A joint Initiative of Ludwig-Maximilians-Universität and Ifo Institute for Economic Research

CESifo Working Papers

EDUCATION, GROWTH AND INCOME INEQUALITY

Coen Teulings
Thijs van Rens*

CESifo Working Paper No. 653 (4)

January 2002

Category 4: Labour Markets

CESifo
Center for Economic Studies & Ifo Institute for Economic Research
Poschingerstr. 5, 81679 Munich, Germany
Phone: +49 (89) 9224-1410 - Fax: +49 (89) 9224-1409
e-mail: office@CESifo.de
ISSN 1617-9595

* We would like to thank Mikael Lindahl and Daniele Checchi for sharing their datasets. We are also grateful to Robert Shimer, Giorgio Primiceri and Alan Krueger for helpful comments.

CESifo *Working Paper No. 653*
January 2002

EDUCATION, GROWTH AND INCOME INEQUALITY

Abstract

When types of workers are imperfect substitutes, the Mincerian rate of return to human capital is negatively related to the supply of human capital. We work out a simple model for the joint evolution of output and wage dispersion. We estimate this model using cross-country panel data on GDP and Gini coefficients. The results are broadly consistent with our hypothesis of diminishing returns to education. The implied elasticity of substitution fits Katz and Murphy's (1992) estimate. A one year increase in the stock of human capital reduces the rate of return by about 2 per cent. The combination of imperfect substitution and skill biased technological change closes the gap between the Mincer equation and GDP growth regressions almost completely.

JEL Classification: E20, J24, O10, O15.

Coen Teulings
Tinbergen Institute
Erasmus University
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Teulings@few.eur.nl

Thijs van Rens
Princeton University
Department of Economics
001 Fisher Hall
Princeton NJ 08544-1021
U.S.A.

1 Introduction

If workers with various levels of education were perfect substitutes, relative wages would be independent of the distribution of human capital. However, studies into the substitutability of worker types, for example Katz and Murphy (1992), have shown that this is not the case. Then, a simple economic argument establishes that the Mincerian rate of return should be negatively related to the average years of education among the workforce. Raising the average years of education in the economy makes low-skilled workers more scarce, raising their wages, while at the same time increasing the supply of highly educated workers, thereby reducing their wages. This mechanism reduces the return to human capital.

The relation between GDP and education at the aggregate level is a simple reflection of a Mincerian earnings function at the micro level, when externalities of education can be ignored, as is suggested by a number of recent studies (Heckman and Klenow, 1997; Acemoglu and Angrist, 1999). This simple theory of imperfect substitution between workers with different levels of human capital has joint implications for GDP and income dispersion. The effect of an increase in the mean level of education on GDP should decline with the level of education. Hence, we expect a negative second order effect of increases in the education level on growth. Since wages are the main source of income for most families, measures of income inequality should be positively related to the return to education. The average level of education in the economy affects the return to schooling negatively. Hence, it compresses the wage distribution. The main idea of this paper is to simultaneously estimate the effect of the average education level on GDP and income dispersion.

The application of the Mincerian earning function as the driving force in the relation between GDP and education puts this paper in the extensive stream of research into the cross country relation between education and growth. In Barro and Sala-i-Martin (1999), a higher education level makes the labor force more able to deal with technological innovations, yielding a relation between the level of human capital and the growth of output. Barro and Sala-i-Martin found indeed that the level of education has a strong and significant effect on future GDP growth, as did Benhabib and Spiegel (1994) in an earlier study. The effect of the growth in education on the growth of output, conditional on the effect of the level of education, is insignificant in their regressions. These results cast doubt on the relevance of the Mincer equation for the aggregate level, increasing the popularity of human capital based endogenous growth models.

Following Krueger and Lindahl (2000), we argue that these conclusion are due to a number of misspeci…cations. Measurement error attenuates the coe¢ cient for the growth in education. However, just Krueger and Lindahl's argument does not …ll the whole gap between the Mincer equation and the GDP growth regression. The long run rate of return to education remains above any reasonable estimate. Gallup, Sachs and Mellinger (1999) show that geography matters for GDP. Proximity to the sea for transport and a temperate climate to avoid tropical diseases are great advantages to a country. A combination of …xed e¤ects due to geography, imperfect substitution between types of labor, and skill biased technological progress brings us much closer to a full reconcilliation of the GDP data and the Mincer equation. Countries with a favorable geography are richer and can therefore invest more in human capital. Hence, human capital variables pick up part of the favorable …xed geography e¤ect. The initial advantage in human capital increases in the course of time due to skill biased technological progress. This gives the impression that education yields a higher growth of GDP, not a higher level.

Previous studies on the relation between inequality and growth have focused on the e¤ect of the one upon the other, some papers arguing that growth reduces inequality (the so called Kuznets curve), others highlighting the e¤ect of inequality on growth (see Bénabou 1996 for a survey). Our approach di¤ers from this literature, in that we take both inequality and growth as dependent variables, simultaneously determined by the level of human capital. If the average education level has a negative e¤ect on inequality and a positive e¤ect on growth, as implied by our model, then this provides an explanation for the negative correlation between inequality and growth that has spurred this literature.

The theoretical framework we apply is derived from an assignment model with heterogeneous workers and heterogeneous jobs, see Teulings (2001). Highly educated workers have a comparative advantage in complex jobs. The return to education is therefore higher in more complex jobs. When the supply of highly educated workers increases, there are insu¢ cient complex jobs for them. Some high skilled workers have to do less complex jobs, where their human capital has a lower return. This yields a negative relation between the aggregate supply of education and its Mincerian rate of return. We test this relationship by entering a second order term in education in a GDP regression. Furthermore, education should enter negatively in a regression of the variance of log wages, since a reduction of the Mincerian rate of return compresses wage di¤erentials. The simple model we present in the next section formalizes these ideas. We also use our estimates to derive the compression elasticity: the percentage decline in the return to human capital per percent increase in the value of its stock. This concept relates our results to Katz and Murphy's (1992) estimate of the elasticity of substitution between low- and highly skilled workers, providing a check on the interpretation of our estimation results.

Our empirical work uses Barro and Lee's (1999) panel data on GDP and education and Deininger and Squires' (1996) data on Gini coe¢cients for 100 countries over the period 1960-1990. Although the micro labor literature has shown that the log-linear Mincerian wage equation is strikingly robust (see Card 1999 for a survey), the estimated returns for di¤erent countries vary substantially (Psacharopoulos 1994; Bils and Klenow 1998). This paper exploits this variation to estimate the degree of substitutability between worker types. We will also present direct evidence of diminishing returns to education from a cross section of Mincerian rates of return estimated from micro data for various countries.

Empirical research in this area is troubled by the issue of causality: does a higher education level lead to higher GDP or is it the other way around. The same problem applies to the relation between education and income inequality. Indeed, Bils and Klenow (1998) have argued that the posited causation from education to growth should be reversed. However, their arguments apply to the endogenous growth relation, and not to the Mincerian relation invoked here.[1] Our solution to the endogeneity problem relies on the time-lags in the causation from GDP to average level of schooling of the population. First, the political system has to decide on spending of additional tax revenues on education. Then, new teachers have to be trained and schools have to be built. Only then the …rst new cohort can undergo the improved training. It will then take some years or so before the …rst cohort of better educated students enter the labor market. It takes several new cohorts of better educated workers before there is a noticeable e¤ect on the average level of education of the workforce. We argue therefore, that it is reasonable to assume that GDP only a¤ects education level with a lag of at least 10 years. We explore whether our results are driven by a few countries that experience high growth during the sample period (e.g. Asian tigers).

Our empirical results provide strong support for a negative relation between the supply of human capital and its return. Moreover, the estimation results are also largely mutually consistent quantitatively: a one year increase in the stock of human capital reduces its return by about 2 percentage points. This estimate is consistent with Katz and Murphy's (1992) estimate of the elasticity between low and high skilled workers.

We account for skill biased technological progress by entering cross e¤ects of time

[1] Bils and Klenow (1999) argue that if endogenous growth is due to the role of education di¤using the most recent state of technology, then the education of new cohorts should be more valuable, leading to a negative correlation between growth and the return to experience.

dummies and education. This relates our analysis to O'Neill (1995). He asks the question as to why the huge investments in human capital by LDCs have not contributed to a convergence in GDP between LDCs and the industrialized world. His explanation relies on skill biased technological progress: "The recent shift in production techniques toward high-skilled labor has resulted in a substantial increase in the returns to education. This trend, when combined with the large disparities that still exist in education levels between the developed and less developed countries, has led to an increase in inequality despite the significant reduction in the education gap that has occurred over the last 20 years." (p.1299). Our results confirm his analysis.

The interaction terms of education and time dummies allow inference on the pace of skill biased technological change. The GDP and inequality regressions yield quantitatively similar estimates, suggesting skill biased technological change to account for a 3% to 4% increase in the return to education per decade. This is equivalent to the reduction in the return that would be achieved by a 0.8 year increase in the average level of schooling, about as much as the actual increase in the education level over period covered by our sample. Finally, our analysis reduces the difference between the long and the short run rate of return to education from a factor 6, as in Krueger and Lindahl (2000), to less than 2. One can therefore conclude, with some exaggeration, that Tinbergen's race (1975) between education and technology and Mincer's earnings function rule the world.

The paper is structured as follows. In section 2, we present a simple Walrasian model with imperfect substitution between types of labor. Section 3 discusses the data and presents the estimation results. Section 4 concludes.

2 Theoretical framework

2.1 A simple growth model with emphasis on human capital

Consider the long run growth path of an economy with physical and human capital. All markets are perfectly competitive, so that wages equal marginal productivity. We specify both a simple aggregate production function and a Mincerian earnings function.

First, consider the Mincerian earnings function. Let w_{it} be the log wage of worker i at time t and let s_{it} be the years of schooling she attained; w_{it} is assumed to satisfy the Mincerian earnings function:

$$w_{it} = \omega_0(S_t; t) + \omega_1(S_t; t)\, s_{it} + \sigma u_{it} \equiv w_t(s_{it}; u_{it}) \tag{1}$$

where S_t is the average education level of the workforce in the economy, u_{it} is a mean zero unit variance random variable representing other characteristics of workers (like experience and innate ability) and σ is its standard deviation. Both the intercept $\omega_0(\cdot)$ and the Mincerian rate of return to human capital $\omega_1(\cdot)$ vary over time and with the average education level of the workforce. Equation (1) is constrained to be linear in s_{it}, implying that the rate of return to education at particular point in time t is independent of the years of schooling of an individual worker. This assumption plays an important role in the subsequent analysis.

Next, consider the aggregate production function. Let output per worker be governed by a constant returns to scale Cobb Douglas production function:

$$\begin{aligned} y_t &= \alpha h_t + (1-\alpha) k_t \\ h_t &\equiv \beta_1 S_t - \tfrac{1}{2}\beta_2 S_t^2 + \beta_3 S_t t + \beta_4 t \end{aligned} \tag{2}$$

where y_t is log output per worker, k_t log capital per worker and h_t is log average productivity. We assume $S_t < \frac{\beta_1}{\beta_2}$, so that $\frac{dh_t}{dS_t} > 0$. The first term in the expression for h_t measures the effect of schooling. The second term measures the diminishing returns to education: the higher the mean level of education of the workforce, the smaller the return to additional schooling. The third term captures the effect of skill biased technical progress, while the final term reflects neutral technical progress: other things equal, the return to education increases over time when $\beta_3 > 0$.

First, consider the role of capital in this economy. Firms maximize profits per worker, yielding a first order condition for the optimal capital stock:

$$RK_t = (1-\alpha) Y_t \Rightarrow k_t = y_t + \ln(1-\alpha) - \ln r \tag{3}$$

where Y_t and K_t denote the exponentials of the corresponding lower case variables, and R is the rental rate of capital which we assume to be constant over time. Equation (3) reflects the standard result for a Cobb Douglas technology that the share of capital in output is equal to $1-\alpha$. We assume that firms adjust their capital stock sufficiently fast, so that we can ignore deviations from its equilibrium value. Then, combining the FOC for capital and the production function:

$$\begin{aligned} y_t &= \beta_1 S_t - \tfrac{1}{2}\beta_2 S_t^2 + \beta_4 t + \beta_3 S_t t + \frac{1-\alpha}{\alpha}\left(\ln(1-\alpha) - \ln r\right) \\ k_t &= \beta_1 S_t - \tfrac{1}{2}\beta_2 S_t^2 + \beta_4 t + \beta_3 S_t t + \frac{1}{\alpha}\left(\ln(1-\alpha) - \ln r\right) \end{aligned} \tag{4}$$

The equations for log output and capital are identical, up to a constant term. Estimation of the separate contributions of human and physical capital on the basis of equation (2) is therefore problematic, due to endogeneity of k_t. In the absence of measurement error in both S_t and k_t, equation (2) is unidentified since $\beta_1 S_t - \frac{1}{2}\beta_2 S_t^2 + \beta_4 t + \beta_3 S_t t$ is collinear with k_t. In the presence of measurement error, the relative magnitudes of their coefficients merely reflects the precision of their measurement. Krueger and Lindahl (2000) argue that capital data are correlated to output by construction, since investment data figure in both series. Hence, measurement error in both series are likely to be correlated. This explains why they find $1 - \alpha$ to be much higher than one would expect on the basis of capital's share in output (about 0.35). We shall therefore omit capital from all our regressions and report estimation results for equation (4) only.

Next, consider the role of types of labor in this economy. We have a similar condition for labor as for capital, aggregating over all individuals:

$$\alpha Y_t = \int\int W_t(s; u) f_t(s; u)\, ds du \tag{5}$$

where $f_t(s; u)$ is the joint cross-sectional density of s and u, and $W_t(s; u) \equiv \exp(w_t(s; u))$ is the wage rate of an individual with s years of schooling and characteristics u. Labor gets a share α of total output.

Marginal productivity theory implies that the increase in output from adding one worker with characteristics $(s; u)$ to the workforce of this economy raises output by $W_t(s; u)$. This implication extends to the (marginal) effect of new human capital: a marginal increase in the years of education of worker i will raise output by:

$$\frac{\partial Y_t}{\partial s_{it}} = \frac{\partial W_t(s_{it}; u_{it})}{\partial s_{it}} = W_t(s_{it}; u_{it})\, \omega_1(S_t; t) \tag{6}$$

where Y_t denotes aggregate output, and $W_t(s_{it}; u_{it})$ is given by the Mincer equation (1). Equation (6) states that the increase in output due to an increase in the schooling level of worker i by an amount h, equals the gain in output due to the addition of a worker with characteristics $(s + h; u)$ minus the loss in output due to the removal of a worker with characteristics $(s; u)$.

Consider an increase of the years of education of all workers by an equal amount $ds_{it} = ds$ for all i. By construction, the average years of education S_t changes by that same amount: $dS_t = ds$, thus shifting the marginal distribution of education to the right. Then, each worker's wage increases by an amount $\frac{\partial W_t(s_{it}; u_{it})}{\partial s_{it}} ds = W_t(s_{it}; u_{it})\, \omega_1(S_t; t)\, ds$. The change in total output is obtained from the production function (2):

$$\frac{\partial Y_t}{\partial S_t} ds = (\beta_1 - \beta_2 S_t + \beta_3 t)\, \alpha Y_t ds$$

By equation (6), the effect of this increase in S_t on aggregate output is equal to the sum over all workers of the increases in individual wages. Thus:

$$
\begin{aligned}
\frac{\partial Y_t}{\partial S_t} &= \int\int \frac{\partial W_t(s;u)}{\partial s} f_t(s;u)\,ds\,du \\
(\beta_1 - \beta_2 S_t + \beta_3 t)\,\alpha Y_t &= \omega_1(S_t;t) \int\int W_t(s;u)\, f_t(s;u)\,ds\,du \\
&= \omega_1(S_t;t)\,\alpha Y_t
\end{aligned}
$$

where the third equality follows from equation (5). The second line relies on the linearity of the Mincerian earnings function (1) in s_{it}, for otherwise $\omega_1(S_t;t)$ could not be brought outside the integral.

Dividing through by the labor share, we obtain an expression for the return to education:

$$\omega_1(S_t;t) = \beta_1 - \beta_2 S_t + \beta_3 t \tag{7}$$

The increase in log aggregate output is equal to Mincerian rate of return to education. Or, in other words, the private return to education, as measured in a cross section analysis on individual wages, is equal to the social rate of return, as measured in a time series analysis of log aggregate output. This conclusion does not come as a surprise, since in this Walrasian world, there are no external effects of schooling decisions.

The return to education $\omega_1(\cdot)$ determines relative wages of workers with various levels of education. If β_2 were 0, then the relative wages would be independent of S_t and workers with different levels of education would be perfect substitutes. With $\beta_2 > 0$, an increase in the mean level of education in the economy reduces the rate of return to education. Teulings (2001) provides a production technology that yields this implication.[2]

2.2 Inequality and the compression elasticity

An increase in the level of education reduces the return on further investments in human capital by $\beta_2 dS_t$. This fall in the return on human capital compresses wage differentials. We use this relation to analyze the interaction between the evolution of output and income dispersion D_t. For simplicity, capital income is assumed to be distributed

[2] Because we do not need an expression for $\omega_0(S_t;t)$ for our empirical application, it is not presented here. However, the declining marginal return to education implies that a below average educated worker gains from an increase in the mean level of human capital, whereas an above average worker looses out (in both cases, keeping constant the human capital of that worker).

proportional to labor income, so that the log wage distribution and the log income distribution differ only by their first moment. We assume that s_{it} and u_{it} are jointly normally distributed, with correlation ρ. Furthermore, we assume that the variance of s_{it} is constant over time $V(s_{it}) = V$.[3] We can then derive an expression for the variance of log income $D_t = V(w_{it})$ from the Mincer equation (1).

$$
\begin{aligned}
D_t &= \omega_1(S_t; t)^2 V + 2\omega_1(S_t; t) V^{1/2}\sigma\rho + \sigma^2 \\
&= \theta_{0t} - \theta_{1t} S_t + \theta_2 S_t^2 \qquad (8)
\end{aligned}
$$

where

$$
\begin{aligned}
\theta_{0t} &= (\beta_1 + \beta_3 t)^2 V + 2(\beta_1 + \beta_3 t) V^{1/2}\sigma\rho + \sigma^2 \\
\theta_{1t} &= 2\beta_2(\beta_1 + \beta_3 t) V + 2\beta_2 V^{1/2}\sigma\rho \\
\theta_2 &= \beta_2^2 V
\end{aligned}
$$

The variation in income due to the education component is equal to the variance of years of education, multiplied by the return to education. The second equality follows from substitution of equation (7). Equation (8) establishes cross equation restrictions on the equations for output and income dispersion. When information on $\rho; \sigma$ and V is available, these restrictions can be tested. Notice that if $\beta_2 = 0$, D_t would not depend on S_t.

The coefficient β_2 relates in a simple way to earlier empirical findings, like Katz and Murphy's (1992) estimate of the substitution elasticity between low- and high-skilled workers of 1.4. For this purpose, we define the compression elasticity γ as the percentage reduction in the return to human capital per percent increase in the value of its stock. This elasticity can be calculated from equations (1) and (2) as the relative reduction of the return to human capital per year increase in S_t, divided by the effect of this increase in the level of schooling on the log value of the stock of human capital:

$$
\gamma(S_t; t) \equiv -\frac{\partial \ln \omega_1(S_t; t)}{\partial \ln H_t} = \frac{-\partial \omega_1(S_t; t)/\partial S_t}{\omega_1(S_t; t)\,\partial h_t/\partial S_t} = \frac{\beta_2}{(\beta_1 - \beta_2 S_t + \beta_3 t)^2} \qquad (9)
$$

Equation (9) implies that the compression elasticity is increasing in S_t. This implication is imposed by the quadratic specification for h_t adopted in equation (2) and

[3] This is a crucial assumption for the analysis. If V varies over time, the linear form of Mincerian equation (1) would collapse, see Teulings (2001) for details. An increase in V raises labor supply in both tails of the schooling distribution. This reduces relative wages in the tails. In the empirical sections, we shall adopt a pragmatic approach, by including V_t as an additive control variable in our regressions.

should not be taken at face value. However, Teulings (2001) shows that the compression elasticity is indeed increasing in the level of human capital in the special case of a Leontief production technology over different types of labor.[4]

The compression elasticity relates to the Katz and Murphy elasticity of substitution between high and low-skilled labor $\eta_{\text{low-high}}$ by the following relation, see Teulings (2001):

$$\gamma(S_t; t) = \frac{1}{\eta_{\text{low-high}} D_t} \tag{10}$$

Using Katz and Murphy's (1992) estimate of $\eta_{\text{low-high}} = 1.4$ and using a typical value for wage dispersion in the United States of $D_t \simeq 0.36$, the compression elasticity is of the order of magnitude of 2 for the United States. We will use equations (9) and (10) to compare Katz and Murphy's estimate to our estimation results.

2.3 Why linearity of the Mincer equation is important

The interpretation of the second order effect of years of education on GDP as being caused by imperfect substitutability of worker types relies on the linearity of the Mincer equation in s_{it}. In the subsequent argument, we ignore technological progress and assume u_{it} and s_{it} to be uncorrelated for convenience. Suppose that workers with various levels of schooling are perfect substitutes (so ω_0 and ω_1 do not depend on S_t), but that the Mincerian earnings function (1) is concave in the years of education:

$$w_{it} = w_t(s_{it}; u_{it}) = \omega_0 + \omega_1 s_{it} - \tfrac{1}{2}\omega_2 s_{it}^2 + \sigma u_{it} \tag{11}$$

Then, repeating the derivation of equation (7), we get:

$$\begin{aligned} \frac{\partial Y_t}{\partial S_t} &= \int\int \frac{\partial W_t(s;u)}{\partial s} f_t(s;u)\, ds\, du \\ (\beta_1 - \beta_2 S_t)\, \alpha Y_t &= \int\int (\omega_1 - \omega_2 s)\, W_t(s;u)\, f_t(s;u)\, ds\, du \end{aligned}$$

In appendix A we show that the integral has an analytic solution, and the above expression can be written as:

$$(\beta_1 - \beta_2 S_t)\, \alpha Y_t = \left(\frac{\omega_1}{\omega_2 V + 1} - \frac{\omega_2}{\omega_2 V + 1} S_t \right) \alpha Y_t \tag{12}$$

[4] In that case, the compression elasticity satisfies (dropping the time dependence for convenience)

$$\gamma(S) = \gamma(0) \exp[\gamma(0)\, \omega_1(0)\, S]$$

Hence, the model implies:

$$¯_2 = \frac{!_2}{!_2 V + 1}$$

This expression yields an alternative interpretation for $¯_2 > 0$. Instead of imperfect substitution between types of labor, the negative second order e¤ect of education on output is now interpreted as declining marginal returns to human capital for each individual worker. In this case the aggregate return to human capital also declines when the human capital stock increases since every worker moves along its schedule of declining marginal returns. We can derive an equation for income inequality D_t for this interpretation which is observationally equivalent to equation (8). Again, this yields an alternative interpretation of a negative e¤ect of S_t on income inequality. In fact, any combination of concavity of the Mincerian earnings function and imperfection in the substitutability of worker types can explain $¯_2 > 0$. Data on output and the variance of log income alone do allow to disentangle both models. However, as observed by Krueger and Lindahl (2000), the abundant empirical evidence on the Mincerian earnings function does not suggest any systematic non-linearities in the relation between log wages and years of schooling. We shall therefore interpret the second order e¤ect in the log output equation as evidence that di¤erent types of labor are imperfect substitutes.

3 Empirical evidence

3.1 Data sources

Our empirical analysis is largely based on data from two sources: the Barro and Lee (1996, 1993) data on educational attainment and the Deininger and Squire (1996) data on income inequality. These datasets were supplemented with data on real GDP per worker from the Penn World Table (Summers and Heston 1991) mark 5.6a.

The Barro and Lee dataset contains detailed data on educational attainment for 114 countries for the period 1960-1990 in intervals of 5 years. Barro and Lee report the fraction of the population that attained a certain education level, as well as the average duration of this education level. They use these data to construct the average education level of the population in years. We also calculate a rough estimate of the variance of the education distribution.[5]

[5] Barro and Lee calculate average years of education from attainment data (percentage of the population that have attained a certain level of schooling) combined with data on the typical duration of

Deininger and Squire (1996) use results from a large number of studies and assess their comparability. Their dataset contains Gini coefficients of the income distribution for 115 countries from 1947 to 1996. We use only the 'high quality' data for the period 1960-1990. The 'high quality' label is provided by Deininger and Squire on the basis of three criteria: data are (i) based on a national household survey, (ii) which is representative of the population, and (iii) in which all sources of income have been counted. The total number of observations in the high quality sample is 693. The data contain missing values due to limitations to the time period of data availability, and due to missing observations within that time period. For virtually all countries, data are available only every two or five years or at irregular intervals. We construct data for 5 year intervals from 1960 to 1995 by linear inter- and extrapolation.[6] This method yields a dataset containing 370 observations for 98 countries. Only for 58 countries we have three or more observations. We calculated the variance of log income from the Gini coefficients, assuming that log income is distributed normally. The details of this calculation can be found in appendix B.

Table 1 summarizes the main variables in the combined dataset.[7]

3.2 Direct estimates of diminishing returns to education

Before presenting the estimation results for our main dataset, we present some estimates of the effect of the mean years of schooling on the return to human capital as measured directly from individual data. In table 2 we have ranked a large number of countries

---

each level of schooling (1996, p.218). We can express the calculation as:

$$S = f_{pri}S_{pri} + f_{sec}(D_{pri} + S_{sec}) + f_{high}(D_{pri} + D_{sec} + S_{high})$$

where S is average years of schooling in the total population, f_{level} is the fraction of the population that has attained a certain education level (no education, primary education, secondary education or higher education), D_{level} is the typical duration of the different education levels, and S_{level} is the average duration of a certain education level for those people that have not continued to attain a higher education level. Intuitively $S_{level} < D_{level}$ due to early drop-out.

The calculation of average years of schooling in this expression is just an expected value, which suggests the following proxy for the variance in education within each country (cf. Checchi 1999):

$$V(S) = f_{pri}S_{pri}^2 + f_{sec}(D_{pri} + S_{sec})^2 + f_{high}(D_{pri} + D_{sec} + S_{high})^2 - S^2$$

[6] For interpolation we use $\hat{x}_t = \frac{n}{n+p}x_{t-p} + \frac{p}{n+p}x_{t+n}$, where n is the time span till the next observations and $p \leq 2$ is the time span since the previous observation. For extrapolation we use the observation that is closest by. This procedure is efficient if the Gini follows a random walk, as is almost true empirically.

[7] The data are available at http://www.princeton.edu/~tvanrens/paper.

for which such estimates of the return to schooling are available. The data are obtained from Bils and Klenow (1998) and include estimates from Psacharopoulos (1994) and other authors (sources in the table). We have plotted the return to education against the average schooling level in figure 1, panel A. Apart from Jamaica, there is a clear negative relationship between the two. The return to education is plotted against income inequality in Panel B. This relation documents that inequality is indeed strongly related to the return to education.

Table 3 presents the results for some simple regressions on these data. Obviously, these estimates should be interpreted with some care. The data in table 3 provide the best estimates that are available for many countries, but it is not clear to which extend these estimates are comparable across countries. In particular, the underlying studies differ in whether and how they account for ability bias and measurement error. Nevertheless, the estimates are informative. They show that the return to education is about 16% for countries with an education level of zero, and decreases by about 0.7% for every year of education. For the average education level of 5.3 years in our sample, this would correspond to a return to schooling of 12%. In the US, with an average education level of 12 years of schooling in 1990, the return to education would be about 7.5%. This simple cross section analysis provides therefore first evidence of the negative relation between the return to education and the mean years of schooling in the economy. The time dummies suggest that there has been skill biased technological progress from 1985 to 1990, raising the return to human capital by 4%. However, there is little action before 1985. The estimation results even suggest a negative skill bias in that period, but the results are insignificant. Weighting countries by log GDP per worker or log population size does not affect these conclusions.

3.3 Estimation results for GDP

We apply an error correction version of equation (4) for output as a starting point for our empirical analysis. We replace the time trends in skill biased and neutral technological progress by dummies to allow for variations in their pace. Indexing countries by j, the equation we estimate is:

$$\begin{aligned} \Delta y_{jt} &= -\delta y_{jt-1} + \Delta\beta_{0t} + \delta\beta_{0t-1} + \beta_{1t}\Delta S_{jt} - \tfrac{1}{2}\beta_2 \Delta S_{jt}^2 \\ &\quad + \left(\Delta\beta_{1t} + \delta\beta_{1t-1}\right) S_{jt-1} - \tfrac{1}{2}\delta\beta_2 S_{jt-1}^2 + v_{jt} \\ &= -\delta y_{jt-1} + \gamma_{0t} + \gamma_{1t}\Delta S_{jt} + \gamma_2 \Delta S_{jt}^2 + \gamma_{3t} S_{jt-1} + \gamma_4 S_{jt-1}^2 + v_{jt} \end{aligned} \tag{13}$$

where v_t is an error term. The short run return to human capital is $\gamma_{1t} + 2\gamma_2 S_{jt-1}$, while the long run return is $(\gamma_{3t} + 2\gamma_4 S_{jt-1})/\delta$. Krueger and Lindahl (2000) have shown that estimates of the return to human capital from this type of model are strongly affected by attenuation bias because of measurement error when using short time intervals. However, the longer the time interval, the greater the risk of reverse causality. As argued in the introduction, we take it to be unlikely that shocks to GDP have a major impact on the mean level education within 10 years. Hence, we apply a 10 year observation interval. This implies that we have at most 3 observations on the change in education for each country, 1960 till 1990.

Estimation results for equation (13) are reported in Table 4. Column (1) replicates Krueger and Lindahl (2000, Table 3). The results differ slightly because we use GDP per worker rather than GDP per capita. The short run effect of 8% additional GDP per year education is roughly consistent with the micro literature on the Mincerian earnings function. The long run effect takes a long time to materialize, as can be seen from the low coefficient of the level of GDP lagged. However, the long run effect is 6 times larger than the short run effect (0.00297/0.00616 = 48 log points increase in GDP per additional year of education), exceeding by far any estimate of the Mincerian rate of return.

Column (2) of Table 4 adds the crucial second order effect in education. Its coefficient has the expected negative sign and is significant at the 5% level. The second order term is about 1/20 of the first order term, both for the short and the long run terms. This ratio of one over 20 will be a recurrent theme in all our estimates. This regression implies a return to education in the range of -1.7% to 10% for an average education level of 12 to 4 years. When we allow for skill biased technological change as in column (3) the coefficient γ_1 seems to increase substantially, but this is because the reference category for the time dummy interactions is 1990. Although the short run cross-effects of time dummies and education are not very precisely measured, they provide some information regarding the nature of technological progress. Their negative sign is evidence of skill biased technological progress: keeping constant the average education level, the return to education has gone up over the period. The pace of skill biased technological progress increased dramatically during the eighties, raising the return by as much as 6.7%. To get some idea about the size of the impact of skill biased technological progress, we can use the second order term for education to calculate the increase in average years of education that is required to offset this increase: $\frac{0.067}{2 \times 0.0085} = 4$ years. The effect of skill biased technological progress on the return to schooling in the eighties (one decade) was about twice as high as the effect of the increase in the average education level over the whole sample period (three decades). The long run coefficients yield a similar picture. Note however that the long run return to education is still 6 times higher than the short run return.

We report some specification tests in columns (4) through (6). Column (4) adds the variance in the years of education. This does not affect the results. In columns (5) and (6) observations are weighted by log GDP per worker and log population size respectively. Again, this does not make much difference. The WLS estimates show that our results are not driven by a few very poor or very small countries, and are consistent with tests that show that there is no heteroskedasticity in the residuals.

We also estimated the model with random and fixed effects. These regressions strongly suggest the presence of country specific fixed effects. This does not come as a surprise. Gallup, Sachs and Mellinger (1999) have shown the importance of geography for growth and GDP. Access to open sea or navigable rivers is an important advantage. Countries with a temperate climate do much better than countries in the tropical zone. The authors present evidence that the effect of climate is likely to be due to tropical diseases, in particular malaria. Where these factors are largely fixed (there is some reduction in the number of countries where malaria is endemic), we should allow for fixed effects in our estimation.

OLS estimation of equation (13) is inconsistent in the presence of fixed effects as y_{t-1} is correlated with the fixed effect. Also, OLS in first differences would be inconsistent because of the lagged dependent variable. We therefore use the methodology set out in Blundell and Bond (1998). We respecify equation (13) as:

$$\begin{aligned} y_{jt} = {} & \gamma_{0t} + (1-\delta) y_{jt-1} + \gamma_{1t} S_{jt} - (\gamma_{1t} - \gamma_{3t}) S_{jt-1} \\ & + \gamma_2 S_{jt}^2 - (\gamma_2 - \gamma_4) S_{jt-1}^2 + f_j + \varepsilon_{jt} \end{aligned} \quad (14)$$

where we assume:

$$\begin{aligned} E[\varepsilon_{jt} f_j] &= 0 \\ E[\varepsilon_{jt} \varepsilon_{jt-s}] &= 0 \quad \text{for } s \neq 0 \\ E[\varepsilon_{jt} S_{jt-s}] &= 0 \quad \text{for } s \geq 0 \end{aligned}$$

The third assumption reflects our identifying assumptions that shocks ε_{jt} in log GDP take at least ten years to have a significant effect on S_t. The efficient GMM estimator of equation (14) uses the following moment conditions (Arellano and Bond 1991)

$$\begin{aligned} E[\Delta\varepsilon_{jt} y_{jt-s}] &= 0 \quad \text{for } s \geq 2 \\ E[\Delta\varepsilon_{jt} S_{jt-s}] &= 0 \quad \text{for } s \geq 1 \end{aligned}$$

which follow directly from our assumptions on the error term above. We estimated this model using the DPD98 for Gauss package (Arellano and Bond 1998). Table 5 gives the estimation results. Column (1) is identical to column (3) in table 4, but now presented in levels as in equation (14). Column (2) repeats column (1) in first differences. Both estimators are inconsistent. Column (3) presents the GMM estimation results using the above moment conditions. The results are insignificant, as was to be expected given the small number of observations and short time dimension of our data.

Blundell and Bond (1998) suggest jointly estimating equation (14) in first differences and in levels. This results in an efficiency gain, particularly in panels with a short time dimension, because the estimator uses additional moment conditions. We have to make one additional assumption:

$$E[f_j \cdot S_{jt}] = 0$$

Then, two additional moment conditions are available:

$$\begin{aligned} E[(f_j + \varepsilon_{jt}) \cdot y_{jt-s}] &= 0 \quad \text{for } s \geq 1 \\ E[(f_j + \varepsilon_{jt}) \cdot S_{jt-s}] &= 0 \quad \text{for } s \geq 0 \end{aligned}$$

Columns (4) and (5) present the estimation results using all moment conditions, where we use a two step procedure to account for the covariance structure in the error terms in column (5). We take column (5) as the benchmark for our discussion. The Sargan test-statistic for the validity of instruments is 12.65 with 12 degrees of freedom (p-value is 39.5%), accepting the over-identifying restrictions. The ratio of the first and second order term of S_t is still about 20, both for the contemporaneous and the lagged effect. However, the long run effect is now much closer to the short run effect than in Table 4: the short run effect of the first order term is 0.46 and the long run effect $\frac{0.46-0.13}{1-0.63} = 0.89$, less than 2 times the short run return. This is as close as our analysis will bring us to the Mincerian wage equation. Finally, there is clear evidence of skill biased technological progress, raising the return to education by about 4.5% during the eighties and about 3.5% during the nineties (keeping constant the mean level of education).

The estimate for the diminishing returns to education $\beta_2 = 2\gamma_2 = -0.048$ is about 7 times higher than the direct estimate in table 3. The combination of allowing for fixed effects and skill biased technological change is crucial for this result. There is a clear intuition for this. Geography gives some nations an initial advantage over others. These countries can afford a higher level of investment in human capital, raising their level of S_t. Hence, S_t is correlated with the fixed effect and is likely to pick up some of the effects of geography in a regression without …xed e¤ects. Next, countries with a high level of S_t see their initial advantage increased by skill biased technological progress. When we do not allow for this type of technological progress by including time dummies crossed with S_t, this e¤ect shows up as endogenous growth due to a high initial level of education. A combination of Tinbergen's race between education and technology, Mincer's return to human capital and Gallup, Sachs and Mellinger's geography gives therefore a …ne description of the evolution of GDP between 1960 and 1990.

The returns to education by decade, evaluated at the average education level across countries in our sample are as follows.

	1970	1980	1990
Average education level S_t	3.83	4.56	5.32
Return to Education	19.9%	20.8%	20.6%

Notice that the numbers are not strictly comparable over time because some countries do not have data on education for the whole sample period. The number for 1980 is about twice times Krueger and Lindahl's estimate of 8.5%. However, the return is much lower in the OECD countries. It is even negative for the country with the highest education level, the United States (S_t = 12 in 1990).[8] A 0.8 year increase in the mean value of S_t during eighties su¢ces to o¤set the e¤ect of skill biased technological progress, which seems to be a more realistic number than the 4 years calculated on the basis of table 3. The race between education and technology has no clear winner: the upward e¤ect of technology is o¤set by the increase in the average education level across the world.

From equation (9) we can calculate the compression elasticity evaluated at the average education level in 1990 using the estimates of column (5): ° (5:3; 1990) = 1:14. This is lower than the value of 2 implied by Katz and Murphy's (1992) estimate of the elasticity of substitution between highly and low-skilled workers. However, their estimate applies to the United States. We cannot calculate the complexity dispersion parameter for the United States due to its estimated negative rate of return to human capital, but theory suggests that the complexity dispersion parameter is increasing in S_t, see the discussion in Section 2.2. Hence, our estimation results are reasonably consistent with Katz and Murphy's elasticity of substitution.

As pointed out by Krueger and Lindahl (2000), a shorter observation period exacerbates the consequences of measurement error in ¢S_t. In table 6 we report the estimation

[8] One expects this result to be due to the restricted functional form of the model, using only a quadratic in education. We tried including a third order term, but the data contain insu¢cient variation to allow reliable estimation.

results for Krueger and Lindahl's specification and for our baseline regression (table 4, column 3) using 5, 10 and 20 year changes. Reading the table horizontally, we see that the coefficient estimates for ΔS_t and $\Delta \left(S_t^2 \right)$ increase as we use longer time intervals. From column (1) to column (3) the number of observations drops from 607 to 292. Nevertheless the significance of the parameter estimates increases substantially. The long run coefficients do not change much. Moving from a 10 to a 20 year observation period raises the coefficients even further, though not by far as much as in Krueger and Lindahl's specification. This result is problematic for the conclusions of Krueger and Lindahl. Measurement error provides a justification for using long time intervals, but there is no clear rule as to how long the interval should be. Whereas the long run return is 6 times higher than the short run return when measured by using 10 year intervals, one can increase the estimate of the short run return to almost any level by using longer and longer time difference intervals. Therefore, the smaller difference between long and short run return and the lower sensitivity of the estimation results to the differencing interval applied, makes one feel more comfortable about the interpretation of the results. Columns (7) and (8) repeat the estimations for 20 year time interval with the Kyriacou (1991) data for education. The results are largely similar to the Barro and Lee education data.

Table 7 presents a robustness check. Our results might be driven by a few countries with exceptionally high growth rates and exceptionally high investment in human capital, both persisting over the whole 30 year period covered. This would open a channel for reverse causality by the following story: some countries grow fast over prolonged period, and use their additional revenues to invest in education. In that case, the increase in the average level of education in this observation period is just a predictor of the raise in education during the previous observation period. Hence, we exclude first the 10 highest and lowest observations on $\Delta y_t; \Delta S_t; y_t$ and S_t in a number of regressions. Obviously, this compression of the variation in the data reduces the significance of the coefficients. However, the crucial coefficient γ_2 never changes sign and is quite stable.

3.4 Estimation results for inequality

As starting point, we estimate an extended version of equation (8):

$$D_{jt} = \mu_{0t} + \mu_{1t} S_{jt} + \mu_2 S_{jt}^2 + \mu_3 V_{jt} + \varepsilon_{jt} \qquad (15)$$

where we added V_{jt} as a control variable as discussed in section 2.2. Again, we use a ten year observation period. The data on income inequality are less comparable across countries than the data on GDP growth and education level. In particular, the Gini coefficients in the Deininger and Squire dataset are based on different definitions: some use income and others expenditure data, some are based on the household as a reference unit and others on the individual, some are based on gross and others on net income. As suggested by Deininger and Squire (1996) we include dummy variables in the regressions to control for changes in the definition of the income variable.

The OLS estimation results for equation (15) are reported in table 8. Columns (1) to (3) present results for the model in levels. Column (1) presents the full model. The main variables S_t and S_t^2 have the expected sign, though the latter is not significant. Note however, that just the significance of θ_{1t} is sufficient evidence for $\beta_2 > 0$, since neither S_t nor S_t^2 would have any effect on income dispersion if $\beta_2 = 0$. If the correlation ρ between u_{it} and s_{it} were zero, the model would imply that the first and second order effects in this regression differ by the same ratio as the first and second order effects in the GDP equation: $\frac{\theta_2}{\theta_{1t}} = \frac{-\beta_2}{2\beta_{1t}} = \frac{\gamma_2}{\gamma_{1t}}$, see equation (8). In our estimates of the GDP equation we found a rather robust ratio of one over 20 between the second and first order effects. In column (1), this ratio is much lower. This would be consistent with a positive correlation between years of schooling and other worker characteristics, $\rho > 0$, but due to the lack of precision in the measurement of θ_2, we cannot draw strong conclusions. This is documented by the results in column (2): dropping the time variation in θ_{1t} raises θ_2 by a factor 2. We take column (2) as a benchmark.

Testing cross equation restrictions between (8) and (13) requires information on $V, \rho,$ and σ. An estimate for V can be found in table 1: $V \simeq 12.6$. Since we do not have a reliable estimate for ρ, the subsequent calculations are based on $\rho = 0$.[9] The estimation results in column (5) of table 5 for 1990 imply:

$$\begin{aligned} \theta_{1t} &= 2\beta_2(\beta_1 + \beta_3 t)V = 4\gamma_2\gamma_{1t}V = 0.57 \\ \theta_2 &= \beta_2^2 V = 4\gamma_2^2 V = 0.03 \end{aligned}$$

The estimated values for θ_{1t} in column (2) of table 8 are a factor 7 smaller than what one would expect on the basis of estimate of the GDP growth equation. The estimate for θ_2 is a factor 18 too small.

[9] This provides a lower bound on the effect of education on wage dispersion

$$\theta_{1t} = 2\beta_2(\beta_1 + \beta_3 t)V + 2\beta_2 V^{1/2}\sigma\rho = 2\beta_2\left(\beta_1 + \beta_3 t + V^{-1/2}\sigma\rho\right)V$$

An upper bound can be found by setting $\rho = 1$ and σ^2 equal to the total variance of log wages: $\sigma = D_t^{1/2} \approx 0.75$ from table 1. In that case $V^{-1/2}\sigma\rho = 0.21$, about half the size of $\beta_1 + \beta_3 t$ which is between 0.38 and 0.46, see Table 5. Hence, setting $\rho = 0$ will not greatly affect the conclusions in the text.

Two remarks are in place here. First, the estimates for θ_1 and θ_2 (in absolute value) are positively correlated: a low estimate for θ_1 generates a low estimate for θ_2 as well. Constraining the ratio between the first and second order effect to 20, the estimate goes up to $\theta_1 = -0.15$ ($t-$value: 9.39), reducing the difference with its expected value on the basis of the GDP model to a factor 4.

Second, in the derivation of equation (8) we assumed that capital income is distributed proportionally to labor income. This assumption is clearly incorrect. Since capital income accounts for a large share on income inequality and since inequality is unrelated to the return to human capital, the empirical effect of S_t on inequality can be expected to be smaller than predicted by equation (8).

The proxy for the variance of the schooling distribution that we include as a control variable in the regressions is insignificant. This suggests that the direct effect of schooling on the income distribution (a more homogeneous human capital distribution leads to less income dispersion) is less important than the indirect, general equilibrium effect (a higher average education level reduces the return to human capital and therefore compresses the income distribution). However, since we only have a crude proxy for the variance of education, we may expect its coefficient to be attenuated towards zero. In any case its inclusion does not affect the other coefficient estimates.

Column (3) enters fixed effects as a robustness check. Though the sign of the coefficients remains consistent with the model, they are no longer significant. An alternative way to eliminate country specific effects is by first differencing equation (15). Estimation results for this model are presented in columns (4) trough (7). Column (4) presents the results when both S_t and S_t^2 are included. Both θ_{1t} and θ_2 are insignificant, but have the expected sign. Column (5) presents the most robust test of the model: testing $\beta_2 > 0$ by entering only S_t while allowing for fixed country effects by first differencing. The coefficient for S_t is significant.

The positive and significant intercept documents a rising trend in income inequality, keeping education constant. This trend can be explained by the effect of skill biased technological progress. Using the results in column (5) we can evaluate the size of this effect. From equation (8) we have $\frac{\partial D_t/\partial S_t}{\partial D_t/\partial t} = \beta_2/\beta_3$ (again setting $\rho = 0$). Hence, we can estimate β_2/β_3 as the ratio of the coefficient for ΔS_t and the constant term, yielding $\beta_2/\beta_3 = 5.6$. From the GDP regression in table 5, column (5) we can retrieve β_2/β_3 as $2 \times$ the coefficient on $\Delta\left(S_t^2\right)$ divided by estimate for skill biased technological progress, that is, 3.5 % per decade. Hence $\beta_2/\beta_3 = 1.4$. Based on the estimates for GDP one would have expected a four times higher intercept in the inequality regression. This calculation indicates that there are other factors compressing inequality, which offset the effect of skill biased technological progress.

Columns (6) and (7) present results when we weigh observations by log GDP per worker and log population size. Like in the GDP growth equation, this does not make a lot of difference. We present a final robustness check in column (8). As pointed out by Atkinson and Brandolini (1999), additive dummy variables may be insufficient to control for changes in definitions of the Gini coefficient. We therefore dummied all 21 observations with a definitional change separately. This correction is clearly asking too much from the data (the number of observations is only 77), and all coefficient estimates become insignificant, though the coefficient for ΔS_t still has the expected sign.

3.5 Inequality and growth

The positive effect of education on GDP and its negative effect on inequality imply a negative correlation between inequality and GDP. We estimated the global average return to education at around 21%, and the effect of education on the variance of the log income distribution at around -8% (evaluated at the average education level $S_t = 4.56$ in 1980). These estimates imply a correlation between GDP and the variance of log wages of

$$Corr(y_{jt}; D_{jt}) = \frac{-0.08 \cdot 0.21 \cdot V(S_{jt})}{V(y_{jt})^{1/2} V(D_{jt})^{1/2}} = -0.42$$

where we used the variance of the average education level across countries and time, and the standard deviations of y_{jt} and D_{jt} from table 1. The observed correlation between y_{jt} and D_{jt} in our sample is -0.20, and the correlation between Δy_{jt} and ΔD_{jt} is -0.29.

Most of the existing literature has focused on the relation between inequality and GDP growth (see Bénabou 1996 for a survey). However, since GDP growth is correlated with the level of GDP (correlation coefficient 0.24), the negative correlation between Δy_{jt} and D_{jt} (correlation is -0.13) that has spurred this literature, may very well be due to the negative correlation between y_{jt} and D_{jt} caused by education and possible other third factors. Instead, the literature has focused on a causal relation between inequality and growth, an approach that has recently been questioned by Quah (2001). Quah argues that because most of the variation in inequality is across countries and most of the variation in growth is across time, it is unlikely that inequality has an empirically relevant effect on growth. Our results offer support for this argument. Modelling GDP and inequality as being jointly determined by education implies an even larger negative correlation than is observed in the data. This approach seems more promising than looking for a causal relation between inequality and growth or vice versa.

4 Concluding remarks

We have shown that the evolution of GDP, the Gini coefficient and the rate of return to education can be captured by a simple Walrasian model of imperfect substitution between workers with various levels of education in the presence of skill-biased technological progress. Human capital enters as a factor of production in this simple constant returns to scale Cobb-Douglas economy. We derived easy to interpret relations between educational attainment, GDP and income inequality that can be estimated from cross-country panel data.

Our empirical results provide strong support for the negative relation between the supply of human capital and its return. The implied return to schooling in different countries is well in line with evidence from micro data. Our estimates provide a simple explanation for the negative correlation between inequality and growth based on the comovement of these variables with the average education level. Our results suggest that this mechanism is quantitatively more important than a causal relationship between inequality and growth.

A Non-linear Mincer equation

To get expression (12) in the text, we first used the assumption that s_{it} and u_{it} are uncorrelated to integrate out over u

$$
\begin{aligned}
& \int\int (\omega_1 - \omega_2 s) W_t(s;u) f_t(s;u)\, ds du \\
= & \int\int (\omega_1 - \omega_2 s) \exp\left(\omega_0 + \omega_1 s - \tfrac{1}{2}\omega_2 s^2 + \sigma u\right) f_t(s) f(u)\, ds du \\
= & \int (\omega_1 - \omega_2 s) \exp\left(\omega_0 + \omega_1 s - \tfrac{1}{2}\omega_2 s^2 + \tfrac{1}{2}\sigma^2\right) f_t(s)\, ds
\end{aligned}
$$

Second, notice that since $f_t(s)$ is the pdf of a normal (with mean S_t and variance V), $\exp\left(\omega_0 + \omega_1 s - \frac{1}{2}\omega_2 s^2 + \frac{1}{2}\sigma^2\right) f_t(s)$ can be rewritten as a constant A_t^* times the pdf of a normal with mean μ_t^* and variance V^*

$$
\begin{aligned}
& \exp\left(\omega_0 + \omega_1 s - \tfrac{1}{2}\omega_2 s^2 + \tfrac{1}{2}\sigma^2\right) f_t(s) \\
= & \frac{1}{\sqrt{2\pi V}} \exp\left(\omega_0 + \omega_1 s - \tfrac{1}{2}\omega_2 s^2 + \tfrac{1}{2}\sigma^2 - \tfrac{1}{2}\frac{(s - S_t)^2}{V}\right) \\
= & \frac{A_t^*}{\sqrt{2\pi V^*}} \exp\left(\tfrac{1}{2}\frac{(s - \mu_t^*)^2}{V^*}\right)
\end{aligned}
$$

where

$$\mu_t^* = \frac{\omega_1 V + S_t}{\omega_2 V + 1}$$

Furthermore, from equation (5) we have

$$\alpha Y_t = \int\int W_t(s;u)\, f_t(s;u)\, ds du = \int \exp\left(\omega_0 + \omega_1 s - \tfrac{1}{2}\omega_2 s^2 + \tfrac{1}{2}\sigma^2\right) f_t(s)\, ds = A_t^*$$

Hence

$$\int\int (\omega_1 - \omega_2 s)\, W_t(s;u)\, f_t(s;u)\, ds du = \omega_1 \alpha Y_t - \omega_2 \mu_t^* \alpha Y_t$$

B Gini coefficient and the variance of log income

Let $W \in \left[\underline{W}; \overline{W}\right]$ denote income with density $f(W)$, distribution function $F(W)$ and mean M. $F(W)$ measures the share of the population with income lower than W. Let $Z(W)$ denote the cumulative share of total income earned by people with income lower than W. By definition:

$$Z(W) = \frac{1}{M}\int_{\underline{W}}^{W} x f(x) dx \tag{16}$$

The graph of the Lorenz curve has $F(W)$ on the horizontal and $Z(W)$ on the vertical axis. The Gini coefficient $G \in [0; 1]$ is given by twice the area between the Lorentz curve and the 45-degree line.

$$G = 1 - 2\int_0^1 Z dF = 2\int_0^1 F dZ - 1$$

By change of variables, using $dZ = \frac{1}{M} W f(W) dW$, this expression can be written as:

$$G = \frac{2}{M}\int_{\underline{W}}^{\overline{W}} W f(W) F(W)\, dW - 1$$

Assume income to be log normally distributed so that $F(W) = \Phi\left(\frac{w-\mu}{\sigma}\right)$ and $M = e^{\mu+\frac{1}{2}\sigma^2}$, where $w \equiv \ln W$ and μ and σ^2 are the mean and variance of w. By change of variables $v = \frac{w-\mu}{\sigma} \Rightarrow dW = \sigma e^{\sigma v+\mu} dv$, the Gini coefficient can written as:

$$G = \frac{2}{M}\int_0^{\infty} W \frac{\phi\left(\frac{w-\mu}{\sigma}\right)}{\sigma W} \Phi\left(\frac{w-\mu}{\sigma}\right) dW - 1 = 2e^{-\frac{1}{2}\sigma^2}\int_{-\infty}^{\infty} e^{\sigma v}\phi(v)\Phi(v)\,dv - 1$$

which maps the Gini coefficient to the variance of the log income distribution σ^2. Numerically evaluating this expression for different values of σ shows that the relationship is virtually linear in the relevant range. Variances of log income of 0, 0.1, 0.2, 0.3 and 0.4 correspond to Gini coefficients of 52.05, 56.33, 60.39, 64.20 and 67.78 respectively.

References

[1] Acemoglu, Daron & Joshua Angrist (1999). How large are Social Returns to Education? Evidence from Compulsory Schooling Laws. NBER Working Paper No.7444.

[2] Arellano, Manuel and Steve Bond (1991). Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. Review of Economic Studies, vol.58 no.2, pp. 277-297.

[3] Arellano, Manuel and Stephen Bond (1998). Dynamic Panel Data Estimation Using DPD98 for Gauss. mimeo, available at: http://www.cemfi.es/~arellano/

[4] Atkinson A.B. and A. Brandolini (1999). Promise and Pitfalls in the Use of "Secondary" Data-Sets: Income Inequality in OECD Countries. mimeo, available at: http://www.nuff.ox.ac.uk/economics/people/atkinson.htm or: http://www.bancaditalia.it/pubblicazioni/temidi

[5] Barro, Robert J. and Xavier Sala-i-Martin (1999). Economic Growth. Cambridge MA: MIT Press (first MIT Press edition, originally published by McGraw-Hill, 1995).

[6] Barro, Robert J. and Jong Wha Lee (1993). International Comparisons of Educational Attainment. Journal of Monetary Economics, vol.32 no.3, pp.363-394 (dataset available at: http://www.worldbank.org/research/growth/ddbarlee.htm).

[7] Barro, Robert J. and Jong Wha Lee (1996). International Measures of Schooling Years and Schooling Quality. American Economic Review, vol.86 issue 2 Papers and Proceedings May 1996, pp.218-223 (dataset available at: http://www.nber.org/data or http://www.worldbank.org/research/growth/ddbarle2.htm).

[8] Bénabou, Roland (1996). Inequality and Growth. In: NBER Macroeconomics Annual 1996, Ben S. Bernanke and Julio Rotemberg (eds.). Cambridge MA: MIT Press. Also available as NBER Working Paper No.5658.

[9] Benhabib, Jess and M. Spiegel (1994). The Role of Human Capital in Economic Development: Evidence from Aggregate Cross-Country Data. Journal of Monetary Economics, vol.34 no.2, pp.143-174.

[10] Bils, Mark and Peter J. Klenow (1998) Does Schooling Cause Growth or the Other Way Around? NBER Working Paper No.6393.

[11] Blundell, Richard and Stephen Bond (1998). Initial Conditions and Moment Restrictions in Dynamic Panel Data Models. Journal of Econometrics, vol. 87, pp.115-143.

[12] Card, David (1999). The Causal E¤ect of Education on Earnings. Chapter 30 in the Handbook of Labor Economics, Volume 3, O. Ashenfelter and D. Card (eds.). Amsterdam: Elsevier.

[13] Checchi, Daniele (1999). Does Educational Achievement Help to Explain Income Inequality? mimeo, Università degli Studi di Milano Bicocca. Available at http://www.eco-dip.unimi.it/pag_pers/checchi/checchi1.htm.

[14] Checchi, Daniele (2000). Inequality in Incomes and Access to Education. A Cross-Country Analysis (1960-1995). mimeo, Università degli Studi di Milano Bicocca. Available at http://www.eco-dip.unimi.it/pag_pers/checchi/checchi1.htm.

[15] Checchi, Daniele and Luca Flabbi (1999). Income and Educational Distribution Dataset: Various Countries in Panel Format (description of the dataset, data available at: http://www.eco-dip.unimi.it/pag_pers/checchi/checchi.htm)

[16] Deininger, Klaus and Lyn Squire (1996). A New Data Set Measuring Income Inequality. World Bank Economic Review, vol.10 no.3, pp.565-591 (dataset available at: http://www.worldbank.org/research/growth/dddeisqu.htm).

[17] Gallup, J.L., J.D. Sachs and A.D. Mellinger (1999). Geography and economic development. NBER Working Paper No.6849.

[18] Heckman, James J. and Peter J. Klenow (1997). Human Capital Policy. mimeo, University of Chicago. Available at: http://www.klenow.com/.

[19] Katz, Lawrence F. and Kevin M. Murphy (1992). Changes in Relative Wages, 1963-1987: Supply and Demand Factors. Quarterly Journal of Economics, vol.107 issue 1, pp.35-78.

[20] Krueger, Alan B. and Mikael Lindahl (2000). Education for Growth: Why and For Whom? NBER Working Paper No.7591. Forthcoming in Journal of Economic Literature, vol.39 nr.4 (Dec 2001).

[21] Kyriacou, George (1991). Level and Growth E¤ects of Human Capital: A Cross-Country Study of the Convergence Hypothesis. New York University C.V. Starr Center for Applied Economics Working Paper RR 9126 (www.econ.nyu.edu/working) (dataset available on request).

[22] O'Neill, Donal (1995). Education and Income Growth: Implications for Cross-Country Inequality. Journal of Political Economy, vol.103 no.6, pp.1289-1301.

[23] Psacharopoulos, G. (1994). Returns to Investment in Education: A Global Update. World Development, vol.22 no.9, pp.1325-1343.

[24] Quah, Danny (2001). Some simple arithmetic on how income inequality and economic growth matter. mimeo, LSE: 11 June 2001. Available at: http://econ.lse.ac.uk/sta¤/dquah/.

[25] Summers, Robert and Alan Heston (1991). The Penn World Table (Mark 5): An Expanded Set of International Comparisons, 1950-1988. Quarterly Journal of Economics, vol.106 no.2, pp. 327-368 (the new dataset, mark 5.6a, is available at: http://pwt.econ.upenn.edu or http://www.nber.org/data).

[26] Teulings, Coen N. (2001). Comparative Advantage, Relative Wages, and the Accumulation of Human Capital. mimeo, Erasmus University Rotterdam / Tinbergen Institute.

[27] Tinbergen, Jan (1975). Income Distribution: Analysis and Policies. Amsterdam: North Holland.

Table 1. Description of the main variables in the dataset

Variable	Obs	Mean	Std.Dev.	Min	Max	Description and source
y_t	1060	8.611	1.037	6.122	11.172	Log real GDP per worker, 1985 intl. prices, Chain index (PWT 5.6a).
Δy_t	429	0.021	0.027	-0.066	0.101	10 year changes in real GDP per worker. (annualized)
D_t	370	0.560	0.319	0.100	1.552	Variance of log income. Calculated from Gini coefficient income distribution (Deininger and Squire).
ΔD_t	92	0.000	0.017	-0.052	0.051	10 year changes in variance of income. (annualized)
S_t	775	4.240	2.848	0.040	12.000	Average years of education attained by the population over 25 years of age (Barro and Lee).
ΔS_t	328	0.066	0.066	-0.225	0.387	10 year changes in average years of education. (annualized)
V_t	662	12.657	5.834	1.043	35.823	Variance of the education distribution (rough estimate constructed on the basis of Barro and Lee data).
ΔV_t	273	0.249	0.297	-0.888	1.361	10 year changes in variance of education. (annualized)

Table 2. Return to education in several countries

PWT 5.0 country code	Country		Average years of schooling population over 25: year	educ. level	Return to Education: year	ret. to educ
123	Poland	POL	85	8.7	86	.024
126	Sweden	SWE	80	9.45	81	.026
114	Greece	GRC	85	6.89	85	.027
118	Italy	ITA	85	5.75	87	.028
107	Austria	AUT	85	7.17	87	.039
115	Hungary	HUN	85	7.93	87	.039
50	Canada	CAN	80	10.23	81	.042
83	China	CHN	85	4.04	85	.045
110	Denmark	DNK	90	11.21	90	.047
89	Israel	ISR	80	9.11	79	.057
85	India	IND	80	2.72	81	.062
131	Australia	AUS	80	10.02	82	.064
121	Netherlands	NLD	85	8.29	83	.066
41	Tanzania	TZA	80	.	80	.067
127	Switzerland	CHE	85	8.99	87	.072
68	Bolivia	BOL	90	4.11	89	.073
113	Germany West	DEU	90	8.83	88	.077
53	Dom. Rep.	DOM	90	3.76	89	.078
117	Ireland	IRL	85	7.87	87	.079
78	Venezuela	VEN	90	4.89	89	.084
75	Peru	PER	90	5.5	90	.085
21	Kenya	KEN	80	2.46	80	.085
77	Uruguay	URY	90	6.69	89	.09
104	Thailand	THA	70	3.54	71	.091
66	USA	USA	90	12	89	.093
94	Malaysia	MYS	80	4.49	79	.094
124	Portugal	PRT	85	3.45	85	.094
29	Morocco	MAR	70	.	70	.095
54	El Salvador	SLV	90	3.4	90	.096
129	UK	GBR	70	7.66	72	.097
97	Pakistan	PAK	80	1.74	79	.097
61	Nicaragua	NIC	80	2.83	78	.097
109	Cyprus	CYP	85	7.56	84	.098
72	Ecuador	ECU	85	5.36	87	.098
74	Paraguay	PRY	90	4.72	89	.103
51	Costa Rica	CRI	90	5.4	89	.105
92	Korea	KOR	85	8.03	86	.106
67	Argentina	ARG	90	7.77	89	.107
100	Singapore	SGP	75	4.38	74	.113
98	Philippines	PHL	90	6.73	88	.119
70	Chile	CHL	90	6.16	89	.121
4	Botswana	BWA	80	2.29	79	.126
62	Panama	PAN	90	7.55	89	.126
125	Spain	ESP	90	6.25	90	.13
60	Mexico	MEX	85	4.34	84	.141
56	Guatemala	GTM	90	2.56	89	.142
71	Colombia	COL	90	4.25	89	.145
69	Brazil	BRA	90	3.56	89	.154
86	Indonesia	IDN	80	3.09	81	.17
58	Honduras	HND	90	3.68	89	.172
20	Cote d'Ivoire	CIV	85	.	85	.207
59	Jamaica	JAM	90	4.51	89	.28

Education data from Barro and Lee. Return to education data from Bils and Klenow (1998).
Original sources return to education: Rosholm and Smith 1996 (Denmark), Calan and Reilly 1993 (Ireland), Armitage and Sabot 1987 (Kenya and Tanzania), Alba-Ramirez and San Segundo 1995 (Spain), Arai 1994 (Sweden), Chiswick 1977 (Thailand), Krueger and Pischke 1992 (USA and Germany) and Psacharopoulos 1994 (all other countries); see Bils and Klenow for full references.

Table 3. Direct estimates of diminishing returns to schooling (OLS estimates)

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	OLS excl. Jamaica	WLS (GDP/w)	WLS (GDP/w) excl. Jamaica	WLS (population)	WLS (population) excl. Jamaica
S_t	-0.00708	-0.00638	-0.00721	-0.00649	-0.00673	-0.00614
	(3.23)	(3.68)	(3.41)	(3.86)	(3.18)	(3.49)
(year=70)	-0.02297	-0.01538	-0.02100	-0.01382	-0.02247	-0.01620
	(0.81)	(0.69)	(0.75)	(0.62)	(0.85)	(0.74)
(year=80)	-0.03538	-0.02759	-0.03542	-0.02819	-0.03556	-0.02902
	(2.49)	(2.44)	(2.52)	(2.51)	(2.55)	(2.49)
(year=85)	-0.04061	-0.03381	-0.04012	-0.03365	-0.04270	-0.03700
	(3.06)	(3.21)	(3.13)	(3.28)	(3.30)	(3.42)
Constant	0.15663	0.14513	0.15725	0.14591	0.15451	0.14490
	(10.33)	(11.95)	(10.50)	(12.07)	(10.34)	(11.54)
Observations	49	48	49	48	49	48
R-squared	0.36	0.40	0.37	0.41	0.36	0.39

Absolute value of t statistics in parentheses. Dependent variable is the Return to Education as in table 2. WLS regressions are weighted by log GDP per worker or log population size. The dummy for 1975, and the dummy for 1985 in column (6), was dropped because there were no observations.

Table 4. GDP growth equation

	(1) OLS	(2) OLS	(3) OLS (baseline model)	(4) OLS with V[educ]	(5) WLS (GDP/w)	(6) WLS (population)
ΔS_t	0.08546	0.17025	0.24335	0.24508	0.24717	0.24814
	(4.11)	(3.25)	(3.84)	(3.09)	(3.94)	(3.99)
$\Delta (S_t^2)$		-0.00780	-0.00848	-0.00881	-0.00840	-0.00898
		(2.07)	(2.16)	(1.75)	(2.18)	(2.32)
ΔS_t (year=70)			-0.09705	-0.07495	-0.09901	-0.09956
			(1.87)	(1.34)	(1.95)	(1.89)
ΔS_t (year=80)			-0.06732	-0.07423	-0.07728	-0.06933
			(1.35)	(1.42)	(1.60)	(1.39)
ΔV_t				-0.00461		
				(0.73)		
S_{t-1}	0.00297	0.00857	0.01217	0.00902	0.01231	0.01218
	(4.31)	(4.45)	(5.42)	(3.21)	(5.51)	(5.57)
S_{t-1}^2		-0.00045	-0.00058	-0.00034	-0.00058	-0.00059
		(2.67)	(3.29)	(1.60)	(3.36)	(3.46)
S_{t-1}(year=70)			-0.00349	-0.00325	-0.00386	-0.00323
			(2.80)	(2.39)	(3.14)	(2.65)
S_{t-1}(year=80)			-0.00300	-0.00391	-0.00339	-0.00276
			(2.63)	(3.14)	(3.02)	(2.48)
V_{t-1}				0.00037		
				(1.03)		
y_{t-1}	-0.00616	-0.00787	-0.00839	-0.00723	-0.00848	-0.00812
	(2.99)	(3.74)	(4.04)	(2.81)	(4.07)	(4.08)
(year=70)	0.03449	0.03506	0.05590	0.05516	0.05769	0.05427
	(10.21)	(10.34)	(8.17)	(7.25)	(8.21)	(8.05)
(year=80)	0.02120	0.02179	0.04017	0.04659	0.04269	0.03832
	(6.54)	(6.82)	(5.77)	(6.12)	(5.99)	(5.61)
Constant	0.03816	0.04033	0.02715	0.02040	0.02735	0.02601
	(2.34)	(2.51)	(1.67)	(1.04)	(1.66)	(1.66)
Observations	292	292	292	250	292	292
R-squared	0.32	0.34	0.37	0.38	0.37	0.37
F-statistic[1]	11.29	9.56	11.08	4.65	11.27	11.29
p-value	0.0009	0.0022	0.0010	0.0321	0.0009	0.0009

Absolute value of t statistics in parentheses.

[1] H_0: Long-run effect (coefficient S_{t-1} divided by minus coefficient y_{t-1}) equals short-run effect (coefficient ΔS_t). The F-tests reject the null when the p-value is smaller than 0.05.

Table 5. GDP growth equation: Dynamic panel data estimates

	(1) OLS in levels		(2) OLS in first difs (incons.)	(3) Arellano-Bond	(4) Blundell-Bond, 1-step	(5) Blundell-Bond, 2-step
S_t	0.24335	ΔS_t	0.21467	0.71161	0.37104	0.46365
	(3.84)		(2.48)	(1.03)	(4.26)	(6.33)
S_t^2	-0.00848	$\Delta (S_t^2)$	-0.00744	-0.06484	-0.02025	-0.02420
	(2.16)		(1.31)	(1.07)	(4.09)	(5.94)
S_t^{70}	-0.09705	ΔS_t^{70}	-0.06567	0.07700	-0.06592	-0.07970
	(1.87)		(0.82)	(0.04)	(1.12)	(1.59)
S_t^{80}	-0.06732	ΔS_t^{80}	-0.05795	-0.10613	-0.02990	-0.03461
	(1.35)		(0.95)	(0.10)	(0.48)	(0.66)
S_{t-1}	-0.05954	ΔS_{t-1}	-0.00040	-0.31410	-0.05333	-0.12747
	(1.04)		(0.01)	(0.32)	(0.70)	(2.28)
S_{t-1}^2	0.00272	$\Delta (S_{t-1}^2)$	0.00002	-0.01429	0.00225	0.00778
	(0.64)		(0.00)	(0.32)	(0.41)	(1.84)
S_{t-1}^{70}	-0.02478	ΔS_{t-1}^{70}	0.00240	0.31557	-0.01259	-0.02127
	(0.48)		(0.04)	(0.37)	(0.29)	(0.77)
S_{t-1}^{80}	-0.06210	ΔS_{t-1}^{80}	-0.02123	0.31499	-0.00615	-0.01405
	(1.15)		(0.26)	(0.17)	(0.10)	(0.28)
y_{t-1}	0.91608	Δy_{t-1}	0.11605	1.05351	0.71236	0.62961
	(44.07)		(1.37)	(1.54)	(7.62)	(7.53)
(yr=70)	0.55900				-0.20276	-0.20986
	(8.17)				(2.87)	(3.57)
(yr=80)	0.40168	(yr=80)	0.36002	0.22577	-0.59898	-0.61523
	(5.77)		(3.92)	(0.30)	(7.50)	(9.50)
Const.	0.27154	Const.	-0.26536	-0.64783	2.17565	2.81648
	(1.67)		(3.03)	(1.21)	(3.18)	(4.73)
Obs.	292	Obs.	184	184	286	286
R-sq	0.95	R-sq	0.26			
Nr of countries		Nr of countries		102	102	102

Absolute value of t statistics in parentheses, based on robust standard errors.

Table 6. GDP growth equation: the effect of measurement error

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	5 year changes		10 year changes (baseline model)		20 year changes		20 year changes, Kyriacou data	
ΔS_t	0.03991	0.06276	0.08546	0.24335	0.15236	0.29273	0.13828	0.24317
	(2.74)	(1.12)	(4.11)	(3.84)	(3.00)	(2.52)	(4.37)	(2.46)
$\Delta (S_t^2)$		-0.00293		-0.00848		-0.01655		-0.00989
		(1.02)		(2.16)		(1.77)		(1.26)
ΔS_t (year=65)		0.09728						
		(1.35)						
ΔS_t (year=70)		-0.00882		-0.09705				
		(0.18)		(1.87)				
ΔS_t (year=75)		0.01557						
		(0.28)						
ΔS_t (year=80)		-0.01051		-0.06732				
		(0.22)		(1.35)				
ΔS_t (year=85)		0.04885						
		(0.82)						
S_{t-1}	0.00349	0.01441	0.00297	0.01217	0.00368	0.01176	0.00526	0.01074
	(5.48)	(6.21)	(4.31)	(5.42)	(3.88)	(4.21)	(4.47)	(3.15)
S_{t-1}^2		-0.00064		-0.00058		-0.00062		-0.00042
		(3.89)		(3.29)		(2.29)		(1.32)
S_{t-1}(year=65)		-0.00526						
		(3.09)						
S_{t-1}(year=70)		-0.00510		-0.00349				
		(3.03)		(2.80)				
S_{t-1}(year=75)		-0.00447						
		(2.89)						
S_{t-1}(year=80)		-0.00534		-0.00300				
		(3.50)		(2.63)				
S_{t-1}(year=85)		-0.00263						
		(1.77)						
y_{t-1}	-0.00706	-0.00913	-0.00616	-0.00839	-0.01179	-0.01306	-0.01294	-0.01354
	(3.79)	(4.80)	(2.99)	(4.04)	(4.42)	(4.96)	(4.44)	(4.61)
(year=65)	0.03189	0.05489						
	(7.02)	(6.08)						
(year=70)	0.03398	0.05876	0.03449	0.05590				
	(7.71)	(6.62)	(10.21)	(8.17)				
(year=75)	0.02259	0.04379						
	(5.22)	(4.87)						
(year=80)	0.01977	0.04715	0.02120	0.04017				
	(4.62)	(5.32)	(6.54)	(5.77)				
(year=85)	-0.00457	0.00631						
	(1.08)	(0.66)						
Constant	0.04808	0.03376	0.03816	0.02715	0.09750	0.09286	0.09354	0.08605
	(3.25)	(2.17)	(2.34)	(1.67)	(4.87)	(4.81)	(4.48)	(4.06)
Observations	607	607	292	292	97	97	79	79
R-squared	0.22	0.26	0.32	0.37	0.22	0.29	0.28	0.31

Absolute value of t statistics in parentheses.
Estimates in columns 1, 3 and 5 correspond to Krueger and Lindahl (2001) table 3. The results differ slightly because we use GDP per worker rather than GDP per capita as the dependent variable.

Table 7. Subsample robustness of the GDP growth equation

	(1) Without 10 countries with highest growth in education	(2) Without 10 countries with highest growth in GDP	(3) Without 10 countries with highest education level	(4) Without 10 countries with highest GDP	(5) Without 10 countries with lowest education level	(6) Without 10 countries with lowest GDP
ΔS_t	0.23695	0.18019	0.20674	0.22825	0.21525	0.23387
	(3.34)	(2.86)	(2.81)	(3.31)	(3.26)	(3.56)
$\Delta (S_t^2)$	-0.01001	-0.00981	-0.00391	-0.00653	-0.00574	-0.00696
	(2.44)	(2.47)	(0.76)	(1.47)	(1.38)	(1.71)
ΔS_t (year=70)	-0.07701	-0.00388	-0.11270	-0.10193	-0.10266	-0.11475
	(1.30)	(0.07)	(1.99)	(1.81)	(1.94)	(2.16)
ΔS_t (year=80)	-0.05366	0.00191	-0.06663	-0.06577	-0.08831	-0.07981
	(0.95)	(0.04)	(1.20)	(1.20)	(1.72)	(1.56)
S_{t-1}	0.00993	0.00926	0.00900	0.01131	0.01084	0.01118
	(4.28)	(4.11)	(2.92)	(4.42)	(4.16)	(4.71)
S_{t-1}^2	-0.00044	-0.00042	-0.00037	-0.00056	-0.00048	-0.00048
	(2.50)	(2.47)	(1.28)	(2.73)	(2.46)	(2.64)
S_{t-1}(year=70)	-0.00284	-0.00266	-0.00161	-0.00288	-0.00359	-0.00364
	(2.27)	(2.17)	(0.94)	(1.95)	(2.61)	(2.75)
S_{t-1}(year=80)	-0.00204	-0.00183	-0.00163	-0.00231	-0.00358	-0.00353
	(1.78)	(1.62)	(1.11)	(1.74)	(2.84)	(2.90)
y_{t-1}	-0.00690	-0.00590	-0.00766	-0.00815	-0.00798	-0.00874
	(3.11)	(2.75)	(3.40)	(3.63)	(3.80)	(3.76)
(year=70)	0.05228	0.04693	0.05104	0.05416	0.05703	0.05828
	(7.57)	(6.94)	(6.64)	(7.36)	(7.21)	(7.60)
(year=80)	0.03481	0.03057	0.03589	0.03822	0.04533	0.04420
	(4.92)	(4.36)	(4.65)	(5.18)	(5.60)	(5.67)
Constant	0.02054	0.01657	0.02804	0.02746	0.02665	0.03153
	(1.21)	(1.01)	(1.60)	(1.58)	(1.55)	(1.64)
Observations	269	268	265	268	272	266
R-squared	0.36	0.36	0.38	0.37	0.35	0.37
Countries excluded from the sample	Congo Egypt China Hong Kong Jordan Korea Taiwan Austria Cyprus Romania	Botswana Swaziland Hong Kong Japan Korea Singapore Taiwan Malta Bulgaria Romania	Canada USA Denmark Finland Sweden Australia New Zealand Czechoslovakia East Germany Soviet Union	Canada USA Bahrain Kuwait Belgium France Germany Netherlands Switzerland Australia	Benin Centr. Afr. Rep. Gambia Mali Mozambique Niger Sierra Leone Sudan Afghanistan Nepal	Centr. Afr. Rep. Lesotho Malawi Mali Niger Rwanda Togo Uganda Zaire Myanmar

Absolute value of t statistics in parentheses.

Table 8. Income inequality

	(1) OLS in levels	(2) OLS in levels	(3) FE in levels		(4) OLS in first difs	(5) OLS in first difs	(6) WLS (GDP/w)	(7) WLS (popul)	(8) OLS with dums for def. ch.
S_t	-0.07192	-0.08573	-0.05534	ΔS_t	-0.09820	-0.05611	-0.05718	-0.05394	-0.01934
	(2.47)	(3.05)	(1.62)		(1.40)	(1.96)	(2.01)	(1.94)	(0.77)
S_t^2	0.00085	0.00170	0.00365	$\Delta (S_t^2)$	0.00320				
	(0.38)	(0.78)	(1.56)		(0.66)				
S_t^{60}	-0.03155								
	(0.95)								
S_t^{70}	-0.02715								
	(1.51)								
S_t^{80}	0.00623								
	(0.41)								
V_t	0.00065	0.00105	-0.00070	ΔV_t	0.00094	-0.00176	-0.00169	-0.00269	0.00065
	(0.20)	(0.33)	(0.20)		(0.13)	(0.29)	(0.29)	(0.47)	(0.13)
(yr=60)	0.11027	-0.05257	-0.00012						
	(0.60)	(0.76)	(0.00)						
(yr=70)	0.13346	0.00491	-0.01062	(yr=70)	-0.00801	-0.00846	-0.00835	-0.00779	-0.00474
	(1.37)	(0.10)	(0.44)		(1.49)	(1.59)	(1.57)	(1.53)	(1.03)
(yr=80)	-0.06782	-0.02754	-0.03376	(yr=80)	-0.00554	-0.00562	-0.00560	-0.00459	-0.00351
	(0.67)	(0.68)	(1.79)		(1.28)	(1.30)	(1.30)	(1.11)	(0.88)
1{inc}	0.09302	0.09840	0.25144	Δ1{inc}	0.04095	0.04155	0.04169	0.04030	
	(1.70)	(1.81)	(4.08)		(3.89)	(3.98)	(3.89)	(3.93)	
1{hh}	-0.04313	-0.03647	-0.00107	Δ1{hh}	-0.00059	0.00007	-0.00008	-0.00065	
	(1.20)	(1.03)	(0.04)		(0.11)	(0.01)	(0.01)	(0.12)	
1{gr}	0.26680	0.26782	0.00693	Δ1{gr}	-0.00487	-0.00527	-0.00550	-0.00458	
	(6.98)	(7.00)	(0.13)		(0.42)	(0.46)	(0.48)	(0.42)	
				dumms					yes
Const.	0.73888	0.76879	0.55529	Const.	0.01011	0.01056	0.01039	0.01008	0.00571
	(10.48)	(11.35)	(5.48)		(2.72)	(2.90)	(2.90)	(2.85)	(1.71)
Obs.	262	262	262	Obs.	77	77	77	77	77
R-sq	0.47	0.46	0.21	R-sq	0.29	0.29	0.28	0.27	0.63
Nr of countries			71	Nr of countries					
				F-stat[1]					4.34
				p-value					0.0000

Absolute value of t statistics in parentheses.

[1] H_0: Dummies for definitional changes jointly insignificant. The F-tests reject the null when the p-value is smaller than 0.05.

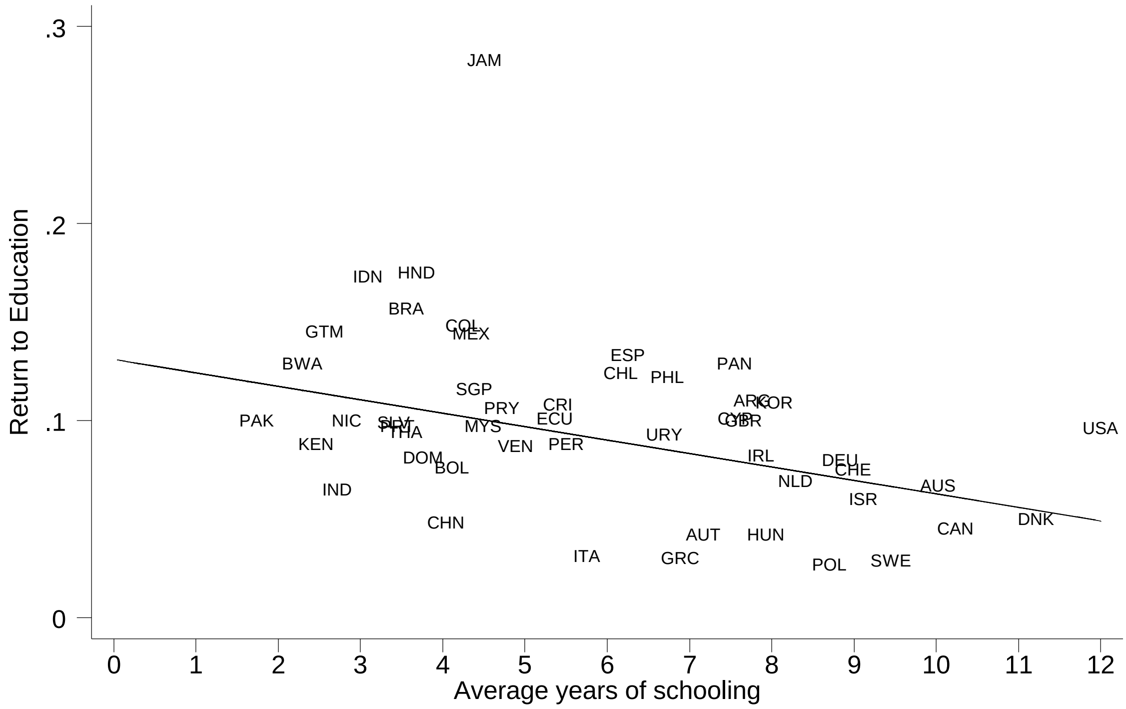
Table 9. Subsample robustness of the inequality equation

	(1) Without 10 countries with highest growth in education	(2) Without 10 countries with highest growth in GDP	(3) Without 10 countries with highest inequality growth	(4) Without 10 countries with highest education level	(5) Without 10 countries with highest GDP	(6) Without 10 countries with highest inequality	(7) Without 10 countries with lowest inequality
ΔS_t	-0.07871	-0.07214	-0.00494	-0.06751	-0.06705	-0.04840	-0.05793
	(2.22)	(2.00)	(0.17)	(1.89)	(2.03)	(1.71)	(1.82)
ΔV_t	-0.00315	-0.00174	0.00083	-0.00717	-0.00005	-0.00095	-0.00111
	(0.45)	(0.26)	(0.16)	(0.86)	(0.01)	(0.16)	(0.16)
(year=70)	-0.00933	-0.00874	-0.00724	-0.01193	-0.00853	-0.00967	-0.00831
	(1.63)	(1.44)	(1.57)	(1.87)	(1.33)	(1.84)	(1.36)
(year=80)	-0.00487	-0.00559	-0.00423	-0.00892	-0.00661	-0.00465	-0.00548
	(1.02)	(1.13)	(1.05)	(1.82)	(1.36)	(1.09)	(1.16)
Δ(def=inc)	0.04134	0.04150	0.02620	0.04344	0.04132	0.03678	0.04167
	(3.88)	(3.79)	(2.99)	(3.90)	(3.34)	(3.60)	(3.79)
Δ(def=hh)	-0.00038	-0.00020	0.01535	-0.00052	0.00088	0.00538	0.00028
	(0.07)	(0.03)	(2.88)	(0.08)	(0.15)	(0.94)	(0.05)
Δ(def=gr.)	-0.00527	-0.00516	-0.00723	-0.00419	-0.00292	-0.00624	-0.00549
	(0.45)	(0.43)	(0.77)	(0.34)	(0.17)	(0.56)	(0.45)
Constant	0.01152	0.01121	0.00252	0.01474	0.01241	0.00909	0.01068
	(2.98)	(2.84)	(0.73)	(2.98)	(2.81)	(2.55)	(2.61)
Obs.	69	66	61	64	64	73	70
R-squared	0.32	0.31	0.39	0.33	0.32	0.30	0.28
Countries excluded from the sample	Congo Egypt China Hong Kong Jordan Korea Taiwan Austria Cyprus Romania	Botswana Swaziland Hong Kong Japan Korea Singapore Taiwan Malta Bulgaria Romania	Guatemala Brazil Chile Venezuela China Hong Kong Thailand Australia New Zealand Soviet Union	Canada USA Denmark Finland Sweden Australia New Zealand Czechoslov. E. Germany Soviet Union	Canada USA Bahrain Kuwait Belgium France Germany Netherlands Switzerland Australia	Gabon Guinea Biss. Lesotho Malawi Sierra Leone South Africa Zimbabwe Guatemala Honduras Brazil	Belgium Hungary Uk Bulgaria Czechoslov. Romania Latvia Slovak Rep. Slovenia Ukraine

Absolute value of t statistics in parentheses.

Figure 1. Return to education, education and inequality

A. Diminishing returns to education



B. Returns to education and inequality