

# The Dynamics of Age Structured Human Capital and Economic Growth

Jesús Crespo Cuaresma  
University of Innsbruck  
Innsbruck, Austria  
International Institute for Applied Systems Analysis,  
Laxenburg, Austria

Wolfgang Lutz  
Vienna University of Economics and Business,  
Vienna, Austria  
Vienna Institute of Demography, Austrian Academy of  
Sciences, Vienna, Austria  
International Institute for Applied Systems Analysis,  
Laxenburg, Austria

Warren Sanderson  
Stony Brook University  
Stony Brook, New York, USA  
International Institute for Applied Systems Analysis,  
Laxenburg, Austria

Paper presented at: New Directions in the Economic Analysis of Education  
Milton Friedman Institute, University of Chicago  
November 20-21, 2009

# The Dynamics of Age Structured Human Capital and Economic Growth

Jesús Crespo Cuaresma

Wolfgang Lutz

Warren Sanderson

## Abstract

We show that the age structure of human capital matters for economic growth. This question had not been tackled empirically until very recently due to the lack of comparable cross-country data on age-specific educational attainment. We present the newly developed IIASA/VID dataset, which provides consistent information about educational attainment by age and sex for 120 countries for the period 1970-2000 at five year intervals. The dataset has been created making use of the information contained in censuses, labor market surveys, and the Demographic and Health Surveys, coupled with a reconstruction exercise based on multistate demographic methods of backward projection. Using the framework of the Benhabib-Spiegel (1994) model, we show that improvements in education in young age groups are particularly important for explaining economic growth in developing countries. Furthermore, we show that the use of education data disaggregated by age improves out-of-sample economic growth predictions significantly.

## 1. Introduction

Education affects economic growth, but not just through the average number of years of schooling experienced by the adult population. The thesis of this paper is that the demography of the stock of human capital – its age and attainment structure - also matters for economic growth.

Past studies of the effects of education on economic growth have been hamstrung by noisy and incomplete data. The Barro and Lee (2001) data provided information on the distributions of national populations by sex, mean years of schooling and levels of educational attainment for the age groups 15+ and 25+. Two subsequent datasets (de la Fuente and Domenech (2006) and Cohen and Soto (2007)) were created for the purpose of improving on the Barro and Lee data. All three use a variety of surveys and censuses supplemented with information on school enrollment rates. A detailed analysis of the three data sets, however, finds that there is little basis for choosing among them (Bosworth and Collins (2003)). The IIASA/VID educational attainment dataset (Lutz et al. (2007)) uses a different methodology to reconstruct educational attainment data and provides the joint distribution of population by age, sex, and educational attainment for 120 countries from 1970 to 2000 in five-year intervals. We show in this paper that making use of the information contained in this joint distribution is crucial for understanding the relationship between education and economic growth and we invite others to use this additional detail in their empirical studies as well.

The study of economic growth must start with the study of the people who produce it. They work with their own hands, design, build and operate the machines of production, and structure and run the institutions and markets that make growth possible. Julian Simon concluded that the size of the human population together with the technologies these people produce is the root cause of economic growth (Simon 2000). He rightly argues that people are the carriers of knowledge but then goes on to the more controversial assertion that since the discoveries of the past were produced by people the rate of discoveries must have been influenced by human numbers.

But people do not come as an amorphous mass. Not every member of a given population makes the same contribution to the economy. People differ greatly in terms of their ability and willingness to contribute, with labor force participation rates and skill levels being the two individual properties that have been most systematically recognized in the economics literature. Both of these factors are closely related to age. Here we define human capital – the topic of this seminar – as the number of people who are participating in the work force, are in sufficiently good health and are differentiated by their skill levels which is conveniently approximated by level of educational attainment. In the following global level analysis, where data on health and labor force participation are difficult to obtain, we only focus on the educational attainment aspect of human capital by age and sex.

Since the level of educational attainment is primarily a property associated with individuals, any change in the distribution of attainment categories in an aggregate population is the consequence of changes of the individuals over these categories. In a society closed to migration such changes can happen either through the process of educating more individuals (which typically happens at younger ages) or through the process of younger people moving up the age pyramid and replacing the elderly as they die off. An important complicating factor in this process is the fact that mortality rates tend to differ significantly by level of education. The extent of this differential varies from country to country. Torrey and Kingkade (1990), for instance, use Russian data to show how average education during the 1980s quickly approached that of the US primarily due to the fact that the less educated Russians died at much higher rates than the more educated ones.

When it comes to projecting populations by levels of educational attainment, one also has to consider the fact that almost universally more educated women have fewer children. For Ethiopia, for

example, recent DHS (Demographic and Health Surveys) data show that women without formal education have on average six children while those with secondary or higher have only two. This rather complex dynamics of changes in the population structure by age, sex and level of educational attainment can be appropriately addressed by demographic models and in particular the multi-state cohort component model which has been explicitly designed for such questions (Keyfitz (1985), Rogers (1975) and Rogers and Land (1982)). Since the object of interest here is the changing composition of people by level of education these demographic methods specifically designed for dealing with people seem to be clearly preferable to other methods – such as the perpetual inventory method – that were designed for dealing with the physical capital stock and have difficulty dealing with differential attrition rates, among other problems.

In this paper we present the data IIASA/VID data on educational attainment by age and sex and describe the methods used for the reconstruction exercise. Using demographic multi-state methods to model the dynamics of human capital formation, we can explain why earlier studies based on coarser definitions of human capital resulted in such a wide variety of findings and point the way toward more consistent findings based on a full consideration of the dynamics of age-structured human capital. In the second part of the paper, we use the model put forward by Benhabib and Spiegel (1994) to show empirically that the age structure of human capital matters for subsequent economic growth. Furthermore, we also present empirical evidence that out-of-sample economic growth predictions which exploit the age structure dimension of human capital data tend to be systematically more accurate than predictions based on educational attainment data aggregated by age.

Section 2 describes the IIASA/VID dataset and its extensions for forecasting. In Section 3, we use these data to provide new insights into the relationship between education and economic growth and assess the quality of education-based out-of-sample growth projections. Section 4 contains our concluding thoughts.

## **2. Modeling the dynamics of the educational composition of the population**

Traditionally, demographic analysis in general and modeling of population dynamics in particular has been confined to the two demographic dimensions age and sex. More recently, it has been suggested educational attainment should be added as a third dimension, since education seems to be such an important source of measurable population heterogeneity (in terms of differential fertility and mortality rates) that disregarding it would produce distorted results. Moreover there is increasing interest in the educational composition of the population per se since education is considered a key determinant of many things ranging from health to economic growth to the quality of institutions and democracy (Lutz (2009)).

When it comes to measuring education it is important to distinguish between stock and flow variables. The most common flow variables are school enrollment rates, which are published for most countries by UNESCO and other education agencies. When trying to derive any stock measures from such flow data one also has to be aware of the problem that in the settings of many developing countries it is advantageous for schools to report exaggerated enrollment figures to national authorities, which may lead to biased estimates of educational attainment for poorer countries.

The stock of human capital can be measured in several different ways depending on whether one also wants to include aspects of the quality of the education and actual skills held. Historical data on such quality assessments based on actual testing of adults are unfortunately only available for a limited number of (mostly OECD) countries. For a global level analysis, the stock can be measured by categories of highest educational attainment (here the UNESCO-ISCED categories have become the standard) as well as by mean years of schooling. The original empirical data mostly come in terms of

attainment distributions. The calculation of mean years of schooling requires additional country-specific assumptions.

For virtually every country in the world there is now enough information from censuses and surveys to estimate the educational structure of the population by age and sex for at least one recent point in time and from there to reconstruct the past dynamics of educational attainment change and to project and the likely paths of future changes. This is precisely how the IIASA/VID dataset was created. The starting point of the analysis is a distribution such as the one depicted for the Republic of Korea for the year 2000 in Figure 1. It gives a multi-state age pyramid with women on the right side and men on the left in 5-year age groups above age 15. The colors in each age group show the numbers of men and women without any formal education as well as with some primary, at least completed junior secondary and completed tertiary education. It is evident from the figure that younger cohorts are much better educated than older ones and among the older ones women are clearly less educated than men. This reflects the history of educational improvement in South Korea over the past decades.

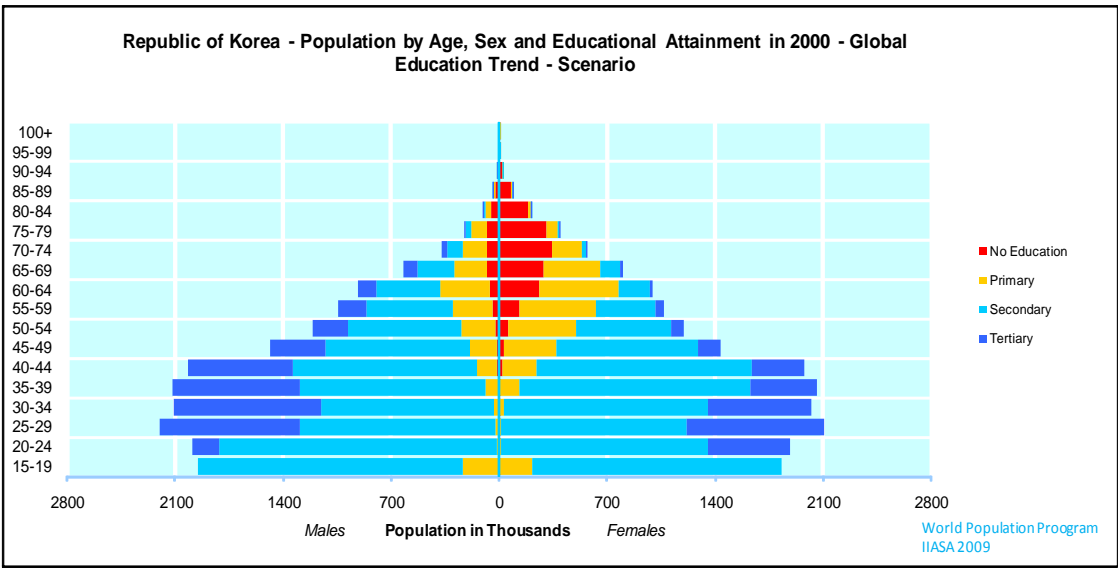


Figure 1: Population pyramid of South Korea in 2000 with colors indicating different educational attainment categories

Using this information by age and sex we can go backwards in time (in five year steps) until 1970, reaching the distribution depicted in Figure 2. The cohort aged 65-69 in 2000 is now aged 35-39 in 1970. It is bigger in size than in 2000 because some people have died over the course of the years but the educational attainment distribution is very similar because most people had received their final attainment level by age 35. In order to obtain the data depicted in Figure 2, we only have to adjust for the facts that people with higher education tend to have lower mortality and —where necessary— for the existence of educational migration differentials. For the highest age groups, for example those 80+ in 2000 and hence aged 50+ in 1970, additional assumptions have to be made about the trend of improving attainment, if we want more more age detail for the 50+ cohort in 1970.

Following this principle, the International Institute for Applied Systems Analysis (IIASA), in collaboration with the Vienna Institute of Demography (VID) of the Austrian Academy of Sciences, has recently produced a unique new dataset which applies demographic multi-state projection techniques to reconstruct the population by age, sex and level of educational attainment from empirical data from around 2000 back to 1970 in five-year steps. This has been done for 120 countries and is in detail documented in Lutz et al. (2007). This new data set is superior to other existing data sets for four reasons: (1) because of its detail (four educational categories for 5-year age groups of men and women), (2) because of the consideration of differential mortality, (3) because of

the strict consistency of the definition of educational categories over time, which is a major problem in empirical historical data sets, in which the underlying educational definitions often change, and (4) because of its natural extension to forecasting.

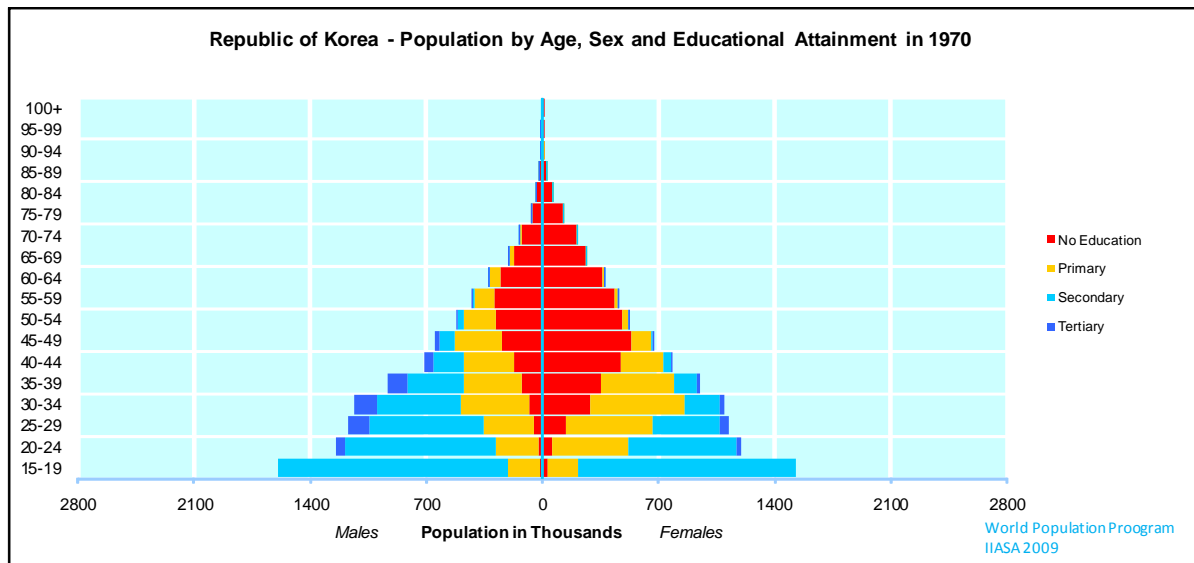


Figure 2: Reconstructed population pyramid of South Korea in 1970 with colors indicating different educational attainment categories

The same demographic multi-state approach can be applied to project the populations for all countries by age, sex and level of educational attainment. This follows the same principle based on the assumption that the highest educational attainment is stable after a certain age. The forward projection method requires assumptions on more parameters than in the case of reconstruction: in addition to the educational mortality differentials that had already been used for the reconstructions, future trends in education-specific fertility (and migration in the case of an open population) need to be assumed, as well as future trends in the age- and sex-specific probabilities of moving to higher education categories. By defining alternative scenarios concerning these future educational progression probabilities, IIASA has recently produced such projections of the population by levels of educational attainment to 2050 for the same set of countries in the world (KC et al. 2008).

The results of the global education trend scenario for South Korea for 2030 are depicted in Figure 3. This scenario assumes that the age-specific transition probabilities to higher educational categories for each country follow the trends that were observed over the last decade. Hence, in practice, this implies that all countries follow paths that were taken by those countries that are a bit further advanced in their education transitions. Other scenarios relying on different age-specific transition probabilities can be found in KC et al. (2008).

The projection for the Republic of Korea clearly shows the significant population aging to be expected as a consequence of low fertility combined with the increase life expectancy. In this context an interesting economic question is to what extent the expected significant improvement in the level of education will be able to compensate for the expected negative consequences of rapid population aging. This is an economic question of great relevance for many rapidly ageing societies around the world and these new projections of populations by age and level of education should be helpful in facilitating such analyses.

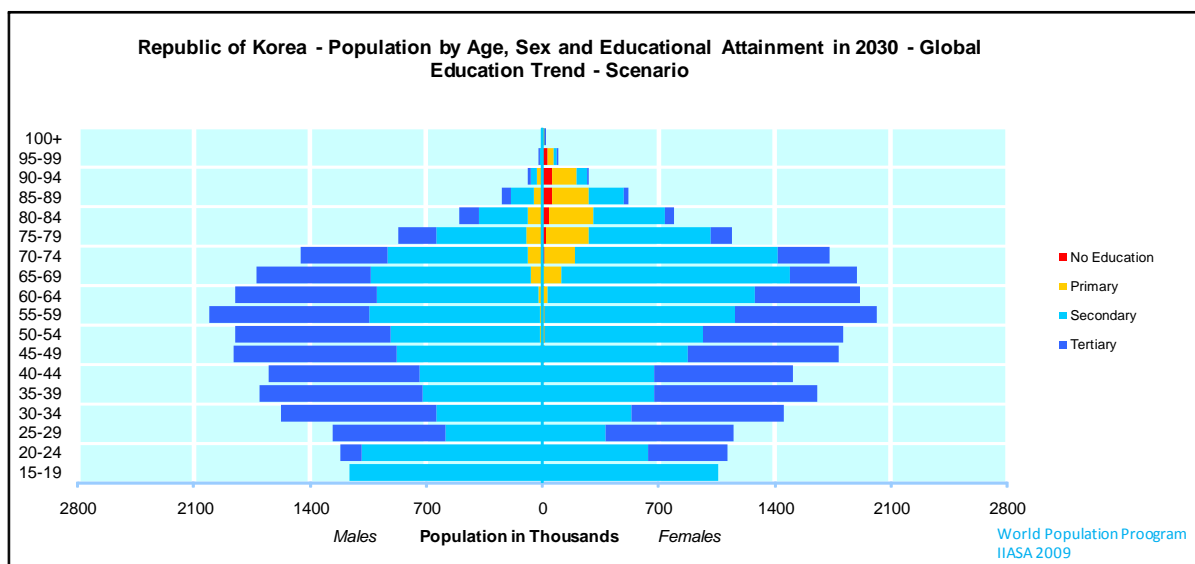


Figure 3: Projected population pyramid of South Korea in 2030 with colors indicating different educational attainment categories

Because of the lack of age structured human capital data, past studies have had to focus on data aggregated across all adult ages, i.e. for the populations 15+ or 25+. The most commonly used variable, mean years of schooling, tries to summarize the entire information provided by the above given pyramids in one single number. This will inevitably result in a certain loss of information. The key question to be discussed here is whether this loss of information is of importance for the study of the effect of changes in human capital on economic growth. Rapid increases in the educational level of the young adult population – as we have seen it for Korea above – are significantly diluted once they are merged into the broad 15+ category, that includes many less educated older adults and hence contain less statistical signal.

There are many ways to use the detailed human capital information that can be seen in these human capital/population pyramids. We used them in Lutz et al. (2008) and found that universal primary education (the second of the Millennium Development Goals) does not seem to be enough to kick start economic growth and break potential poverty traps and that only widespread secondary education among younger adults will be able to achieve this goal. For industrialized countries, as expected, the increase in the young population with tertiary education turned out to be key for economic growth.

### 3. Age structure and education in cross-country growth regressions

#### 3.1 The role of age-structured education in explaining economic growth

In this paper, we focus on the importance of the age structure of human capital as an explanatory variable for economic growth. We are thus empirically assessing the vintage nature of human capital, justified by the fact that the learning process of different generations, as well as its interaction with the prevailing technology, implies that the age composition of the human capital stock plays a prominent role in shaping the effects of education on income growth. From a theoretical point of view, these arguments have been put forward prominently by Chari and Hopenhayn (1991) and Boucekkine et al. (2002), who construct vintage models of human capital accumulation. Recently, Kredler (2009) extends the work by Chari and Hopenhayn (1991) and shows in an infinite horizon overlapping generations model with endogenous human capital formation that younger generations receive a higher premium on technology-specific skills. The incentive created by this premium

induces young individuals to accumulate human capital faster than older generations did, which in turn induces a higher rate of growth in earnings. Other theoretical contributions to the interaction between human capital accumulation and age structure emphasize the role of depreciation of the human capital stock (see Becker, (1964)). Van Immhoff (1988) uses an overlapping generations to show that increases in young-age human capital accumulation are an important determinant of income convergence, particularly in economies with low population growth.

In this contribution, we analyze empirically the relationship between the age structure of human capital and economic growth in the framework of cross-country growth regressions. In a first stage, we need to define the variables which summarize the age/education dynamics. Previous research using the IIASA/VID dataset made several different choices when it comes to summarizing the rich information contained in the age-structured educational attainment data: Crespo Cuaresma and Lutz (2007) use years of schooling by age group, Lutz et al. (2008) use educational attainment proportions aggregated in broad age groups and Crespo Cuaresma and Mishra (2009) use educational attainment at individual 5-year age groups.

In this contribution we use summary measures based on the dataset of mean years of schooling by five-year age groups. We start by constructing synthetic indicators based on extracting the principal components of the data on age-structured mean years of schooling. Let  $X = \left( \text{MYS}_{i,t}^{20} \text{ MYS}_{i,t}^{25} \dots \text{MYS}_{i,t}^{60} \right)$  be a matrix of stacked data on the mean years of schooling for each age group of the adult population (assumed here to be defined by 20-65), where  $\text{MYS}_{i,t}^a$  denotes the mean years of schooling of individuals with in the age group  $(a, a+5)$  for country  $i$  in period  $t$ . We perform a standard principal component analysis by extracting the eigenvalues of the correlation matrix based on  $X$  and reducing the dimension of the data by projecting the original data on the subspace spanned by the first  $L$  eigenvectors. This is a straightforward way of summarizing the common dynamics inherent in the data.

Table 1 presents the loadings corresponding to each one of the education measures for the first two components estimated on the data of mean years of schooling as the proxy for educational attainment by age group. It also presents the corresponding eigenvalues and the (cumulative) proportion of the total variance explained by the factors. The two first components explain more than 99% of the variance in the original data. The resulting component loadings can be easily interpreted and summarize in a simple manner the interplay between the process of aggregate human capital change and the dynamics of educational attainment changes at the level of the age structure. The first component assigns practically equal positive weights to the education measure corresponding to all age groups. Human capital accumulation in a given country is thus reflected in upward trends in this component. The second component, on the other hand, assigns (increasingly) negative loadings to the educational attainment of older age groups. Increases in this component are related to populations in which educational attainment tends to concentrate in young cohorts (we dub this component the *youth concentration* component).

Figure 4 presents a scatterplot of the two components extracted from the IIASA/VID dataset for the full sample of 120 countries in the period 1970-2000. Each observation refers to a country/period combination, with periods measured in 5-year intervals. The visible inverted-U relationship shows the increase of inter-cohort educational attainment concentration in the first stages of the transition to higher levels of educational attainment, which stabilizes or reverts as the young educated cohorts enter the older age groups. The shape of the relationship is found both across and within countries.



Factor loadings		
	First component	Second component
<i>MYS</i> <sup>20</sup>	0.327	0.468
<i>MYS</i> <sup>25</sup>	0.331	0.400
<i>MYS</i> <sup>30</sup>	0.335	0.294
<i>MYS</i> <sup>35</sup>	0.337	0.158
<i>MYS</i> <sup>40</sup>	0.339	0.006
<i>MYS</i> <sup>45</sup>	0.338	-0.143
<i>MYS</i> <sup>50</sup>	0.336	-0.283
<i>MYS</i> <sup>55</sup>	0.332	-0.401
<i>MYS</i> <sup>60</sup>	0.326	-0.499
Eigenvalue	8.659	0.287
% exp. var. (cumulative)	0.962	0.994

Table 1: Principal components analysis: loadings for educational attainment variables by age group

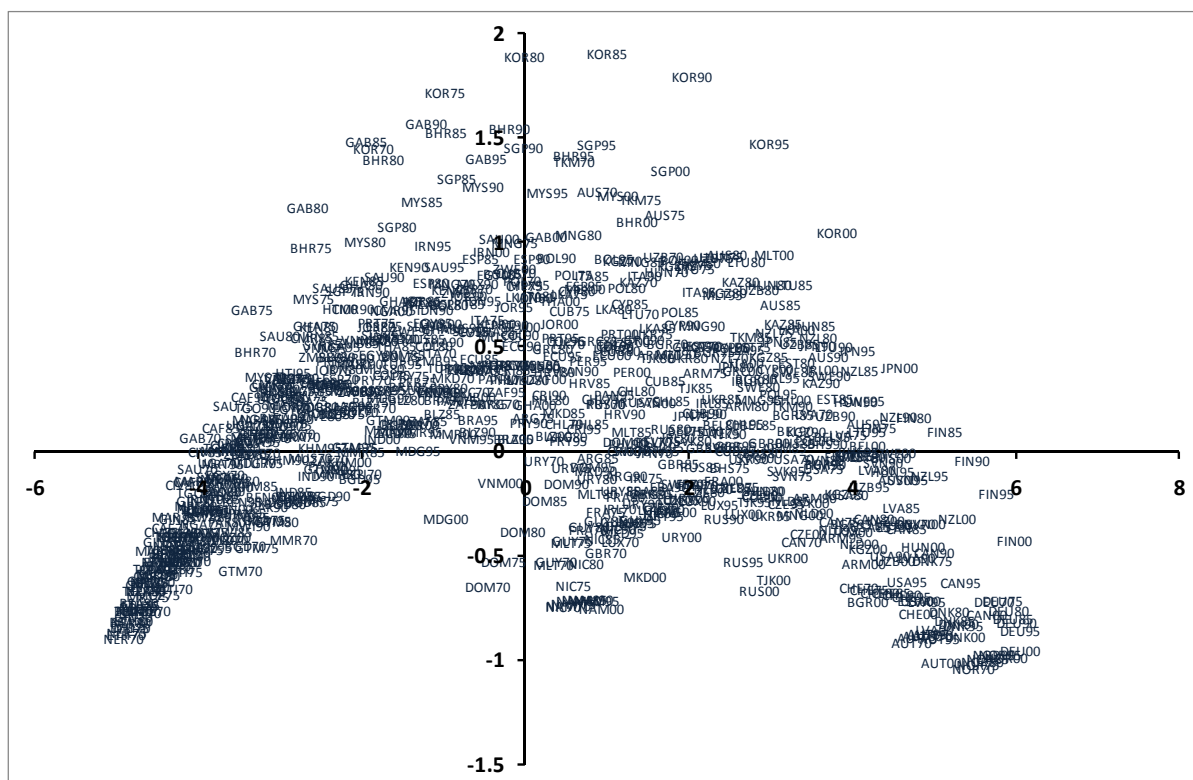


Figure 4: Education level component (x-axis) versus youth concentration component (y-axis), full sample

The information contained in the IIASA/VID dataset concerning the interplay of overall educational developments and the distribution of education across age groups allows us to assess the effects of human capital accumulation on economic growth by exploiting differences in the timing and extent of education expansions. In particular, this and other summary measures of the data can help us estimate the potentially differential effect of human capital accumulation in different cohorts on subsequent economic growth.

We start by comparing the results of cross-country growth regressions based on variables which integrate away the full age detail and those which retain information about the age structure of human capital. We estimate cross-country growth regressions using the standard modelling framework put forward by Benhabib and Spiegel (1994). Under the assumption of a Cobb-Douglas production function with constant returns to scale on labor and capital, the preferred empirical specification in Benhabib and Spiegel (1994) can be written in the form of the following model,

$$(\ln y_{it} - \ln y_{i0}) = \beta_0 + \beta_1(\ln k_{it} - \ln k_{i0}) + \beta_2 MYS_{i0} + \beta_3 \ln y_{i0} + \beta_4 MYS_{i0} \ln y_{i0} + \varepsilon_i, \quad (1)$$

where  $y_{it}$  is GDP per capita in country  $i$  at time  $t$ ,  $k_{it}$  measures physical capital stock per worker,  $MYS_{i0}$  is the human capital variable (mean years of schooling of the adult population in Benhabib and Spiegel (1994) or Cohen and Soto (2007)).

The interaction between education and the initial level of income is included to model the effect of human capital in the process of technology adoption. The theoretical background to the inclusion of such a variable, based on the Nelson-Phelps paradigm (Nelson and Phelps, 1966, and Romer, 1990), is based on the assumption that the effect of human capital on economic growth through technology adoption depends on the distance of the catching-up country to the technological frontier. Human capital in the model is assumed to affect technological progress by having an influence both on domestic innovation (through the linear term in (1)) and on the diffusion of foreign technology (through the interaction term). We estimate the specification given by (1) using the data on income per capita and capital-labor ratios in Cohen and Soto (2007), as well as their data on mean years of schooling and the measures obtained using the IIASA-VID dataset. The original source for the GDP per capita data is Heston et al. (1991) and for the capital-labor ratio is Easterly and Levine (2001). As in Cohen and Soto (2007), we use a cross-section of countries for the period 1970-1990 and evaluate the independent variables at the initial year, while the growth rates are annualized rates over the 20-year period. Descriptive statistics of the data used in the regressions are found in the Appendix.

The estimation results are presented in Table 2 for the sample of seventy countries which are common to the Cohen-Soto and IIASA/VID datasets. The first and second column present the results based on aggregated educational data for the adult population (mean years of schooling for population between 15 and 64 years of age in the case of the IIASA/VID dataset and over 25 years of age for the Cohen-Soto data). The third column uses an ad-hoc indicator of the age distribution of human capital from the IIASA-VID dataset, namely the difference in mean years of schooling between the age group 20-25 and the age group 60-65, in addition to the aggregate mean years of schooling variable. This variable is a simple measure of youth concentration and has a relationship with the age-aggregated measure of educational attainment which is relatively similar to that between the total education and youth concentration component in Figure 4. In the fourth column of Table 2, we use the value of the two principal components presented above as measures of total educational attainment and youth concentration.

Despite the fact that the sign of the estimated parameters coincide with our theoretical expectations, the results of the estimation with aggregated data (columns 1 and 2) imply that initial levels of educational attainment have no significant effect on economic growth, independently of whether the Cohen-Soto or the IIASA-VID dataset are used as source.<sup>1</sup> When age-structured variables are used

---

<sup>1</sup> As in Benhabib and Spiegel (1994) and Cohen and Soto (2007), we used OLS to estimate our cross-country growth regressions. Alternatively, we also estimated the models using methods which are robust against outliers. The results

(columns 3 and 4), however, the results indicate that overall educational attainment can explain differences in income growth for countries at the highest percentiles of the distribution of income, while the measures of educational youth concentration appear as significant determinants of economic growth with an effect which is decaying with the level of development of the country. This result is independent of the proxy used to account for youth concentration (differences in mean years of schooling or the second principal component extracted from the IIASA-VID dataset).<sup>2</sup>

	Cohen-Soto	IIASA-VID	IIASA-VID age structure	IIASA-VID components
$\ln k_{iT} - \ln k_{i0}$	0.580*** [0.0645]	0.597*** [0.0647]	0.558*** [0.0568]	0.560*** [0.0573]
$MYS_{i0}$	0.0106 [0.00670]	0.00236 [0.00635]	-0.0153* [0.00857]	
$y_{i0}$	-0.00483 [0.00316]	-0.0017 [0.00313]	-0.00179 [0.00377]	0.00156 [0.00439]
$MYS_{i0} \times y_{i0}$	-0.000796 [0.000657]	-0.000113 [0.000630]	0.00171* [0.000868]	
$MYSdiff_{20-60, i0}$			0.0212** [0.00872]	
$MYSdiff_{20-60, i0} \times y_{i0}$			-0.00204** [0.000910]	
Total education component				-0.0202* [0.0106]
Youth concentration component				0.0722** [0.0302]
Total education component $\times y_{i0}$				0.00224** [0.00107]
Youth concentration component $\times y_{i0}$				-0.00666** [0.00319]
Constant	0.0269 [0.0243]	0.00592 [0.0251]	0.00161 [0.0285]	-0.0173 [0.0426]
Observations	70	70	70	70
R-squared	0.655	0.609	0.664	0.671

\*(\*\*)[\*\*\*] indicates significance at the 10%(5%)[1%] level. Robust standard errors in brackets.

Table 2: Estimation results: Benhabib-Spiegel model using age-aggregated and age-disaggregated data

The Benhabib-Spiegel specification implies that, due to the role that education plays in the income convergence process, its effect on economic growth depends on the level of development of the country, a feature which is modeled through the inclusion of the interaction between the human

presented in Table 2 were qualitatively and quantitatively unchanged. These results are available from the authors upon request.

<sup>2</sup> It should be noticed that, despite the fact that the Nelson-Phelps framework has been used to justify the empirical specification, the theoretical interpretation of these results is ambiguous, as noticed explicitly by Krueger and Lindhal (2001). The effect of education on income may be interpreted as affecting steady state income levels, the speed of income convergence and the growth rate in the steady state.

capital variable and the initial level of income. Figures 5 and 6 show the estimated partial effect on economic growth corresponding to the variable summarizing mean years of schooling of the adult population, as well as for the youth concentration variable measured as difference in mean years of schooling between the 60-65 age group and the 20-25, obtained from the regression presented in the third column of Table 2. The estimates are presented in the form of a scatterplot, with initial income in the x-axis, and we present the parameter estimates together with a confidence interval of two standard deviations computed using the delta method. The results indicate that educational improvements taking place at young age groups are positively related to economic growth. This effect, in turn, is negatively related to initial income per capita, with poorer countries benefiting most of human capital accumulation at the lower end of the age distribution. On the other hand, after controlling for the level of youth concentration of educational attainment, the overall level of human capital (as measured by mean years of schooling of the adult population) only appears significantly related to economic growth for the subsample of richest countries in our dataset. These results indicate thus that by concentrating on the initial relative educational attainment of younger age groups as compared to older age groups, the growth experience of the countries in the sample for 1970-1990 can be better explained than using aggregated measures. Furthermore, relatively poor countries tended to profit more from educational improvements at the lower level of the age structure than rich countries.

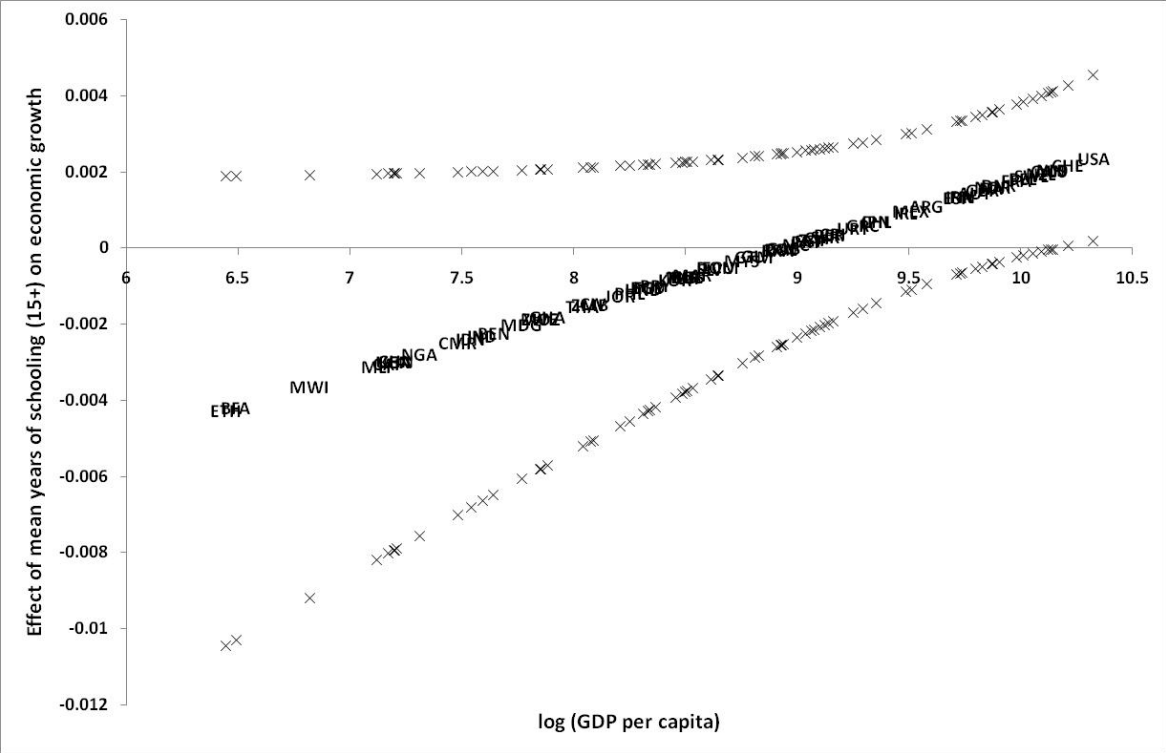


Figure 5: Estimated parameter associated with mean years of schooling (adults over 15) ( $\pm 2 \times$  standard deviation)

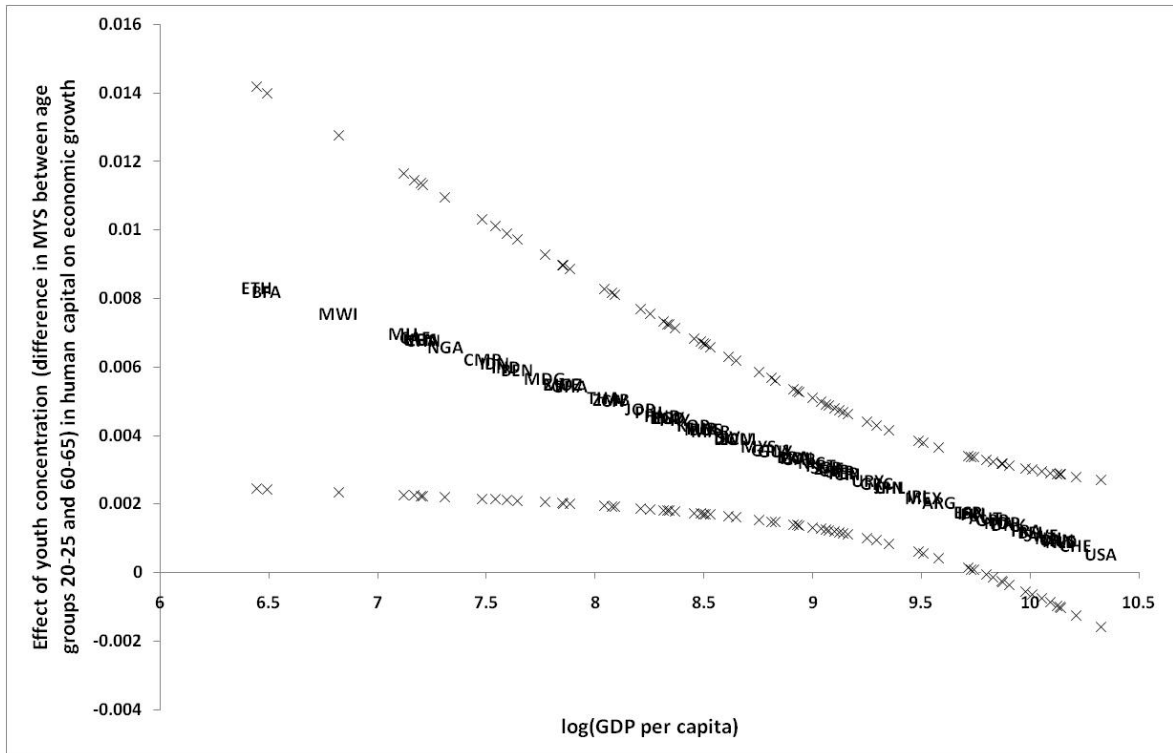


Figure 6: Parameter associated to the difference in mean years of schooling between the 20-25 age group and the 60-65 age group ( $\pm 2 \times$  standard deviation)

Figure 7 shows the growth rates of income per capita attributable by the model to the human capital variables and their interaction with initial income, together with the mean years of schooling of the adult population. They correspond to the part of the fitted growth rates of income which include human capital variables ( $\hat{\beta}_2 MYS_{i0} + \hat{\beta}_4 MYS_{i0} \ln y_{i0} + \hat{\beta}_5 MYSdiff_{20-60,i0} + \hat{\beta}_6 MYSdiff_{20-60,i0} \ln y_{i0}$ ). The estimated model indicates that the contribution of human capital to economic growth has been particularly large in absolute value for rich economies, where the innovation effect plays the most important role, and for East Asian countries (Korea, Indonesia, Malaysia and Singapore). These economies experienced large educational pushes at the beginning of the period considered, which resulted in sizable differences in educational attainment between the young and old cohorts.

Our results shed light on the contradictory results obtained hitherto in the empirical literature relating measures of human capital distribution across individuals and economic growth (see for example Castelló and Domenech (2002), who find a negative relationship between human capital inequality and economic growth, or Park (2006), who reports a positive effect of human capital dispersion on income growth). These studies are not able to identify directly the sources of such dispersion in educational attainment across individuals. By concentrating exclusively on the dispersion caused by the differences in the pair formed by age and educational attainment, we are able to identify the direct positive effect of improvements in educational attainment of young individuals.

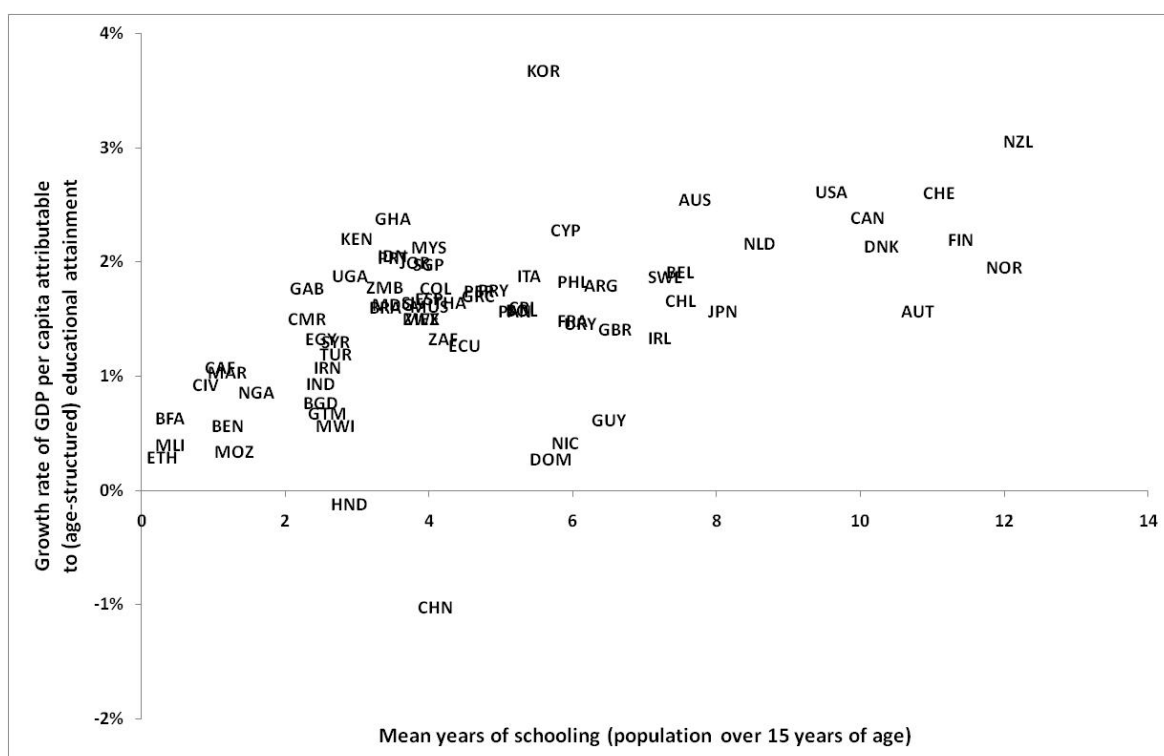


Figure 7: Growth rate of GDP per capita attributable to age-structured educational attainment against mean years of schooling

### 3.2 On the predictive ability of age-structured human capital data for income growth

The method of reconstruction of educational attainment data based on multi-state projection techniques provides a natural framework to generate future projections of age-structured human capital measures. In order to evaluate the benefits of using age-structured educational attainment data for projecting GDP per-capita paths across countries, we carry out a simple forecasting exercise based on the cross-country growth regressions estimated above. Using the estimates presented in Table 2, we obtain (out-of-sample) predictions for the annual growth rate of income per-capita based on data at the final year of the dataset used in the regressions. We compare the results with the actual annual growth rate of GDP per capita in the period 1990-2000 in order to evaluate the predictive ability of models based on different proxies for educational attainment.<sup>3</sup>

For all models, we obtain the projections of GDP per capita growth by assuming that the growth rate of capital per worker remains at the average level observed in the period 1970-1990. We obtain predictions of GDP per capita growth using the values of the corresponding educational variables for each model estimated in Table 2 for 1990.

The mean squared prediction errors for the growth rate of GDP per capita in the period 1990-2000 using each specification are presented in Table 3 for the full sample, as well as for subsamples of non- OECD countries, low and medium income (LMIC) countries and low-income countries. The results clearly indicate the superior predictive ability of models based on age-structured mean years of schooling data, which systematically attain lower prediction errors than any of the models based on aggregate measures of human capital for the full sample, as well as each one of the subsamples

<sup>3</sup> Although the income data used for the estimation corresponds to the Penn World Tables Mark 5.6, (Heston et al., 1991), we compute these growth rates of GDP per capita using the latest available version of the Penn World Tables (PWT 6.3, Heston et al. 2009).

evaluated. Predictions, which come from models which include the direct measurement of the interaction between human capital accumulation and the youth concentration effect, are systematically better than predictions based on models with aggregated data.

	Mean Squared Prediction Error, GDP per capita growth, 1990-2000			
	Full sample	Non-OECD	LMIC	LIC
Cohen-Soto dataset	0.1289	0.1260	0.0962	0.2309
MYS 15+	0.1197	0.1137	0.0958	0.1642
Principal components	0.1084*	0.1058*	0.0871*	0.1469
MYS 15+ and MYS difference	0.1044**	0.1014*	0.0872	0.1363*
Obs.	70	63	25	12

Note: Mean squared prediction error (multiplied by 100) based on annualized GDP per capita growth for 1990-2000. \* (\*\*) indicates that the MSPE corresponding to that model/sample is significantly different from the MSPE of the model using the Cohen-Soto dataset at the 10% (5%) level using a paired t-test.

Table 3: Mean square out-of-sample prediction error (MSPE): GDP per capita growth, 1990-2000

In the next step, we perform unbiasedness tests based on the predictions obtained from each one of the models. The tests for unbiasedness are based on the following regression

$$(\ln y_{i2000} - \ln y_{i1990}) = \alpha + \theta E_{1990}(\ln y_{i2000} - \ln y_{i1990} | M_j) + \varepsilon_i, \quad (2)$$

where  $E_{1990}(\ln y_{i2000} - \ln y_{i1990} | M_j)$  is the prediction of GDP per capita growth in the period 1990-2000 for country  $i$  implied by model  $j$ , using data up to 1990. The null hypothesis of unbiasedness corresponds to the restriction  $\alpha=0, \theta=1$ . Table 4 presents the results of the regressions for each model, as well as the unbiasedness tests. The test cannot reject that the predictions from the models which use age-structured data are unbiased, while significant deviations from unbiasedness appear in the models estimated using aggregated data. These results imply that age-structured human capital data do not only help understand historical economic growth experiences better, but also provide a useful tool to obtain reliable future projections of income per capita.

	$\alpha$	$\theta$	$R^2$	Unbiasedness test [p-value]
Cohen-Soto dataset	0.015** [0.006]	0.650** [0.282]	0.073	3.564** [0.034]
MYS 15+	0.013** [0.005]	0.960*** [0.262]	0.165	4.734** [0.012]
Principal components	0.009 [0.006]	0.931*** [0.235]	0.187	2.044 [0.138]
MYS 15+ and MYS difference	0.009 [0.005]	0.963*** [0.219]	0.221	2.227 [0.116]

Note: Estimates from equation (2) for the full sample of 70 countries. Unbiasedness tests refer to tests of the null hypothesis  $\alpha=0, \theta=1$ .

Table 4: Forecast unbiasedness tests

## 4. Conclusions

The IIASA/VID dataset provides detailed data on the demography of human capital that previously had not been available. We have focused in this paper on a particular aspect of the new data – the implications of the age structure of human capital on economic growth. This is only one of a wide variety of studies that can be done, using the rich detail of the IIASA/VID data.

The effects of education on economic growth has been a controversial topic with different specifications and different datasets yielding different results. In this paper, we demonstrate that the misspecification of models, because of the omission of details of the age structure of the stock of human capital, can be one of the causes why human capital appeared statistically insignificant in previous cross-country growth regressions.

There are a number of open questions that need to be explored. The basis of the IIASA/VID data are joint distributions of people by age, sex, and educational attainment. To move from these distributions to means years of school requires additional assumptions about the average number of years to associate with each level of attainment. This is not straightforward because some people who have a given level of attainment have additional years of schooling, but not enough to reach the next attainment threshold. A second alternative is to weight attainment levels by a fixed set of weights that reflect the usual number of years of school associated with each attainment level. Whether or how we should weight school attainment levels to obtain synthetic measures of years of schooling is a matter for future research.

In this paper, we have emphasized the age structure of human capital. In doing so, we have ignored sex. The interactions of sex with both age and educational attainment deserves the additional study that is now made possible with the new data.

In an article that synthesized the previous research findings Kelley and Schmidt(2005) considered the effects of age structure on economic growth and found that changes in age structure could account for around one-fifth of the changes in economic growth rates across a wide variety of countries from 1960 to 1990. They also found that aggregated mean years of school variables did not have a statistically significant effect on economic growth. It is important to revisit the literature on the effects of population age structure on economic growth with the new IIASA/VID detailed data to see the extent to which age disaggregated education data matter.

Recently quite a bit of progress have been made in estimating indicators of educational quality (Hanushek and Kimko (2000), Altinok and Murseli (2007), and Hanushek and Woessmann (2008)). A challenging project for future research would be to use this information to create indicators of educational quality by age and attainment level.



## Appendix: Descriptive statistics, variables in cross-country regressions

Variable description	Observations	Mean	Std. Dev.	Min	Max
Growth, GDP per capita 1970-1990	70	0.012	0.020	-0.042	0.061
Initial income per capita (log)	70	8.788	0.952	6.492	10.324
Growth rate of capital-to-worker ratio 1970-1990	70	0.024	0.026	-0.039	0.096
Mean years of schooling, 1970 (25+, Cohen-Soto)	70	4.312	2.997	0.055	11.182
Mean years of schooling, 1970 (15+, IIASA/VID)	70	4.947	2.890	0.400	12.200
Total education component, 1970	70	-1.457	2.717	-4.857	5.480
Youth concentration component, 1970	70	-0.190	0.461	-1.138	1.347

## References

- Altinok, N. and H. Murseli. 2007. International Database on Human Capital Quality. *Economic Letters*. 96(2): 237-244.
- Barro, R. and JW Lee. 2001. International Data on Educational Attainment: Updates and Implications. *Oxford Economic Papers*. 53(3): 541-563.
- Becker, G. 1964. *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*. Chicago, University of Chicago Press.
- Benhabib, J. and Spiegel, M. 1994. The Role of Human Capital in Economic Development: Evidence from Aggregated Cross-Country Data. *Journal of Monetary Economics*. 34(2): 143-173.
- Bosworth, B. and S. Collins. 2003. The Empirics of Economic Growth: An Update. *Brookings Papers on Economic Activity*. 2003(2): 113-179.
- Boucekkine, R., D. de la Croix and O. Licandro. 2002. Vintage human capital, demographic trends, and endogenous growth. *Journal of Economic Theory* 104, 340-375.
- Castelló, A. and R. Domenech. 2002. Human capital inequality and economic growth: Some new evidence. *The Economic Journal* 112, C187-C200.
- Chari, V. V. and H. Hopenhayn. 1991. Vintage human capital, growth, and the diffusion of new technology. *Journal of Political Economy*, 99: 1142-1165.
- Cohen, D. and M. Soto. 2007. Growth and Human Capital: Good Data, Good Results. *Journal of Economic Growth*. 12(1): 51-76.
- Crespo Cuaresma, J. and W. Lutz. 2007. Human capital, age structure and economic growth: Evidence from a new dataset. IIASA Interim Report IR-07-011, IIASA, Laxenburg.
- Crespo Cuaresma, J. and T. Mishra. 2009. The role of age-structured education data for economic growth forecasts. *Journal of Forecasting*. forthcoming.

- de la Fuente, A. and R. Domenech. 2006. Human Capital and Growth Regressions: How Much Difference Does Data Quality Make? *Journal of the European Economic Association*. 4(1): 1-36.
- Easterly, W. and R. Levine. 2001. It's Not Factor Accumulation: Stylized Facts and Growth Models. *World Bank Economic Review*. 15(2): 177-219.
- Hauk, W and R. Wacziarg. 2009. A Monte Carlo Study of Growth Regressors. *Journal of Economic Growth*. 14(2): 103-147.
- Hanushek, E. and D. Kimko. 2000. Schooling, Labor Force Quality, and the Growth of Nations. *American Economic Review*. 90(5): 1184-1208.
- Hanushek, E. and L. Woessmann. 2008. The Role of Cognitive Skills in Economic Development. *Journal of Economic Literature*. 46(3): 607-668.
- Heston, A. R. Summers and B. Aten. 1991. Penn World Table Version 5.6, Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania.
- Heston, A. R. Summers and B. Aten. 2009. Penn World Table Version 6.3, Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania, August 2009.
- Kelley, A. and R. Schmidt. 2005. Evolution of Recent Economic-Demographic Modeling: A Synthesis. *Journal of Population Economics*. 18: 275-300.
- KC, S., B. Barakat, A. Goujon, V. Skirbekk, and W. Lutz. 2008. Projection of Populations by Level of Education, Age and Sex for 120 Countries for 2005-2050. IIASA Interim Report. IR-08-038.
- Keyfitz, Nathan. 1985. *Applied Mathematical Demography*, 2nd Edition. New York: Springer-Verlag.
- Kredler, M. 2009. Experience vs. Obsolescence: A Vintage Human Capital Model. Mimeo, Universidad Carlos III de Madrid.
- Krueger, A. and L. Lindahl. 2001. Education for Growth: Why and for Whom? *Journal of Economic Literature*. 39(4): 1101-1136.
- Lutz, W. 2009. Sola Schola et Sanitate: Human Capital as the Root Cause and Priority for International Development? *Philosophical Transactions of the Royal Society B-Biological Sciences*. 364(1532): 3031-3047.
- Lutz, W., J. Crespo Cuaresma and W. Sanderson. 2008. The Demography of Educational Attainment and Economic Growth. *Science*. 319(5866): 1047-1048.

- Lutz, W., A. Goujon, S. KC, and W. Sanderson. 2007. Reconstuction of Population by Age, Sex and Level of Educational Attainment of 120 Countries for 1970-2000. Vienna Yearbook of Population Research. 2007: 193-235.
- Nelson, R., E. Denison, K. Sato, and E. Phelps. 1966. Investment in Humans, Technological Diffusion, and Economic Growth. American Economic Review. 56(2): 69-82.
- Park, J. 2006. Dispersion of Human Capital and Economic Growth. Journal of Macroeconomics. 28: 520-539.
- Rogers, A. 1975. Introduction to Multiregional Mathematical Demography. New York: Wiley.
- Rogers, A. and K. Land, Eds. 1982. Multidimensional Mathematical Demography. London: Academic Press.
- Romer, P. 1990. Endogenous Technical Change. Journal of Political Economy. 98(5) Part 2: S71-S102.
- Simon, J. 1981. The Ultimate Resource. Princeton, N.J.: Princeton University Press.
- Summers, R. and A. Heston. 1991. The Penn World Table (Mark-5) – An Expanded Set of International Comparisons, 1950-1988. Quarterly Journal of Economics. 106(2): 327-368.
- Torrey, B. and W. Kingkade. 1990. Population Dynamics of the United States and Soviet Union. Science. 247(4950): 1548-1552.
- van Imhoff, E. 1988. Age structure, education, and the transmission of technical change. Journal of Population Economics. 1: 167–181