WORKING PAPER #429
INDUSTRIAL RELTIONS SECTION
PRINCETON UNIVERSITY
JANUARY, 2000

Education for Growth: Why and For Whom?

Alan B. Krueger
Princeton University and NBER

and

Mikael Lindahl
Stockholm University

This Draft: December 1999
First Draft: August 1998

*We thank Angus Deaton, Richard Freeman, Zvi Griliches, Gene Grossman, Larry Katz, Kjetil Storesletten, and seminar participants at Berkeley, the London School of Economics, MIT, Princeton University, University of Texas at Austin, Uppsala University, IUI, FIEF, SOFI, and the Tinbergen Institute for helpful discussions, and Peter Skogman, Mark Spiegel and Bob Topel for providing data. Alan Krueger thanks the Princeton University Industrial Relations Section for financial support; Mikael Lindahl thanks the Swedish Council for Research in the Humanities and Social Science for financial support.

EDUCATION FOR GROWTH: WHY AND FOR WHOM?

ABSTRACT

This paper tries to reconcile evidence from the microeconometric and empirical macro growth literatures on the effect of schooling on income and GDP growth. Much microeconometric evidence suggests that education is an important causal determinant of income for individuals within countries. At a national level, however, recent studies have found that increases in educational attainment are unrelated to economic growth. This finding appears to be a spurious result of the extremely high rate of measurement error in first-differenced cross-country education data. After accounting for measurement error, the effect of changes in educational attainment on income growth in cross-country data is at least as great as microeconometric estimates of the rate of return to years of schooling. Another finding of the macro growth literature -- that economic growth depends positively on the initial stock of human capital -- is shown to result from imposing linearity and constant-coefficient assumptions on the estimates. These restrictions are often rejected by the data, and once either assumption is relaxed the initial level of education has little effect on economic growth for the average country.

Alan B. Krueger
Economics Department
Princeton University
Princeton, NJ 08540

Mikael Lindahl
Economics Department
Stockholm University
Stockholm, Sweden

1. Introduction

Interest in the rate of return to investment in education has been sparked by two independent developments in economic research in the 1990s. On the one hand, the micro labor literature has produced several new estimates of the monetary return to schooling that exploit natural experiments in which variability in workers' schooling attainment was generated by some exogenous and arguably random force, such as quirks in compulsory schooling laws or students' proximity to a college. On the other hand, the macro growth literature has investigated whether the level of schooling in a cross-section of countries is related to the countries' subsequent GDP growth rate. This paper summarizes and tries to reconcile these two disparate but related lines of research.

The next section reviews the theoretical and empirical foundations of the Mincerian human capital earnings function. Our survey of the literature indicates that Mincer's (1974) formulation of the log-linear earnings-education relationship fits the data rather well. Each additional year of schooling appears to raise earnings by about 10 percent in the United States, although the rate of return to education varies over time as well as across countries. There is surprisingly little evidence that omitted variables (e.g., inherent ability) that might be correlated with earnings and education cause simple OLS estimates of wage equations to significantly overstate the return to education. Indeed, consistent with Griliches's (1977) conclusion, much of the modern literature finds that the upward "ability bias" is of about the same order of magnitude as the downward bias caused by measurement error in educational attainment.

Section 3 considers the empirical macro growth literature. First, we relate the Mincerian wage equation to the macro growth model. The Mincer model implies that the change in a country's average level of schooling should be the key determinant of income growth. The macro growth literature, by contrast, typically specifies growth as a function of the initial level of education. Moreover, we show that if the return to education changes over time (e.g., because of exogenous skill-biased technological change), the macro growth models are unidentified. Much of the empirical growth literature has eschewed the Mincer model because studies such as Benhabib and Spiegel (1994) find that the change in education is not a determinant of economic growth.[1] We show, however, that Benhabib and Spiegel's finding that increases in education are unrelated to economic growth results because there is virtually no signal in the education data they use, conditional on the growth of capital.

In contrast to the micro education literature, the macro growth literature has devoted only cursory attention to potential problems caused by measurement errors in education. Despite their aggregate nature, available data on average schooling levels across countries are poorly measured, in large part because they are often derived from enrollment flows. The reliability of country-level education data is no higher than the reliability of individual-level education data. For example, the correlation between Barro and Lee's (1991) and Kyriacou's (1991) measures of average education across 68 countries in 1985 is 0.86, and the correlation between the *change* in schooling between 1965 and 1985 from these two sources is only 0.34. Additional estimates of the reliability of country-level education data based on our analysis of comparable micro data from the World Values Survey for 34 countries suggests that measurement error is particularly prevalent for secondary and higher schooling. The measurement errors in schooling are positively correlated over time, but not as highly correlated as true years of schooling. We find that

[1]There are also notable exceptions that have embraced the Mincer model, such as Hall and Jones (1998) and Klenow and Rodriguez-Clare (1997).

measurement errors in education severely attenuate estimates of the effect of the change in schooling on GDP growth. Nonetheless, we show that measurement errors in schooling are unlikely to cause a spurious positive association between the initial level of schooling and GDP growth across countries, *conditional on* the change in education. Thus, like Topel (1999), we conclude that both the change and initial level of education are positively correlated with economic growth.

Finally, we explore whether the significant effect of the initial level of schooling on growth continues to hold if several restrictive assumptions in the growth equation are relaxed. For example, we estimate a variable-coefficient model that allows the coefficient on education to vary across countries (as is found in the micro data), and we relax the linearity assumption of the initial level of education. These extensions show that the positive effect of the initial level of education on economic growth is sensitive to econometric restrictions that are rejected by the data.

Our main conclusion is that while support for the micro Mincerian wage equation is strong, the evidence of a positive effect of the stock of education on a country's growth rate is tenuous. Even if one accepts the strong assumptions necessary to interpret the coefficient on the initial level of education in cross-country growth regressions as identifying externalities from education, the results most likely do not apply to the OECD countries, or to the average country in the world: the positive effect of the initial level of education applies primarily to countries with very low levels of education.

2. *The Microeconomics of the Private Return to Education*

Since at least the beginning of the century, economists and other social scientists have sought to estimate the economic rewards individuals and society gain from completing higher levels of schooling.[2] It has long been recognized that workers who attended school longer may possess other characteristics that would lead them to earn higher wages irrespective of their level of education. If these other characteristics are not accounted for, then simple comparisons of earnings across individuals with different levels of schooling would overstate the return to education. Early attempts to control for this "ability bias" included the analysis of data on siblings to difference-out unobserved family characteristics (e.g., Gorseline, 1932), and regression analyses which included as control variables observed characteristics such as IQ and parental education (e.g., Griliches and Mason, 1972). This literature is thoroughly surveyed in Griliches (1977), Rosen (1977), Willis (1986), and Card (1999). We briefly review evidence on the Mincerian earnings equation, emphasizing recent studies that exploit exogenous variations in education in their estimation.

2.1 The Mincerian wage equation

Mincer (1974) showed that if the only cost of attending school an additional year is the opportunity cost of students' time, and if the proportional increase in earnings caused by this additional schooling is constant over the lifetime, then the log of earnings would be linearly related to individuals' years of schooling, and the slope of this relationship could be interpreted as the rate of return to investment in schooling.[3] He augmented this model to include a quadratic

---

[2]Early references are Gorseline (1932), Walsh (1935), Miller (1955), and Wolfle and Smith (1956).

[3]This insight is also in Becker (1964) and Becker and Chiswick (1966), who specify the cost of investment in human capital as a fraction of earnings that would have been received in the absence of the investment. There are, of course, other theoretical models that yield a log-linear earnings-schooling relationship. For example, if the

term in work experience to allow for returns to on-the-job training, yielding the familiar Mincerian wage equation:

(1) $\ln W_i = \beta_0 + \beta_1 S_i + \beta_2 X_i + \beta_3 X_i^2 + \epsilon_i$,

where $\ln W_i$ is the natural log of the wage for individual i, S_i is years of schooling, X_i is experience, X_i^2 is experience squared, and ϵ_i is a disturbance term. With Mincer's assumptions, the coefficient on schooling, β_1, equals the discount rate, because schooling decisions are made by equating two present value earnings streams: one with a higher level of schooling and one with a lower level. An attractive feature of Mincer's model is that time spent in school (as opposed to degrees) is the key determinant of earnings, so data on years of schooling can be used to estimate a comparable return to education in countries with very different educational systems.

Equation (1) has been estimated for most countries of the world by OLS, and the results generally yield estimates of β_1 ranging from .05 to .15, with slightly larger estimates for women than men (see Psacharopoulos, 1994). The log-linear relationship also provides a good fit to the data, as is illustrated by the plots for the U.S., Sweden, West Germany, and East Germany in Figure 1.[4] These figures display the coefficient on dummy variables indicating each year of schooling, controlling for experience and gender, as well as the OLS estimate of the Mincerian return. It is apparent that the semi-log specification provides a good description of the data even

production function relating earnings and human capital is log-linear, and individuals randomly choose their schooling level (e.g., optimization errors), then estimation of equation (1) would uncover the educational production function.

[4]The German figures are from Krueger and Pischke (1995). The American and Swedish figures are based on the authors' calculations using the 1991 March Current Population Survey and 1991 Swedish Level of Living Survey. The regressions also include controls for a quadratic in experience and sex.

in countries with dramatically different economic and educational systems.[5]

Much research has addressed the question of how to interpret the education slope in equation (1). Does it reflect unobserved ability and other characteristics that are correlated with education, or the true reward that the labor market places on education? Is education rewarded because it is a signal of ability (Spence, 1973), or because it increases productive capabilities (Becker, 1964)? Is the social return to education higher or lower than the coefficient on education in the Mincerian wage equation? Would all individuals reap the same proportionate increase in their earnings from attending school an extra year, or does the return to education vary systematically with individual characteristics? Definitive answers to these questions are not available, although the weight of the evidence clearly suggests that education is not merely a proxy for unobserved ability. For example, Griliches (1977) concludes that instead of finding the expected positive ability bias in the return to education, "The implied net bias is either nil or negative" once measurement error in education is taken into account.

The more recent evidence from natural experiments also supports a conclusion that omitted ability does not cause upward bias the return to education (see Card, 1999 for a survey). For example, Angrist and Krueger (1991) observe that the combined effect of school start age cutoffs and compulsory schooling laws produces a natural experiment, in which individuals who are born on different days of the year start school at different ages, and then reach the compulsory

[5]Evaluating micro data for states over time in the U.S., Card and Krueger (1992) find that the earnings-schooling relationship is flat until the education level reached by the 2nd percentile of the education distribution, and then becomes log-linear. There is also some evidence of sheep-skin effects around college and high school completion (e.g., Park, 1994). Although statistical tests often reject the log-linear relationship for a large sample, the figures clearly show that the log-linear relationship provides a good approximation to the functional form. It should also be noted that Murphy and Welch (1990) find that a quartic in experience provides a better fit to the data than a quadratic.

schooling age at different grade levels. If the date of the year individuals are born is unrelated to their inherent abilities, then, in essence, variations in schooling associated with date of birth provide a natural experiment for estimating the benefit of obtaining extra schooling in response to compulsory schooling laws. Using a sample of nearly one million observations from the U.S. Censuses, Angrist and Krueger find that men born in the beginning of the calendar year, who start school at a relatively older age and can dropout in a lower grade, tend to obtain less schooling. This pattern only holds for those with a high school education or less, consistent with the view that compulsory schooling is responsible for the pattern. They further find that the pattern of education by quarter-of-birth is mirrored by the pattern of earnings by quarter-of-birth: in particular, individuals who are born early in the year tend to earn less, on average.[6] Instrumental variables (IV) estimates that are identified by variability in schooling associated with quarter-of-birth suggest that the payoff to education is slightly higher than the OLS estimate.[7] Angrist and Krueger conclude that the upward bias in the return to schooling is about the same order of magnitude as the downward bias due to measurement error in schooling.

Other studies have used a variety of other sources of arguably exogenous variability in schooling to estimate the return to schooling. Harmon and Walker (1995), for example, more directly examine the effect of compulsory schooling by studying the effect of changes in the compulsory schooling age in the United Kingdom, while Card (1995a) exploits variations in

---

[6]Again, no such pattern holds for college graduates.

[7]Bound, Jaeger and Baker (1995) argue that Angrist and Krueger's IV estimates are biased toward the OLS estimates because of weak instruments. However, Staiger and Stock (1997), Donald and Newey (1997), Angrist, Imbens and Krueger (1999), and Chamberlain and Imbens (1996) show that weak instruments do not account for the central conclusion of Angrist and Krueger (1991).

schooling attainment owing to families' proximity to a college in the U.S. Duflo (1998) uses variation in educational attainment related to school building programs across islands in Indonesia. These three papers find that the IV estimates of the return to education that exploit a "natural experiment" for variability in education exceed the corresponding OLS estimates, although the difference between the IV and OLS estimates often is not statistically significant. Ashenfelter, Harmon and Oosterbeek (1999) compile estimates from 27 studies, and find that the conventional OLS return to schooling is .066, on average, whereas the average IV estimate is .093.[8]

We interpret this evidence as suggesting that the return to an additional year of education obtained for reasons like compulsory schooling or school-building projects is more likely to be greater, than lower, than the conventionally-estimated return to schooling. Because the schooling levels of individuals who are from more disadvantaged backgrounds tend to be those who are most affected by the interventions examined in the literature, Lang (1993) and Card (1995b) have inferred that the return to an additional year of schooling is higher for individuals from disadvantaged families than for those from advantaged families, and suggest that such a result follows because disadvantaged individuals have higher discount rates.

Other related evidence for the U.S. suggests the payoff to investments in education are higher for more disadvantaged individuals. First, while studies of the effect of school resources on student outcomes yield mixed results, there is a tendency to find more beneficial effects of school resources for disadvantaged students (see, for example, Summers and Wolfe, 1977, Krueger, 1999 and Rivkin, Hanushek and Kain, 1998). Second, evidence suggests that pre-school

[8]Ashenfelter, Harmon and Oosterbeek explore whether publication bias accounts for the tendancy of IV estimates to exceed the OLS estimates. After adjusting for selection bias, however, they still find that the return is higher, on average, in the IV estimates than in the OLS estimates (.081 versus .064).

programs have particularly large, long-term effects for disadvantaged children in terms of reducing crime and welfare dependence, and raising incomes (see, Barnett, 1992). Third, several studies have found that students from advantaged and disadvantaged backgrounds make equivalent gains on standardized tests during the school year, but children from disadvantaged backgrounds fall behind during the summer while children from advantaged backgrounds move ahead (see Entwisle, Alexander, and Olson, 1997). And fourth, evidence suggests that college students from more disadvantaged families benefit more from attending elite colleges than do students from advantaged families (see Dale and Krueger, 1998).

It is unclear whether this evidence of a higher return to human capital for disadvantaged students applies outside the U.S. But in all regions of the world, Psacharopolous (1994) concludes that there is a higher return to primary schooling than to secondary or tertiary schooling, which also suggests that disadvantaged children benefit more from additional human capital investments.

2.2 Social versus Private Returns to Education

The social return to education can, of course, be higher or lower than the private return. The social return can be higher because of externalities from education, which could occur, for example, if higher education leads to technological progress that is not captured in the private return to that education, or if more education produces positive externalities, such as a reduction in crime and welfare participation, or more informed political decisions. The former is more likely if human capital is expanded at higher levels of education while the latter is more likely if it is expanded at lower levels. It is also possible that the social return to education is less than the private return. For example, Spence (1973) and Machlup (1970) note that education could just be a credential, which does not raise individuals' productivities. It is also possible that in some developing countries, where the incidence of unemployment may rise with education (e.g., Blaug, Layard and Woodhall, 1969) and where the return to physical capital may exceed the return to human capital (e.g., Harberger, 1965), increases in education may reduce total output.

Most of the micro human capital literature focuses on the private rather than social return to education, but the finding of little ability bias in the Mincerian wage equation casts doubt on at least some forms of credentialling arguments. The possibility of externalities to education motivates much of the macro growth literature, to which we now turn.

3. *Macro growth equations*

Thirty years ago, Fritz Machlup (1970, p. 1) observed, "The literature on the subject of education and economic growth is some two hundred years old, but only in the last ten years has the flow of publications taken on the aspects of a flood." The number of cross-country regression studies on education and growth has surged even higher in recent years. The literature is voluminous, and a new journal has been devoted solely to economic growth. Rather than exhaustively review the literature, we summarize the main models and findings, and explore the impact of several econometric issues.

The macro growth literature yields three principally different conclusions from the micro literature. First, the initial stock of human capital matters, not the change in human capital. Second, secondary and post-secondary education matter more for growth than primary education. Third, female education has an insignificant and sometimes negative effect on economic growth.

3.1 From the Mincer Model to the Macro Growth Model

To compare the Mincer model to the macro growth literature, first consider a Mincerian wage equation for each country j and time period t:

(1') $\ln W_{ijt} = \beta_{0jt} + \beta_{1jt}S_{ijt} + \epsilon_{ijt}$,

where we have suppressed the experience term for convenience. This equation can be aggregated across individuals each year by taking the means of each of the variables, yielding what Heckman and Klenow (1997) call the "Macro-Mincer" wage equation:

(2) $\ln Y^{g}_{jt} = \beta_{0jt} + \beta_{1jt}S_{jt} + \epsilon_{jt}$,

where Y^{g}_{jt} denotes the geometric mean wage and S_{jt} is mean education. Heckman and Klenow (1997) compare the coefficient on education from cross-country log GDP equations to the coefficient on education from micro Mincer models. Once they control for life expectancy to proxy for technology differences across countries, they find that the macro and micro regressions yield similar estimates of the effect of education on income.[9] They conclude from this exercise that the "macro versus micro evidence for human capital externalities is not robust."

The macro Mincer equation can be differenced between year t and t-1, giving:

(3) $\Delta\ln Y^{g}_{j} = \beta'_{0} + \beta_{1jt}S_{jt} - \beta_{1jt-1}S_{jt-1} + \Delta\epsilon'_{jt}$,

where Δ signifies the change in the variable from t-1 to t, β'_{0} is the mean change in the intercepts, and $\Delta\epsilon'_{jt}$ is a composite error that includes the deviation between each country's intercept change and the overall average. Differencing the equation removes the effect of any additive, permanent

---

[9]When they omit life expectancy, however, education has a much larger effect in the macro regression than micro regression. Whether longer life expectancy is a valid proxy for technology differences, or a result of higher income, is an open question (see Smith, 1999).

differences in technology. If the return to schooling is constant over time, we have:

(4) $\Delta \ln Y^g_j = \beta'_0 + \beta_{1j}\Delta S_j + \Delta\epsilon'_{jt}$.

Notice that this formulation allows the time-invariant return to schooling to vary across countries. If β_{1j} does vary across countries, and a constant-coefficient model is estimated, then $(\beta_1-\beta_{1j})\Delta S_j$ will add to the error term.

Also notice that if the return to schooling varies over time, then by adding and subtracting $\beta_{1jt}S_{jt-1}$ from the right-hand-side of equation (3), we obtain:

(5) $\Delta \ln Y^g_j = \beta'_0 + \beta_{1jt}\Delta S_j + \delta S_{jt-1} + \Delta\epsilon'_{jt}$,

where δ is the change in the return to schooling ($\Delta\beta_{1j}$). If the return to schooling has increased (decreased) secularly over time, the initial level of education will enter positively (negatively) into equation (5). An implicit assumption in much of the macro growth literature therefore is that the return to education is either unchanged, or changed endogenously, by the stock of human capital.

Although the empirical literature for the U.S. clearly shows a fall in the return to education in the 1970s and a sharp increase in the 1980s (e.g., Levy and Murnane, 1992), the findings for other countries are mixed. For example, Psacharopolous (1994; Table 6) finds that in the average country the Mincerian return to education fell by 1.7 points over periods of various lenghths (average of 12 years) since the late 1960s. By contrast, O'Neill (1995) finds that between 1967 and 1985 the return to education measured in terms of its contribution to GDP rose by 58 percent in developed countries and by 64 percent in less developed countries.

The typical macro growth model estimated in the literature is motivated by the convergence literature. This leads to interest in estimating parameters of an underlying model such as $\Delta Y_j = \alpha_j - \beta(Y_{jt-1}-Y^*_j) + \mu_j$, where ΔY_j denotes the annualized change in log GDP per capita in country j between t-1 and t, α_j denotes country j's steady-state growth rate, Y_{jt-1} is the log of initial GDP per-capita, Y^*_j is steady-state log GDP per capita, and β measures the speed of convergence to steady-state income. The intuition for this equation is straightforward: countries that are below their steady-state income level should grow quickly, and those that are above it should grow slowly. A typical estimating equation is:

(6) $\Delta Y_j = \beta_0 + \beta_1 Y_{jt-1} + \beta_2 S_{j,t-1} + \beta_3 Z_{j,t-1} + \epsilon_j$

where ΔY_j is the change in log GDP per capita from year t-1 to t, S_{t-1} is average years of schooling in the population in the initial year, Y_{t-1} is the log of initial GDP per capita, and Z_{t-1} includes variables such as inflation, capital, or the "rule of law index."[10] Also note that schooling is sometimes specified in logarithmic units in equation (6). Barro and Sala-i-Martin (1995), Benhabib and Spiegel (1994), and others conclude that the change in schooling has an insignificant effect if it is included in a GDP growth equation, even though this variable is predicted to matter in the Mincer model and in some endogenous economic growth models (e.g., Lucas, 1988). Equation (6) is typically estimated with data for a cross-section or pooled sample of countries spanning a 5, 10, or 20 year period.

The first-differenced macro-Mincer equation (4) differs from the macro growth equation in several respects. First, the macro growth models use the change in log GDP per capita as the dependent variable, rather than the change in the mean of log earnings. If income has a log normal distribution with a constant variance over time, and if labor's share is also constant, then

[10]Henceforth we use the terms GDP per capita and GDP interchangeably.

aggregating GDP in this way would not matter.[11] Without micro data for a large sample of countries over time, the impact of aggregation is difficult to assess, and we do not pursue this issue further. Second, and probably more importantly, the macro growth literature typically omits the change in schooling. Third, because the macro models are motivated by issues of convergence they include the initial level of GDP, capital, and correlates for steady-state income. Indeed, a primary motivation for including human capital variables in these equations is to control for steady state income, Y*.

3.1.1 Interpretation

There are at least six ways to interpret the coefficient on the initial level of schooling in equation (6).[12] First, schooling may be a proxy for steady-state income. Countries with more schooling would be expected to have a higher steady-state income, so conditional on GDP in the initial year, we would expect more educated countries to grow faster ($\beta_2 > 0$).[13] If this were the case, higher schooling levels would not change the steady-state growth rate, although it would raise steady-state income. Second, schooling could change the steady-state growth rate by enabling the work force to develop, implement and adopt new technologies (see Nelson and Phelps, 1966 Welch, 1970 and Romer, 1990), again leading to the prediction $\beta_2 > 0$. Third, countries with low initial stocks of human capital could have greater opportunities to grow by

---

[11]Heckman and Klenow (1997) point out that half the variance of log income will be added to the GDP equation if income is log normal.

[12]See Aghion and Howitt (1998) and Temple (1999a) for excellent overviews of growth models.

[13]Transitory measurement error in GDP will have the same effect.