Measurement error in education and growth regressions*

Miguel Portela
University of Minho, Portugal
mangelo@eeg.uminho.pt

Rob Alessie
University of Groningen, The Netherlands
r.j.m.alessie@rug.nl

Coen Teulings
CPB - Netherlands Bureau of Economic Policy Analysis, The Netherlands
C.N.Teulings@cpb.nl

28th November 2009

Abstract

The use of the perpetual inventory method for the construction of education data per country leads to systematic measurement error. This paper analyses its effect on growth regressions. We suggest a methodology for correcting this error. The standard attenuation bias suggests that using these corrected data would lead to a higher coefficient. Our regressions reveal the opposite. We discuss why this is the case.

Keywords: growth regressions, education, perpetual inventory method, systematic measurement error.

JEL codes: I2, O4.

*The first author gratefully acknowledges financial support by the Portuguese Foundation of Science and Technology (FCT; ref. SFRH/BD/5114/2001). We would like to thank Thijs van Rens, Ana Rute Cardoso, and the participants at the Tinbergen Institute Seminar, the 2003 Microeconometrics Conference at CEMFI, the IXth Spring Meeting of Young Economists, the 2004 European Economic Association and Econometric Society Conference, for comments on previous versions of the paper. We also thank anonimous referees for their helpful comments.

I. Introduction

Measurement error in education is widely recognized as an important source of bias in growth regressions, see for example Krueger and Lindahl (2001). Barro and Lee (1993, 2001) constructed education data from census information where available, and for missing information used enrolment data and the perpetual inventory method for updating. We show that this updating yields a systematic measurement error, as it yields an underestimation of the growth of education during the period.[1] Classical errors in variables would lead to an underestimation of the coefficient for education in a growth regression. The opposite holds in this case, because the underestimation of the growth of education is compensated by an overestimation of its return. Previous attempts to correct for this error have either only been successful for a limited number of countries or were based on some case-by-case corrections made by the researchers (see de la Fuente and Doménech, 2006, and Cohen and Soto, 2007). We propose a simple correction procedure for data points based on the perpetual inventory method that does not require any *ad hoc* decisions.

The issue of the measurement error in education data is of practical relevance for the interpretation of the relation between education and GDP. The estimated effect of education on economic growth depends on the reliability of education data. Benhabib and Spiegel (1994) and Barro and Sala-i-Martin (1999) conclude that it is the level of education, not its change, that has an impact on economic growth, which is evidence in favour of Nelson and Phelps' (1966) argument that growth is driven by the stock of human capital. Krueger and Lindahl (2001) argue that these conclusions are affected by measurement

[1] Standard errors-in-variables models focus on random measurement errors, but do not solve the issue of systematic measurement errors.

error in education, and that the problem of measurement error is exacerbated by first differencing, since that reduces the signal-noise ratio. Krueger and Lindahl's solution to this problem is to lengthen the differencing period, thereby increasing the signal. They show that indeed the coefficient on the change in education increases by taking a longer differencing period. The authors conclude that "the change in education is positively associated with economic growth once measurement error in education is accounted for" (Krueger and Lindahl, 2001, p.1130). Thus they find empirical evidence in favour of Lucas' (1988) argument that human capital should be interpreted as a normal input in the production process. The analysis in this paper provides a natural explanation for why the coefficient for the change in education is larger when using a 10 instead of a 5 year differencing period that does not rely on measurement error. Our results indicate that an additional year of education has only a moderate effect on GDP of about $4-7\%$ in the short run, but a huge effect of about $50-60\%$ in the long run, which however takes a long time to materialize (up to a century). The explanation we propose compares to Teulings and van Rens (2008). Lengthening the differencing period yields an estimate that looks more like the long run estimate, and is therefore higher.

The most used data set on international education attainment is the one released by Barro and Lee (2001).[2] They build their data on educational attainment from census or survey data.[3] When this information is not available, the authors use a perpetual inventory method based on enrolment data for the inter- and extrapolation of missing observations

[2] Alternative sources are Kyriacou (1991) and Nehru et al. (1995). Kyriacou's work, along with the research by Psacharopoulos and Arriagada (1986), made an important contribution to the field. The Kyriacou data are only available for the period 1965–1985. Nehru et al., although a relevant contribution, ignore census data. de la Fuente and Doménech (2002, p.6) criticise this choice, and argue that it is difficult to justify "discarding the only direct information available on the variables of interest."

[3] Educational attainment corresponds to the highest number of years of schooling achieved.

from census data points. For intermediate observations, the constructed data point is a weighted average of the forward-perpetual inventory method from the last available census observation and the interpolation between two census observations. For the observations before the first and after the last census observation, interpolation is infeasible. Then, the constructed data are based either on the forward- or the backward-perpetual inventory method to or from the closest census observation.

Barro and Lee's data received criticism. de la Fuente and Doménech (2006) provide a revised version of the Barro and Lee (1996) data set for a sample of 21 OECD countries. The authors use "previously unexploited sources [and follow] a heuristic approach to obtain plausible time profiles for attainment levels by removing sharp breaks in the data that seem to reflect changes in classification criteria" (de la Fuente and Doménech, 2002, p.1). In order to circumvent unreliable changes in the data the authors have chosen plausible education values for specific years. Missing observations are filled in, if possible by interpolation, or otherwise by back- and forward projections. de la Fuente and Doménech refrain from using flow estimates based on enrolment data as they seem to produce unlikely time profiles. The authors state that "the construction of our series involves a fair amount of guesswork" (de la Fuente and Doménech, 2002, p.14), and argue that their data seem more plausible when compared to other series.[4]

Similarly to de la Fuente and Doménech (2006), the concern with the quality of the data led Cohen and Soto (2007) to build an alternative data set on countries' education. Their methodology seeks to minimize the extrapolations and keep the data as close as

[4]These two data sets are not directly comparable since Barro and Lee's data measures education completed, while de la Fuente and Doménech's data measures education attended.

possible to information available from national censuses. For that they use more census information than Barro and Lee, and use a different methodology for extrapolating the missing data. An important difference to de la Fuente and Doménech is that Cohen and Soto allow for the use of enrolment data when needed.

This paper analyses the difference between observations based directly on census information and data updated with the perpetual inventory method more deeply. We find large and statistically significant differences. The average downward bias when using the perpetual inventory method is about 60% of the average increase in education during a five year period. One would expect that these differences have a large effect on growth regressions, in particular when using 5 year differences.[5] This turns out not to be the case. Using our corrected measure of education in a growth regression yields a lower effect of education. The reason for this unexpected outcome mainly lies in the way Barro and Lee have imputed missing values before the first and after the last census observation. As we said before, these missing values have been filled in by using the perpetual inventory method. By using this method, Barro and Lee overstate educational attainment for the initial observations of a time series while for the final observations they understate it. As a result, Barro and Lee have underestimated the variance in years of education. This effect is offset in the estimation procedure. The corrected coefficient for the short run return is some 6% lower. It is therefore important to use a dataset that corrects for this bias in a consistent and objective way.[6]

The paper is organized as follows. In the next section we show that there is a sys-

[5] Many countries hold a census every ten years, so that 5 year differences switch back and forth between direct census information and updating by the perpetual inventory method.

[6] Our dataset is available at http://www.eeg.uminho.pt/economia/mangelo/education/.

tematic difference between census and non-census education data. In Section III we will concentrate on the interaction between education and growth using the corrected data on education, comparing the results with known figures. Finally, concluding remarks are presented in Section IV.

II. The extent to which the difference between census and non-census data is systematic

Origins and identification of the systematic difference

Our hypothesis is that the methodology used by Barro and Lee to impute education data when census information is missing yields a systematic underestimation of the growth of education. Barro and Lee assume a uniform mortality rate across educational levels and assume that the educational level of those retiring is equal to the mean educational level among the workforce. In reality older people are on average less educated and have higher mortality rates. Therefore, this procedure overestimates the educational level of those who are retiring and underestimates the survival of more educated individuals. This leads to an underestimation of the growth of educational attainment for the country in periods when census data is not available. This underestimation caused by the imputation procedure is particularly relevant for countries where schooling rises rapidly, because there the difference between the average educational level in the population and the level attained by retirees is substantial.[7] If this hypothesis is true, we should observe in the data that: (i)

[7] See the Appendix for a detailed exposition.

the increase in education between two consecutive census observations should be higher than the increase between non-census observations, (ii) the education level jumps upward between a non-census to a census observation, and that this jump is larger, the larger the period since the previous census. Figure 1 shows the argument for a hypothetical country with 9 observations. At the horizontal axis we have the periods, while the vertical axis plots the average education level in each period. The fat line represents the evolution of true educational level. An empty square on this line denotes the observations for which census information is available. The circle dots represent constructed data points using the enrolment data. The lighter line represents this. The filled square dots represent the data points as would have been constructed if no census information was available. The longer the period since the last census, the greater the measurement error. Available data on education is represented by empty squares and circle dots.[8] Figure 1 shows the effect of the bias such that observed changes in education are lower when the observation is based on non-census data.

Table 1: Example of census variables for Figure 1

Period	$Census$	$Before$	$Last$	$LastC$	$After$
1	0	1	0	0	0
2	1	0	0	0	0
3	0	0	1	0	0
4	0	0	2	0	0
5	1	0	3	3	0
6	1	0	1	1	0
7	0	0	1	0	0
8	1	0	2	2	0
9	0	0	0	0	1

We test the simple idea outlined in this section by constructing four variables. $Before$ applies to observations before the first census observation; it measures the time interval till

[8] The variables $PEdu$ and Edu are discussed in more detail later on.

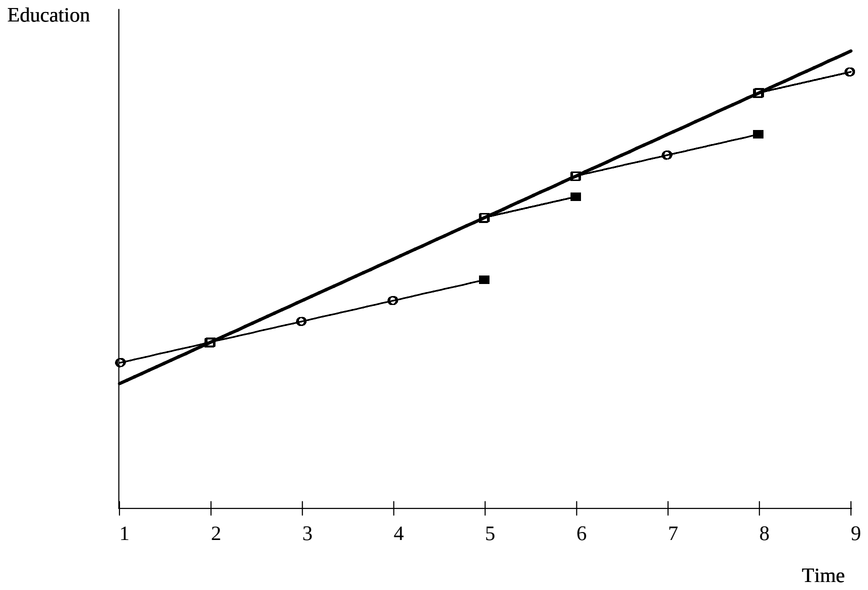

Figure 1: Plot of education with census and non-census data.
Notes: This plot refers to a hypothetical country. The empty square is used for census information; the circle dots represent the estimated points using the enrolment data and the benchmark census information; and the filled square dots represent the values of education that would be estimated for periods in which we have census data.

the first census. $Last$ and $LastC$ apply to observations between two census observations; $Last$ records the number of periods since the previous census, while $LastC$ records the same number, but just for census data points, being zero otherwise. $After$ applies to observations after the last available census; it measures the number of periods since the last census. Table 1 gives the value of these variables for the example in Figure 1. We include these regressors in the following model

$$Edu_{it} = \gamma_t + \beta_B Before_{it} + \beta_L Last_{it} + \beta_{LC} LastC_{it} + \beta_A After_{it} + \eta_i + \varepsilon_{it}, \qquad (1)$$

where Edu_{it} is the education level of country i in period t as measured by Barro and Lee; γ_t is the specific effect for period t, η_i is a country specific effect, and ε_{it} is an idiosyncratic error term. If our hypothesis is correct, we would expect

$$\begin{aligned} -\beta_L &= \beta_{LC} > 0, \\ \beta_B &> \beta_{LC}, \\ -\beta_A &> \beta_{LC}. \end{aligned}$$

The coefficients for $Before$ and $After$ should be larger (in absolute value) than the coefficient on $Last$ and $LastC$ since Barro and Lee use a weighted average of a simple interpolation based on adjacent census observations and the estimate based on the perpetual inventory method for interpolation, while they can only use the perpetual inventory method for extrapolation. If we had used just a dummy for non-census observations, then its coefficient would have been a weighted average of the changes associated with different time intervals till the previous census. Moreover, it would not have differentiated between the positive measurement error for observations before the first census and the negative measurement error for all other non-census observations. In equation (1) census observations are the reference point: $Before$ and $After$ are zero, $Last$ and $LastC$ have the same value and $\beta_L = -\beta_{LC}$.

Data description

Table 2 provides a description of the data. We will focus our attention on population aged 15 and over. The dummy variable $Census$ assumes the value 1 for observations based on a census or survey, and 0 otherwise. The variables $Before$, $Last$, $LastC$, and $After$ are constructed from the $Census$ variable as described above. The income variable is real Log GDP per worker, $LGDP$, and is obtained from the Penn World Table 6.1 (Heston et al., 2002).[9] All variables are available on five year intervals, between 1960 and 2000. Average income increased by 18% per decade, while average education increased by 0.70 year of education per decade, achieving 6.33 years in 2000. Its dispersion has been relatively stable over time, with a slight increase in the beginning of the sample period. With the exception of 1985, income dispersion has steadily increased. Only 32% of the observations with information on education are based on census/survey data. There is a concentration of census information at the start of each decade, 1970, 1980, and 1990. This is a particularly relevant feature when first differencing the data using a 5 year time frame. 46% of the countries have 2 or fewer census observations, and only 26% have 4 or more. Finally, the distribution of countries per period is relatively balanced.[10]

Empirical evidence

Taking first-differences in equation (1) eliminates the fixed country effect:

$$\Delta Edu_{it} = \Delta\gamma_t + \beta_B \Delta Before_{it} + \beta_L \Delta Last_{it} + \beta_{LC} \Delta LastC_{it} + \beta_A \Delta After_{it} + \Delta\varepsilon_{it}, \quad (2)$$

where Δ is the first difference operator, and $\Delta\gamma_t$ are period specific changes. Estimation results are presented in Table 3.[11] In column 1 we report the estimation of equation (1) using the fixed-effects estimator. Column 2 reports the results for equation (2), where we use OLS. For the model in levels in column 1 the hypothesis of the absence of coun-

[9] We use the variable "Real GDP Chain per worker".
[10] See Table A.1 in the Appendix for a list of countries used in our analysis.
[11] The standard errors are robust to heteroskedasticity and error correlation within countries.

Table 2: Summary statistics and the distribution of census data for 5 year interval data

Variable	Statistic	1965	1970	1975	1980	1985	1990	1995	2000	Total
Observations		104	106	110	111	112	116	111	111	985
Census	Mean	0.20	0.54	0.35	0.60	0.20	0.39	0.10	0.00	0.32
Before	Mean	0.61	0.30	0.12	0.02	0.00	0.00	0.00	0.00	0.23
	% zeros	66	81	90	98	100	100	100	100	88
Last	Mean	0.46	1.04	1.03	1.29	0.85	0.91	0.17	0.00	0.64
	% zeros	54	39	32	32	49	61	90	100	62
LastC	Mean	0.03	0.68	0.42	0.95	0.30	0.84	0.17	0.00	0.38
	% zeros	97	62	76	49	83	65	90	100	80
After	Mean	0.03	0.07	0.16	0.37	0.83	1.38	2.27	3.27	0.95
	% zeros	97	96	90	79	54	42	10	0	62
Education	Mean	3.90	4.28	4.52	4.99	5.31	5.84	6.07	6.33	5.03
	Std.Dev.	2.56	2.70	2.75	2.86	2.80	2.84	2.80	2.82	2.88
ΔEdu	Mean	0.13	0.44	0.35	0.52	0.32	0.47	0.32	0.26	0.35
	Std.Dev.	0.29	0.57	0.40	0.60	0.35	0.53	0.32	0.14	0.44
Observations		104	104	106	110	111	112	111	111	869
LGDP	Mean	8.98	9.09	9.16	9.24	9.27	9.33	9.39	9.50	9.20
	Std.Dev.	0.95	0.99	1.00	1.04	1.02	1.07	1.11	1.13	1.04
Observations		85	89	92	94	95	97	97	89	821
$\Delta LGDP$	Mean	0.17	0.16	0.11	0.10	0.04	0.04	0.05	0.08	0.09
	Std.Dev.	0.12	0.14	0.15	0.17	0.17	0.17	0.18	0.10	0.16
Observations		83	85	89	92	94	95	97	87	722
		Distribution of census data								
Number of census		1	2	3	4	5	6	7	8	
Number of countries		25	28	33	22	6	0	1	1	116

Note: The summary statistics for ΔEdu and $\Delta LGDP$ are for changes over five year periods.

try specific effects is rejected. Also for this model we reject the null hypothesis that the idiosyncratic level error terms are not serially correlated, as indicated by the AR(1) test. In column 2 we do not reject the null hypothesis that the error terms are not serially correlated, as indicated by the AR(1) test. Combining this set of results, the OLS applied to first-differences without fixed effects (column 2) is the preferred estimation. The subsequent discussion is restricted to this model.

The estimation results strongly confirm our hypothesis regarding the biases in non-census observations. All four variables have the expected sign and are highly significant. The coefficients on *Last* and *LastC* are nearly identical in absolute value, as predicted.

Table 3: Education regressions

	Levels	First-differences
Variable	Fixed Effects	OLS
Before	0.391**	0.250**
	(0.073)	(0.055)
Last	-0.200**	-0.198**
	(0.032)	(0.028)
LastC	0.199**	0.202**
	(0.033)	(0.029)
After	-0.214**	-0.272**
	(0.057)	(0.057)
F-Stat.: census variables	20.322**	17.176**
F-Stat.: time dummies	74.741**	3.353**
F-Test	251.61**	
Within R^2	0.824	
AR(1)	159.194**	-1.070
Observations	985	869
Countries	116	116

Notes: Significance levels: $*$: 5% $**$: 1%. We allow for heteroskedasticity and within country correlation when computing the standard errors (in parentheses). In the first column the dependent variable is Edu, while in the second it is ΔEdu. All regressions include time effects. F-Stat. stands for the F statistic for the test on joint significance of a set of variables. F-Test reports the F statistic for the test of absence of country specific effects. For the first column the null is stated as 'H_0: all $\eta_i = 0$'. AR(1) is the test for first order serial correlation in the errors.

Furthermore, the coefficient on $After$ and $Before$ are larger in absolute value than the coefficient on $Last$. The magnitude of the measurement error is huge, some 0.20 education-years per 5 year period, or about 60% of the total average increase of education per 5 year period. The fill-in procedure of the observations for which no census information is available introduces therefore a large and systematic measurement error in the data. Given the fact that many countries hold a census every 10 years (usually at the beginning of a decade), the systematic measurement error in the non-census observations yields a particular erratic time series of first-differences when using a 5 year period.

How to correct for the systematic difference?

How can we use this information to improve the quality of the data? Our idea is to use the regression results to correct the original data using the following expression:

$$PEdu_{it} \equiv Edu_{it} - \hat{\beta}_B Before_{it} - \hat{\beta}_L Last_{it} - \hat{\beta}_{LC} LastC_{it} - \hat{\beta}_A After_{it}, \qquad (3)$$

where $PEdu_{it}$ is the corrected education variable.[12] Having estimated the measurement error, we propose to correct those measurement errors on Barro and Lee's data using the coefficients shown in column 2 of Table 3 and the data on the census variables.

Table 4: Correlations among education measures in levels

	Edu	PEdu	EduCS	EduDD	Mean	Variance
Edu	1				5.028	8.299
	(985)				(985)	(985)
PEdu	0.987	1			5.281	8.883
	(985)	(985)			(985)	(985)
EduCS	0.956	0.956	1		5.683	9.957
	(420)	(420)	(420)		(420)	(420)
EduDD	0.892	0.888	0.933	1	9.567	4.464
	(155)	(155)	(80)	(155)	(155)	(155)

Notes: The reported numbers are correlations between pairs of the four education variables. Number of observations in parentheses. The last two columns report the mean and the variance of each variable.

Tables 4 and 5 give the overall correlations between the various education variables; Table 4 in levels and Table 5 in first-differences. The correlation between Barro and Lee education level and the corrected education variable is high, 0.99. The correlation between these two variables and the series constructed by Cohen and Soto (2007) ($EduCS$) and

[12] With this formulation and by using the same coefficients we impose the same bias correction across countries. Although we acknowledge that this is not the most realistic assumption, sample size limitations restrict the available alternatives to implement corrections specific to countries.

Table 5: Correlations among education measures in first-differences

	DEdu	DPEdu	DEduCS	DEduDD	Mean	Variance
DEdu	1				0.350	0.192
	(869)				(869)	(869)
DPEdu	0.888	1			0.501	0.147
	(869)	(869)			(869)	(869)
DEduCS	0.369	0.348	1		0.843	0.187
	(335)	(335)	(335)		(335)	(335)
DEduDD	0.068	0.019	0.391	1	0.376	0.020
	(135)	(135)	(60)	(135)	(135)	(135)

Notes: Number of observations in parentheses. First-differences are computed over a 5 year interval, except for $DEduCS$ where first-differences are computed over a 10 year interval. Correlations with the other variables account for this adjustment. See note to Table 4.

de la Fuente and Doménech (2006) ($EduDD$) is only slightly lower. The mean of Barro and Lee data is the lowest of all four, while $EduDD$ presents the lowest variance. For the data by de la Fuente and Doménech, this comparison does not make much sense, since they consider only the very selective sample of 21 OECD countries. To a lesser extent, a similar objection can be raised against a comparison to the Cohen and Soto data, where the difference in the number of observations is mainly due to the fact that they have data once every 10 years. However, the comparison with our corrected data is highly informative. The measurement error in the fill in procedure in the Barro and Lee data leads to an underestimation of the average education level by 0.25 education-years. The measurement error understates the final observations, but overstates the initial observations, which leads to a compression of the "true" variance. So contrary to the classical model, where measurement error is orthogonal to the signal and therefore increases the variance of the observed data, here it compresses the variance.

The assumption on classical measurement error is that it is uncorrelated to the true variable. In our case, the measurement error $Edu - PEdu$ turns out to be negatively correlated with the corrected education variable $PEdu$, as countries with lower values of education have higher measurement errors, which could imply an upward bias on the estimate of the coefficient on education. The covariance between $PEdu$ and the measurement error is -0.41 for the 5 year data and -0.50 for the 10 year data. The variances of the measurement error are 0.23 and 0.28 for the 5 and the 10 year data respectively. The negative overall correlation between the measurement error and the corrected education variable explains why we obtain a lower estimate for the short run return to education when using the corrected data, see Section III.[13]

In first-differences, the correlations between education variables are much lower. The correlation between Barro and Lee and our corrected variable is still high, 0.89. For alternative sources of information, the correlations drop significantly. Once more, the correlations are higher with Cohen and Soto's data. Again, a comparison of the mean and variance of the changes between Barro and Lee and our corrected variable is revealing. The measurement error in Barro and Lee compresses the measured average growth of education substantially, from 1.00 education-years per decade to 0.70 education-years. The variance of the changes is however overestimated in the Barro and Lee data, as one would expect with all the erroneous changes back and forth from census to non-census based observations.[14]

[13]Let y_t denote economic growth and $PEdu_t$ the "true" value of education. Suppose that economic growth can be described by the following regression: $y_t = \kappa_1 + \kappa_2 PEdu_t + \varphi_t$ where φ_t is an error term and $\kappa_2 > 0$. Education is measured with error, i.e. $Edu_t = PEdu_t + m_t$, where Edu_t is the observed value. Suppose we regress y_t on Edu_t, i.e. we consider the following model $y_t = \kappa_1 + \kappa_2 Edu_t + \varphi_t - \kappa_2 m_t$. The plim for the OLS estimator $\hat{\kappa}_2$ is $\kappa_2 - \kappa_2 \frac{\sigma_{PEdu,m}}{\sigma^2_{Edu}} - \kappa_2 \frac{\sigma^2_m}{\sigma^2_{Edu}}$. The term $-\kappa_2 \frac{\sigma^2_m}{\sigma^2_{Edu}}$ reflects the standard attenuation bias. The other term $-\kappa_2 \frac{\sigma_{PEdu,m}}{\sigma^2_{Edu}}$ has the opposite sign because the imputation procedure of Barro and Lee induces a negative correlation between the measurement error m_t and the true value of education $PEdu_t$: $\sigma_{PEdu,m} < 0$ (obviously $\sigma^2_m > 0$). As a result we can have a higher coefficient when estimating with Edu, as compared with the estimation with $PEdu$, as we later observe in our estimations.

[14]As well pointed out by a referee, an alternative strategy to estimate how important measurement error is for the time series variation in education is to regress Edu on $PEdu$ using a fixed effects procedure,

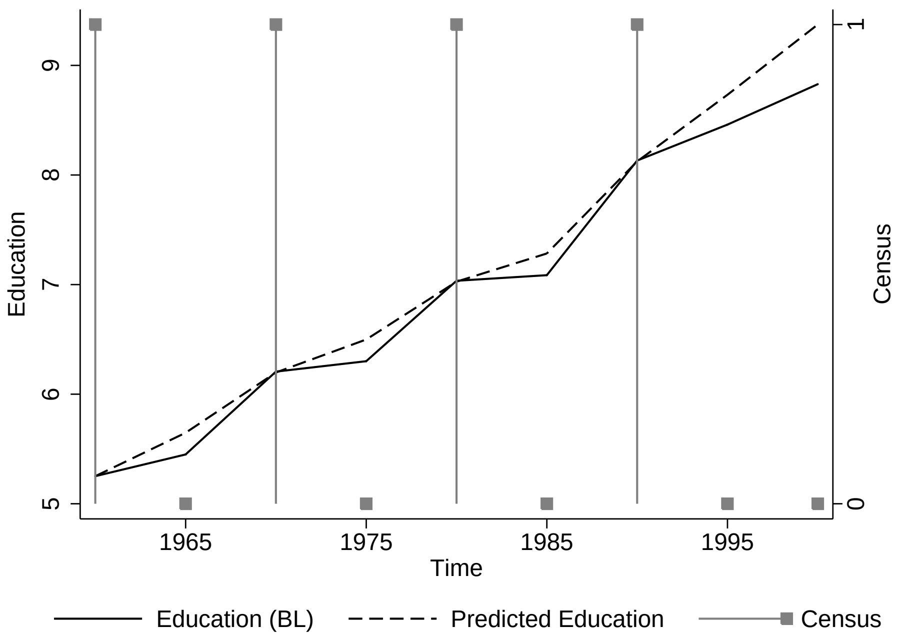

Figure 2: Education information for Argentina

These ideas are well documented by the data on Argentina, as shown in Figure 2. We observe spikes at each census observation for the data estimated by Barro and Lee (Education (BL)). Between census (1960–1990), our procedure (Predicted Education) smoothens the data. However, for observations after the last census available (1990), the correction increases the variance. When the variables are analysed in changes, $PEdu$ has a higher mean, but a smaller variance than Edu. The data also documents the dramatic difference between the measured changes in education when using 5 or 10 year time period. The 5 year differentials are entirely dominated by the difference between census and non-census observations.

---

and retrieve the within R^2. We obtain a value of 0.93 for this goodness of fit measure, which highlights the significance of measurement error for the within variation in Edu. These results indicate that the time series variation in the measurement error should be taken into account when one estimates a model explaining the effects of education on economic growth.

III. Growth regressions: what changes?

Having analysed the difference in education data according to its source, we will now re-evaluate growth regressions. First, we estimate the macro-Mincerian growth equation as defined by

$$\Delta LGDP_{it} = \delta_t + \alpha_0 LGDP_{i,t-\tau} + \alpha_1 \Delta Edu_{it} + \alpha_2 Edu_{i,t-\tau} + \upsilon_{it}, \tag{4}$$

where δ_t are time specific effects, $LGDP_{it}$ stands for log real income per worker in country i in period t, Edu_{it} is the average education level, τ is the time span of the data, and υ_{it} is an i.i.d. error term. All variables in changes are annualised. The functional form adopted here is comparable to the one used in Topel (1999) and Krueger and Lindahl (2001).[15]

The first two columns of Table 6 reproduce estimations from Topel (1999, Table 4), and Krueger and Lindahl (2001, Table 3) [K&L(2001)]. The other columns are our estimations of equation (4) using the two measures of education, Edu and $PEdu$, at different time spans of the data, 5 and 10 years, respectively. In the last column of Table 6 we use the Edu measure of education and introduce Before, Last, LastC, and After directly as regressors in the equation instead of fixing their coefficients a priori at the value as reported in Table 3. The estimation procedure is OLS, and we report standard errors robust to heteroskedasticity and error correlation within countries.[16]

Similarly to Topel (1999) and Krueger and Lindahl (2001), we conclude that con-

[15] A discussion on the functional form can be found in Ferreira et al. (2004) and Pritchett (2006).

[16] The estimation procedure of $PEdu$, combined with the availability of more recent data, lead to the use of a different sample of those used by Topel (1999) and Krueger and Lindahl (2001). Compared to Krueger and Lindahl, we use 97 countries, with a total of 722 (353) observations for the 5 year data (10 year data), instead of their 110 countries and 607 (292) observations. Within each time span we use the same sample throughout columns three to five.

Table 6: The effect of education on growth - annualised OLS estimations

	5 year data				
Variable	Topel(1999)	K&L(2001)	Edu	PEdu	Edu
ΔEdu	0.041**	0.039**	0.0517**	0.0488**	0.0462**
	(0.014)	(0.014)	(0.0137)	(0.0146)	(0.0150)
LagEdu	0.004**	0.004**	0.0035**	0.0037**	0.0036**
	(0.001)	(0.001)	(0.0009)	(0.0010)	(0.0009)
LagLGDP	-0.007**	-0.006*	-0.0060**	-0.0063**	-0.0062**
	(0.002)	(0.003)	(0.0022)	(0.0023)	(0.0023)
Census variables	-	-	-	-	Yes
R^2	0.218	0.197	0.1315	0.1296	0.1414
Observations	608	607	722	722	722
Countries	111	110	97	97	97
	10 year data				
ΔEdu	0.085**	0.086**	0.0882**	0.0789**	0.0758**
	(0.020)	(0.024)	(0.0213)	(0.0222)	(0.0215)
LagEdu	0.004**	0.004**	0.0039**	0.0041**	0.0040**
	(0.001)	(0.001)	(0.0009)	(0.0009)	(0.0009)
LagLGDP	-0.007**	-0.005†	-0.0073**	-0.0076**	-0.0075**
	(0.002)	(0.003)	(0.0021)	(0.0022)	(0.0022)
Census variables	-	-	-	-	Yes
R^2	0.315	0.284	0.2336	0.2240	0.2472
Observations	290	292	353	353	353
Countries	111	110	97	97	97

Notes: Significance levels: † : 10% ∗ : 5% ∗∗ : 1%. Robust standard errors in parentheses. The results under Topel(1999) reproduce part of Table 4 in Topel (1999). The results under K&L(2001) reproduce part of Table 3 in Krueger and Lindahl (2001). In this case the number of countries is the maximum number of countries reported by the authors. All variables in changes were divided by the time span in each data. The dependent variable is annualised first-difference real Log GDP per worker, ΔLGDP. All regressions include time effects. In the last column we re-estimate the model using Edu and include the census variables as regressors.

temporaneous changes in education have a positive and statistically significant effect on economic growth, which contradicts the findings of Benhabib and Spiegel (1994) and Barro and Sala-i-Martin (1999). For the five year data, the short-run return to education is about 5% per year of education, while Topel (1999) and Krueger and Lindahl (2001) report approximately 4%. The short-run return to education is lower for our corrected data than for the original data.[17] In the classical errors in variables model, measurement error re-

[17] In the long-run the returns to education are similar when using the two education variables. For the

duces the coefficient (in absolute value) due to attenuation bias. One reason why we get a different result in this case is that the measurement error is systematic.[18] It reduces the mean of the change in education by 0.15 education-years per 5 year period, see Table 5. Although the estimated coefficient for the corrected variable is lower, the total estimated effect of education on GDP is larger, because the positive effect on the growth of education outweighs the negative effect on the coefficient. So, the effect of the bias on the coefficient is a balance between two forces: introducing the spurious component in ΔEdu reduces the coefficient (standard attenuation bias), while understatement of the average level of ΔEdu pushes up the coefficient. For the 5 year time frame, both forces almost cancel. For the 10 year time frame, the first component is less important (since many census observations are located at the beginning of a decade), so the latter force clearly dominates.

When we use Barro and Lee's 10 year original data, returns to changes in education are 8.8%, and very similar to the two comparison studies. However, using our corrected value for education the estimated return is only 7.9%.[19] The systematic measurement error on education identified in the previous section could lead to the overestimation of its coefficient in a growth regression, which is clearly corroborated by the 10 year results. While Topel (1999) and Krueger and Lindahl (2001) find that the coefficient more than doubles with the doubling of the time span, the change in our coefficient is smaller, which facilitates the reconciliation between the results for different time spans.

---

5 year data the long-run returns to Edu are 58.3% ($\simeq 0.0035/0.0060 \times 100$), while for $PEdu$ we obtain 58.7% ($\simeq 0.0037/0.0063 \times 100$). The figures for the 10 year data are 53.4% and 53.9%, respectively.

[18] Since we have in our growth regressions more than one explanatory variable measured with error, ΔEdu_{it} and $Edu_{i,t-\tau}$, there are no general results for the sign of the bias.

[19] The bias caused by measurement error (m) depends on the following statistics: $Cov(PEdu, m)/Var(Edu)$ (-0.049 for the 5 year data and $-0.50/8.49 = -0.059$ for the 10 year data) and $Var(m)/Var(Edu)$ ($0.23/8.299 = 0.028$ for the 5 year data and $0.28/8.49 = 0.033$ for the 10 year data). Notice that the values of those ratios are rather small. This might explain our finding that the effect of our correction procedure on the estimates of the growth equations is rather small.

The bias introduced by the perpetual inventory method has a specific structure as described by equation (1). The omitted variable bias thus generated suggests an alternative way to account for it in growth regressions. A simple solution is to introduce the census information in our income regression.[20] The last column of Table 6 reports the results. The growth regression is estimated using Barro and Lee data jointly with the census variables $Before$, $Last$, $LastC$, and $After$. The coefficients on ΔEdu and $LagEdu$ are now slightly smaller than the ones we obtained with $PEdu$. For the 5 year interval data the joint significance test on the census variables yields an $F-statistic$ of 2.81, with a $p-value$ of 0.03, while for the 10 year data the $F-statistic$ is 1.88, with a $p-value$ of 0.12.[21]

A second result, which is identical among the different studies and time spans, indicates that the initial level of education is relevant for economic performance. While the result on ΔEdu supports the human capital interpretation of the role of education in economic growth, this empirical evidence gives also support to the externalities interpretation of the returns to education. Based on our corrected data, $PEdu$, the long-run return to education is $54-59\%$.[22] Although this return seems (too) large, we should keep in mind that the effect takes a long time to materialize. The return is at 50% of its long-run

[20] As pointed out by a referee, when estimating the growth regression an additional justification for this procedure is that this way we are also taking into account that the systematic measurement error might cause a serial correlation in the error term of the growth equation (4).

[21] A further factor that yields overestimation of the effect of education based on the Barro and Lee data is that the variable $Before$ turns out to be a predictor of future growth. The most likely explanation is that holding a census is not an exogenous variable. So countries that initially do not have a census, and later on have, are countries that are likely to have grown faster than average. As pointed out by a referee, these results also indicate that, at least partially, the extent of systematic measurement error associated with the perpetual inventory method depends on the size of schooling changes over time. In regions like sub-Saharan Africa or East Asia this issue is quite relevant. In the first case the average value of the variable Before is 0.36 and the average change in schooling is 0.25, while in the second these figures are 0.09 and 0.46, respectively. For the whole sample the figures are 0.23 and 0.35, respectively.

[22] That is, 0.0041/0.0076 or 0.0037/0.0063, respectively.

value after $75 - 99$ years. The immediate return is 4.2% for the 5 year time period and 6.5% for the 10 year period.[23] The numbers for the 5 and 10 year time interval are very similar. This puts into question Krueger and Lindahl's interpretation of this difference as being due to an increase in the signal to noise ratio when lengthening the observation period. Lengthening the observation period makes the short return look much like the long-run return, which happens to be substantially higher than the short-run return. In Figure 3 the return to education over the first 110 years is depicted. The time path of the cumulated returns to education is very similar for the two time spans, and for the two education variables.

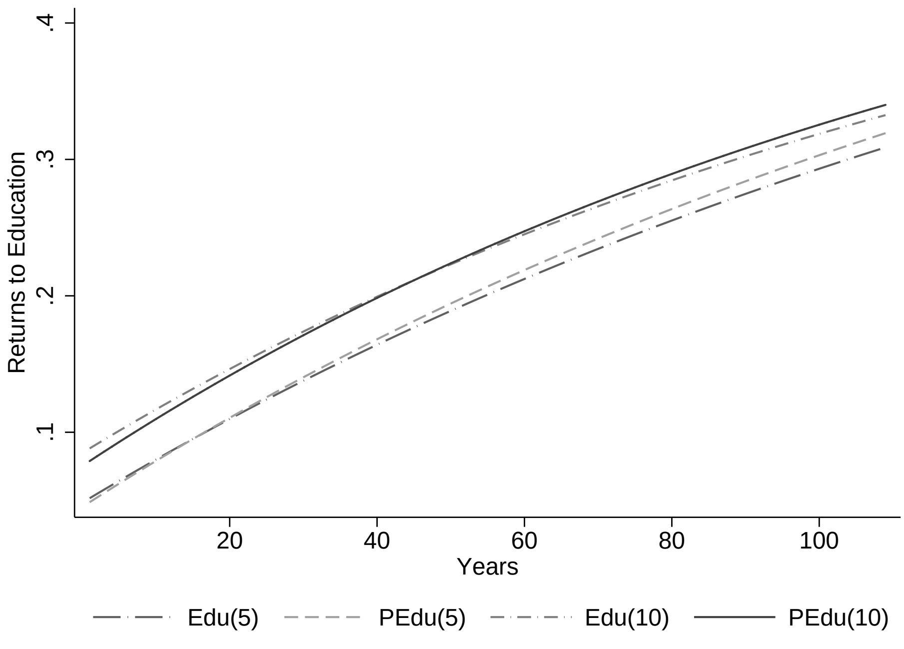

Figure 3: Returns to education for the different time spans and education variables

The results indicate that the GDP half–life adjustment ranges between 91 and 110

[23] Please see the Appendix for details on the computation of the immediate return.

years,[24] which stresses the idea that whichever externalities are associated with permanent changes in education, they will take a long time before benefiting a given country. The results indicate that the time length of the data sets currently available is too short to identify in a precise way the long-run returns to the investment in education.

Finally, in Table 7 we replicate the estimations using Cohen and Soto's (2007) education variable. The differences with Barro and Lee's data occur essentially on the coefficient on changes in education. In this case, the contemporaneous returns to education are around 11%, more than 4 percentage points above our corrected estimates.[25] In the long-run, using $EduCS$ indicates a return to education of about 49%, more than 8 percentage points below the values we obtain using Barro and Lee's data.

Table 7: Growth regressions - comparison with Cohen and Soto data

	10 year data			
Variable	Edu	PEdu	Edu	EduCS
ΔEdu	.0838**	.0664*	.0664*	.1107**
	(.0255)	(.0285)	(.0271)	(.0342)
LagEdu	.0036**	.0039**	.0037**	.0032**
	(.0010)	(.0010)	(.0010)	(.0009)
LagLGDP	-.0063**	-.0066**	-.0065**	-.0065*
	(.0024)	(.0025)	(.0025)	(.0026)
Census variables	-	-	Yes	-
R^2	.2245	.2155	.2474	.2204

Notes: Significance levels: $*$: 5% $**$: 1%. Robust standard errors in parentheses. The dependent variable is annualised first-difference real Log GDP per worker, ΔLGDP. ΔEdu stands for annualised changes in education. All regressions include time effects. The sample includes 300 observations and 79 countries.

[24] That is, $\ln(2)/0.0076 \simeq 91$ and $\ln(2)/0.0063 \simeq 110$, respectively.

[25] In our analysis, we are missing 18 countries in Cohen and Soto's data, which are in Barro and Lee sample. The countries are Barbados, Botswana, Congo, Gambia, Guinea-Bissau, Hong Kong, Iceland, Israel, Lesotho, Pakistan, Papua New Guinea, Poland, Rwanda, Slovakia, Slovenia, Sri Lanka, Taiwan, and Togo.

IV. Final remarks

Our analysis of Barro and Lee (2001) education data reveals a systematic difference between the observations derived from census data and observations that inter- or extrapolate from these data points using the perpetual inventory method. On average, Barro and Lee's data underestimate the growth of education by about one fifth of a year every five year period. Contrary to previous work by Cohen and Soto (2007) and de la Fuente and Doménech (2006), our results suggest that this systematic measurement error can lead to overstatement of the short-run return. Furthermore, our results suggest that the interpretation of Krueger and Lindahl (2001) of the higher short-run return to education when using a 10 instead of a 5 year time period for differencing is not necessarily correct. Krueger and Lindahl claim that the higher short-run return when using the 5 year time period is due to attenuation bias. Our calculations show that the increase in the short-run return when using a 10 year differencing period can be well explained by the fact that the immediate return is a factor ten smaller than the long-run return. The short-run return as measured in a discrete time framework is always a mixture of the immediate return and the long return. When using a 10 year time period, the short-run return looks more like the long-run return than when using a 5 year time period, which offers a more plausible explanation for Krueger and Lindahl's empirical results than their interpretation of attenuation bias. However, the long-run effect takes a long time to materialize, the half-time period being around a century. Therefore, current data do not cover a long enough period of time for a precise estimate of the long-run effect of education on GDP. We have to wait for another century.

References

Aghion, P. and Howitt, P. (1998), *Endogenous Growth Theory*, MIT Press, Cambridge, MA.

Barro, R. and Lee, J.-W. (1993), 'International comparisons of educational attainment', *Journal of Monetary Economics* **32**(3), 363–394.

Barro, R. and Lee, J.-W. (1996), 'International measures of schooling years and schooling quality', *American Economic Review* **86**(2), 218–223.

Barro, R. and Lee, J.-W. (2001), 'International data on educational attainment: Updates and implications', *Oxford Economic Papers* **53**(3), 541–563.

Barro, R. and Sala-i-Martin, X. (1999), *Economic Growth*, MIT Press, Cambridge, Massachusetts.

Benhabib, J. and Spiegel, M. (1994), 'The role of human capital in economic development: Evidence from aggregate cross-country data', *Journal of Monetary Economics* **34**(2), 143–173.

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

de la Fuente, A. and Doménech, R. (2002), 'Human capital in growth regressions: How much difference does data quality make? An update and further results', CEPR Discussion Paper 3587, London.

de la Fuente, A. and Doménech, R. (2006), 'Human capital in growth regressions: How much difference does data quality make?', *Journal of the European Economic Association* **4**(1), 1–36.

Ferreira, P. and Issler, J. and de Abreu Pessôa, S. (2004), 'Testing production functions used in empirical growth studies', *Economics Letters* **83**(1), 29–35.

Heston, A. and Summers, R. and Aten, B. (2002), 'Penn World Table version 6.1'. Center for International Comparisons at the University of Pennsylvania (CICUP), October, Philadelphia.

Krueger, A. and Lindahl, M. (2001), 'Education for growth: Why and for whom?', *Journal of Economic Literature* **39**(4), 1101–1136.

Kyriacou, G. (1991), 'Level and growth effects of human capital: A cross-country study of the convergence hypothesis'. New York University Economic Research Report: 91-26.

Lucas, R. (1988), 'On the mechanics of economic development', *Journal of Monetary Economics* **22**(1), 3–42.

Mankiw, N. and Romer, D. and Weil, D. (1992), 'A contribution to the empirics of economic growth', *Quarterly Journal of Economics* **107**(2), 407–438.

Nehru, V. and Swanson, E. and Dubey, A. (1995), 'A new database on human capital stocks in developing and industrial countries: Sources, methodology and results', *Journal of Development Economics* **46**(2), 379–401.

Nelson, R. and Phelps, E. (1966), 'Investment in humans, technology diffusion and economic growth', *American Economic Review* **56**(2), 69–75.

Pritchett, L. (2001), 'Where has all the education gone?', *The World Bank Economic Review* **15**(3), 367–391.

Pritchett, L. (2006), Does learning to add up add up? The returns to schooling in aggregate data, *in* 'Handbook of Education Economics', Vol. 1, Erik Hanushek and Finish Welch, eds, Amsterdam: North Holland, chapter 11, pp. 637–695.

Psacharopolos, G. and Ana Maria Arriagada (1986), 'The educational composition of the labour force: An international comparison', *International Labour Review* **125**(5), 561–574.

Summers, R. and Heston, A. (1991), 'The Penn World Table (mark 5): An expanded set of international comparisons, 1950-1988', *Quarterly Journal of Economics* **106**(2), 280–315.

Temple, J. (1999), 'A positive effect of human capital on growth', *Economics Letters* **65**(1), 131–134.

Teulings, C. and Van Rens, T. (2008), 'Education, growth, and income inequality', *The Review of Economics and Statistics* **90**(1), 89–104.

Topel, R. (1999), Labor markets and economic growth, *in* 'Handbook of Labor Economics', Vol. 3C, Orley Ashenfelter and David Card, eds, Amsterdam: North Holland, chapter 44, pp. 2943–2984.

Appendix

Educational attainment and mortality rates

Let N_ς = number of persons in the workforce at the beginning of period ς, $\varsigma = t, t+1$, and let $\overline{PEdu}_{t+1}$ be the average education level at the beginning of period $t+1$.

The PIM method can be summarized by means of the following equation:

$$\overline{PEdu}_{t+1} = \frac{N_t}{N_{t+1}}\overline{PEdu}_t - \frac{ND_t}{N_{t+1}}\overline{PEdu\_die}_t + \frac{Nnew_{t+1}}{N_{t+1}}\overline{Enroll}_{t+1}, \tag{A1}$$

where ND_t number of people who left the workforce (due to dying or retirement) during period t; $Nnew_{t+1}$ = number of people who entered the workforce during period t; $\overline{PEdu\_die}_t$ = average education of those who died; $\overline{Enroll}_{t+1}$ average education of those who entered the workforce. Notice that $N_{t+1} = N_t - ND_t + Nnew_{t+1}$.

In terms of changes, equation (A1) becomes

$$\begin{aligned}\overline{PEdu}_{t+1} - \overline{PEdu}_t &= \frac{N_t - N_{t+1}}{N_{t+1}}\overline{PEdu}_t - \frac{ND_t}{N_{t+1}}\overline{PEdu\_die}_t + \frac{Nnew_{t+1}}{N_{t+1}}\overline{Enroll}_{t+1} = \\ &\frac{ND_t}{N_{t+1}}\left(\overline{PEdu}_t - \overline{PEdu\_die}_t\right) + \frac{Nnew_{t+1}}{N_{t+1}}\left(\overline{Enroll}_{t+1} - \overline{Pedu}_t\right).\end{aligned} \tag{A1'}$$

Barro and Lee[26] take equation (A1) as a starting point of their imputation procedure. The problem is that $\overline{PEdu\_die}_t$ is not observed by Barro and Lee. They assume that $\overline{PEdu\_die}_t = \overline{PEdu}_t$. In the paper we point out that it is rather likely that $\overline{PEdu\_die}_t <$

[26] In this appendix we do not take into account that the imputation procedure of Barro and Lee is partly based on interpolations (see the end of section II.1).

$\overline{PEdu}_t$ in practice. Barro and Lee predict average education $\overline{Edu}$ in the following way:

$$\overline{Edu}_{t+1} = \frac{N_t}{N_{t+1}}\overline{Edu}_t - \frac{ND_t}{N_{t+1}}\overline{Edu}_t + \frac{Nnew_{t+1}}{N_{t+1}}\overline{Enroll}_{t+1}, \tag{A2}$$

or in terms of changes

$$\overline{Edu}_{t+1} - \overline{Edu}_t = \frac{N_t - N_{t+1}}{N_{t+1}}\overline{Edu}_t - \frac{ND_t}{N_{t+1}}\overline{Edu}_t + \frac{Nnew_{t+1}}{N_{t+1}}\overline{Enroll}_{t+1} =$$

$$\frac{Nnew_{t+1}}{N_{t+1}}\left(\overline{Enroll}_{t+1} - \overline{Edu}_t\right). \tag{A2'}$$

Suppose that in year t a census has been held, i.e. $\overline{PEdu}_t = \overline{Edu}_t$. Since $\overline{PEdu\_die}_t < \overline{PEdu}_t$, then a comparison between equations (A1) and (A2) reveals that $\overline{Edu}_{t+1} < \overline{PEdu}_{t+1}$, i.e. the growth in the average education level between year t and $t+1$ is underestimated. Along the same line of reasoning, one can easily prove that $\overline{Edu}_{t+\iota} < \overline{PEdu}_{t+\iota}$, $\iota = 2, 3, ...$, i.e. the growth in the average education level between census year t and year $t + \iota$ is underestimated by Barro and Lee.

By subtracting equation (A2') from (A1') we can compare the growth rate in the two education measures

$$\Delta\overline{PEdu}_{t+1} - \Delta\overline{Edu}_{t+1} = \frac{ND_t}{N_{t+1}}\left(\overline{PEdu}_t - \overline{PEdu\_die}_t\right) + \frac{Nnew_{t+1}}{N_{t+1}}\left(\overline{Edu}_t - \overline{PEdu}_t\right).$$

Notice that the first term of the right hand side is positive, under the plausible assumption $\overline{Edu}_t > \overline{PEdu\_die}_t$ the second term is negative, so that we cannot unambiguously sign the left hand side. However, in case of a stationary or decreasing workforce

$(ND_t \geq Nnew_{t+1})$, it holds that $\Delta\overline{PEdu}_{t+1} > \Delta\overline{Edu}_{t+1}$, i.e. due to the imputation procedure of Barro and Lee the growth in the true education level is underestimated. In case of a country with a growing workforce $(ND_t < Nnew_{t+1})$, underestimation takes place if

$$Nnew_{t+1}\left(\overline{Pedu}_t - \overline{Edu}_t\right) < ND_t\left(\overline{PEdu}_t - \overline{PEdu\_die}_t\right).$$

This condition is particularly appropriate for countries where schooling rises rapidly, because there the difference between the average education level and the level attained by those who die is rather large.

Computation of the immediate return

The immediate return and the half–life can be calculated by assuming that innovations in the education variable are uniformly distributed over the observation period. We do the calculations for $PEdu$, and for the 5 year observation period. First, we calculate the raw estimate of half-time

$$\frac{\ln(2)}{0.0063} = 110.0234.$$

Second, we correct for the fact that part of the effect is realized immediately. Since the short-run return, S, can be defined as

$$S = L\left(1 - e^{-\lambda t}\right),$$

where L is the long-run return, and λ is the convergence rate to equilibrium, our results imply that

$$0.0488 = \frac{0.0037}{0.0063}\left(1 - e^{-0.0063t}\right).$$

So, the time needed to reach the immediate effect is

$$\frac{\ln\left(1 - 0.0488\frac{0.0063}{0.0037}\right)}{-0.0063} = 13.7695.$$

Finally, we take into account the fact that the immediate effect is measured imperfectly, by using a five year time interval. Assuming that the innovation is distributed uniformly, we have to add half of the length of the time interval. The estimated half-time is given by

$$110.0234 - 13.7695 + 2.5 = 98.7539.$$

The immediate effect has also to be corrected for the length of the observation period (the longer the observation period, the more the estimated immediate effect will look like the long-run effect). This can be done by taking the time to reach the immediate effect corrected for half the time interval, and using a first order Taylor expansion of the function $1 - e^{-\lambda t}$, λt,

$$0.0063 * (13.7695 - 2.5) = 0.071.$$

Hence, 7.1% of the long-run effect is realized immediately:

$$0.071 \times \frac{0.0037}{0.0063} = 4.2\%.$$

Table A.1: List with the 116 countries used in the analysis

Afghanistan	Cyprus	Iceland	Nepal	Spain
Algeria	Czech Rep.	India	Netherlands	Sri Lanka
Argentina	Czechoslovakia	Indonesia	New Zealand	Sudan
Australia	Denmark	Iran	Nicaragua	Swaziland
Austria	Dominican Rep.	Iraq	Niger	Sweden
Bahrain	Ecuador	Ireland	Norway	Switzerland
Bangladesh	Egypt	Israel	Pakistan	Syria
Barbados	El Salvador	Italy	Panama	Taiwan
Belgium	Fiji	Jamaica	P. New Guinea	Tanzania
Benin	Finland	Japan	Paraguay	Thailand
Bolivia	France	Jordan	Peru	Togo
Botswana	Gambia	Kenya	Philippines	Trinidad & Tobago
Brazil	Germany, East	Korea, South	Poland	Tunisia
Bulgaria	Germany, West	Kuwait	Portugal	Turkey
Cameroon	Ghana	Lesotho	Romania	Uganda
Canada	Greece	Liberia	Russia	United Kingdom
Central African Rep.	Guatemala	Malawi	Rwanda	US of America
Chile	Guinea-Bissau	Malaysia	Senegal	Uruguay
China	Guyana	Mali	Sierra Leone	Venezuela
Colombia	Haiti	Mauritius	Singapore	Yemen
Rep. of the Congo	Honduras	Mexico	Slovakia	Yugoslavia
Costa Rica	Hong Kong	Mozambique	Slovenia	Zambia
Croatia	Hungary	Myanmar	South Africa	Zimbabwe
Cuba				