

NBER WORKING PAPER SERIES

EDUCATION FOR GROWTH: WHY AND FOR WHOM?

Alan B. Krueger
Mikael Lindahl

Working Paper 7591
<http://www.nber.org/papers/w7591>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
March 2000

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. The views expressed herein are those of the authors and are not necessarily those of the National Bureau of Economic Research.

© 2000 by Alan B. Krueger and Mikael Lindahl. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Education for Growth: Why and For Whom?
Alan B. Krueger and Mikael Lindahl
NBER Working Paper No. 7591
March 2000
JEL No. J24, E20

ABSTRACT

This paper tries to reconcile evidence from the microeconomic and empirical macro growth literatures on the effect of schooling on income and GDP growth. Much microeconomic evidence suggest 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 microeconomic 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
and NBER

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

¹There 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

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

³This 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.

⁴The 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

⁵Evaluating 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

⁶Again, no such pattern holds for college graduates.

⁷Bound, 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

⁸Ashenfelter, Harmon and Oosterbeek explore whether publication bias accounts for the tendency 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, Psacharopoulos (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 credentialing 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 :

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

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:

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

where Y_{jt}^g 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:

$$(3) \quad \Delta \ln Y_{jt}^g = \beta'_0 + \beta_{1jt}S_{jt} - \beta_{1j,t-1}S_{j,t-1} + \Delta \epsilon'_{jt},$$

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

⁹When 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:

$$(4) \quad \Delta \ln Y_j^e = \beta'_0 + \beta_{1j} \Delta S_j + \Delta \epsilon'_{jt}$$

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:

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

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, Psacharopoulos (1994; Table 6) finds that in the average country the Mincerian return to education fell by 1.7 points over periods of various lengths (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 $\Delta Y_j = \alpha_j - \beta(Y_{j,t-1} - Y_j^*) + \mu_j$, 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, $Y_{j,t-1}$ 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:

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

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

¹⁰Henceforth 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

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

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

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

implementing technology developed abroad. In this case, one would expect $\beta_2 < 0$. Aghion and Howitt (1998) and Benhabib and Spiegel (1994) suggest, however, that for countries with below-average productivity and income, this "catch-up" effect would be more prominent among those with higher average educational attainment, yielding the prediction $\beta_2 > 0$ for this subset of countries.

Fourth, a positive or negative coefficient on initial schooling may simply reflect an exogenous change in the return to schooling, as shown in equation (5). Fifth, anticipated increases in future economic growth could cause schooling to rise (i.e., reverse causality), as argued by Bils and Klenow (1998). Sixth, the schooling variable may "pick up" the effect of the change in education, which is typically omitted from the growth equation.

3.2 Basic Results and Effect of Measurement Error

Table 1 replicates and extends the "growth accounting" and "endogenous growth" regressions in Benhabib and Spiegel's (1994) influential paper.¹⁴ Their analysis is based on Kyriacou's (1991) measure of average years of schooling for the work force in 1965 and 1985, Summers and Heston's GDP and labor force data, and a measure of physical capital derived from investment flows for a sample of 78 countries. Following Benhabib and Spiegel, the regression in column (1) relates the annualized growth rate of GDP to the log change in years of schooling.

¹⁴We were not able to exactly replicate Benhabib and Spiegel's results because we use a revised version of Summers and Heston's GDP data. Nonetheless, our estimates are very close to theirs. For example, Benhabib and Spiegel report coefficients of -.059 for the change in log education and .545 for the change in log capital when they estimate the model in column 1 of Table 1; our estimates are -.072 and .523. Some of the other coefficients differ because of scaling; for comparability with later results, we divided the dependent variable and variables measured in changes by 20.

From this model, Benhabib and Spiegel conclude, "Our findings shed some doubt on the traditional role given to human capital in the development process as a separate factor of production." Instead, they conclude that the stock of education matters for growth (see column 2 and 5) by enabling countries with a high level of education to adopt and innovate technology faster.

Topel argues that Benhabib and Spiegel's finding of an insignificant and wrong-signed effect of schooling changes on GDP growth is due to their log specification of education.¹⁵ The log-log specification follows if one assumes that schooling enters an aggregate Cobb-Douglas production function linearly. Given the success of the Mincer model, however, we would agree with Topel that it is more natural to specify human capital as an exponential function of schooling in a Cobb-Douglas production function, so the change in linear years of schooling would enter the growth equation. In any event, the logarithmic specification of schooling does not fully explain the perverse effect of educational improvements on growth in Benhabib and Spiegel's analysis.¹⁶ Results of estimating a linear education specification in column 4 still show a statistically insignificant (though positive) effect of the linear change in schooling on economic growth.

Columns 3 and 6 show that controlling for capital is key to Benhabib and Spiegel's finding of an insignificant effect of the change in schooling variable. When physical capital is excluded

¹⁵Mankiw, Romer and Weil (1992; Table VI) estimate a similar specification.

¹⁶The log specification is part of the explanation, however, because if the model in column (3) is estimated without the initial level of schooling, the change in log schooling has a negative and statistically significant effect, whereas the change in the level of schooling has a positive and statistically significant effect if it is included as a regressor in this model instead.

from the growth equation, the change in schooling has a statistically significant and positive effect in either the linear or log schooling specification. Why does controlling for capital have such a large effect on education? As shown below, it appears that the insignificant effect of the change in education is a result of the low signal in the education change variable. Indeed, we show below that conditional on the other variables that Benhabib and Spiegel hold constant (especially capital), the change in schooling conveys virtually no signal.¹⁷

Notice also that the coefficient on capital is high in Table 1, around .50 with a t-ratio close to 10. In a competitive, Cobb-Douglas economy, the coefficient on capital growth in a GDP growth regression should equal capital's share of national income. Gollin (1998) estimates that labor's share ranges from .65 to .80 in most countries, after allocating labor's portion of self-employment and proprietors' income. Consequently, capital's share is probably no higher than .20 to .35. Two reasons why the coefficient on capital could be biased upwards are: (1) measured capital is derived from investment flows, and GDP is a direct function of investment, so errors in the investment data will be positively correlated with the dependent variable; (2) countries that experience rapid GDP growth may find it easier to raise investment, creating simultaneity bias. These factors might explain why capital has such a large and significant coefficient in the growth equations. If the coefficient on capital growth in column (5) of Table 1 is constrained to equal .20 or .35 -- a plausible range for capital's share -- the coefficient on the schooling change rises to .09 or .06, and becomes statistically significant.

¹⁷Pritchett (1998) estimates essentially the same model as Benhabib and Spiegel (i.e., column 1 of Table 1), and instruments for schooling growth using an alternative education series. However, if there is no variability in the portion of measured schooling changes that represent true schooling changes conditional on capital, the instrumental variables strategy is inconsistent. This can easily be seen by noting that there would be no variability due to true education changes conditional on capital in the reduced form of the model.

3.2.1 The Extent of Measurement Error in International Education Data

We disregard errors that arise because years of schooling are an imperfect measure of human capital, and focus instead on the more tractable problem of estimating the extent of measurement error in cross-country data on average years of schooling. Benhabib and Spiegel's measure of average years of schooling for the work force was derived by Kyriacou (1991) as follows. First, survey-based estimates of average years of schooling for 42 countries in the mid 1970s were regressed on the countries' primary, secondary and tertiary school enrollment rates. Coefficient estimates from this model were then used to predict years of schooling from enrollment rates for all countries in 1965 and 1985. This method is likely to generate substantial noise since the fitted regression may not hold for all countries and time periods, enrollment rates are frequently mismeasured, and the enrollment rates are not properly aligned with the workforce. Changes in education derived from this measure are likely to be particularly noisy. Benhabib and Spiegel use Kyriacou's education data for 1965, as well as the change between 1965 and 1985.

The widely-used Barro and Lee (1993) data set is an alternative source of education data. For 40 percent of country-year cells, Barro and Lee measure average years of schooling by survey- and census-based estimates reported by UNESCO. The remaining observations were derived from historical enrollment flow data using a "perpetual inventory method."¹⁸ The Barro-Lee measure is undoubtedly an advance over existing international measures of educational attainment, but errors in measurement are inevitable because the UNESCO enrollment rates are of doubtful quality in many countries (see Behrman and Rosensweig, 1993 and 1994). For

¹⁸Each country has a survey- and census-based estimate in at least one year, which provides an anchor for the enrollment flows.

example, UNESCO data are often based on beginning of the year enrollment. Additionally, students educated abroad are miscounted in the flow data, which is probably a larger problem for higher education. More fundamentally, secondary and tertiary schooling is defined differently across countries in the UNESCO data, so years of secondary and higher schooling are likely to be noisier than overall schooling. Notice also that because errors cumulate over time in Barro and Lee's stock-flow calculations, the errors in education will be positively correlated over time.

As is well known, if an explanatory variable is measured with additive white noise errors, then the coefficient on this variable will be attenuated toward zero in a bivariate regression, with the attenuation factor, R , asymptotically equal to the ratio of the variance of the correctly-measured variable to the variance of the observed variable (see, e.g., Griliches, 1986). A similar result holds in a multiple regression (with correctly-measured covariates), only now the variances are conditional on the other variables in the model. To estimate attenuation bias due to measurement error, write a nation's measured years of schooling, S_j , as it's true schooling, S_j^* , plus a measurement error denoted e_j : $S_j = S_j^* + e_j$. It is convenient to start with the assumption that the measurement errors are "classical"; that is, errors that are uncorrelated with S^* , other variables in the growth equation, and the equation error term. Now let S^1 and S^2 denote two imperfect measures of average years of schooling for each country, with measurement errors e^1 and e^2 respectively (where we suppress the j subscript).

If e^1 and e^2 are uncorrelated, the fraction of the observed variability in S^1 due to measurement error can be estimated as $R_1 = \text{cov}(S^1, S^2) / \text{var}(S^1)$. R_1 is often referred to as the reliability ratio of S^1 , and has probability limit equal to $\text{var}(S^*) / \{\text{var}(S^*) + \text{var}(e^1)\}$. Assuming constant variances, the reliability of the data expressed in changes ($R_{\Delta S^1}$) will be lower than the

cross-sectional reliability if the serial correlation of the true variable is higher than the serial correlation of the measurement errors because $R_{\Delta S^1} = \text{var}(S^*) / \{\text{var}(S^*) + \text{var}(e)(1-r_e)/(1-\rho_{s^*})\}$, where r_e is the serial correlation of the errors and ρ_{s^*} is the serial correlation of true schooling. In practice, the reliability ratio for changes in S^1 can be estimated by: $R_{\Delta S^1} = \text{cov}(\Delta S^1, \Delta S^2) / \text{var}(\Delta S^1)$. Note that if the errors in S^1 and S^2 are positively correlated, the estimated reliability ratios will be biased upward.

We can calculate the reliability of the Barro-Lee and Kyriacou data if we treat the two variables as independent estimates of educational attainment. It is probably the case, however, that the measurement errors in the two data sources are positively correlated because, to some extent, they both rely on the same mismeasured enrollment data.¹⁹ Consequently, the reliability ratios derived from comparing these two measures probably provide an upper bound on the reliability of the data series.

Panel A of Table 2 presents estimates of the reliability ratio of the Kyriacou and Barro-Lee education data. Appendix Table A.1 reports the correlation and covariance matrices for the measures. The reliability ratios were derived by regressing one measure of years of schooling on the other.²⁰ The cross-sectional data have considerable signal, with the reliability ratio ranging

¹⁹Another complication is that the Kyriacou data pertain to the education of the work force, whereas the Barro-Lee data pertain to the entire population age 25 and older. If the regression slope relating true education of workers to the true education of the population is one, the reliability ratios reported in the text are unbiased. Although we do not know true education of workers and the population, in the Barro-Lee data set a regression of the average years of schooling of men (who are very likely to work) on the average education of the population yields a slope of .99, suggesting that workers and the population may have close to a unit slope.

²⁰Barro and Lee (1993) compare their education measure with alternative series by reporting correlation coefficients. For example, they report a correlation of .89 with Kyriacou's education data and .93 with Psacharopoulos's. Our cross-sectional correlations are not very different. They do not report correlations for changes in education.

from .77 to .85 in the Barro-Lee data and exceeding .96 in the Kyriacou data. The reliability ratios fall by 10 to 30 percent if we condition on the log of 1965 GDP per capita, which is a common covariate. More disconcerting, when the data are measured in changes over the 20 year period, the reliability ratio for the data used by Benhabib and Spiegel falls to less than 20 percent. By way of comparison, note that Ashenfelter and Krueger (1994) find that the reliability of self-reported years of education is .90 in micro data on workers, and that the reliability of self-reported differences in education between identical twins is .57.

These results suggest that if there were no other regressors in the model, the estimated effect of schooling changes in Benhabib and Spiegel's results would be biased downward by 80 percent. But the bias is likely to be even greater because their regressions include additional explanatory variables that absorb some of the true changes in schooling. The reliability ratio conditional on the other variables in the model can be shown to equal $R'_{s1} = (R_{s1} - R^2)/(1 - R^2)$, where R^2 is the multiple coefficient of determination from a regression of the measured schooling change variable on the other explanatory variables in the model. A regression of the change in Kyriacou's education measure on the covariates in column (4) of Table 1 yields an R^2 of 23 percent. If the covariates are correlated with the signal in education changes and not the noise, then there is *no* variability in true schooling changes left over in the measured schooling changes conditional on the other variables in the model. Instead of rejecting the traditional Mincerian role of education on growth, a reasonable interpretation is that Benhabib and Spiegel's results shed no light on the role of education changes on growth.

The Barro and Lee data convey more signal than Kyriacou's data when expressed in changes. Indeed, nearly 60 percent of the variability in observed changes in years of education

in the Barro-Lee data represent true changes. This makes the Barro-Lee data preferable to use to estimate the effect of educational improvements. Despite the greater reliability of the Barro-Lee data, there is still little signal left over in these data conditional on the other variables in the model in column 4 of Table 1; a regression of the change in the Barro-Lee schooling measure on the change in capital, change in population, and initial schooling yields an R^2 of .28. Consequently, conditional on these variables about 40 percent of the remaining variability in schooling changes in the Barro-Lee data is true signal.

As mentioned, we suspect the estimated reliability ratios are biased upward because the errors in the Kyriacou and Barro-Lee data are probably positively correlated. To derive a measure of education with independent errors, we calculated average years of schooling from the World Values Survey (WVS) for 34 countries. The WVS contains micro data from household surveys that were conducted in nearly 40 countries in 1990 or 1991. The survey was designed to be comparable across countries. In each country, individuals were asked to report the age at which they left school. With an assumption of school start age, we can calculate the average number of years that individuals spent in school. We also calculated average years of secondary and higher schooling by counting years of schooling obtained after 8 years of schooling as secondary and higher schooling. Notice that these measures will not be error free either. Errors could arise, for example, because some individuals repeated grades, because we have made an erroneous assumption about school start age or the beginning of secondary schooling, or because of sampling errors. But the errors in this measure should be independent of the errors in Kyriacou's and Barro and Lee's data. The appendix provides additional details of our calculations with the WVS.

Panel B of Table 2 reports the reliability ratios for the Barro-Lee data and WVS data for 1990. The reliability ratio of .90 for the Barro-Lee data in 1990 is slightly higher than the estimate for 1985 based Kyriacou's data, but within one standard error. Thus, it appears that correlation between the errors in Kyriacou's and Barro-Lee's data is not a serious problem. Nonetheless, another advantage of the WVS data is that they can be used to calculate upper secondary schooling using a constant (if imperfect) definition across countries. As one might expect given differences in the definition of secondary schooling in the UNESCO data, the reliability of the secondary and higher schooling (.72) is lower than the reliability of all years of schooling.

Lastly, it should be noted that the measurement errors in schooling are highly serially correlated in the Barro-Lee data. This can be seen from the fact that the correlation between the 1965 and 1985 schooling levels across countries is .97 in the Barro-Lee data, while less than 90 percent of the variations in the cross-sectional data across countries appear to represent true signal. If the reliability ratios reported in Table 2 are correct, the only way the time-series correlation in education could be so high is if the errors are serially correlated. The correlation of the errors can be estimated as:

$$[\text{cov}(S_{85}^{\text{BL}}, S_{65}^{\text{BL}}) - \text{cov}(S_{85}^{\text{BL}}, S_{65}^{\text{K}})] / [(1 - R_{85}^{\text{BL}}) \text{var}(S_{85}^{\text{BL}}) (1 - R_{65}^{\text{BL}}) \text{var}(S_{65}^{\text{BL}})]^{1/2},$$

where the superscript BL stands for Barro-Lee's data and K for Kyriacou's data. Using the reliability ratios in Table 2, the estimated correlation of the errors in Barro-Lee's schooling measure between 1965 and 1985 is .61. The correlation between true schooling in 1965 and 1985

is estimated at .97.²¹ Since the serial correlation of true schooling is higher than the serial correlation of the errors, the reliability of the first-differenced education data is lower than the reliability of the cross-sectional data.

3.3 Additional Growth Models

Measurement errors aside, one could question whether physical capital should be included in a GDP growth equation because it is potentially an endogenous variable. Fast growing countries have more access to capital (Blomström, Lipsey, and Zejan, 1993). Additionally, considerations of the low signal in schooling changes conditional on capital growth, and the mechanical correlation between measured capital and GDP (since capital is typically derived from investment), lead us to prefer parsimonious models that omit capital. Barro (1997) also excludes capital, so there is precedent for a parsimonious specification in the growth literature. We first report models without controlling for capital, and then focus on the effect of capital in long-difference models in Section 3.6.

Table 3 reports "stylized" macro growth models without controlling for physical capital for samples spanning 5, 10 or 20 year periods. The dependent variable is the annualized change in the log of real GDP per capita per year based on Summers and Heston's (1991) Penn World Tables, Mark 5.6. Results are generally similar if GDP per worker is used instead. We use GDP per capita because it reflects labor force participation decisions and because it has been the focus of much of the previous literature. The schooling variable is Barro and Lee's measure of average

²¹We estimate the serial correlation between true schooling levels in 1985 and 1965 using the formula: $\rho_s = [\text{cov}(S_{85}^{BL}, S_{65}^{K})\text{cov}(S_{65}^{BL}, S_{85}^{K}) / \text{cov}(S_{85}^{BL}, S_{85}^{K})\text{cov}(S_{65}^{BL}, S_{65}^{K})]^{1/2}$.

years of schooling for the population age 25 and older. When the change in average schooling is included as a regressor in these models, we divide it by the number of years in the time span so the coefficients are comparable across columns. The equations were estimated by OLS, but the standard errors reported in the table allow for a country-specific component in the error term.²² We exclude other variables (e.g., rule of law index) that are sometimes included in macro growth models to focus on education, and because those other variables are probably influenced themselves by education.²³

Our findings confirm the results in Topel (1999). The change in schooling has little effect on GDP growth when the growth equation is estimated with high frequency changes (i.e., 5 years). However, increases in average years of schooling have a positive and statistically significant effect on economic growth over periods of 10 or 20 years. The magnitude of the coefficient estimates on both the change and initial level of schooling over long periods are large, probably too large to represent the causal effect of schooling.

The finding that the time span matters so much for the change in education suggests that measurement error in schooling influences these estimates. Over short time periods, there is little change in a nation's true mean schooling level, so the transitory component of measurement error in schooling would be large relative to variability in the true change. Over longer periods, true education levels are more likely to change, increasing the signal relative to the noise in measured

²²An alternative approach would be to estimate a restricted seemingly unrelated system or random effects model. Absent measurement error, these estimators are more efficient. But because bias due to measurement errors in the explanatory variables is exacerbated with these estimators, we elected to estimate the parameters by OLS and report robust standard errors.

²³If we control for the initial fertility rate, the education variable becomes much weaker and insignificant. See Krueger and Lindahl (1999).

changes. Measurement error bias appears to be greater over the 5 and 10 year horizons, but it is still substantial over 20 years. Since the change in schooling and initial level of GDP are essentially uncorrelated, the coefficient on the 20-year change in schooling in column 8 is biased downward by a factor of $1-R_{\Delta S}$, which is around 40 percent according to Table 2. Thus, adjusting for measurement error would lead the coefficient on the change in education to increase from .18 to $.30 = .18/(1-.4)$. This is an enormous return to investment in schooling, equal to three or four times the private return to schooling estimated within most countries. The large coefficient on schooling suggests the existence of quite large externalities from educational changes (Lucas, 1988) or simultaneous causality in which growth causes greater educational attainment. It is plausible that simultaneity bias is greater over longer time intervals, so some combination of varying measurement error bias and simultaneity bias could account for the time pattern of results displayed in Table 3.

Like Benhabib and Spiegel, Barro and Sala-i-Martin (1995) conclude that contemporaneous changes in schooling do not contribute to economic growth. There are four reasons to doubt their conclusion, however. First, Barro and Sala-i-Martin analyze a mixed sample that combines changes over both 5-year (1985-90) and 10-year (1965-75 and 1975-85) periods; examining changes over such short periods tends to exacerbate the downward bias due to measurement errors. Second, they examine changes in average years of secondary and higher schooling. As was shown in Table 2, the cross-sectional reliability of secondary and higher schooling is lower than the reliability of all years of schooling, and the changes are likely to be less reliable as well. Third, they include separate variables for changes in male and female years of secondary and higher schooling. These two variables are highly correlated ($r = .85$), which would exacerbate

measurement error problems if the signal in the variables is more highly correlated than the noise. If average years of secondary and higher schooling for men and women combined, or years of secondary and higher schooling for either men or women, is used instead of all years of schooling in the 10-year change model in column 6 of Table 3, the change in education has a sizable, statistically significant effect. Fourth, they estimate a restricted Seemingly Unrelated Regression (SUR) system, which exacerbates measurement error bias because asymptotically this estimator is equivalent to a weighted average of an OLS and fixed-effects estimator.

Barro (1997) stresses the importance of male, secondary and higher education as a determinant of GDP growth. In his analysis, female secondary and higher education is negatively related to growth. We have explored the sensitivity of the estimates to using different measures of education: namely, primary versus higher education, and male versus female education. When we test for different effects of years of primary and secondary and higher schooling in the model in column 6 of Table 3, we cannot reject that all years of schooling have the same effect on GDP growth (p-value equals .40 for initial levels and .12 for changes). We also find insignificant differences between primary and secondary schooling if we just use male schooling. We do find significant differences if we further disaggregate schooling levels by gender, however. The initial level of primary schooling has a positive effect for women and a negative effect for men, the initial level of secondary school has a negative effect for women and a positive effect for men, the change in primary schooling has a positive effect for women and a negative effect for men, and the change in secondary schooling has a negative effect for women and a positive effect for men.

Caselli, Esquivel and Lefort (1996) also examine the differential effect of male and female education on growth over five year intervals. They estimate a fixed effects variant of equation

(6), and instrument for initial education and GDP with their lags. Contrary to Barro, they find that female education has a positive and statistically significant effect on growth, while male education has a negative and statistically significant effect. This result appears to stem from the introduction of fixed effects: if we estimate the model with fixed effects but without instrumenting for education, we find the same gender pattern, whereas if we estimate the model without fixed effects and instrument with lags the results are similar to Barro's. Although country fixed effects arguably belong in the growth equation, it is particularly difficult to untangle any differential effects of male and female education in such a specification because measurement error is exacerbated.²⁴ But Caselli, Esquivel and Lefort's findings are consistent with the micro-econometric literature, which finds education has a higher return for women than men.

We conclude that because schooling levels are highly correlated for men and women, one needs to be cautious interpreting the effect of education in models that disaggregate education by gender and level of schooling. For this reason, and because the total number of years of education is the variable specified in the Mincer model, we have a preference for using the average of all years of schooling for men and women combined in our econometric analysis.

3.4 Initial Level of Education

The effect of the initial level of education on growth has been widely interpreted as an indication of large externalities from the stock of a nation's human capital on growth. Benhabib and Spiegel (1994; p. 160), for example, conclude, "The results suggest that the role of human

²⁴Note that instrumenting with lagged education does not solve the measurement error problem because we find that measurement errors in education are highly correlated over time.

capital is indeed one of facilitating adoption of technology from abroad and creation of appropriate domestic technologies rather than entering on its own as a factor of production." And Barro (1997, p. 19) observes, "On impact, an extra year of male upper-level schooling is therefore estimated to raise the growth rate by a substantial 1.2 percentage points per year." Topel (1999), however, argues that "the magnitude of the effect of education on growth is vastly too large to be interpreted as a causal force." Indeed, Topel calculates that the present value of a one percentage point faster growth rate from an additional year of schooling would be about four times the cost, with a 5 percent real discount rate. He concludes that externalities from schooling may exist, but they are unlikely to be so large. One possibility -- which we explore and end up rejecting -- is that level of schooling is spuriously reflecting the effect of the change in schooling on growth.

Countries with higher initial levels of schooling tended to have larger increases in schooling over the next 10 or 20 years in Barro and Lee's data, which is remarkable given that measurement error in schooling will induce a negative covariance between the change and initial level of schooling. We initially suspected that the base level of schooling spuriously picks up the effect of schooling increases, either because schooling changes are excluded from the growth equation or because the included variable is noisy. The following calculations make clear that this is unlikely, however.

To proceed, it is convenient to write the cross-country growth equation as:

$$(7) \Delta Y = \beta_0 + \beta_1 S^*_{t-1} + \beta_2 S^*_t + \epsilon$$

where asterisks signify the correctly measured initial and ending schooling variables, and we have

suppressed the country j subscript.²⁵ We have also ignored covariates, but they could easily be "pre-regressed out" in what follows. If all that matters for growth is the change in schooling, we would find $\beta_1 = -\beta_2$. A test of whether the initial level of schooling has an independent, positive effect on growth *conditional on the change in schooling* turns on whether $\beta_1 + \beta_2 > 0$.

In practice, equation (7) is estimated with noisy measures of schooling that have serially correlated errors, as previously documented. Under the assumption of serially correlated but otherwise classical measurement errors, it can be shown that the limit of the coefficient on initial schooling is:

$$(8) \quad \text{plim } b_1 = \frac{R_{S_{t-1}} - \lambda r^2}{1 - r^2} \beta_1 + \frac{\lambda - R_{S_t}}{1 - r^2} \times \frac{\text{cov}(S_{t-1}, S_t)}{\text{var}(S_{t-1})} \beta_2$$

where $R_{S_{t-1}}$ and R_{S_t} are the reliability ratios for S_{t-1} and S_t , r is the correlation between S_{t-1} and S_t , and $\lambda = \text{cov}(S_{t-1}^*, S_t^*) / \text{cov}(S_{t-1}, S_t)$ is less than one if the measurement errors are positively correlated. An analogous equation holds for b_2 . Some algebra establishes that the sum $b_1 + b_2$ has probability limit:

$$(9) \quad \text{plim } (b_1 + b_2) = \beta_1 \frac{R_{S_{t-1}}(1 - \psi r) + \lambda r(\psi - r)}{1 - r^2} + \beta_2 \frac{R_{S_t}(1 - \psi^{-1} r) + \lambda r(\psi^{-1} - r)}{1 - r^2}$$

where $\psi = \text{var}(S_t) / \text{var}(S_{t-1})$.

Notice that if the variance in the measurement errors and the variance in true schooling are constant, then:

²⁵Notice that the scaling differs here from that in Table 1 and 3: namely, we do not divide any explanatory variable by the number of years in the period in Table 4.

$$(10) \quad \text{plim} (b_1 + b_2) = (\beta_1 + \beta_2) \frac{R_S + \lambda r}{1+r}$$

where R_S is the time-invariant reliability ratio of the schooling data.²⁶ Since $(R_S + \lambda r)/(1+r)$ is bounded by zero and one, in this case the sum of the coefficients is necessarily attenuated toward zero, so we would underestimate the effect of the initial level of education. Hence, measurement error in schooling is unlikely to drive the significance of the initial effect of education.

Table 4 presents estimates of equation (7) over 5, 10 and 20 year periods. The bottom of the table reports $b_1 + b_2$, as well as the measurement-error-corrected estimate of $b_1 + b_2$. We estimated the numerator of λ using the covariance between the 1990 WVS data and lagged Barro-Lee data (either 5, 10 or 20 year lags), and we estimated R_S from the WVS data as well.²⁷ Over each time interval, the results indicate that the negative coefficient on initial education is not as large in magnitude as the positive coefficient on second-period education, consistent with our earlier finding that the initial level has a positive effect on growth conditional on the change in education. Moreover, the correction for measurement error tends to raise $b_1 + b_2$ by .0004 to .0021 log points.

Finally, as an alternative approach to the measurement error problem, in columns 7 and

²⁶Griliches (1986) derives the corresponding formula if measurement errors are serially uncorrelated.

²⁷In models that include initial GDP, we first remove the effect of initial GDP before calculating $R_{S_{t-1}}$, R_{S_t} and λ . In the models without initial GDP, we assume $R_{S_{t-1}} = R_{S_t}$.

8 we use Kyriacou's schooling measures as instruments for Barro and Lee's schooling data.²⁸ If the measurement errors in the two data sets are uncorrelated, one set of measures can be used as an instrument for the other. Although the IV model can only be estimated for a subset of countries, these results also suggest that measurement error in schooling is not responsible for the positive effect of the initial level of schooling on economic growth.²⁹ Moreover, if Barro and Lee's data are used to instrument for Kyriacou's data in this equation, the sum of the schooling coefficients in column (8) nearly doubles.

3.5 Measurement Error in GDP

Another possibility is that transitory measurement errors in GDP explain why initial schooling matters in the growth equation. Intuitively, this would work as follows: If a country has a low level of education for its measured GDP, it is likely that its true GDP is less than its measured GDP. If the error in GDP is transitory, then subsequent GDP growth will appear particularly strong for such a country because the negative error in the GDP is unlikely to repeat in the second period. One indication that this may contribute to the strong effect of the level of education comes from including second period GDP instead of initial GDP in the growth equation. In this situation, measurement errors in GDP would be expected to have the opposite effect on the initial level of education. And indeed, if second period GDP is included instead of initial GDP

²⁸Sometimes the growth literature uses lagged variables as instruments. Since the measurement errors are highly serially correlated in education (and other variables), this is unlikely to completely solve measurement error problems.

²⁹If the model in column (8) is estimated by OLS with the subsample of 67 countries, the results are virtually identical to those for the full sample in column (6).

in the model in column (7) of Table 3, the coefficient on initial education becomes negative and statistically insignificant.

For two reasons, however, we conclude that measurement error in GDP is unlikely to drive the significant effect of the initial schooling variable. First, in Table 3 and Table 4 it is clear that the initial level of education has a significant effect even when initial GDP is not held constant. Second, using the WVS, we calculated the reliability of the Summers and Heston GDP data for 1990. Specifically, to estimate the reliability of log GDP, R_Y , we regressed the log of real income per person in the WVS on the log of real GDP per capita in the Summers and Heston data. The resulting coefficient was .92 (t-ratio=12.3), indicating substantial signal. Both measures were deflated by the same PPP measure in these calculations, which may inflate the reliability estimate, but if we add log PPP as an additional explanatory variable to the regression the reliability of the GDP data is .89 (t-ratio=11.9). Although the WVS income data neglect non-household income and these estimates are based on just 17 countries, the results indicate that Summers and Heston's data convey a fair amount of signal, and that the errors in GDP are highly serially correlated. If we assume that R_Y is .92 and the serial correlation in the errors is .5, the coefficient on initial education in the 10-year GDP growth equation would be biased upward by about a third.³⁰

³⁰With constant variances, the limit of the coefficient on initial log GDP is $R_Y\beta - (1-R_Y)(1-r)$, where R_Y is the reliability of log GDP, β is the population regression coefficient with correctly-measured GDP, and r is the serial correlation in the measurement errors. To estimate the effect of measurement error in GDP on the schooling coefficient, we constrained the coefficient on initial GDP to equal $\{b + (1-R_Y)(1-r)\}/R_Y$, where b is the coefficient on initial GDP obtained by OLS without correcting for measurement error, and re-estimated the growth equation.

3.6 *The Effect of Physical Capital*

The level and growth rate of capital are natural control variables to include in the GDP growth regressions. First, initial log GDP can be substituted for capital in a Solow growth model only if capital's share is constant over time and across countries (e.g., a Cobb-Douglas production function). Second, and more importantly for our purposes, capital-skill complementarity would imply that some of the increased output attributed to higher education in Table 3 should be attributed to increased capital (see, e.g., Goldin and Katz, 1997 on capital-skill complementarity). As mentioned earlier, however, systematic correlation between measurement errors in capital and GDP, as well as endogeneity of capital, are reasons to be wary about including the growth of capital in a GDP equation. Here we examine the robustness of our results to controlling for physical capital.

Column (1) of Table 5 reports an estimate of the same 20-year growth model as in column 9 of Table 3, augmented to include the growth of capital per worker. We use Klenow and Rodriguez-Clare's (1997) capital data because they appear to have more signal than Benhabib and Spiegel's capital data.³¹ The coefficient on the change in education falls by more than 50 percent when capital growth is included, although it remains barely statistically significant at the .10 level. In column (2) we add the initial log capital per worker, and in column (3) exclude the initial log capital from the column (2) specification. Including initial log capital drives the coefficient on the change in schooling to close to zero. Notice also that the log of initial capital per worker has little

³¹A regression of Benhabib and Spiegel's change in log capital on the corresponding variable from Klenow and Rodriguez-Clare yields a regression coefficient (and standard error) of .95 (.065). The reverse regression yields a coefficient of .69 (.05). Hence, Klenow and Rodriguez-Clare's measure appears to have a high signal-to-noise ratio.

effect in columns (2) and (3).³² The growth of capital per worker has an enormous effect on GDP growth. With Cobb-Douglas technology and competitive factor markets, the coefficient on the growth in capital in Table 5 would equal capital's share; instead, the coefficient is at least double capital's share in most countries (see Gollin, 1998). This finding is consistent with the errors in capital being systematically related to GDP, since both are functions of investment. To explore the sensitivity of the results, in column (4) we constrain the coefficient on the growth in capital to equal 0.35, which is on the high end of the distribution of non-labor's share around the world. These results indicate that both the change and initial level of schooling are associated with economic growth. Moreover, the coefficient on the change in education is similar to microeconomic estimates.

As mentioned earlier, including capital could exacerbate the measurement error in schooling. Indeed, we find that the reliability of Barro-Lee's 20-year change in schooling data falls from .58 to .46 once we condition on the change in capital, suggesting that the coefficient on the change in schooling in columns 1-3 of Table 5 should be roughly doubled.³³ In column (5), to try to overcome measurement error we estimate the growth equation by instrumental variables, using Kyriacou's schooling data as excluded instruments for the change and level of schooling. This is the same estimation strategy previously used by Pritchett (1998), but we

³²If the change in log capital per work is dropped from the model in column (3), then initial log capital per worker does have a statistically significant, negative effect, and the schooling coefficients are similar to those in column 9 of Table 3.

³³Temple (1999b) finds that eliminating observations with large residuals causes the coefficient on the growth in education in Benhabib and Spiegel's data to rise and become statistically significant, conditional on the growth in capital. We find a similar result with Benhabib and Spiegel's data, although similarly eliminating outliers has little effect on the results in Table 5 which use the Barro and Lee education data.

employ different schooling data as instruments, and use a different measure of capital. Unfortunately, because there is so little signal in education conditional on capital, the IV results yield a huge standard error (.167). Pritchett similarly finds a large standard errors from his IV estimates, although his point estimates are negative.³⁴ One final point on these estimates is that, to be comparable to the Mincerian return to schooling, the coefficient on the change in education should be divided by labor's share if the aggregate production function is Cobb-Douglas and human capital is an exponential function of schooling. This would raise the cross-country estimate of the return to schooling even further.

We draw four main lessons from this investigation of the role of capital. First, the change in capital has an enormous effect in a GDP growth equation, probably because of a mechanical relationship between the errors in measuring capital and GDP or reverse causality. Second, the impact of both the level and change in schooling on economic growth is sensitive to whether the change in capital is included in the growth equation and allowed to have a coefficient that greatly exceeds capital's share. Third, controlling for capital exacerbates measurement error problems in schooling. Instrumental variables estimates designed to correct for measurement error in schooling yield such a large standard error on the change in schooling that the results are consistent with schooling changes having no effect on growth or a large effect on growth. Fourth, when the coefficient on capital growth is constrained to equal a plausible value, changes in years of schooling are positively related to economic growth. Overall, unless measurement error problems in schooling and capital are overcome, we doubt the cross-country growth equations that

³⁴ Aside from the different data sources, the difference between our IV results and Pritchett's appears to result from his use of log schooling changes. If we use log schooling changes, we also find negative point estimates.

control for capital growth will be very informative insofar as the benefit of education is concerned.

4. Less Restrictive Macro Growth Model

The macro growth equations impose the restriction that all countries have the same relationship between growth and initial education, and that the relationship is linear. The first assumption is particularly worrisome because the micro evidence indicates that the return to schooling varies considerably across countries, and even across regions within countries. For example, institutional factors that compress the wage structure in some countries result in lower returns to schooling in those countries (see, e.g., the essays in Freeman and Katz, 1995). Moreover, one might expect externalities from education to be greater in countries where the private return is depressed below the world market level. Thus, we first allow the effect of the stock of education on growth to vary by country. Next, we examine the effect of the linearity assumption. Both of these extensions to the standard growth specification suggest that the constrained specification estimated in the literature should be viewed with caution; qualitatively different results are obtained when the constraints are relaxed.

4.1 Heterogeneous Country Effects of Education

Consider the following variable-coefficient version of the macro growth equation:

$$(11) \Delta Y_j = \beta_0 + \beta_{1j} S_{j,t-1} + \epsilon_j$$

where we allow each country to have a separate schooling coefficient (β_{1j}) and ignore other covariates. If there is more than one observation per country, equation (11) can be estimated by

interacting education with a set of dummy variables indicating each country. It is instructive to note that the coefficient on education (β_1) estimated from the restricted OLS model with a homogenous education slope can be written as a weighted average of the country-specific slopes:

$$(12) \quad \beta_1 = \sum_j w_j \beta_{1j} = \sum_{j,t-1} \frac{(S_{j,t-1} - \bar{S})^2}{\sum_{j,t-1} (S_{j,t-1} - \bar{S})^2} \beta_{1j}$$

where the weights are the country-specific contributions to the variance in schooling. Of course, if the assumptions of the constant-coefficient model hold (and the other Gauss-Markov assumptions hold), the w_j weights are the most efficient weights. But if a variable-coefficient model is more appropriate, there is no *a priori* reason to prefer the OLS weights over other weights. Indeed, it is rather odd to weight the country-specific slopes by the OLS weights if the slopes differ across countries. Probably a more relevant single estimate would be the unweighted-average coefficient, which represents the expected value of the education coefficient for countries in the world. Lastly, notice that if equation (11) is augmented to include covariates the simple weighted average interpretation of the constant-coefficient model in (12) does not apply, but the average of the country-specific coefficients is still informative.³⁵

Table 6 summarizes estimates of a variable-coefficient model using 5-year and 10-year changes in GDP. Panel A reports results of regressing GDP growth on average years of schooling for the population age 25 and older, initial GDP and time dummies. Columns 1 and 3 report the

³⁵If country fixed effects are included as covariates in equation (11), however, the OLS constant coefficient can still be decomposed as a weighted average of the country-specific coefficients even if there are other covariates. But we exclude country fixed effects so that these estimates are comparable to the earlier ones, and because including fixed effects would exacerbate measurement error bias.

constant-coefficient model, whereas columns 2 and 4 report the mean of the country-specific education coefficients. The constant education slope assumption is overwhelmingly rejected by the data for each time period (p -value < 0.0001). Indeed, the R^2 of the equations more than doubles when the education slopes are unconstrained. Of more consequence, the average slope coefficient is negative, though not statistically significant, in the variable-coefficient model. These results cast doubt on the interpretation of initial education in the constrained macro growth equation common in the literature.

Panel B of Table 6 reports results in which average years of secondary and higher schooling for males is used instead of average years of all education for the entire adult population. This variable has been emphasized as a key determinant of economic growth in Barro's work. Again, however, the results of the constant-coefficient model are qualitatively different than those of the variable-coefficient model. Indeed, for the average country in the sample, a greater initial level of secondary and higher education has a statistically significant, negative association with economic growth over the ensuing 10 years.

The estimates reported in Table 6 exclude the change in education to focus on the effect of the initial stock of education. We have, however, experimented with a variable-coefficient model for the 10-year change in education variable. These estimates were quite fragile. For example, if we regress annual GDP growth over ten years on a set of interactions between the 10-year change in education and country dummies, the initial log of GDP per capita, and time dummies, the average coefficient for the change in education interactions is negative. But if we also include a set of interactions between initial education and country dummies in this model, the average coefficient on the change in education swells to .18, and the average coefficient on initial

education is negative. Since the latter model uses three observations to estimate two parameters for each country, we are reluctant to stress these results.

It is clear from Table 6 that education has a heterogeneous effect on economic growth across countries. What bearing does this result have on the convergence literature? Lee, Pesaran and Smith (1998) show that country heterogeneity in technological progress that is assumed homogeneous across countries in a fixed-effects model with a lagged dependent variable will generate a spurious correlation between the lagged dependent variable and the error term. A similar result will follow if heterogeneous education coefficients are constrained to equal a constant coefficient, so we would regard the convergence coefficient with some caution since it depends on controlling for S_{t-1} . Nonetheless, it is worth emphasizing that we still obtain a negative average coefficient on education if we drop initial GDP from the variable-coefficient model. Because we are interested in understanding the role of education in economic growth, we do not pursue the convergence issue further, but we think the results of the variable-coefficient model reinforce Lee, Pesaran, and Smith's skeptical interpretation of the conventional estimate of the convergence parameter.

4.2 The Importance of Linearity

It is common in the empirical growth literature to assume that initial education has a linear effect on subsequent GDP growth. Although Mincer (1974) provides conditions under which education has a linear relationship with log earnings, these conditions do not necessarily imply that the level of initial education has a linear relationship with income growth. The impact of the linearity assumption has not been rigorously tested in growth models. As a first approach to test

this assumption, we created a set of dummy variables which indicated whether each country's initial average years of education fell in the 0-1 range, 1-2 range, 2-3 range, and so on. We then included these dummies in lieu of linear education in the growth regression in column (4) of Table 3. Figure 2 displays the dummy coefficients (i.e., the unrestricted relationship) and the estimated linear relationship between GDP growth and education. The education-GDP growth relationship appears upward sloping at low levels of education, then plateaus between 4 and 8 years of education, and becomes downward sloping above 8 years of education.

The increment to GDP growth at each level of average education is shown on the bottom of Figure 2 (i.e., change in the unrestricted relationship). Loosely speaking, the OLS regression coefficient estimate for education in the linear specification can be thought of as a weighted average of the slope increments at each level of education.³⁶ The OLS estimate places more weight on the slope increment at a particular schooling level if: (1) there is a greater gap in average years of schooling between countries above and below that schooling level; (2) the schooling level is close to the median schooling level. Figure 3A displays the weights applied to the increments by OLS for the linear specification in a model without other covariates. For comparison, Figures 3B and 3C show histograms of years of schooling for all countries and for OECD countries.

As in the previous subsection, the OLS weighting scheme is optimal if there is a constant slope at each schooling level. But if the relationship is not linear over the range of schooling, the OLS weights may distort the relationship, and other weighting schemes may be more appropriate.

³⁶This discussion is only approximate for two reasons. First, average years of education is not a categorical variable; we have transformed it to one. Second, if there are covariates, the relationship is more complicated. For more details, see Yitzhaki (1996) and Angrist and Krueger (1999).

For example, it seems natural to weight the increments by the distribution of schooling for countries in the world. Alternatively, we could weight by the OECD's distribution of education. The weighted slope is greater than the OLS slope if the distribution of education throughout the world is used to weight the increments in Figure 2, whereas the average slope becomes negative if the OECD distribution of education is used as weights.

An alternative, simpler approach to explore the impact of linearity is to estimate a nonlinear function of education. Specifically, we included initial education and its square in the model in column 4 of Table 3. The data seem to prefer a quadratic specification, as the square term is statistically significant.³⁷ More importantly, the relationship is inverted-U shaped, with a peak at 7.5 years of education. Since the mean education level for OECD countries in 1990 was 8.4 years in Barro and Lee's data, the average OECD country is on the downward-sloping segment of the education-growth profile. We similarly find an inverted-U shaped relationship between education and GDP growth which peaks below the level of education of developed countries when we examine 5 and 20 year changes in GDP, or male upper secondary schooling.³⁸ These findings underscore W. Arthur Lewis's (1964) observation that, "it is not possible to draw a simple straight line relating secondary education to economic growth."

³⁷See Krueger and Lindahl (1999).

³⁸Although these findings may appear surprising in light of the macro growth literature, they are consistent with results in Barro (1997; Table 1.1, column 1). In particular, the interaction between male upper-secondary education and log GDP has a negative effect on growth, and the results imply the effect of schooling on growth becomes negative for countries whose GDP exceeds the average by 1.9 log units.

5. *Conclusion*

The micro and macro literatures both emphasize the role of education in income growth. An accumulation of research using individual-level education and income data since the beginning of the 20th century provides robust evidence of a substantial payoff to investment in education, especially for those who traditionally complete low levels of schooling. From the micro evidence, it is unclear whether the social return to schooling exceeds the private return, although available evidence suggests that positive externalities in the form of reduced crime and reduced welfare participation are more likely to be reaped from investments in disadvantaged than advantaged groups (e.g., Heckman and Klenow, 1997). The macroeconomic evidence of externalities in terms of technological progress from investments in higher education seems to us more fragile, resulting from imposing constant-coefficient and linearity restrictions that are rejected by the data.

Our findings also help resolve an important inconsistency between the micro and macro literatures on education: Contrary to Benhabib and Spiegel's (1994) and Barro and Sala-i-Martin's (1995) conclusions, the cross-country regressions indicate that the change in education is positively associated with economic growth once measurement error in education is accounted for. Indeed, after adjusting for measurement error, the change in average years of schooling often has a greater effect in the cross-country regressions than in the within-country micro regressions. The larger return to schooling found in the cross-country models suggests that reverse causality or omitted variables create problems at the country level of analysis, or that increases in average educational attainment generate nationwide externalities. Although the microeconomic evidence in several countries suggests that within countries the causal effect of education on earnings can be estimated reasonably well by taking education as exogenous, it does not follow that cross-

country differences in education can be taken as a cause of income as opposed to a result of current income or anticipated income growth. Moreover, countries that improve their educational systems are likely to concurrently change other policies that enhance growth, possibly producing a different source of omitted-variable bias in cross-country analyses.

Education, as Harbison and Myers (1965) stress, "is both the seed and the flower of economic development." It is difficult to separate the causal effect of education from the positive income demand for education in cross-country data. Mankiw (1997) describes the presumed exogeneity of school enrollment as the "weak link" in the empirical growth literature. In our opinion, this link is unlikely to be strengthened unless researchers can identify natural experiments in schooling attainment similar to those that have been exploited in the microeconomic literature, and unless measurement errors in the cross-country data are explicitly taken into account in the econometric modelling. In view of the difficulties in obtaining accurate country-level data on changes in educational attainment, it might be preferable to examine growth across regions of countries with reliable data. Acemoglu and Angrist (1999), who look across U.S. states, and Moretti (1998), who looks across U.S. cities, provide good starts down this path, although they reach conflicting conclusions regarding any deviation between the social and private returns to education.

References

- Acemoglu, Daron and Joshua D. Angrist, "How Large are the Social Returns to Education? Evidence from Compulsory Schooling Laws," mimeo, MIT, September 17, 1999.
- Aghion, Phillippe, and Howitt, Peter. Endogenous Growth Theory, Cambridge, MA: MIT Press, 1998.
- Angrist, Joshua D., Guido W. Imbens and Alan B. Krueger, "Jackknife Instrumental Variables Estimation," Journal of Applied Econometrics, January-February 1999.
- Angrist, Joshua D. and Alan B. Krueger, "Does Compulsory School Attendance Affect Schooling and Earnings?," Quarterly Journal of Economics, 106(4), November 1991, pp. 979-1014.
- Angrist, Joshua D. and Alan B. Krueger, "Empirical Strategies in Labor Economics," in Orley Ashenfelter and David Card, eds., Handbook of Labor Economics, (Amsterdam: North Holland, 1999).
- Ashenfelter, Orley A., Colm Harmon and Hessel Oosterbeek, mimeo., "Empirical Estimation of the Schooling/Earnings Relationship – A Review," forthcoming Labour Economics, 1999.
- Ashenfelter, Orley A. and Alan B. Krueger, "Estimates of the Economic Return to Schooling from a New Sample of Twins," American Economic Review, 84(5), December 1994, pp. 1157-73.
- Barnett, W. Steven, "Benefits of Compensatory Preschool Education," Journal of Human Resources 27(2), Spring 1992, pp. 279-312.
- Barro, Robert J. and Xavier Sala-i-Martin, Economic Growth (New York: McGraw-Hill, 1995).
- Barro, Robert J., Determinants of Economic Growth: A Cross-Country Empirical Study, Lionel Robbins Lectures, (Cambridge, MA: MIT Press, 1997).
- Barro, Robert J. and Jong-Wha Lee, "International Comparisons of Educational Attainment," Journal of Monetary Economics, 32(3), December 1993, pp. 363-394.
- Behrman, Jere and Mark Rosensweig, "Caveat Emptor: Cross-country Data on Education and the Labor Force," Journal of Development Economics 44, June 1994, pp. 147-72.
- Behrman, Jere and Mark Rosensweig, "Adult Schooling Stocks: Comparisons Among Aggregate Data Series," mimeo., University of Pennsylvania, July 19, 1993.

Benhabib, J. and Spiegel, M., "The Role of Human Capital in Economic Development: Evidence from Aggregate Cross-Country Data", Journal of Monetary Economics, 34(2), October 1994, pp.143-74.

Bils, Mark and Peter J. Klenow, "Does Schooling Cause Growth or the Other Way Around?" NBER Working Paper 6393, 1998.

Blaug, Mark, Richard Layard and Maureen Woodhall, The Causes of Graduate Unemployment in India (London: Allen Lane, Penguin Press, 1969).

Blomström, M., R. Lipsey, and M. Zejan, "Is Fixed Investment the Key to Economic Growth," NBER Working Paper No. 4436, Cambridge, MA, 1993.

Bound, John, David Jaeger and Regina Baker, "Problems with Instrumental Variables Estimation When the Correlation Between the Instruments and the Endogenous Explanatory Variable is Weak," Journal of the American statistical Association, 90(430), June 1995, pp. 443-50.

Card, David, "Using Geographic Variation in College Proximity to Estimate the Returns to Schooling," in L. Christofides, E. Grant, and R. Swidinsky, eds., Aspects of Labor Market Behaviour: Essays in Honour of John Vanderkamp (Toronto: University of Toronto Press, 1995a), pp. 201-222.

Card, David, "Earnings, Schooling and Ability Revisited," in Solomon W. Polachek, ed., Research in Labor Economics. (Greenwich, CT: JAI Press, 1995b).

Card, David, "The Causal Effect of Schooling on Earnings," Orley Ashenfelter and David Card, eds., Handbook of Labor Economics (Amsterdam: North Holland, 1999).

Card, David and Alan B. Krueger, "Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States", Journal of Political economy, 100(1), February 1992, pp. 1-40.

Chamberlain, Gary and Guido Imbens, "Hierarchical Bayes Models with Many Instrumental Variables," mimeo., Cambridge, MA, Harvard University, August 1996.

Dale, Stacy and Alan Krueger, "The Payoff to Attending a More Selective College: An Application of Selection on Observables and Unobservables," Princeton University, Industrial Relation Section, Working Paper No. 409, 1998.

Donald, Steven and Whitney Newey, "Choosing the Number of Instruments," MIT, mimeo., 1997.

Duflo, Esther, "Evaluating the Schooling and Labor Market Consequences of a School

- Construction Program: An Analysis of the Indonesian experience," mimeo., MIT, February 1998.
- Entwisle, Doris, Karl Alexander, Linda Olson, Children, Schools and Inequality (Boulder, CO: Westview Press, 1997).
- Freeman, Richard and Lawrence Katz, eds., Differences and changes in wage structures (Chicago and London: University of Chicago Press, 1995), pp. 25-66.
- Goldin, Claudia and Lawrence Katz, "The Origins of Technology-Skill Complementarity," NBER Working Paper No. 5657, National Bureau of Economic Research, Cambridge, MA.
- Gollin, Douglas, "Getting Income Shares Right: Self Employment, Unincorporated Enterprise, And the Cobb-Douglas Hypothesis," mimeo., Williams College, June 1998.
- Gorseline, Donald E., The Effect of Schooling Upon Income (Bloomington, IN: University of Indiana Press, 1932).
- Griliches, Zvi, "Estimating the Returns to Schooling: Some Econometric Problems," Econometrica, 45(1), January 1977, pp. 1-22.
- Griliches, Zvi, "Economic Data Issues," in Zvi Griliches and Michael D. Intriligator, eds., Handbook of Econometrics (Amsterdam: North-Holland, 1986).
- Griliches, Zvi and William M. Mason "Education, Income, and Ability," Journal of Political Economy, 80(3), Part II, May/June 1972, pp. S74-S103.
- Hall, Robert and Charles I. Jones, "Why Do Some Countries Produce So Much More Output per Worker than Others?" Forthcoming, Quarterly Journal of Economics, 1998.
- Harberger, Arnold, "Investment in Men Versus Investment in Machines: The Case of India," in C. Arnold Anderson and Mary Jean Bowman, eds., Education and Economic Development (Chicago: Aldine Publishing Co., 1965), pp. 11-50.
- Harbison, Frederick, and Charles Myers, eds., Manpower and Education (New York: McGraw-Hill Book Company, 1965), p. xi.
- Harmon, Colm and Ian Walker, "Estimates of the Economic return to Schooling for the United Kingdom," American Economic Review, 85(5), December 1995, pp. 1278-86.
- Heckman, James and Peter Klenow, "Human Capital Policy," mimeo., University of Chicago, December 1997.
- Inglehart, Ronald, Codebook for World Values Survey (Ann Arbor, MI: Institute for Social

Research 1994).

Klenow, Peter, and Andrés Rodríguez-Clare, "The Neoclassical Revival in Growth Economics: Has it Gone Too Far?" In Ben Bernanke and Julio Rotemberg, eds., NBER Macroeconomics Annual (Cambridge, MA: MIT Press, 1997), pp. 73-103.

Krueger, Alan B., "Experimental Estimates of Education Production Functions," Quarterly Journal of Economics, May 1999.

Krueger, Alan B. and Mikael Lindahl, "Education for Growth in Sweden and the World," Swedish Economic Policy Review, Autumn 1999, vol. 6, pp. 289-339.

Krueger, Alan B. and Jorn-Steffen Pischke "A Comparative Analysis of East and West German Labor Markets: Before and After Unification," in Richard Freeman and Lawrence Katz, eds., Differences and changes in wage structures. (Chicago and London: University of Chicago Press, 1995), pp. 405-45.

Kyriacou, G., "Level and Growth Effects of Human Capital", Working Paper, C.V. Starr Center, New York University. 1991.

Lang, Kevin "Ability Bias, Discount Rate Bias and the Return to Education," mimeo., Boston University, Department of Economics, 1993.

Lee, Kevin, M. Hashem Pesaran, and Ron Smith, "Growth Empirics: A Panel Data Approach -- A Comment," Quarterly Journal of Economics, 113 (1), February 1998, pp. 319-324.

Levy, Frank and Richard Murnane, "Earnings Levels and Earnings Inequality: A Review of Recent Trends and Proposed Explanations," Journal of Economic Literature (30) (3), September 1992, pp. 1333-81.

Lewis, W. Arthur, "Secondary Education and Economic Structure," Social and Economic Studies, 10(2), 1964, pp. 219-232. Reprinted in Mark Gersovitz, ed., Selected Economic Writings of W. Arthur Lewis (New York: New York University Press, 1983), pp. 509-522.

Lucas, Robert, "On the Mechanics of Economic Development," Journal of Monetary Economics 22(1), July 1988, pp. 3-42.

Machlup, Fritz, Education and Economic Growth (Lincoln, NE: University of Nebraska Press, 1970).

Mankiw, N.G., "Comment," in Ben Bernanke and Julio Rotemberg, eds., NBER Macroeconomics Annual (Cambridge, MA: MIT Press, 1997), pp. 103-106.

Mankiw, N.G., David Romer and D.M. Weil, "A Contribution to the Empirics of Economic Growth," Quarterly Journal of Economics, 107(2), May 1992, pp. 407-437.

Miller, Herman P., Income of the American People (Washington, DC: United States Government Printing Office, 1955).

Mincer, Jacob, Schooling, Earnings, and Experience (New York: Columbia University Press, 1974).

Moretti, Enrico, "Social returns to Education and Human Capital Externalities: Evidence from Cities," mimeo, UC Berkeley, December 1998.

Murphy, Kevin M and Finis Welch, "Empirical Age-Earnings Profiles," Journal of Labor Economics, 8(2), April 1990, pp. 202-229.

Nelson, R. and E. Phelps, "Investment in Humans, Technological Diffusion, and Economic Growth," American Economic Review, 56(), 1966, pp. 69-75.

O'Neill, Donal. "Education and Income Growth: Implications for Cross-Country Inequality," Journal of Political Economy, 103, no 6, December 1995, pp. 1289-1301.

Park, Jin Huem. "Returns to Schooling: A Peculiar Deviation from Linearity," Princeton University, Industrial Relations Section Working Paper No. 335, October 1994.

Pritchett, Lant. "Where Has All the Education Gone?" Policy Research Working Paper No. 1581, The World Bank, Washington, D.C., June 1997.

Psacharopoulos, G., "Returns to Investment in Education: A Global Update", World Development, 22(9), September 1994, pp.1325-43.

Rivkin, Steven Eric Hanushek, and John Kain, "Teachers, Schools and Academic Achievement," mimeo., University of Rochester, July 1998.

Romer, Paul, "Endogenous Technological Change," Journal of Political Economy, 89 (5), Part 2, October 1990, pp. S71-S102.

Rosen, Sherwin, "Human Capital: A Survey of Empirical Research," in R. Ehrenberg, ed., Research in Labor Economics, (Greenwich, Conn, JAI Press, vol. 1. 1977).

Smith, James. "Healthy Bodies and Thick Wallets: The Dual Relation Between Health and SES," Journal of Economic Perspectives, 1999.

Spence, A. Michael, "Job Market Signaling," Quarterly Journal of Economics, 87(3), August

1973, pp. 355-74.

Staiger, Douglas and James H. Stock, "Instrumental Variables Regressions with Weak Instruments," Econometrica, 65(3), May 1997, pp. 557-86.

Summers, R. and Alan Heston, "The Penn World Table (Mark 5): An Expanded Set of International Comparisons, 1950-1988," Quarterly Journal of Economics, 106(2), May 1991, pp. 327-368.

Summers, Anita A. and Barbara L. Wolfe, "Do Schools Make a Difference?" American Economic Review, 67(4) Sept. 1977, pp. 639-52.

Temple, Jonathan, "The New growth Evidence," Journal of Economic Literature, Vol. XXXVII, March 1999a, pp. 112-156.

Temple, Jonathan, "A Positive Effect of Human Capital on Growth," Economic Letters, October 1999b, 65(1), pp. 131-34.

Topel, Robert, "Labor Markets and Economic Growth," in Orley Ashenfelter and David Card, eds., Handbook of Labor Economics, (Amsterdam: North Holland, 1999).

UNESCO, Statistical Yearbook, 1967 (Paris: UNESCO, 1967).

Walsh, J.R. "Capital Concept Applied to Man," Quarterly Journal of Economics, XLIX, February 1935, pp. 255-85.

Welch, F., "Education in Production," Journal of Political Economy, 78(1), Jan.-Feb. 1970, pp. 35-59.

Willis, Robert J., "Wage Determinants: A Survey and Reinterpretation of Human Capital Earnings Functions," in Ashenfelter, Orley A. and Richard Layard, eds. Handbook of Labor Economics, Vol. 1 (Amsterdam: North Holland, 1986).

Wolfe, Dael and Joseph G. Smith, "The Occupational Value of Education for Superior High-School Graduates," Journal of Higher Education, 27(4), April 1956, 201-232.

Yitzhaki, Shlomo, "On Using Linear Regressions in Welfare Economics," Journal of Business and Economics Statistics, 14, 1996, pp. 478-86.

Data Appendix

The second wave of the World Values Survey (WVS) was conducted in 42 countries between 1990 and 1993. The sampled countries represented almost 70% of the world population, including several countries where micro data normally are unavailable. The survey was designed by the World Values Study Group (1994), and conducted by local survey organizations (mainly Gallup) in each of the surveyed countries. In most countries, a national random sample of adults (over age 18) was surveyed. For 12 of the countries in our sample (Belgium, Brazil, Canada, China, India, Italy, Netherlands, Portugal, Spain, Switzerland, West Germany and U.K.), sampling weights were available to make the survey representative of the country's population; the other samples are self-weighting. A feature of the survey is that the questionnaire was designed to be similar in all countries to facilitate comparisons across countries. There are, however, drawbacks to using the WVS for our purposes. The primary purpose of the WVS was to compare values and norms across different societies. Although questions about income and education were included, they appear to have been a lower priority than the normative questions. For example, family income was collected as a categorical variable in ten ranges, and some countries failed to report the currency values associated with the ranges. We were able to derive comparable data from the WVS on mean years of schooling for 34 countries and on mean income for 17 countries.

Mean years of schooling is calculated from question V356 in the WVS, which asked, "At what age did you or will you complete your formal education, either at school or at an institution of higher education? Please exclude apprenticeships." The variable is typically bottom coded at 12 years of age and top coded at 21 years of age. Although there are some benefits of formulating the question this way, for our purposes it also creates some problems. First, we do not know the age at which respondents started their education. For this reason we have used data from UNESCO (1967) on the typical school starting age in each country. Second, the top and bottom coding is potentially a problem. For almost one third of the countries (Austria, Brazil, Denmark, India, Norway, Poland, South Korea, Sweden, Switzerland and Turkey), however, a question was asked concerning formal educational attainment. Since, as mentioned above, one of the benefits with the WVS is that the same questions are asked in all the countries, we used this variable only to solve the bottom and top coding problem.¹ We have coded illiterate/no schooling as 0 years of schooling and incomplete primary schooling as 3 years. In the two countries where graduate studies is a separate category, we have set this to 19. For the countries in which the educational attainment variable does not exist, we set years of schooling for those in the bottom-

¹ For South Korea and Switzerland, however, we exclusively used this variable to derive years of schooling because the question about school leaving age is not asked in these countries. For Turkey, school leaving age is only coded as three possible ages, so we use both the educational attainment and school leaving age variable to derive years of schooling.

coded category equal to the midpoint of 0 and the bottom coded years of schooling.² Similarly, we set years of schooling for the highest category equal to the midpoint of 18 and the top coded years of schooling.

As mentioned, the family income variable in WVS is reported in 10 categories. We coded income as the midpoints of the range in each category. This variable is also censored from below and above. For simplicity, we set income for those who were bottom coded at the midpoint between zero and the lower income limit. We handled top coding by fitting a Pareto distribution to family income above each country's median income. Assuming that this distribution correctly characterizes the highest category, we calculated the mean of the censored distribution. We converted the family income variable to a dollar-value equivalent by multiplying the family income variable in each country by the ratio of the purchasing power parity in dollars to the corresponding local currency exchange, using Summers and Heston's (1991) data.

The logarithm of mean income per capita was calculated as the logarithm of the sum of family income in common currency divided by the total number of individuals in all households in the sample. The total number of individuals in each household is calculated as the sum of the number of children living at home and the number of adults present. (Two adults were assumed to be present if the respondent was married; otherwise, one adult was assumed to be present.) Appendix Table A2 reports the weighted mean years of schooling and log income per capita derived from the WVS. The weights for these calculations were the sampling weights reported in the WVS. The sample size used to calculate mean schooling is also reported.

²For East and West Germany the bottom code was 14, and for Finland it was 15. Because school start age was 7 in Finland and East Germany, and 6 in West Germany, we set years of schooling equal to 6 in West Germany and Finland and 5 in East Germany for those who were bottom coded.

Appendix Table A1: Correlation and Covariance Matrices for Barro-Lee and Kyriacou Years of Schooling Data

A. Correlation Matrix

	S_{65}^{BL}	S_{85}^{BL}	S_{65}^K	S_{85}^K	ΔS^{BL}	ΔS^K
S_{65}^{BL}	1.00					
S_{85}^{BL}	0.97	1.00				
S_{65}^K	0.91	0.92	1.00			
S_{85}^K	0.81	0.86	0.88	1.00		
ΔS^{BL}	0.23	0.46	0.36	0.51	1.00	
ΔS^K	-0.12	-0.03	-0.17	0.33	0.34	1.00

B. Covariance Matrix

	S_{65}^{BL}	S_{85}^{BL}	S_{65}^K	S_{85}^K	ΔS^{BL}	ΔS^K
S_{65}^{BL}	6.65					
S_{85}^{BL}	7.07	8.01				
S_{65}^K	5.66	6.29	5.88			
S_{85}^K	5.27	6.19	5.38	6.41		
ΔS^{BL}	0.42	0.93	0.62	0.92	0.51	
ΔS^K	-0.39	-0.10	-0.50	1.02	0.30	1.52

Notes: Sample size is 68. A superscript BL refers to the Barro-Lee data and a superscript K refers to the Kyriacou data. The subscript indicates the year. Unlike the other tables, the change in schooling is not annualized.

Appendix Table A2: International Data Derived from the World Values Survey

Country	Mean years of schooling (Std. deviation)	Log family income per capita	Survey Year	Sample Size
Argentina	10.23 (4.88)	--	1991	766
Austria	8.69 (4.88)	8.78	1990	1,296
Belgium	11.53 (3.29)	8.92	1990	2,328
Bulgaria	11.29 (3.83)	--	1990	877
Brazil	4.04 (3.04)	--	1991-92	1,154
Canada	12.60 (3.19)	9.48	1990	1,483
Czechoslovakia	11.78 (2.86)	--	1990	1,190
Chile	10.48 (4.37)	7.81	1990	1,137
China	10.32 (3.51)	--	1990	745
Denmark	12.50 (3.66)	--	1990	862
East Germany	9.12 (3.90)	--	1990	1,175
Finland	12.61 (3.81)	9.00	1990	534
France	11.12 (3.62)	--	1990	830
Hungary	9.79 (3.57)	7.91	1990	895
Iceland	12.16 (3.74)	--	1990	575
Ireland	10.20 (2.81)	--	1990	847
India	2.97 (4.48)	7.07	1990	1,908
Italy	7.88 (4.90)	9.02	1990	1,616
Japan	12.29 (2.85)	9.10	1990	855
Mexico	8.44 (5.47)	--	1990	835
Netherlands	11.89 (3.65)	9.17	1990	876
Norway	13.43 (4.46)	9.20	1990	1,063
Poland	10.11 (3.56)	--	1989	803
Portugal	6.12 (4.79)	8.41	1990	823
Romania	10.50 (4.36)	--	1993	933
Russia	12.35 (3.67)	--	1991	1,551
Spain	8.61 (4.49)	8.24	1990	2,991
South Korea	12.00 (3.58)	--	1990	1,040
Sweden	12.79 (3.40)	--	1990	848
Switzerland	8.63 (2.61)	--	1988-89	1,154
Turkey	6.13 (4.65)	8.09	1990-91	805
U.K.(excl. N.I.)	11.20 (2.50)	9.17	1990	1,288
USA	13.26 (2.96)	9.49	1990	1,477
West Germany	9.78 (3.34)	9.19	1990	1,770

Note: Sample size pertains to the number of observations used to calculate years of schooling.

Table 1: Replication and Extension of Benhabib and Spiegel (1994)
Dependent Variable: Annualized Change in Log GDP, 1965-85

Variable	<u>Log Schooling</u>			<u>Linear Schooling</u>		
	(1)	(2)	(3)	(4)	(5)	(6)
$\Delta \text{Log S}$	-.072 (.058)	.178 (.112)	.614 (.162)	---	---	---
Log S_{65}	---	.010 (.004)	.026 (.005)	---	---	---
ΔS	---	---	---	.012 (.023)	.039 (.024)	.151 (.034)
S_{65}	---	---	---	---	.003 (.001)	.004 (.001)
Log Y_{65}	-.009 (.002)	-.012 (.002)	-.015 (.003)	-.008 (.002)	-.014 (.002)	-.014 (.004)
$\Delta \text{Log Capital}$.523 (.048)	.461 (.052)	---	.521 (.051)	.465 (.052)	---
$\Delta \text{Log Work Force}$.175 (.164)	.232 (.160)	---	.110 (.160)	.335 (.167)	---
R^2	.694	.720	.291	.688	.726	.271

Notes: All change variables were divided by 20, including the dependent variable. Sample size is 78 countries. Standard errors are in parentheses. All equations also include an intercept. S_{65} is Kyriacou's measure of schooling in 1965; $\Delta \text{Log S}$ is the change in log schooling between 1965 and 1985, divided by 20; and Y_{65} is GDP per capita in 1965. Mean of the dependent variable is .039; standard deviation of dependent variable is .020.

Table 2. Reliability of Various Measures of Years of Schooling

A. Estimated Reliability Ratios for Barro-Lee and Kyriacou Data

	Reliability of <u>Barro-Lee Data</u>	Reliability of <u>Kyriacou Data</u>
Average years of schooling, 1965	.851 (.049)	.964 (.055)
Average years of schooling, 1985	.773 (.055)	.966 (.069)
Change in years of schooling, 1965-85	.577 (.199)	.195 (.067)

B. Estimated Reliability Ratios for Barro-Lee and World Values Survey Data

	Reliability of <u>Barro-Lee Data</u>	Reliability <u>of WVS Data</u>
Average years of schooling, 1990	.903 (.115)	.727 (.093)
Average years of secondary and higher Schooling, 1990	.719 (.167)	.512 (.119)

Notes: The estimated reliability ratios are the slope coefficients from a bivariate regression of one measure of schooling on the other. For example, the .851 entry in the first row is the slope coefficient from a regression in which the dependent variable is Kyriacou's schooling variable and the independent variable is Barro-Lee's schooling variable. The .964 ratio in the second column is estimated from the reverse regression. In panel B, the reliability ratios are estimated by comparing the Barro-Lee and WVS data. In the WVS data set, secondary and higher schooling is defined as years of schooling attained *after 8 years of schooling*.

Sample size for panel A is 68 countries. Sample size for panel B is 34 countries. Standard errors are reported in parentheses.

Table 3: The Effect of Schooling on Growth
Dependent Variable: Annualized Change in Log GDP per Capita

	5-year changes			10-year changes			20-year changes		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
S_{t-1}	.004 (.001)	---	.004 (.001)	.003 (.001)	---	.004 (.001)	.005 (.001)	---	.005 (.001)
ΔS	---	.031 (.015)	.039 (.014)	---	.075 (.026)	.086 (.024)	---	.184 (.057)	.182 (.051)
$\text{Log } Y_{t-1}$	-.005 (.003)	.004 (.002)	-.006 (.003)	-.003 (.003)	.004 (.001)	-.005 (.003)	-.010 (.003)	-.001 (.002)	-.013 (.003)
R^2	.197	.161	.207	.242	.229	.284	.184	.103	.281
N	607	607	607	292	292	292	97	97	97

Notes: First six columns include time dummies. Equations were estimated by OLS. The standard errors in the first six columns allow for correlated errors for the same country in different time periods. Maximum number of countries is 110. Columns 1-3 consist of changes for 1960-65, 1965-70, 1970-75, 1975-80, 1980-85, 1985-90. Columns 4-6 consist of changes for 1960-70, 1970-80, 1980-90. Columns 7-9 consist of changes for 1965-85. $\text{Log } Y_{t-1}$ and S_{t-1} are the log GDP per capita and level of schooling in the initial year of each period. ΔS is the change in schooling between $t-1$ and t divided by the number of years in the period. Data are from Summers and Heston and Barro and Lee. Mean (and standard deviation) of annualized per capita GDP growth is .021 (.033) for columns 1-3, .022 (.026) for columns 4-6, and .022 (.020) for columns 7-9.

Table 4: The Effect of Measurement Error on the Sum of Schooling Coefficients
Dependent Variable: Annualized Change Log GDP per Capita

	OLS						IV	
	5-year changes		10-year changes		20-year changes		20-year changes	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
S_{t-1}	-.004 (.003)	-.004 (.003)	-.005 (.002)	-.005 (.002)	-.005 (.003)	-.004 (.003)	-.020 (.010)	-.023 (.011)
S_t	.007 (.003)	.008 (.003)	.008 (.002)	.009 (.002)	.007 (.003)	.009 (.003)	.020 (.009)	.028 (.010)
Log Y_{t-1}	---	-.006 (.003)	---	-.005 (.003)	---	-.013 (.003)	---	-.020 (.006)
b_1+b_2	.0026 (.0005)	.0042 (.0009)	.0023 (.0005)	.0037 (.0009)	.0020 (.0008)	.0052 (.0011)	.0001 (.0015)	.0055 (.0024)
Meas. Error Corrected b_1+b_2	.0030	.0052	.0028	.0047	.0041	.0067	---	---
R^2	.197	.207	.272	.284	.159	.281	---	---
N	607	607	292	292	97	97	67	67

Notes: All regressions also include time dummies and an intercept. The standard errors in the first four columns allow for correlated errors within countries over time. The time periods covered are the same as in Table 4. In columns 7 and 8, Kyriacou's education data are used as instruments for Barro and Lee's education data. All other columns only use Barro and Lee's education data. See text for description of the measurement error correction. Mean (and standard deviation) of dependent variable are .021 (.033) for columns 1-2, .022 (.026) for columns 3-4, .022 (.020) for columns 5-6, and .019 (.019) for columns 7-8.

Table 5: The Effect of Schooling and Capital on Economic Growth
Dependent Variable: Annualized Change in Log GDP per Capita, 1965-85

	OLS				IV
	(1)	(2)	(3)	(4)	(5)
ΔS	.066 (.039)	.017 (.032)	.015 (.042)	.083 (.043)	.069 (.167)
S_{65}	.004 (.001)	.0013 (.0008)	.0005 (.0010)	.002 (.001)	-.001 (.002)
Log Y_{65}	-.009 (.003)	-.026 (.003)	---	---	---
Δ Log Capital per Worker	.598 (.062)	.795 (.058)	.648 (.073)	.35*	.597 (.119)
Log Capital per Worker 1960	---	.016 (.002)	.002 (.002)	-.002 (.002)	.001 (.004)
R^2	.63	.76	.58	.12	.55
Sample Size	92	92	92	92	66

Notes: Change variables have been divided by the number of years spanned by the change (20 years for schooling and log GDP, 25 years for capital). Schooling data used in the regressions are from Barro and Lee. The instrumental variables model in column 6 uses Kyriacou's schooling data as excluded instruments for the level and change in Barro-Lee's schooling variables. Capital data are from Klenow and Rodriguez-Clare (1997), and pertain to 1960-85. The coefficient on the change in log capital in column 4 is constrained to equal .35, which is roughly capital's share.

Table 6: Mean Estimates from a Random coefficient Specification
Dependent Variable: Annualized Change Log GDP per Capita

	5-Year Changes		10-Year Changes	
	Constant Coefficient (1)	Mean Variable- Coefficient Estimate (2)	Constant Coefficient (3)	Mean Variable- Coefficient Estimate (4)
A. All Years of Schooling				
Coefficient Estimate for S_{t-1}	.0040 (.0007)	-.0033 (.0036)	.0033 (.0008)	-.0064 (.0059)
p-value	---	.000	---	.000
R ²	.197	.481	.242	.690
B. Male Secondary and Higher Schooling				
Coefficient Estimate for Male Secondary+ S_{t-1}	.0088 (.0017)	-.0196 (.0114)	.0081 (.0020)	-.0353 (.0179)
p-value	---	.0000	---	.0000
R ²	.190	.469	.242	.658

Notes: All regressions control for initial Log GDP per capita and time dummies. The number of countries is 110 for the 5-year change equations and 108 for the 10-year change models. The p-value is for test of equality of country-specific education coefficients in the variable coefficient model. Sample size is 607 in columns 1-2 and 292 in columns 3-4.

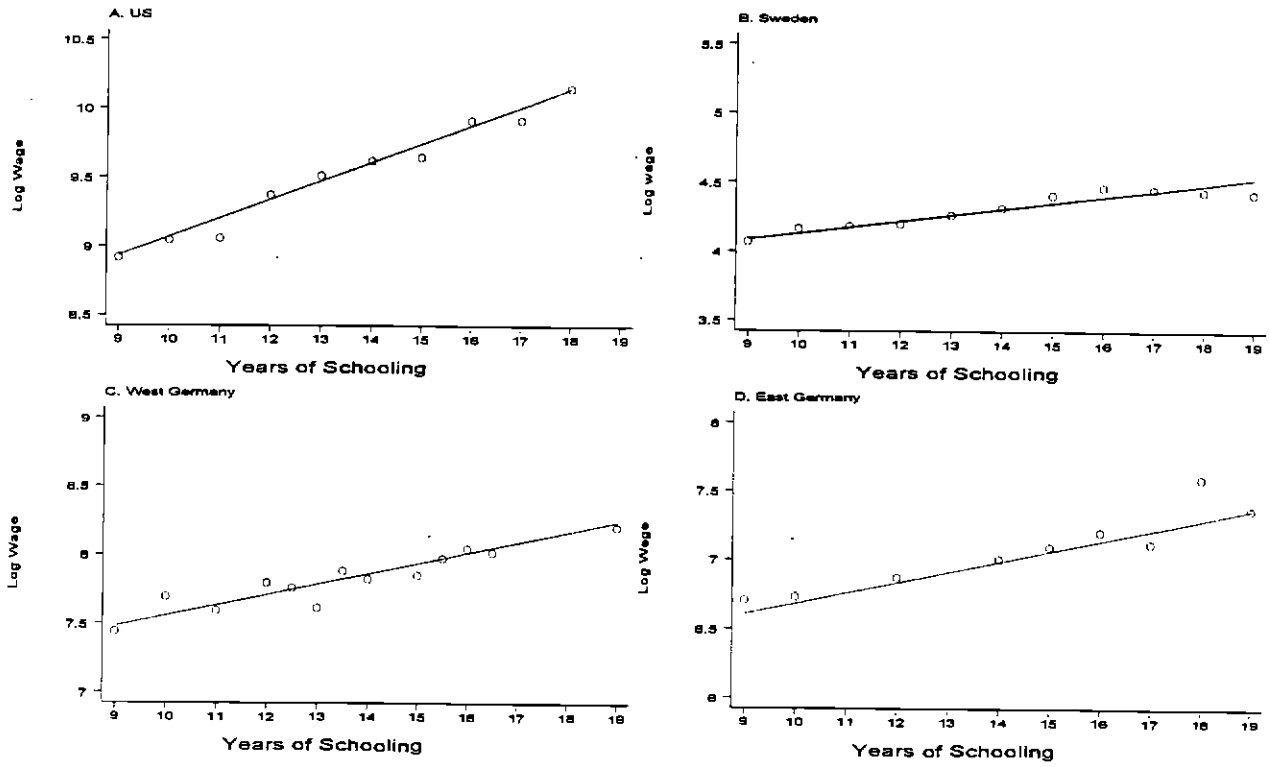


Figure 1: Unrestricted Schooling-Log Wage Relationship and Mincer Earnings Specification

Figure 2
 Linear and Unrestricted Schooling-GDP Growth Relationship, and Change in Unrestricted Relationship

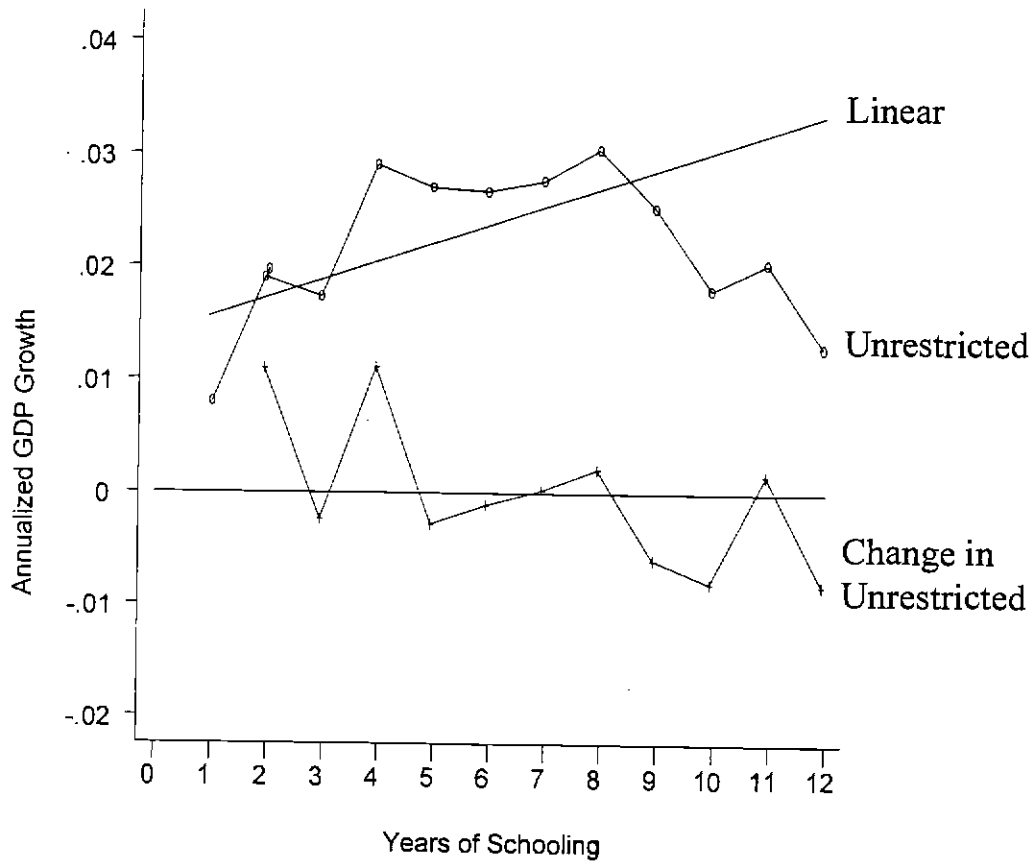
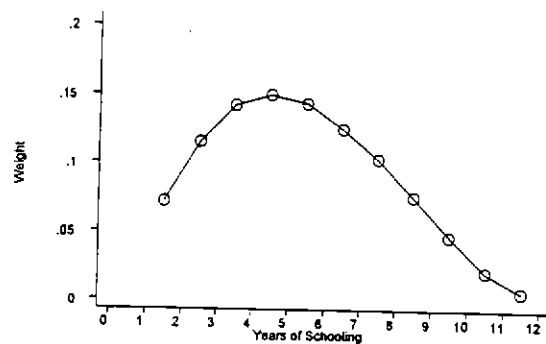
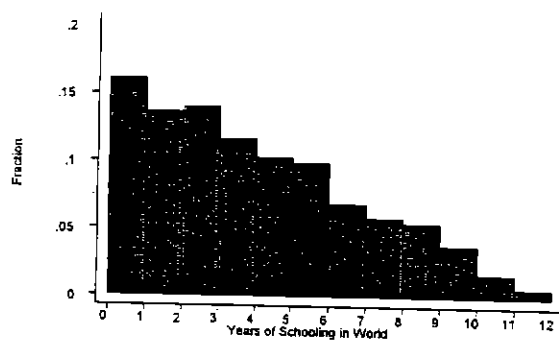


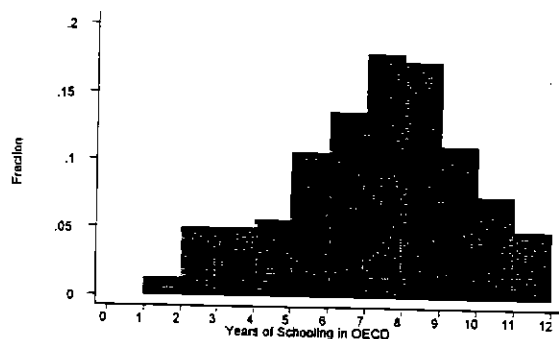
Figure 3
OLS Weighting Function and Education Histograms



A. OLS Weighting Function



B. World Distribution of Education



C. OECD Distribution of Education