Tjalling C. Koopmans Research Institute

Discussion Paper Series nr: 04-14

Measurement Error in Education and Growth Regressions

Miguel Portela
Coen Teulings
Rob Alessie

Universiteit Utrecht

Utrecht School of Economics

Tjalling C. Koopmans Research Institute
Utrecht School of Economics
Utrecht University

Vredenburg 138
3511 BG Utrecht
The Netherlands
telephone +31 30 253 9800
fax +31 30 253 7373
website www.koopmansinstitute.uu.nl

The Tjalling C. Koopmans Institute is the research institute and research school of Utrecht School of Economics. It was founded in 2003, and named after Professor Tjalling C. Koopmans, Dutch-born Nobel Prize laureate in economics of 1975.

In the discussion papers series the Koopmans Institute publishes results of ongoing research for early dissemination of research results, and to enhance discussion with colleagues.

Please send any comments and suggestions on the Koopmans institute, or this series to P.vanDriel@econ.uu.nl

ontwerp voorblad: WRIK Utrecht

How to reach the authors

Miguel Portala
Tinbergen Institute
Roeterstraat 31
1018 WB Amsterdam
the Netherlands
Minho University,
Porto, Portugal
Email: portela@tinbergen.nl

Rob Alessie
Utrecht University
Utrecht School of Economics
Vredenburg 138
3511 BG Utrecht
The Netherlands.
E-mail: r.alessie@econ.uu.nl

Coen Teulings
Tinbergen Institute
Roeterstraat 31
1018 WB Amsterdam
the Netherlands
Erasmus University
Rotterdam, the Netherlands
E-mail: teulings@few.eur.nl

Utrecht School of Economics
Tjalling C. Koopmans Research Institute
Discussion Paper Series 04-14

Measurement Error in Education and Growth Regressions

Miguel Portela[a]
Coen Teulings[a]
Rob Alessie[b]

[a]Tinbergen Institute

[b]Utrecht School of Economics
Utrecht University

April 2004

Abstract
The perpetual inventory method used for the construction of education data per country leads to systematic measurement error. This paper analyses the effect of this measurement error on GDP regressions. There is a systematic difference in the education level between census data and observations constructed from enrolment data. We discuss a methodology for correcting the measurement error. The standard attenuation bias suggests that using these corrected data would lead to a higher coe.cient. Our regressions reveal the opposite. We discuss why the measurement error yields an overestimation. Our analysis contributes to an explanation of the difference between regressions based on 5 and on 10 year first—differences.

Keywords: growth, education, measurement error

JEL classification: I2, O4

Acknowledgements
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, the participants of the Tinbergen Institute Seminar, and also the participants at the Microeconometrics Conference at CEMFI for helpful comments

1 Introduction

Measurement error in education is widely recognized has an important source of bias in growth regressions; see for example Krueger and Lindahl (2001). This paper shows that the way Barro and Lee (2001) constructed the education data yields a systematic error. Some data points are directly derived from census observations, other are derived from this previous census information, using 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. Previous attempts to correct for this error have either only been successful for a limited number of countries or were based on arbitrary corrections made by the researcher (see de la Fuente and Doménech, 2002, and Cohen and Soto, 2001). Our analysis leads to 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 great practical relevance for the interpretation of the relation between education and GDP. There are two main approaches: (i) Nelson and Phelps (1966), and (ii) Lucas (1988). In the former, human capital is crucial to innovate and adopt new technologies. Hence, the *growth* rate of output is determined by the *level* of human capital. In the latter, human capital is interpreted as a normal input in the production process. Hence, *changes* in output are determined by *changes* in the human capital stock. 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' argument that growth is driven by the stock of human capital catch up. Krueger and Lindahl (2001) argue that these conclusions are highly affected by measurement error in the average education of countries. The problems of measurement error are exacerbated when taking first differences. First differences reduce the signal and increase the noise. Hence, the signal - noise ratio falls dramatically by first differencing. Krueger and Lindahl's solution to this problem is to increase the differencing period from 5 to 10, or 20, years, 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," finding empirical evidence in favour of Lucas' argument. However, the problem with this conclusion is that Nelson and Phelps (1966) model would lead to exactly the same conclusion. When the *level* of education affects output *growth*, then the effect of the *level* of education on output increases linearly with the differencing period, and hence, so does the effect of the *change* in education.

From our analysis, first we conclude that measurement error in education data is important. We find large and highly significant differences between data points based directly on census information and data updated with the perpetual inventory method. One would expect that these differences have large effect on growth regressions, in particular where differencing exacerbates the problem, in particular when using 5 year differences. 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. This turns out not to be the case. Using our corrected measure of education reduces the coefficient on changes in education. This runs counter to the standard argument of contamination bias, which is supposed to lead to lower coefficients when using data spoiled by measurement error. The reason for this paradox is that, in the standard model, measurement error increases the variance of the explanatory variable, since the measurement error is supposed to be orthogonal to the signal. In this case, the measurement error decreases the variance, since the perpetual inventory method smoothes observations at the beginning and the end of the observation period, thereby compressing the data. However, our exercise contributes to the explanation of the differences in the coefficient of the change in education based on a 5 and a 10 year differencing period, and compares to Krueger and Lindahl (2001). All this leads to the inevitable conclusion, previously obtained by Teulings and Van Rens (2003), that education has a moderate immediate effect on GDP of about $4.2 - 6.5\%$, but a huge long run effect of about $50 - 66\%$, which however takes ages to materialize, the half value time being $70 - 100$ years. This conclusion is obviously conditional on the identifying assumption that is used in whole this literature till so far, that current innovations in GDP have no effect on current innovations in education.

The paper is organized as follows. In the next section we describe different sources of data on education. Then we will analyse how systematic is the difference between census and non-census data. In Section 4 we will concentrate on the interaction between education and growth using the new knowledge on education, comparing the results with known figures. Finally we conclude.

2 Sources of data on education

The most used data set on international education attainment is the one released by Barro and Lee (2001).[1] They build their data on educational attainment from census or survey data. When this information is not available, the authors use a perpetual inventory method based on enrolment data in order to generate either a forward-flow, or a backward-flow. The flows are constructed from the benchmark stocks defined by the census or survey data. For intermediate observations, the constructed data point is a weighted average of the forward-flow and the interpolation between two benchmarks. For the observations before the first and after the last census or survey, interpolation is infeasible. Then, the constructed data apply by either the forward- or the backward-flow to the closest available census or survey data point. The enrolment data are adjusted for repeaters and changes in the duration of years of schooling.

Barro and Lee's data received criticism. de la Fuente and Doménech (2002) construct a revised version of the Barro and Lee (1996) data set for a sample of 21 OECD countries "using previously unexplored sources and following a heuristic approach to obtain plausible time profiles by removing sharp breaks in the data that seem to reflect changes in classification criteria"(de la Fuente and Doménech, 2002) to "avoid unreasonable jumps in the series by choosing the most plausible figure when several are available for the same year, and by reinterpreting some of the data"(de la Fuente and Doménech, 2002).[2] Missing

[1]Alternative sources are Kyriacou (1991) and Nehru *et al.* (1995). The latter ignores census data. De la Fuente and Doménech (2002) criticise this choice, and argue that it is difficult to justify "discarding the only direct information available on the variables of interest."

[2]These two data sets are not directly comparable since Barro and Lee's data is based on people having completed some educational level, while de la Fuente and Doménech's data applies to people who have

observations are filled in, if possible by interpolation, or otherwise by back- and forward projections. The authors "avoided the use of flow estimates based on enrolment data because they seem to produce implausible time profiles"(de la Fuente and Doménech, 2002). The authors state that "the construction of our series involves a fair amount of guesswork," and that their data "look more plausible than most existing series, at least in terms of their time profile."

Cohen and Soto (2001) extend the work of de la Fuente and Doménech (2002) to several other countries. An important difference to de la Fuente and Doménech is that Cohen and Soto allow for the use of enrolment data when needed. The authors have constructed a data set for 95 countries with information on education achievement from 1960 to 2000, for ten year interval, plus a projection for 2010. Their methodology is to "minimize the extrapolations and keep the data as close as possible to those directly available from national censuses"(Cohen and Soto, 2001). They argue that some of the differences between their data and the one provided by Barro and Lee (2001) can be explained by: (i) divergences in classification; (ii) the use of more census information than Barro and Lee; (iii) the use of a different methodology for extrapolating the missing data; (iv) errors in Barro and Lee data.

The conclusion is that, in spite of the improvements in data, so far measurement error in education data remains a problem. The Barro and Lee data are highly erratic. For example, in many cases, the average education level decreases over time within countries, which does not fit casual observation. de la Fuente and Doménech's data is a valuable effort, but requires a large amount of ad hoc decisions and is only available for a sample of 21 countries. Cohen and Soto's data increases the countries sample size, but is only available on 10-year intervals. Also, both these data sets face the criticism that measurement error problems were not entirely solved. Finally, Kyriacou data is very problematic given the estimation procedure used,[3] and the fact that it is only available for the period 1965–1985.

3 How systematic is the difference between census and non-census data?

3.1 Origins and identification of the systematic difference

The hypothesis we will test is that the methodology used to impute missing values in the Barro and Lee data underestimates the true values of education. This underestimation results from the assumption that the survival rate is independent of the educational level. In their own words, Barro and Lee (1993, p.374) state that "some error is introduced (...) if educational attainment is growing rapidly, because the older people then have less human capital and a greater probability of dying." If average education within a country is rising, as it seems to be the case for an important portion of the countries, the implication would be an underestimation of the educational attainment. The increase

attended some educational level.

[3]Kyriacou assumes that the relationship between average years of schooling in the labour force and the enrolment ratios in primary, secondary and higher education is relatively constant over time and across countries.

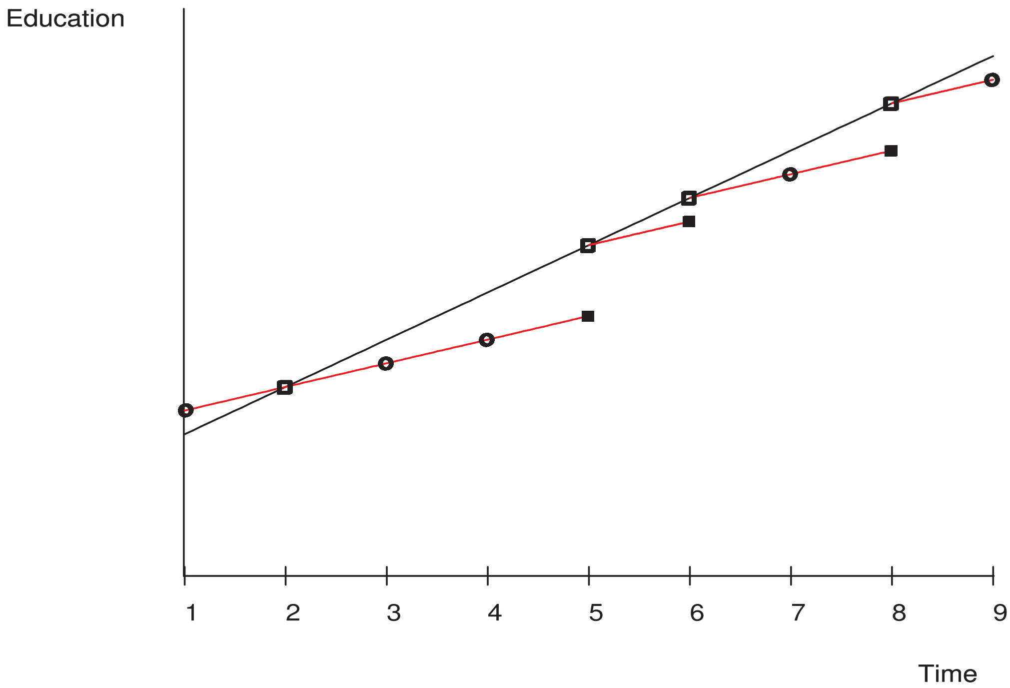

Figure 1: Plot of education with census and non-census data

in the schooling level of a population occurs mainly because the younger generations are more educated. In this case, the estimation procedure underestimates the survival of more educated individuals, resulting in a lower attainment for the country as a whole. The same idea is identified in Barro and Lee (2001, p.545), when they say that "in a typical country in which educational attainment is growing, mortality would be higher for the older people who are less educated. Then the assumption of uniform mortality can cause a downward bias in the estimation of the total educational stock."

If this is true, we should observe in the data that: (i) the increase in education between two consecutive census observations should be higher than the between non-census observations, and (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 time dimension, while the vertical axis plots the average education level in each period. The steeper and darker line represents the evolution of true education. For simplicity, we assume that true education follows a constant trend. The observations represented by an empty square, located in this line, represent the census information available. The circle dots represent the estimated points using the enrolment data and the benchmark census information. We also assume that the estimation process leads to a constant trend, that underestimates the true value. The lighter line represents this. The filled square dots represent the values of education that would be estimated for periods in which we have census data.

The change in education from period 4 (the empty circle in period 4) to period 5 (the empty square in period 5) can be decomposed as the variation predicted by the perpetual inventory method (the filled square dot over the lighter line in period 5), plus the accumulated errors since period 2, originated by the underestimation. The jump between the non-census (hypothetical) and census data points in period 5 (the difference between the filled square dot and the empty square dot) is proportional to the time elapsed since the previous census. In period 3 the error is given by the distance between the empty circle and the steeper line. In period 4 the difference between the empty circle and the steeper line gives the accumulated error in period 3 and 4. The error specific to period 4 can be retrieved if we imagine a non-census trend line departing from the true education value in period 3.

Barro and Lee's procedure implies that missing values are constructed differently according to the type of observation: (i) observations before the first census; (ii) observations between two census observations; (iii) observations after the last census. Our empirical strategy test for systematic differences between census and non-census observations, where we take into account for the differences between these three types. We constructed four variables. $Before$ applies to type (i) observations; it measures the lag till the first census. *Last* and *LastC* apply to type (ii) observations; the first records the number of periods since the previous census, while the second also records the lag till the previous census, but just for census data points, being zero otherwise. $After$ applies to type (iii) and measures the lag till the last census.[4] These variables adequately cover the hypothesised bias introduced by Barro and Lee's procedure as depicted in Figure 1. If we used just a dummy for non-census observation instead, then its coefficient would be a weighted average of the changes associated with different lags till the previous census. Moreover, it would not have differentiated between the positive bias for type (i) observations and the negative bias for type (ii) and (iii) observations.

3.2 Data description

Table 1 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 *Census* variable as described above. The income variable is Real GDP per worker, and is obtained from the Penn World Table 6.1. All variables are available on five year intervals, between 1960 and 2000. Average income increased by 4% per decade, while average education increased by 0.70 year, achieving 6.33 years in 2000. Its dispersion as been relatively stable over time, with a slight increase in the beginning of the sample period. Only 32% of the information on education is based on census/survey data. There is a concentration of census information at the start of each decades, 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 less census observations, and only 26% have 4 or more. Finally, the distribution of countries per period is relatively balanced.

3.3 Empirical evidence

Consider the following model:

$$Edu_{it} = \gamma_t + \beta_1 Before_{it} + \beta_2 Last_{it} + \beta_3 LastC_{it} + \beta_4 After_{it} + \eta_i + \varepsilon_{it} \quad (1)$$

[4] See Table 8 for an example of these variables corresponding to Figure 1.

Table 1: Summary Statistics – 5 year 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	64	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
Δ Edu	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	9.21
	Std.Dev.	0.95	0.99	1.00	1.04	1.02	1.07	1.11	1.13	1.05
Observations		85	89	92	94	95	97	97	89	821
Δ LGDP	Mean	0.03	0.03	0.02	0.02	0.01	0.01	0.01	0.02	0.02
	Std.Dev.	0.02	0.03	0.03	0.03	0.03	0.03	0.04	0.02	0.03
Observations		83	85	89	92	94	95	97	87	722
		Distribution of Census per Country								
Census		1	2	3	4	5	6	7	8	
Countries		25	28	33	22	6	0	1	1	116

Note: the summary statistics for Δ Edu are for changes over five year periods, while for Δ LGDP the statistics are for annualised changes.

where Edu_{it} is the education level of country i, in period t, γ_t is the specific effect for period t, η_i is country i's specific effect, and ε_{it} is a white noise error term. Taking first-differences eliminates the fixed country effect:

$$\Delta Edu_{it} = \gamma_t + \beta_1 \Delta Before_{it} + \beta_2 \Delta Last_{it} + \beta_3 \Delta LastC_{it} + \beta_4 \Delta After_{it} + \Delta\varepsilon_{it} \qquad (2)$$

where Δ is the first difference operator. Estimation results are presented in Table 2. In column 1 we report the estimation of equation (1), using the fixed-effects estimator. Columns 2 and 3 report the results for equation (2), where we use OLS in column 2 and fixed effects in column 3. For the model in levels in column 1 the hypothesis of the absence of country specific effects is rejected. We also reject the hypothesis that the level error terms are not serially correlated. When we apply the fixed-effects estimator to the first-differences specification (column 3) we do not reject the hypothesis that all the country specific effects are equal to zero. Hence, the first differences without fixed effects (column 2) is the preferred estimation. The subsequent discussion is restricted to this model.

Table 2: Education Regressions

	Levels	First-differences	
Variable	Fixed Effects	OLS	Fixed Effects
Before	0.391**	0.250**	0.140*
	(0.072)	(0.055)	(0.065)
Last	-0.200**	-0.198**	-0.186**
	(0.032)	(0.027)	(0.031)
LastC	0.199**	0.202**	0.193**
	(0.032)	(0.029)	(0.029)
After	-0.214**	-0.272**	-0.316**
	(0.057)	(0.056)	(0.085)
Wald joint	83.008**	70.190**	57.522**
Wald time	610.577**	566.373**	39.747
F-Test; H_0: all $\eta_i = 0$	251.61**		1.09
AR(1) test	5.849**	-1.070	-4.611**
Observations	985	869	869
Countries	116	116	116

Significance levels: † : 10% ∗ : 5% ∗∗ : 1%. (Robust standard errors in parentheses). All regressions include time effects.

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 identical in absolute value, as predicted. Furthermore, the coefficient on $After$ and $Before$ are larger in absolute value than the coefficient on *Last*. This too fits our hypothesis. Since the type (ii) observations are a weighted average of interpolation between neighbouring census observations and a forward perpetual inventory method, while type (i) and (iii) are fully based on the perpetual inventory method, the bias is larger for the latter group of observations. The magnitude of the bias is huge, some 0.20 year 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 bias in the data. Given the fact that many countries hold a census every 10 years (usually at the beginning of a decade), the systematic bias in the non-census observations yields a particular erratic time series of first differences when using a 5 year period.

3.4 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 by the subsequent expression

$$PEdu_{it} = Edu_{it} - \beta_1 Before_{it} - \beta_2 Last_{it} - \beta_3 LastC_{it} - \beta_4 After_{it} \tag{3}$$

where $PEdu_{it}$ is the corrected education variable. The coefficients are the ones from column 2 of Table 2.

Table 3: 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)

Number of observations in parentheses.

Tables 3 and 4 give the correlations between the various education variables, Table 3 in levels and Table 4 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 (2001) ($EduCS$) and de la Fuente and Doménech (2002) ($EduDD$) is only slightly lower. The mean and the variance of Barro and Lee data are the lowest of all four. 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 less extent, a similar objection can be raised against a comparison to the Cohen and Soto data, where 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 bias in the fill in procedure in the Barro and Lee data leads to an underestimation of the average education level by 0.25 year. Even more importantly, it leads to an underestimation. An eye on Figure 1 reveals why this is the case. The bias understates the final observations, but overstates the initial observations, leading 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 the measurement error compresses the variance.

In first differences, the correlations between education variable are much lower. The correlation between Barro and Lee and our corrected variable is still high, 0.88. 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 bias in Barro and Lee compresses the measured average growth of education substantially, from 1.00 year per decade to 0.70 year. 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.

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. Between census (1960–1990), our procedure (corrected Education) smoothes the data. However, for observations after the last census available (1990), the constant correction induces an higher 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

Table 4: Correlations among Education Measures in First-Differences

	DEdu	DPEdu	DEduCS	DEduDD	Mean	Variance
DEdu	1				0.350	0.192
	(869)				(869)	(869)
DPEdu	0.883	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	0.020	0.391	1	0.376	0.020
	(135)	(135)	(60)	(135)	(135)	(135)

Number of observations in parentheses.

year differentials are entirely dominated by the difference between census and non-census observations.

4 Growth regressions: what changes?

4.1 OLS estimations

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

$$\Delta LGDP_{it} = \gamma_t + \alpha LGDP_{i,t-\tau} + \beta_1 \Delta Edu_{it} + \beta_2 Edu_{i,t-\tau} + \varepsilon_{it} \tag{4}$$

where γ_t are time specific effects, $LGDP_{it}$ stands for real log 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 ε_{it} is an i.i.d. error term. All variables in changes are annualised.

The first two columns of Table 5 reproduce estimations from Topel (1999, Table 4), and Krueger and Lindahl (2001, Table 3) [K&L(2001)]. The remaining results 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. Just for the 10 year data, the last column of Table 5 reproduces the estimations using the data from Cohen and Soto (2001). The estimation procedure is OLS, and we report standard errors robust to heteroskedasticity and error correlation within countries.

Similarly to Topel (1999) and Krueger and Lindahl (2001), we also conclude that contemporaneous 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, our results indicate that the short run return do changes in education is around 5%, while Topel (1999) and Krueger and Lindahl (2001) results points approximately to 4%. We also conclude that the returns to changes in schooling increases with the time span of the data. Krueger and Lindahl suggest that "the finding that the time span matters so much for the change in education suggests that measurement error in schooling influence these estimates" (p.1119). Our interpretation is that it is the measurement error introduced by the estimation procedure implemented by Barro and Lee that leads to this variation, not the measurement error

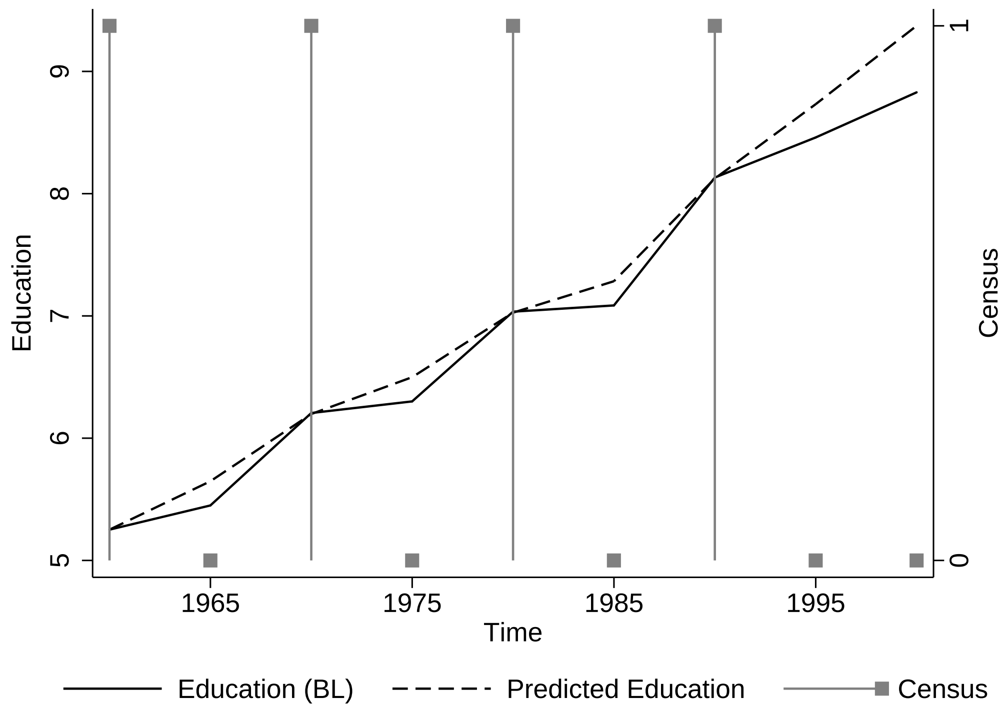

Figure 2: Education information for Argentina

inherent to the census observations. Approximately 42% of the observations on education in the 10 year data are obtained from census or surveys, while in the 5 year data this figure is only 32%. This difference in data quality is associated with a smaller spurious variation in education in the 10 year data, which may explain way the coefficient on ΔEdu increases with the time span.

The most remarkable feature is that the return to education is lower for our corrected data than for the original data, that is shown to be systematically biased. The standard attenuation bias argument tells that measurement error in an explanatory variable reduces its coefficient, *quod non* in this case. The coefficients for our corrected variable are lower instead of higher, for the 5 year estimation, but in particular for the 10 year estimation. A second thought reveals the reason for this phenomenon. The measurement error reduces the mean of the change in education by some 15% (see Table 4). Where the estimated coefficient for the corrected variable is some 6% lower (for the 5 year interval), the estimated effect of education on GDP is $15 - 6 = 9\%$ larger. So, the effect of the bias on the coefficient is a balance between two forces: introducing the spurious component in ΔEdu reduces the coefficient, 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.[5] When we use Barro

[5] 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

Table 5: The Effect of Education on Growth – annualised OLS estimations

	5 year data				
Variable	Topel(1999)	K&L(2001)	Edu	PEdu	EduCS
ΔEdu	0.041**	0.039**	0.0517**	0.0486**	
	(0.014)	(0.014)	(0.0135)	(0.0144)	
LagEdu	0.004**	0.004**	0.0035**	0.0037**	
	(0.001)	(0.001)	(0.0009)	(0.0010)	
LagGDP	-0.007**	-0.006*	-0.0060**	-0.0063**	
	(0.002)	(0.003)	(0.0022)	(0.0023)	
R^2	0.218	0.197	0.1315	0.1295	
Observations	608	607	722	722	
Countries	111	110	97	97	
	10 year data				
ΔEdu	0.085**	0.086**	0.0882**	0.0781**	0.1107**
	(0.020)	(0.024)	(0.0210)	(0.0218)	(0.0336)
LagEdu	0.004**	0.004**	0.0039**	0.0041**	0.0032**
	(0.001)	(0.001)	(0.0009)	(0.0009)	(0.0009)
LagGDP	-0.007**	-0.005†	-0.0073**	-0.0076**	-0.0065*
	(0.002)	(0.003)	(0.0021)	(0.0022)	(0.0026)
R^2	0.315	0.284	0.2336	0.2236	0.2204
Observations	290	292	353	353	300
Countries	111	110	97	97	79

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 Log Real GDP per Worker. All regressions include time effects. The results under EduCS use data for education from Cohen and Soto (2001).

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 7.8%. 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.[6] While for Topel, and Krueger and Lindahl, 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

later on have, are countries that are likely to have grown faster than average.

[6]Suppose our model is defined as $y_t = \beta_1 + \beta_2 x_t^* + \varepsilon_t$, where the observed value of x is $x_t = x_t^* + \phi t + u_t$, and x is underestimated at a constant rate ϕ over time. The model that will be estimated based on the observed variables is $y_t = \beta_1 + \beta_2 x_t + \varepsilon_t - \beta_2 \phi t - \beta_2 u_t$. A general result from error in variables models is that the inconsistency in the estimation of β_2 is given by $\frac{cov(x_t, \varepsilon_t - \beta_2 \phi t - \beta_2 u_t)}{var(x_t)} = \frac{-\beta_2 \phi \sigma_t - \beta_2 \sigma_u}{\sigma_x}$. Given that ϕ is negative by definition, the traditional downward bias will be smaller, and it may even be an upward bias.

for different time spans.

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, the long run return to education is $54-59\%$.[7] 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 value after $76-99$ years. The immediate return is 4.1% for the 5 year time period and 6.4% for the 10 year period.[8] 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

[7] 0.0041/0.0076 or 0.0037/0.0063, respectively.

[8] 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–life

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

Second, we correct for the fact that part of 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

$$0.0486 = \frac{0.0037}{0.0063}\left(1 - e^{-0.0063t}\right).$$

So, the time needed to reach the immediate effect is

$$\frac{\ln\left(1 - 0.0486\frac{0.0063}{0.0037}\right)}{-0.0063} = 13.7106$$

Finally, we take into account for 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-life is given by

$$110.0234 - 13.7106 + 2.5 = 98.8128$$

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,

$$0.0063 * (13.7106 - 2.5) = 0.0706.$$

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

$$0.0706 \times \frac{0.0037}{0.0063} = 4.1\%.$$

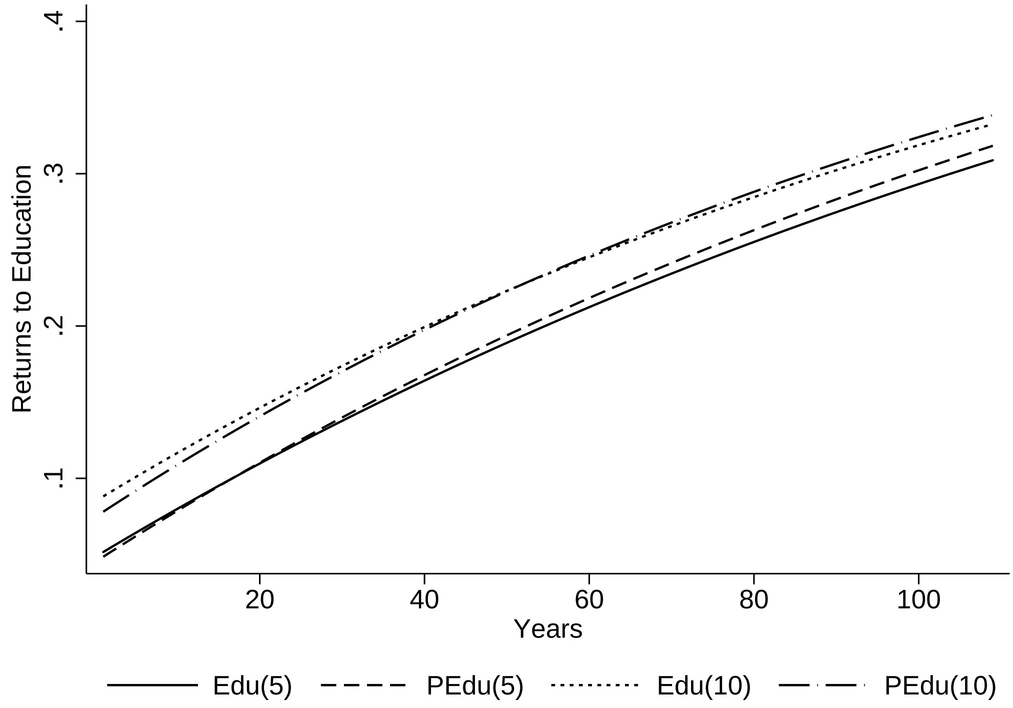

Figure 3: Returns to education for the different time spans and education variables

education is very similar for the two time spans, and for the two education variables.

The results indicate that the GDP half–life adjustment ranges between 91 and 110 years,[9] 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.

Taking the last column of Table 5, the results for the 10 year data of Cohen and Soto differ from the results with Barro and Lee's data essentially on the coefficient on changes in education. In this case, the contemporaneous returns to education are around 11%. However, we need to take into account for the fact that the countries used for the estimations under $EduCS$ are a subsample of those used in Edu.[10]

4.2 Sensitivity analysis: the dynamic model

Adjusting the annualised equation (4), we can define the following dynamic model

$$LGDP_{it} = \tau\gamma_t + (1+\tau\alpha)\, LGDP_{i,t-\tau} + \beta_1 \Delta Edu_{it} + \tau\beta_2 Edu_{i,t-\tau} + \tau\eta_i + \tau\varepsilon_{it}, \qquad (5)$$

[9] $\ln(2)/.0076$ and $\ln(2)/.0063$, respectively.

[10] In our analysis, we are lacking 18 countries in Cohen and Soto 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.

where η_i is the country's specific effect.

In the presence of country's specific effect in equation (5), its estimation by OLS and by the usual panel models, fixed or random effects, is inconsistent. The reason is that, by definition, $LGDP_{i,t-\tau}$ in equation (5) is always correlated with η_i. One possible solution to overcome this problem is to take first differences in equation (5) to eliminate η_i. Arellano and Bond (1991) first-differenced generalized method of moments (GMM) is one of the most applied solutions. Using their procedure avoids the bias introduced by omitted time-invariant variables. However, this solution has poor finite sample properties on bias and precision when "the lagged levels of the series are only weakly correlated with subsequent first-differences, so that the instruments available for the first-differenced equations are weak" (Bond *et al.*, 2001). Blundell and Bond (1998) show that, in this case, the solution of Arellano and Bond (1991) has a large downward finite-sample bias. This problem occurs when the time series are persistent and the number of time series observations is small. An alternative solution would be to implement a system GMM estimation, for first-differences and levels, as argued by Blundell and Bond (1998). Bond *et al.* (2001) argue that this is the best solution to estimate growth regressions.

Table 6: Dynamic Income Regressions – 10 year data

Model	Sys	FD	No-FE	Sys	FD	No-FE
Variable		Edu			PEdu	
LagLGDP	0.780**	0.843**	0.947**	0.903**	0.839**	0.941**
	(0.116)	(0.201)	(0.026)	(0.115)	(0.171)	(0.028)
ΔEdu	0.132**	-0.079	0.073**	0.098**	-0.092	0.072**
	(0.036)	(0.163)	(0.020)	(0.035)	(0.125)	(0.023)
LagEdu	0.133**	-0.097	0.038**	0.098*	-0.109	0.039**
	(0.041)	(0.163)	(0.011)	(0.049)	(0.131)	(0.012)
Wald joint	637.482**	43.602**	6660.310**	655.610**	36.842**	6195.831**
Wald time	104.004**	24.954**	116.953**	91.046**	22.564**	104.380**
Sargan	23.822	13.490	23.623	28.411†	13.467	25.268
Sargan-df	19	12	21	19	12	21
DifSargan	10.333		10.134	14.944*		11.801
AR(1) test	-3.166**	-2.450*	-3.778**	-3.349**	-2.658**	-3.698**
AR(2) test	1.557	0.850	1.482	1.227	1.000	1.303
Observations	350	256	350	350	256	350
Countries	94	94	94	94	94	94

Significance levels: † : 10% ∗ : 5% ∗∗ : 1%. (Robust standard errors in parentheses). The dependent variable is LGDP. All regressions include time effects. The estimation details, including the instrumens used, are described in Section 4.2.

Table 6, and Table 7 in appendix, present the results of the estimation of equation (5) using the data for 10 and 5 year intervals, respectively.[11] For each education variable,

[11]The results in Tables 6 and 7 are directly comparable with the results in Table 5 once we control for the time span of the data. The transformations of the dynamic estimates follow from equation (5).

Edu and $PEdu$, and for each time span, we estimate equation (5) using the system procedure (Sys), the first-differenced procedure (FD), and a system procedure which assumes the absence of a fixed country effect (No-FE). The dependent variable is real log income per worker ($LGDP$), all regressions include time dummies, and education is treated as a predetermined variable. The instruments for the first-difference equations are the level of $LGDP$ lagged two periods and earlier, and levels of education lagged one period and earlier. For both variables we use at most 5 lags, following Bowsher (2002)'s suggestion. For the level equations we use first-difference of $LGDP$ lagged one period, and contemporaneous first-difference of education. The estimation of No-FE is similar to system estimation, but in this case the instruments for the equations in levels are not in first-differences but in levels. The reported results are for the 2-step GMM estimation procedure, following the correction proposed by Windmeijer (2000).[12]

We test for the presence of the specific effect following the procedure described in Arellano (2003, p.124).[13] The statistic of the test is the difference in the Sargan test associated with the estimations FD and No-FE, which follows a chi-squared distribution with the number of degrees of freedom given by the difference in the number of instruments in the referred two estimations. We are testing the validity of the additional set of instruments, when compared with the FD estimation. Our results indicate that we do not reject the null hypothesis; i.e., we do not reject the hypothesis that there is no specific country effect.[14] This implies that the results we are lead to interpret are the ones in Table 5. OLS estimates are consistent in the absence of unobserved heterogeneity, and they are more efficient. Using equation (5) to compare the estimates, we observe that our results for the NO-FE model using $PEdu$ and 10 year data are very similar to the corresponding results reproduced in Table 5. The returns to contemporaneous changes in education are 7.2%, while the coefficient on lag education is 0.004, and the coefficient on lag income is 0.006. The comparable figures from the OLS estimation are 7.8%, 0.004 and 0.008, respectively.

Although the results for the estimation of the model Sys are very different when we use Edu and $PEdu$, they become identical when we estimate the model NO-FE. Using $PEdu$ matters when we compare the Sys and the NO-FE estimation, since the results are more similar. In the 5 year data, Table 7 in appendix, the results for the system estimation are unreliable with a coefficient on lag income above one. Again, the results for the estimation of NO-FE are similar between the two education variables, with the exception of the coefficient on ΔEdu. Using Edu indicates that the impact multiplier of one year change in education is 4.7%, while using $PEdu$ the equivalent value is 3.7%. As before, the results for the 10 year data seem to be more stable.

5 Final remarks

Our analysis of Barro and Lee (2001) education data reveals that there is a systematic difference between the information collected from census or surveys, and the education data

[12] We used the Ox version of DPD (Doornik et al, 2002) to obtain the results in Tables 6 and 7.

[13] Further details on the test for specific effects are discussed in Holtz-Eakin (1988) and Arellano (1993).

[14] The No-FE type of regression has two more instruments when compared with the system estimation. The reason is that in the first case we use an extra period in the level equations.

that results from the perpetual inventory method. On average, this method underestimates education by about one fifth of a year every five year period. This has an impact on the results for the growth regressions. Once we control for the source of information, and we take into account for measurement error, we conclude that both the level and the change in education are relevant for the growth process. However, alternative specifications and data intervals makes a difference for the size of the effects. Further research is need in order to make proper use of the knowledge on the systematic difference between census and non-census data.

Following Teulings and van Rens (2003), it would be important to take into account for second order effects on education. The re-estimation of the data on education is another alternative for future work. Using both the backward and the forward flow, the missing values can be reestimated using the average of both predictions, not only the weighted average between the linear interpolation and the forward prediction. However, the estimation of the missing values after the last census, and before the first census, would still be estimated the same way. It would also be important to estimate educational values taking into account for different survival rates according to the educational attainment. On this topic, Barro and Lee (2001) state that "the limitation of the data on age-specific education levels and mortality rates by age group do not allow us to compute specific mortality rates of population by levels of education.

"

References

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

Arellano, Manuel (1993), 'On the testing of correlated effects with panel data', *Journal of Econometrics* **59**(1-2), 87–97.

Arellano, Manuel (2003), *Panel Data Econometrics*, Oxford University Press, New York.

Arellano, Manuel and Olimpia Bover (1995), 'Another look at the instrumental variable estimation of error-components models', *Journal of Econometrics* **68**(1), 29–51.

Arellano, Manuel and Stephen Bond (1991), 'Some tests of specification for panel data: Monte carlo evidence and an application to employment equations', *Review of Economic Studies* **58**(2), 277–297.

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

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

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

Barro, Robert J. and Xavier Sala-I-Martin (1999), *Economic Growth*, MIT Press, Cambridge, Massachusetts.

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

Blundell, Richard and Stephen Bond (1998), 'Initial conditions and moment restrictions in dynamic panel data models', *Journal of Econometrics* **87**(1), 115–143.

Bond, Stephen R., Anke Hoeffler and Jonathan Temple (2001), 'GMM estimation of empirical growth models', Centre for Economic Policy Research Discussion Paper No. 3048.

Bowsher, Clive G. (2002), 'On testing overidentifying restrictions in dynamic panel data models', *Economics Letters* **77**(2), 211–220.

Cohen, Daniel and Marcelo Soto (2001), 'Growth and human capital: Good data, good results', OECD Development Centre Technical Paper No. 179.

de la Fuente, Angel and Rafael Doménech (2001), 'Schooling data, technological change, and the neoclassical model', *American Economic Review* **91**(2), 323–327.

de la Fuente, Angel and Rafael Doménech (2002), 'Human capital in growth regressions: How much difference does data quality make?', Working paper, Instituto de Análisis Económico, Barcelona.

Doornik, J.A., M. Arellano and S. Bond (2002), *Panel Data Estimation Using DPD for Ox*, Oxford, London. www.nuff.ox.ac.uk/Users/Doornik.

Harvey, Andrew (1990), *The Econometric Analysis of Time Series*, London School of Economics Handbooks in Economics, second edn, Philip Allan, London.

Heckman, James J. and Peter J. Klenow (1997), 'Human capital policy', University of Chicago Working Paper.

Holtz-Eakin, Douglas (1988), 'Testing for individual effects in autoregressive models', *Journal of Econometrics* **39**(3), 297–307.

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

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

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

Mulligan, Casey B. and Xavier Sala-I-Martin (2000), 'Measuring aggregate human capital', *Journal of Economic Growth* **5**(3), 215–252.

Nehru, V., E. Swanson and A. Dubey (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, Richard and Edmund Phelps (1966), 'Investment in humans, technology diffusion and economic growth', *American Economic Review* **56**(2), 69–75.

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

Summers, Robert and Alan Heston (1991), 'The penn world table (mark 5): An expanded set of international comparisons, 1950-1988', *Quarterly Journal of Economics* **106**(2), 280–315. (the new dataset, mark 6.1, is available at http://pwt.econ.upenn.edu/).

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

Teulings, Coen and Thijs Van Rens (2003), 'Education, growth, and income inequality', Tinbergen Institute Discussion Paper 02-001/3.

Topel, Robert (1999), *Labor Markets and Economic Growth*, Orley Ashenfelter and David Card, eds., Amsterdam: North Holland, chapter 44, pp. 2943–2984.

Verbeek, Marno (2000), *A Guide to Modern Econometrics*, Wiley, Chichester.

Windmeijer, Frank (2000), 'A finite sample correction for the variance of linear two-step GMM estimators'. IFS working paper W00/19 (www.ifs.org), London: The Institute for Fiscal Studies.

Woessmann, Ludger (2003), 'Specifying human capital', *Journal of Economic Surveys* **17**(3), 239–270.

A Appendix

Table 7: Dynamic Income Regressions – 5 year data

Model	Sys	FD	No-FE	Sys	FD	No-FE
Variable		Edu			PEdu	
LagLGDP	1.038**	0.916**	0.973**	1.038**	0.926**	0.973**
	(0.054)	(0.105)	(0.011)	(0.050)	(0.093)	(0.011)
ΔEdu	0.037†	0.025	0.047**	0.035	0.007	0.037*
	(0.020)	(0.034)	(0.016)	(0.021)	(0.044)	(0.016)
LagEdu	0.015	0.008	0.017**	0.015	-0.001	0.017**
	(0.018)	(0.032)	(0.005)	(0.018)	(0.041)	(0.005)
Wald joint	1823.20**	83.72**	25416.69**	1967.18**	104.96**	25952.05**
Wald time	103.707**	19.969**	89.606**	69.546**	22.110**	82.449**
Sargan	76.094	61.642†	73.249	73.116	66.104*	74.052
Sargan-df	63	48	65	63	48	65
DifSargan	14.452		11.607	7.012		7.948
AR(1) test	-3.576**	-3.387**	-3.611**	-3.577**	-3.447**	-3.603**
AR(2) test	1.080	1.068	1.089	1.050	1.030	1.053
Observations	722	625	722	722	625	722
Countries	97	97	97	97	97	97

Significance levels: † : 10% ∗ : 5% ∗∗ : 1%. (Robust standard errors in parentheses). The dependent variable is LGDP. All regressions include time effects. The estimation details, including the instrumens used, are described in Section 4.2.

Table 8: Example of Census Variables for Figure 1

Year	$Census$	$Before$	$Last$	$LastC$	$After$
1960	0	1	0	0	0
1965	1	0	0	0	0
1970	0	0	1	0	0
1975	0	0	2	0	0
1980	1	0	3	3	0
1985	1	0	1	1	0
1990	0	0	1	0	0
1995	1	0	2	2	0
2000	0	0	0	0	1