiyuan, Zhou Yinggang: Revelation of the bond market (Chinese version).

12/19 Zhouying Gang, Chen Shiyuan: On the development strategy of the government bond market in China (Chinese version).

12/20 Ángel de la Fuente and Rafael Doménech: Educational Attainment in the OECD, 1960-2010.


12/22 Santiago Fernández de Lis and Alicia Garcia-Herrero: Dynamic provisioning: a buffer rather than a countercyclical tool?

12/23 Ángel de la Fuente: El nuevo sistema de financiación de las Comunidades Autónomas de régimen común: un análisis crítico y datos homogéneos para 2009 y 2010.

12/24: Beatriz Irene Balmaseda Perez, Lizbeth Necoechea Hasfield: Metodología de estimación del número de clientes del Sistema Bancario en México.


12/26 Oscar Arce, José Manuel Campa y Ángel Gavilán: Macroeconomic Adjustment under Loose Financing Conditions in the Construction Sector.

12/27 Ángel de la Fuente: Algunas propuestas para la reforma del sistema de financiación de las comunidades autónomas de régimen común.

12/28 Amparo Castelló-Climent and Rafael Doménech: Human Capital and Income Inequality: Some Facts and Some Puzzles.

The analysis, opinions, and conclusions included in this document are the property of the author of the report and are not necessarily property of the BBVA Group.

BBVA Research’s publications can be viewed on the following website: http://www.bbvaresearch.com