Should less inequality in education lead to a more equal income distribution?

Péter Földvári*

University of Debrecen, Faculty of Economics and Business Administration

Kassai str 26, 4028 Debrecen, Hungary

e-mail: peter.foldvari@econ.unideb.hu

and

Bas van Leeuwen

University of Warwick, Department of Economics,

CV4 7AL, Coventry, the UK

* Corresponding author

Abstract

In this paper we revisit the question whether inequality in education and human capital should be closely related to income inequality. Using the most popular functional forms describing the relationship between output and human capital and education and human capital, we find that the effect of inequality in schooling on income inequality should be very low, even insignificant in an economic sense. This is confirmed by our empirical analysis since we find that the Gini coefficient of the education yields an insignificant coefficient. We cannot confirm either that a more equal distribution of education leads to higher income per capita, even though this result is sensitive to the choice of data.

1. Introduction

It seems all too logical that human capital inequality should affect income inequality. Numerous studies in the Mincerian (1974) tradition have shown that a higher number of years of education results in higher earnings (Psacharopoulos 1994; Psacharopoulos and Patrinos 2004). Hence, if the inequality of years of education in a country increases, should this not also apply to income inequality? Many studies have answered this question with “yes” (see Checci 2004 for an example), and one may at first arrive at similar conclusion from the fundamental models in this field by Becker (1964) and Mincer (1974). There are dissenting views as well: Knight and Sabot (1983), for example, argue that the effect of the expansion of education may at first increase income inequality, but later it should rather lead to a reduction due to decreasing skill premium.

The empirical evidence for the relationship between inequality of education and that of income is also ambiguous. While Becker and Chiswick (1966) find that inequality in education is correlated with inequality in income in the USA, Ram’s (1984, 1989) results suggest that the variance of the educational attainment is uncorrelated with income inequality. In a more recent work, De Gregorio and Lee (2002) find that there is a small but positive relationship between educational and income inequality. Studies focusing on the effect of inequality in education on economic growth lead to much more uniform results: Lopez *et al.* (1998) and Castello and Domenech (2002) find that more equal distribution of human capital is associated with faster growth. However, in both studies the Barro and Lee dataset is used, which is generally considered to be less reliable than its alternatives.

There are several possible explanations why the empirical results are indecisive. One stream of the literature holds either misspecification or wrong statistical indicators responsible. For example, Morrisson and Murtin (2007) argue that inequality in the 'average years of education' is not the same as inequality in human capital. Following Mincer (1974) they argue that the number of years of education has to be multiplied with the rate of return to obtain a reliable estimate of the human capital stock. This approach, however, makes it very likely that there is a non-linear relationship between human capital and education since the rate of returns to education is expected to decrease as the average education level rises. There are also authors who are generally skeptical regarding the use of the Gini coefficient as a measure of inequality (Frankema and Bolt 2006). Their main argument is that the Gini coefficient is "level-dependent" that is not-translation invariant, i.e. the lower the average years of schooling in a country, the bigger the effect of the gap in average years of education between individuals on the Gini coefficient is. Hence, a higher average level of educational attainment, *ceteris paribus*, lowers the Gini coefficient. Also the quality of the estimated income Gini coefficients can be questioned (François and Rojas-Romagosa 2005).

In this paper we offer another explanation why one is likely to find just a very small, probably insignificant effect of inequality in schooling on inequality in income. We argue that using the standard and popular specifications of the production function and the relationship between human capital and education, it is possible to show that the link between inequalities in education and income are just loosely related. This has serious consequences for the efficiency of development policies seeking the reduction of inequality by a more equal distribution of schooling, because it suggests policy ineffectiveness.

In the next section we describe our data sources and inequality estimation methods. Then, we establish an estimable relationship between inequalities of education and income, which is tested in section 4. Here we find no evidence for any significant relationship between inequality in education and the inequality or the level of income. In Section 5 we summarize our main findings.

2. Data

The educational Gini coefficients were constructed from the Barro and Lee (2001) and the Cohen and Soto (2007) databases. We used the formula as suggested by Thomas, Wang, and Fan (2000), Checchi (2004) and Castelló and Doménech (2000, 4). They started with

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

Where $\overline{H}$ is average years of schooling in the population aged 15 years and over, *i* and *j* are different levels of education, n_i and n_j are the shares of population with a given level of education, and $\hat{x}_i$ and $\hat{x}_j$ are the cumulative average years of schooling at each educational level. Again following Castelló and Doménech (2000), we can rewrite this equation in terms of the Barro and Lee data as:

$$G^h = n_0 \frac{n_1 x_2 (n_2 + n_3) + n_3 x_3 (n_1 + n_2)}{n_1 x_1 + n_2 (x_1 + x_2) + n_3 (x_1 + x_2 + x_3)} \quad (2.)$$

Where $x_0 = 0$, x_1 is "average years of primary schooling in the total population" divided by the percentage of the population with at least primary education; x_2 is "average years of secondary schooling in the total population" divided by the percentage of the population with at least secondary education; x_3 is "average years of higher schooling in the total population" divided by the percentage of the population with at least higher; n_0 is the percentage population with no education; n_1 the percentage in the population with primary education; n_2 the percentage in the population with secondary education, and n_3 the percentage in the population with higher education.

An overview of the results is presented in table 1. Although the patterns of the

Table 1

Gini coefficients seem to be consistent, we need to be aware that some authors have pointed at biases in the Barro and Lee data (De La Fuente ad Doménech 2000; Krueger and Lindahl 2001; Portela et al. 2004; Cohen and Soto 2007). For this reason we also estimated the Gini coefficients for education using the Cohen and Soto dataset. This reduces the total number of observations, but improves the quality and the robustness of our results.

The income inequality Ginis were taken from the World Income Inequality Database (WIID). We took as much as possible consistent series within one country following the three-fold distinction proposed by François and Rojas-Romagosa (2005), thus gross household income, net household income, and expenditure person. This of course means that

Table 2

there may be differences among countries, which need to be removed using a within group transformation or captured by country-dummies. Finally, per capita GDP and investment rates were taken from the Penn World Table (Heston, Summers and Aten 2006).

3. The relationship between inequality in income and in education

As pointed out in the introduction, it is generally expected that inequality in education (and human capital) should affect income inequality. If this is the case, government policies pursuing higher equality in education should be efficient in reducing income inequality as well. This is not necessarily supported by empirical evidence though: one does not need to search far in order to find examples in which an increase in formal education does not necessary lead to a more beneficial economic position. Easterly (2002, 83) mentions Pakistan as an example where politicians use teaching positions as patronage, causing three-quarters of the teachers not being able to pass the exams they administer to their students. The same argument can be applied to international organizations. Heyneman (2003) argues that the manpower planning policy of the World Bank in the 1960s and 1970s, focusing on vocational education, was economically ineffective. Indeed, in Indonesia this often only led to changing the nameplates on the doors of the schools involved. We can conclude that one important explanation for finding no relationship between inequality in schooling and income inequality may sometimes arise from such policy failures. There is another explanation, though. In the followings we will demonstrate that by applying the functional forms that are the most often used to establish relationship between income and human capital, and human capital and schooling, we find that inequality in schooling and inequality in income are just loosely related, and the coefficients from such regressions is expected to be very small, probably insignificant.

For simplicity, we assume that the individual *i*'s output for each year *t* is determined by a Cobb-Douglas production function.

$$y_{i,t} = A_t k_{i,t}^{\alpha} h_{i,t}^{1-\alpha} l_t^{1-\alpha} \varepsilon_{i,t} \quad (3.)$$

With the level of technology (A) and work effort ($0<l<1$) assumed to be uniform across all individuals. The income (and in this case real wage) inequality is caused by the differences in the individual's physical (k_i) and human (h) capital endowment. Finally we include a stochastic error term ε_i, assumed to be i.i.d..

The variance of the logarithm of income in the society for each year *t* can be expressed as follows (under the assumption that the regressors are uncorrelated with the error term):

$$Var(\ln y_i)_t = \alpha^2 Var(\ln k_i)_t + (1-\alpha)^2 Var(\ln h_i)_t + 2\alpha(1-\alpha)Cov(\ln k_i, \ln h_i)_t + Var(\ln \varepsilon_i)_t \quad (4.)$$

From (4.) it is straightforward, that a higher variance of human capital (which we now interpret as a higher inequality of human capital) should indeed increase the variance of income (or income inequality) but this effect depends on the parameter α of the production function. Under the standard assumptions (*α* being around 0.3), we may expect that a unit increase in the variance of human capital translates into a lower (roughly 0.5 unit) increase in the variance of the income. Clearly, inequality in physical capital stock (that is an unequal distribution of the ownership of capital goods or capital incomes) has an effect of similar (most probably smaller) magnitude.

As for the covariance between the logs of physical and human capital in (4.), the relationship between the distributions of different factors of production is very likely to be affected by unknown country-specific, institutional factors. A positive covariance between capital stock and human capital, for example, can be interpreted so that people with more human capital are also more likely to enjoy more of the capital incomes, which is probably true for all societies. Alternatively, one may consider it as a measure of the extent of human and physical capital being employed together in the production process. Even though this correlation is not observable one may either assume that this correlation changes slowly enough that it can be treated as a constant country specific factor (taken care of by the help of fixed effect panel specification) or alternatively, one may assume that a higher correlation between the two kind of capital is associated with higher efficiency and consequently higher GDP. If this is the case, GDP per capita should be included in the regression.

In order to carry out an empirical analysis, we need to establish a relationship between inequality of education and the inequality of the human capital stock. The literature has two main assumptions regarding this: the first, that we call traditional view, assumes that the average human capital stock depends directly on the level of educational attainment. This idea can be operationalized as follows:

$$\ln h_{i,t} = \varphi_0 + \varphi_1 S_{i,t} \quad (5.)$$

where S_i denotes the educational attainment of individual i and φ_1 is a technical parameter between zero and one.[†] Now if we pursue this approach, educational attainment is quite easy to integrate into (4.), since:

$$Var(\ln h_i)_t = \varphi_1^2 Var(S_i)_t \quad (6.)$$

Which can be substituted into (4.)

$$Var(\ln y_i)_t = \alpha^2 Var(\ln k_i)_t + (1-\alpha)^2 \varphi_1^2 Var(S_i)_t + 2\alpha(1-\alpha) Cov(\ln k_i, \ln h_i)_t + Var(\ln \varepsilon_i)_t \quad (7.)$$

For the second view on the relationship between human capital and educational inequality we have to remain closer to the original idea by Mincer (1974) as proposed by studies like Bils and Klenow (1998) and Hall and Jones (1999). Here we need to slightly adjust (5.) as follows:

$$\ln h_{i,t} = r_{i,t} \cdot S_{i,t} \; (8.)$$

where $r_{i,t}$ denotes the returns to education in country i in year t. The main problem now, is that without knowing the joint distribution of r and S, we cannot say anything about the education's effect on inequality in human capital and income. It is very likely, that due to decreasing returns to schooling, at very high levels of education the reduction of r can even offset the effect of any increase in schooling on human capital. We can, however, use Murtin and Morrison's (2007) result about the relationship between r and S to simplify (8.). They use the data by Psacharopulos and Patrinos (2004) to estimate the following relationship between

[†] We can even try to guesstimate the value of φ_1. Let us first assume that, in accordance with the Solow model, the long-run growth of the GDP per capita equals the rate of technological development. This implies that the technology in the USA grew by roughly 2.5% annually in the period 1960-1999 (Penn World Table 6.1 data). At the same time the average years of education grew by 0.09 years per year. Now even under the very unlikely assumption that all the observed growth was caused by improvements in human capital endowment (with α=0.7), the parameter φ_1 should be around just 0.194. This is an upper bound estimate; the real effect must be lower.

the rate of returns to education and the average years of schooling (standard errors are reported in parentheses)[‡]:

$$r_i = \underset{(0.009)}{0.125} - \underset{(0.0017)}{0.004} S_i + u_i \quad (9.)$$

Using this, we can rewrite (8.):

$$\ln h_{i,t} = 0.125 S_{i,t} - 0.004 S_{i,t}^2 \quad (10.)$$

From which we can express the variance of the logarithm of human capital:

$$Var(\ln h_{i,t}) = 0.125^2 \cdot Var(S_{i,t}) - 0.004^2 Var(S_{i,t}^2) \quad (11.)$$

We find that the effect of increasing equality in education has a decreasing effect on the equality in human capital due to the decreasing rate of returns to schooling.

Now substituting (11.) into (4.) yields:

$$\begin{aligned} Var(\ln y_i)_t &= \alpha^2 Var(\ln k_i)_t + (1-\alpha)^2 \left(0.125^2 \cdot Var(S_{i,t}) - 0.004^2 Var(S_{i,t}^2)\right) + \\ &+ 2\alpha\left(1-\alpha\right) Cov(\ln k_i, \ln h_i)_t + Var(\ln \varepsilon_i)_t \end{aligned} \quad (12.)$$

If we assume that $Var(S^2)$ is larger than Var(S), we find that theoretically there might exist such a high level of inequality in educational attainment, at which any further increase in the inequality of education begins to reduce income inequality (leading to a negative relationship between inequality in education and income).

Both equations (7.) and (12.) suggest that if the traditional relationship between formal education and human capital holds, reducing the inequality in education is a quite weak tool of achieving a more equal distribution of income. According to (12.) if we were to estimate this relationship by a regression, we would expect the variance of the average years of education to yield a coefficient of 0.7^2 times 0.125^2 equaling 0.00766. Even if it were

[‡] One could also consider following Hall and Jones (1999) in assuming that the rate of returns to education is uniform in all countries but differ only by education level: 13.4% in the first four years of education, 10.1% in the next four years, and 6.8% for any further years spent with education. This however basically expresses the same non-linearity already captured in equation (9) and would just make the derivation less convenient, but would essentially not change the results.

significant in statistical sense, borrowing the expression from McCloskey and Ziliak (1996), it would not be significant in an economic sense.

This seems to be confirmed by an initial test on data: the linear correlation coefficient between the Gini coefficients of income and average years of education in our pooled dataset is just 0.33, which is indeed quite low. Also, plotting the Gini coefficients of income (*giniy*) against inequality in education (*giniedu*) does not reveal an obvious relationship (see Figure 1 and 2).

Figure 1, Figure 2

4. Empirical tests

Since we have no data on the distribution of physical capital stock or the covariances in (4.), we need to assume that these reflect mostly country-specific, probably institutional factors, which we can assume to be constant over the sample period (the length of sample period varies per country due to the heterogeneous availability of survey data). Our empirical specification is as follows:

$$G_{i,t}^{y} = \beta_0 + \beta_1 G_{i,t}^{S} + \beta_2 \ln y_{i,t} + \beta_2 S_{i,t} + \eta_i + \mu_t + u_{i,t} \quad (13.)$$

Where $G_{i,t}^{y}$ and $G_{i,t}^{S}$ denote the Gini coefficients of the income and the years of education in country i in year t, and y and S denotes the per capita income and the average years of education. η_i and μ_t denote the unobserved country-specific and years specific effects and $u_{i,t}$ is the error term assumed to be i.i.d.. Since it is possible that the relationship between inequality in education and income may be different between developed and developing countries, we also estimated (13.) on a OECD and a non-OECD sub-sample. An important issue to address is that even though we use variances in our derivations as a measure of inequality (which makes calculations convenient), we rely on Gini coefficients in the regression analysis. Fortunately, Milanovic (1997) shows that the Gini coefficient can be expressed as a function of variance.[§] If we include the mean of income and the average years of education in the regression, the coefficient of the educational Gini should reflect the effect of a change in the standard deviation, since the mean is already fixed.

Table 3

The results from Table 3 are indicative that after the time-invariant country-specific effects are captured, inequality in education seems not to affect income inequality at all. This finding does not depend on which dataset (Barro-Lee or Cohen-Soto) we used to estimate the inequality in education and applies to both country groups. Empirics seem to suggest the same we already suspected: there is no significant relationship between the inequality in education and inequality in income. As a result, any policy that expects such an effect is likely to fail.

5. Inequality in education and economic growth

There is another reason for a government to attempt to reduce educational inequality: as we mentioned in the introduction, the empirical evidence on a link between inequality in education and economic growth seems to be sound. In order to arrive at an empirically

---

[§] More precisely, he proves that $G \approx \frac{1}{\sqrt{3}} \frac{\sigma_y}{\bar{y}} \rho(y, r_y)$, where G denotes the Gini coefficient, and $\rho(y,r_y)$ is the correlation coefficient between income and rank. He also reports the rank correlation coefficient for a number of countries, which seem to be rather similar in countries with the same level of development. Our assumption that differences in the rank correlation coefficients can be taken as constant in our sample period by country (another reason for fixed effect panel specification), does not seem unrealistic.

testable theoretical model, we follow Lopez *et al*. (1998), who illustrate that the amount of human capital employed in production is not independent of its distribution.[**]

First, we assume that each individual (i) has the following Cobb-Douglas type production function (we disregard technology for convenience):

$$y_{i.t} = k_{i,t}^{\alpha} h_{i,t}^{1-\alpha} \quad (14.)$$

and the average income per capita can be calculated through aggregating the individual production functions:

$$y_t = \frac{1}{N}\sum_{i=1}^{N} k_{i,t}^{\alpha} h_{i,t}^{1-\alpha} \quad (15.)$$

We apply the Taylor theorem to expand (15.) around its mean (denoted by hat) up to the second order, which yields:

$$Y_t = \sum_{i=1}^{N}\hat{k}_t^{\alpha}\hat{h}_t^{1-\alpha} + \alpha\sum_{i=1}^{N}\hat{k}_t^{\alpha-1}\hat{h}_t^{1-\alpha}(k_{i,t}-\hat{k}_t) + (1-\alpha)\sum_{i=1}^{N}\hat{k}_t^{\alpha}\hat{h}_t^{-\alpha}(h_{i,t}-\hat{h}_t) + \frac{1}{2}\left(\alpha(\alpha-1)\sum_{i=1}^{N}\hat{k}_t^{\alpha-2}\hat{h}_t^{1-\alpha}(k_{i,t}-\hat{k}_t)^2\right) - $$
$$-\frac{1}{2}\left(\alpha(1-\alpha)\sum_{i=1}^{N}\hat{k}_t^{\alpha}\hat{h}_t^{-\alpha-1}(h_{i,t}-\hat{h}_t)^2\right) + \alpha(1-\alpha)\sum_{i=1}^{N}\hat{k}_t^{\alpha-1}\hat{h}_t^{-\alpha}(k_{i,t}-\hat{k}_t)(h_{i,t}-\hat{h}_t) + O(t)$$

(16.)

Where O(t) denotes the higher order derivatives that we omit in the followings from the derivations. We take the average of (16.) to arrive at the per capita income:

$$y_t = \hat{k}_t^{\alpha}\hat{h}_t^{1-\alpha} + \frac{1}{2}\alpha(\alpha-1)\hat{k}_t^{\alpha-2}\hat{h}_t^{1-\alpha}Var(k) - \frac{1}{2}\alpha(1-\alpha)\hat{k}_t^{\alpha}\hat{h}_t^{-\alpha-1}Var(h) + \alpha(1-\alpha)\hat{k}_t^{\alpha-1}\hat{h}_t^{-\alpha}Cov(k,h) \quad (17.)$$

After some simplifications we arrive at:

$$y_t = \hat{k}_t^{\alpha}\hat{h}_t^{1-\alpha}\left[1 + \frac{1}{2}\alpha(\alpha-1)\left(\frac{Var(k)}{\hat{k}_t^2} + \frac{Var(h)}{\hat{h}_t^2}\right) + \frac{Cov(k,h)}{\hat{k}_t\,\hat{h}_t}\right] \quad (18.)$$

Expressed in words, (18.) is indicative that a higher inequality of human capital (and physical capital) should reduce per capita income *ceteris paribus* ($0<\alpha<1$), while a higher relationship between the distribution of human and physical capital (that is an individual with higher

[**] We follow Lopez *et al.* (1998) in their suggestion to use a Taylor approximation, but our production function is different: we omit abilities from our model, while they neglect the inequality of the distribution of physical capital.

human capital also have more physical capital) is beneficial. Again we must face the problem that we do not know the variance of physical capital, neither the covariance between *k* and *h*. We assume therefore that these can be captured by the country dummies or gotten rid of by within group transformation. The relationship in (18) is obviously non-linear, but, with given assumptions regarding the parameters, at least monotonous. For convenience, we approximate the relationship with a linear regression and estimate the following fixed effect panel specification:

$$\ln y_{i,t} = \beta_0 + \beta_1 \ln \frac{I_{i,t}}{Y_{i,t}} + \beta_2 \ln pop_{i,t} + \beta_4 S_{i,t} + \beta_5 G^{y}_{i,t} + \beta_6 G^{S}_{i,t} + \eta_i + \mu_t + \varepsilon_{i,t} \quad (19.)$$

where $I_{i,t}$ denotes investments in country *i* in year *t*. We use the investment to output ratio (*I*/*Y*) taken from the Penn World Table 6.1 to proxy for the capital stock which was not available for most countries in our sample.

Table 4

We find that our results depend on the choice of data: using the Barro-Lee data we arrive at the conclusion that, with the exception of OECD countries, the Gini of the formal education yields the expected negative coefficient significant at 10%. We can interpret this result as follows: using the Barro-Lee data, we find that the average education Gini in the non-OECD sample is 50.9, if this were reduced to the OECD average (22.7) that is by roughly 28, it would – according to our estimation – increase the average GDP per capita in the non-OECD countries by about 25.2%, which is a considerable effect. If we use the better quality Cohen-Soto data, however, we find no significant effect between the inequality in education and the level of per capita income. Since the quality of the Cohen and Soto data is better (even though this comes at the cost of fewer observations) we are tempted to accept the latter results. Nevertheless, we have a good reason to suspect that there is a simultaneity bias present in our estimates in Table 4. Namely, the investment to GDP ratio depends on the level of income. We use the lagged values of the investment to GDP ratio and the population as instruments since the first one is predetermined, and the second one is exogenous (we have no reason to suspect that higher income would immediately, or in the short run, lead to a change in population size). The 2SLS procedure yields the following results:

Table 5

As the over-identification test suggest, we use proper instruments even though in the OECD subsample they prove to be weak and therefore those results are not reported. Hence, after taking care of the possible simultaneity we still find that the effect of the educational inequality strongly depends on the choice of data. Again: with the Barro-Lee data we find that in non-OECD countries a lower educational inequality leads to a higher income per capita *ceteris paribus*, while in case of the Cohen and Soto data, we find no significant effect.

5. Conclusion

It would be a natural reaction to expect a high and significant effect of educational- on income inequality. The many studies that have focused on this relationship have not corroborated this expectation so far. Looking at the two models that are usually used to specify some relationship between human capital and education, we find that theoretically they all should lead to a weak if not completely non-existent relationship between educational- and income inequality. Empirically we also find no evidence for the existence of a significant relationship between educational- on income inequality.

This finding has implications for government policies targeting education. It seems that policies that seek to reduce income inequality through more equal distribution of schooling have quite low chance of success. In addition, even though it is also assumed that the lower educational inequality leads to a higher average income level, our empirical test does not seem to support this hypothesis either. Although, if we rely on the Barron and Lee data, we find a positive effect of a reduction of inequality in education on the income level, but this effect disappears when we use the better quality Cohen and Soto data. Since the latter are generally accepted to be of better quality, we conclude that there is no apparent link between inequality in education and income inequality.

References

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

Becker, G. S. (1964) Human Capital, Columbia University: New York

Becker G. S., Chiswick B. R. (1966) Education and the Distribution of Earnings. The American Economic Review 56 (1/2); 358-369

Bils, M., Klenow, P. (1998) Does Schooling Cause Growth or the Other Way Around? NBER Working Paper No. 6393

Castello, A. and Domenech, R. (2002) Human capital inequality and economic growth: some new evidence. Economic Journal 112; 187-200

Castelló-Climent, A. and Doménech R. (2008) Human Capital Inequality, Life Expectancy and Economic Growth, Economic Journal, Vol. 118 (528), 653-677

Checchi D. (2004) Does educational achievement help to explain income inequality? Chapter 4 in A. Cornia (ed), *Inequality, Growth and Poverty in an Era of Liberalization and Globalization*. Oxford University Press

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

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

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

François, J.F., and H. Rojas-Romagosa (2005) The Construction and Interpretation of Combined Cross Section and Time-Series Inequality Datasets, *World Bank Policy Research Working Paper* 3748

Frankema, Ewout, and Bolt, J. (2006) Measuring and Analysing Educational Inequality: The Distribution of Grade Enrolment Rates in Latin America and Sub-Saharan Africa, Groninen Growth and Development Centre, Research Memorandum GD-86

Easterly, William (2002), *The Elusive Quest for Growth: Economists' Adventures and Misadventures in the Tropics*, The MIT Press: Cambridge, Massachusetts and London.

Heston, A., Summers R. and Aten, B. (2006) Penn World Table Version 6.2, Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania

Heyneman, S.P. (2003) The history and problems in the making of education policy at the World Bank 1960-2000, *International Journal of Educational Development*, Vol. 23, 315-337.

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

Knight, J. B. and Sabot, R.H. (1983) Educational expansion and the Kuznets effect. American Economic Review, Vol. 73 (5), 1132-1136.

Krueger, Alan B., and Lindahl, M. (2001) Education for Growth: Why and For Whom?, *Journal of Economic Literature*, Vol. 39 (4), 1101-1136.

Lopez, R., Thomas, V. and Wang, Y. (1998) Addressing the Education Puzzle. The Distribution of Education and Economic Reforms, Policy Research Working Paper 2031, Washington DC: The World Bank.

McCloskey, D. N. and Ziliak, S. T. (1996) “The Standard Error of Regressions,” *Journal of Economic Literature* 34(1), 97-114.

Meyer, J. W., Ramirez, F.O., Soysal, Y. N. (1992) World Expansion of Mass Education, 1870-1980, *Sociology of Education*, Vol. 65 (2), 128-149.

Milanovic, B. (1997) A simple way to calculate the Gini coefficient and its implications, *Economics Letters*, vol. 56 (1), 45–49.

Mincer, J. (1974), ‘Schooling, Experience and Earnings,’ NBER, New York

Morrisson, C., and Murtin, F. (2007) Education Inequalities and the Kuznets Curves: A Global Perspective Since 1870, PSE Working Papers 2007-12

Psacharopoulos, G. (1994), Returns to Investments in Education: A Global Update, World Development, Vol. 22 (9), 1325-1343.

Psacharopoulos, G. and Patrinos, H. (2004) Returns to investment in education: a further update, Education Economics, vol. 12 (2), 111-134.

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

Ram, R. (1984) Population Increase, Economic Growth, Educational Inequality, and Income Distribution: Some Recent Evidence, *Journal of Development Economics*, 14, 419-428.

Ram, R. (1989) Can Educational Expansion Reduce Income Inequality in Less-Developed Countries?, *Economics of Education Review* 8, 185-189.

Ramirez, F. O. and Boli, J. (1987) The Political Construction of Mass Schooling: European Origins and Worldwide Institutionalization, *Sociology of Education*, Vol. 60 (1), 2-17.

Stock, J. H. and Yogo, M. (2002) Testing for Weak Instruments in Linear IV Regression. NBER Technical Working Paper 284.

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

United Nations (1997) *Demographic Yearbook (DYB),* Special Issues - Historical Supplement.

Table 1
Educational Gini-coefficients 1960-2000

	Educational Gini from the Barro-Lee data					
	Full sample		*OECD countries*		*Non-OECD countries*	
	Average Gini	No. observations	Average Gini	No. observations	Average Gini	No. observations
1960	50.5	99	22.8	20	57.5	79
1965	49.6	99	23.3	21	56.7	78
1970	48.1	101	23.5	21	54.6	80
1975	47.9	106	23.8	23	54.6	83
1980	45.3	106	23.1	23	51.4	83
1985	43.0	107	22.7	23	48.6	84
1990	41.3	109	22.1	23	46.4	86
1995	39.1	105	21.8	25	44.6	80
2000	38.0	105	21.8	27	43.6	78

	Educational Gini from the Cohen-Soto data					
	Full sample		*OECD countries*		*Non-OECD countries*	
	Average Gini	No. observations	Average Gini	No. observations	Average Gini	No. observations
1960	54.3	92	20.2	19	63.1	73
1970	47.9	92	16.1	20	56.7	72
1980	40.1	92	12.9	22	49.8	70
1990	34.4	92	10.3	22	42.0	70
2000	29.4	92	9.4	25	36.9	67

Table 2
Income Gini-coefficients 1960-2000

	Income Gini					
	Full sample		*OECD countries*		*Non-OECD countries*	
	Average Gini	No. observations	Average Gini	No. observations	Average Gini	No. observations
1960	44.5	48	40.6	9	45.4	39
1965	40.7	44	36.1	13	42.6	31
1970	43.3	61	36.6	13	45.2	48
1975	41.6	61	34.5	19	44.8	42
1980	40.2	63	33.1	21	43.7	42
1985	39.8	66	31.5	22	43.9	44
1990	42.1	84	32.0	23	45.9	61
1995	43.7	96	34.0	25	47.1	71
2000	42.6	84	33.6	27	46.9	57

Table 3

Results from regression (13.), fixed effect panel (year dummies are included but not reported), dependent variable is the income Gini

	Income Gini					
	Full sample	*Full sample*	*OECD countries*	*OECD countries*	*Non-OECD countries*	*Non-OECD countries*
Constant	44.11** (2.46)	53.08** (2.58)	108.20** (1.98)	138.25* (1.74)	26.74 (1.35)	28.75 (1.45)
Education Gini (Barro-Lee)	-0.089 (-0.79)	-	-0.193 (-0.97)	-	0.119 (0.85)	-
Average years of education (Barro-Lee)	-1.227* (-1.95)	-	-1.558 (-1.42)	-	-0.471 (-0.62)	-
Education Gini (Cohen-Soto)	-	-0.078 (-0.96)	-	0.162 (0.91)	-	0.023 (0.27)
Average years of education (Cohen-Soto)	-	-1.694 (-1.36)	-	-3.137 (-0.75)	-	-2.205* (-1.87)
ln(y)	1.122 (0.59)	0.343 (0.16)	-5.760 (-1.02)	-8.319 (-0.91)	1.896 (0.92)	3.075 (1.41)
R^2	0.057	0.045	0.380	0.407	0.066	0.084
N	524	278	166	89	358	189
Joint significance test of the year dummies (F-test)	3.03 (p=0.045)	2.28 (p=0.066)	3.80 (p=0.005)	3.07 (p=0.036)	1.92 (p=0.070)	2.27 (p=0.071)

Note: Heteroscedasticity and autocorrelation robust t-statistics are reported in parentheses. The sign ***,**,* indicate the coefficient being significantly different form zero at a level of significance 1, 5, and 10%.

Table 4
The effect of inequality in education on per capita income, fixed effect panel (year dummies are included but not reported)

	Full sample		*OECD countries*		*Non-OECD countries*	
Constant	12.364*** (6.47)	12.197*** (6.78)	10.649*** (5.61)	10.674 (4.17)	13.693*** (4.38)	12.221*** (3.73)
$\ln \frac{I_{i,t}}{Y_{i,t}}$	0.294*** (5.20)	0.120** (2.05)	0.232*** (3.87)	0.077 (1.01)	0.293*** (4.65)	0.088 (1.25)
$\ln pop_{i,t}$	-0.586*** (-3.08)	-0.613*** (-3.36)	-0.263 (-1.47)	-0.169 (-0.71)	-0.786** (-2.34)	-0.645* (-1.84)
Average years of education (Barro-Lee)	0.090*** (2.73)	-	0.047** (2.09)	-	0.096** (1.97)	-
Gini Education (Barro-Lee)	-0.009* (-1.87)	-	0.001 (0.24)	-	-0.0096* (-1.71)	-
Average years of education (Cohen-Soto)	-	0.245*** (4.05)	-	-0.009 (-0.14)	-	0.265*** (4.09)
Gini Education (Cohen-Soto)	-	-0.002 (-0.42)	-	-0.003 (-0.85)	-	-0.002 (-0.31)
Gini Income	0.0047 (1.61)	0.004 (1.14)	-0.0034 (-1.17)	-0.003 (-1.09)	0.0063* (1.84)	0.005 (1.44)
R^2	0.713	0.751	0.938	0.945	0.656	0.671
N	524	278	166	89	358	189
Joint significance test of the year dummies (F-test)	7.27 (p=0.000)	5.19 (p=0.000)	19.15 (p=0.000)	9.15 (p=0.000)	2.06 (p=0.050)	2.74 (p=0.036)

Note: Heteroscedasticity and autocorrelation robust t-statistics are reported in parentheses. The sign ***,**,* indicate the coefficient being significantly different form zero at a level of significance 10, 5, and 1%.

Table 5

The effect of inequality in education on per capita income, 2SLS fixed effect panel

The instruments are the investment rate and the population in the previous available period.

(year dummies are included but not reported)

	Full sample		*Non-OECD countries*	
$\ln \frac{I_{i,t}}{Y_{i,t}}$	0. 685*** (3.71)	0.650** (2.24)	0.658*** (3.88)	0.608** (2.03)
ln $pop_{i,t}$	-0.622*** (-2.70)	-0.435** (-2.33)	-1.030*** (-2.75)	-0.469 (-1.30)
Average years of education (Barro-Lee)	0.072** (2.25)	-	0.065 (1.29)	-
Gini Education (Barro-Lee)	-0.008 (-1.48)	-	-0.013* (1.96)	-
Average years of education (Cohen-Soto)	-	0.150** (2.50)	-	0.022 (0.40)
Gini Education (Cohen-Soto)	-	0.004 (1.20)	-	0.0006 (0.14)
Gini Income	0.007* (1.84)	0.001 (0.35)	0.008* (1.93)	0.098 (1.35)
Joint significance test of the year dummies (F-test)	50.13 (p=0.000)	4.65 (p=0.013)	3.33 (p=0.004)	6.08 (p=0.001)
Exclusion test of the instruments from the first stage (F-test)	9.86 (p=0.000)	1.34 (p=0.269)	13.56 (p=0.000)	42.80 (p=0.000)
Hansen J-statistics of overidentification	1.405 (p=0.236)	0.012 (p=0.913)	0.002 (p=0.969)	0.002 (p=0.969)
N	472	207	314	222

Note: Since the chosen instruments were weak for the OECD subsample we only carry out the estimation for the non-OECD subsample. Heteroscedasticity and autocorrelation robust t-statistics are reported in parentheses. The sign ***,**,* indicate the coefficient being significantly different form zero at a level of significance 0.01, 0.05, and 0.1. The J statistics has the null-hypothesis that the chosen instruments are exogenous.

Figure 1
Scatterplot between income – and educational inequality (Barro-Lee data)

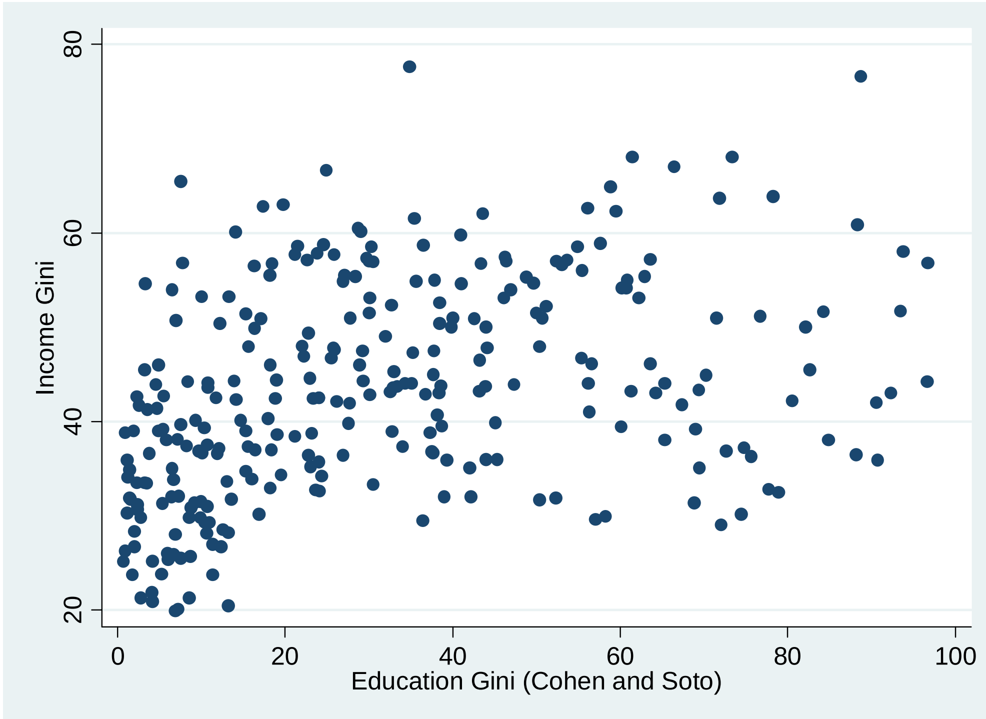

Figure 2
Scatterplot between income – and educational inequality (Cohen and Soto data)