Helsinki
Center
of
Economic
Research

Discussion Papers

Longevity Clubs

Ulla Lehmijoki
University of Helsinki and HECER

Discussion Paper No. 234
September 2008

ISSN 1795-0562

HECER – Helsinki Center of Economic Research, P.O. Box 17 (Arkadiankatu 7), FI-00014 University of Helsinki, FINLAND, Tel +358-9-191-28780, Fax +358-9-191-28781,
E-mail info-hecer@helsinki.fi, Internet www.hecer.fi

HECER
Discussion Paper No. 234

Longevity Clubs*

Abstract

This paper discusses the post-war convergence of life-expectancy in longevity clubs derived from a global sample of countries by the regression tree method. Three clubs are constructed by their literacy rates, whereas low income is typical of the fourth. The convergence of life-expectancy is analyzed both within and between the clubs. The results show that because deficient literacy prevents diffusion of information, the average life-expectancy growth in the club with the lowest literacy was sluggish relative to initial life-expectancy. Furthermore, no convergence appeared within this club. By contrast, the club with medium literacy proceeded steadily and exhibited convergence, presumably due to externalities arising once a critical literacy level was reached. The slow progress in the club with the highest literacy seems to be due the complexity of health information and privatization of health. No convergence appeared within this club either. It also turns out that, since AIDS is closely associated with poverty, most life-expectancy laggards are members of the low-income club.

JEL Classification: J11, J24, I18.

Keywords: Life-expectancy convergence, regression tree, longevity clubs.

Ulla Lehmijoki

Department of Economics
University of Helsinki
P.O. Box 17 (Arkadiankatu 7)
FI-00014 University of Helsinki
FINLAND

e-mail: ulla.lehmijoki@helsinki.fi

* The author is grateful for the funding from the Yrjö Jahnsson Foundation.

1 Introduction

Since World War II there has been a remarkable growth in life-expectancy worldwide. Given that this growth has typically been most rapid where the initial value of life-expectancy was low, global convergence of life-expectancies has taken place. This convergence, however, has not been unequivocal as life-expectancy has increased only slowly in some countries and even decreased in others.

Contradictory findings have subjected life-expectancy convergence to heated debate currently. The optimists claim that most low-longevity countries ultimately catch up with the high-longevity ones (Ram 1998, Neumayer 2003, Becker et al 2005, Riley 2005). The main argument is that the former can adopt the latest technologies relatively easily to reduce infant mortality and infectious diseases, whereas the advanced countries can derive progress only from the slow fight against cardiovascular diseases and aging. Hence, countries with the greatest gaps to health leaders have the largest potential for growth (Omran 1971, Deaton 2003, Vallin and Meslé 2004, Edwards and Tuljapurkar 2005, Cutler et al. 2006, Soares 2007). This optimism, however, may be premature as convergence has stagnated in some cases. The so-called club approach has thus gained acceptance (Mayer-Foulkes 2003). This approach assumes that countries can be classified into longevity clubs with a unique convergence rate in each. Rather than investigating convergence in a broad sample of countries, it would thus be more appropriate to investigate it in sub-samples, i.e., in clubs.

There are several reasons for the appearance of such clubs. Global incomes tend to diverge rather than converge as some countries are lagging behind the generally rising trend (Pritchett 1997). Given that income is an important determinant of life-expectancy (Preston 1975, Fogel 1994, 2004), one expects divergence in life-expectancy as well. The diffusion of health technology is not straightforward, either. The reason is that human capital, which is essential in absorbing of information, has been lowest in low life-expectancy countries, thus making the realization of their high growth potential cumbersome. Given the strong intergenerational heritage in human capital, the families and countries with low human capital tend to face low life-expectancies for several generations, and low life-expectancy, in turn, decreases the incentive to invest in human capital by increasing their risks. Therefore, unfavorable initial conditions are difficult to overcome and long-lasting vicious cycles may follow; some countries may even be caught in a low life-expectancy trap (Becker et al. 1990, Soares 2005). Furthermore, low human capital is associated with high fertility, which tends to decrease life-expectancy through maternal and infant deaths, especially in poor countries. Caldwell (2000) also associates high fertility rates with extramarital sexual activities, increasing the risk of $AIDS$.

In this paper, I concentrate on the need to classify the data into longevity clubs. Both the number and boundaries of these clubs are derived from the data by the regression tree method suggested by Breiman et al. (1984) and Durlauf and Johnson (1995), a data-sorting method which partitions the sample into sub-samples to find the best piecewise linear model. The main advantage of this method is that instead of applying some pre-determined partitioning criterion, it chooses both the partitioning variable and its split value for each split to generate the homogenous sub-samples possible. Successive splits grow up a tree, starting from the root (the full sample) to the leaves (clubs). By supplying several theoretically plausible candidates to serve as the classifying variables one can learn from their importance in terms of the longevity convergence in the sub-samples or clubs.

The regression tree analysis suggests that a global sample of countries should be partitioned into four convergence clubs according to human capital and income. The role of human capital is important, since three clubs are characterized by their initial literacy rates, whereas low initial income is typical of the fourth. A rich convergence analysis is now viable since one can investigate the life-expectancy convergence both *within* and *between* the clubs. The results show that because deficient literacy prevents diffusion of information, the average life-expectancy growth in the club with the lowest literacy was slower than expected from its initial life-expectancy. No convergence arose within this club either. By contrast, the club with medium literacy proceeded rapidly, also exhibiting convergence within the club, presumably because of health externalities arising once some critical literacy level was reached. An explanation for the slow progress in the club with the highest literacy may in turn be in the complexity of health information, making the absorbtion of this information less easy. Since privatization of health eliminates externalities, no convergence has appeared within this club. The analysis here also implies that because the $AIDS$ epidemic was so strongly fueled by poverty, an unwanted re-appearance of the negative association between income and life-expectancy occurs. Given that poverty is still a considerable problem, this paper thus pays attention to the risk of escalation of this epidemic in the future.

The outline of the paper is the following: Section 2 performs a specification test and partitions the data into longevity clubs. Sections 4 and 3 analyze convergence within and between the clubs, and discuss the role of literacy, incomes, and fertility. Section 5 offers conclusions and closes the paper.

2 Regression Tree Analysis

Consider the post-war growth of life-expectancy.[1] In 1960, the average life-expectancy in a sample of 125 countries was only 51.72 years but this increased to 64.63 years by 2001. Since countries with the lowest initial values exhibited the most rapid growth, convergence of life-expectancy appeared in this sample. Figure 1 illustrates this by showing a clear negative association between the initial life-expectancy and its growth, but Figure 1 also shows that convergence was far from universal as many low life-expectancy countries fell behind and, in the most unfortunate cases, life-expectancy even decreased. Examples are Zimbabwe, Zambia, and Botswana with decreases of 10.46, 6.13, and 2.98 years, respectively. It is thus possible that the full sample should be partitioned into sub-samples to derive unbiased estimates for convergence. Hence, we perform a specification test to see whether any sub-samples exist in the data.

The theoretical discussion above suggests that unequal life-expectancy growth appears where considerable disparities exist in initial income, human capital, and fertility so that we take these variables as candidates to classify the data. The income is measured by (the log of) the real per capita GDP and human capital by the literacy rate. Both variables were measured in

[1]life-expectancy at birth, both sexes, denoted by life-expectancy and longevity in the text.

Figure 1: life-expectancy growth regressed against initial life-expectancy (125 countries, period 1960-2001).

1960. The third candidate is fertility. Fertility was already relatively low in developed countries in 1960, but in developing countries it was universally high, discriminating them only poorly. Since the first family planning programs, however, had started to run with the outcome that fertility decreased in some developing countries but was resistant in some others, we choose the fertility change from 1950 to 1960, instead of choosing the initial fertility rate as an indicator of (inherited) fertility behavior,. The demographic data comes from the United Nations (2007), the income data from Maddison (2001) (GPD per capita in constant international dollars), and the literacy data from the World Bank (2000). All variables are presumably exogenous to life-expectancy growth from 1960 to 2001.

To perform the specification test, we divide the full sample into two sub-samples by means of the variables above.[2] Let T stand for the period length (41 years) and $E_{i,t}$ for the life-expectancy in country i at time t. We estimate the standard convergence equation

$$ln(E_{i,t}/E_{i,0})/T = \alpha + \beta \, ln(E_{i,0}) + \epsilon_i, \tag{1}$$

by least squares for the full sample and for the sub-samples separately. The null is that the full sample obeys the common linear model. The Wald test

[2]The first (second) sub-sample consists of countries with higher (lower) than the mean of per capita GDP and literacy, and faster (slower) than the mean decrease in total fertility.

of similarity of α and β in the sub-samples yields $F-$statistics of 3.21 which with $(2, 121)$ degrees of freedom refers to a $p-$value of 0.04, so that the null is rejected in favor of sub-samples. Hence, the club approach is justified here.

Even though the specification test suggests that at least some sub-samples exist in the data, their number and boundaries cannot be properly revealed by mechanical splits. We therefore discover the clubs by the regression tree method. This method partitions the sample to find the best piecewise linear model. For each split, the algorithm chooses the splitting variable and its split value to minimize the sum of squared residuals (SSR) in the sub-samples. Only one-step look ahead and binary splits are used. A detailed description of the regression tree method is available in Breiman et al. (1984) and Durlauf and Johnson (1995).

Figure 2 reports the result of the regression tree analysis. The first split comes in terms of the initial literacy rate, the split value being 90.45%. Countries with an initial literacy rate higher than this value constitute the first longevity club, which we refer to as the "*High-Literacy*" club (symbol h). Figure 2 indicates that this club has 26 members with average life-expectancy growth of 7.62 years. The second split comes in terms of literacy again, separating the countries with an initial literacy rate lower than (or equal to) 35.75% to the "*Low-Literacy*" club (46 members, average life-expectancy growth 14.98 years, symbol l). The initial literacy rate in all remaining countries is between 35.75% and 90.45%, but some of these countries had markedly low incomes initially. The remaining countries are thus partitioned in terms of their incomes, the split value being 6.90. Countries with income below (or equal to) this value constitute the "*Low-Income*" club (14 members, average life-expectancy growth 7.56 years, symbol g), while the remaining countries constitute the "*Medium-Literacy*" club (39 members, average life-expectancy growth 15.90 years, symbol m). For a complete list of countries in these four clubs, see Appendix A.

One can continue the partitioning until the algorithm is unable to find new splits to reduce SSR, or all degrees of freedom are used up (the number of countries in the sub-samples is equal to or smaller than twice the number of regressors). Even though the model fit always increases, every split makes its interpretation more difficult so that an excessive number of clubs

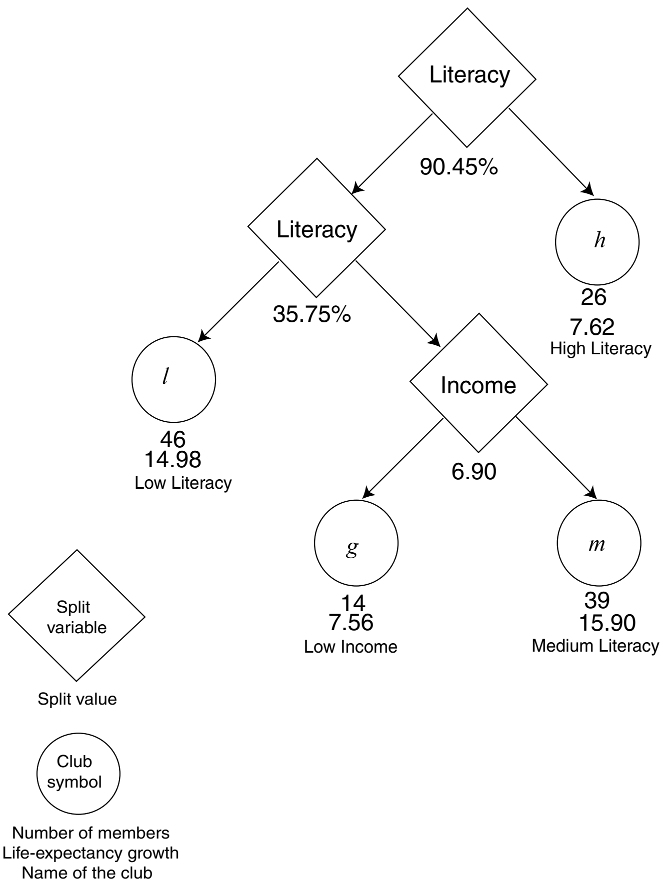

Figure 2: The regression tree. The right arrow indicates the countries for which the split variable takes values larger than the split value. The left arrow indicates the countries for which the split variable takes values smaller than or equal to the split value.

is not desirable. Some "over fitting" may also appear in the tree since it is adapted according to every detail of the present sample, thus decreasing its predictive power in the "population". Therefore, a cross validation exercise is performed for SSR to see, what number of clubs generates the best fit in the "population", with the outcome that the best number of clubs turns out to be four as already depicted in Figure 2.[3] For details in tree pruning, see

[3] In the ten-fold cross validation performed here, the full sample is partitioned into ten approximately equal pieces stochastically. Nine of them are used to develope a tree which is then fitted to the tenth to calculate SSR. By repeating this and averaging, one can calculate a population estimate for SSR for each number of clubs. The club number which minimizes this estimate refers to the best tree size and to the best piecewise linear model.

Breiman et al. (1980) and Durlauf and Johnson (1995).

	LIF60	LIF01	GROWTH	LIT60	DTFR	GDP60	GDP01
Full Sample	51.72	64.63	12.91	54.89	0.168	2644	6380
High-Literacy	69.72	77.34	7.62	98.23	0.089	6611	17264
Medium-Literacy	55.16	71.06	15.90	68.62	0.194	2555	6293
Low-Literacy	40.80	55.79	14.98	20.10	0.150	1075	1539
Low-Income	44.58	52.15	7.57	50.47	0.299	677	2319

Table 1: Longevity club statistics. The notations are: initial and final life-expectancies and life-expectancy growth ($LIF60$, $LIF01$, $GROWTH$), initial literacy rates ($LIT60$), fertility changes ($DTFR$), initial and final incomes ($GDP60$, $GDP01$). All numbers are averages.

The order by which the variables enter to partition the data provides information about their importance. Given that the first two splits come in terms of the literacy rate, the analysis here suggests that literacy (human capital) is the variable which most efficiently discriminates between the clubs. Naturally, it is not the only one. Table 1 shows that the clubs have different initial incomes and fertility changes as well. Since the *High-Literacy* club, for example, had the slowest fertility change and highest initial income, whereas the *Medium-Literacy* club had medium values for these variables as well, not all life-expectancy disparities can be explained by literacy alone. Its role is the most decisive, however, thus supporting the idea that the effects of the initial human capital are intergenerational and long-lasting. The importance of literacy also suggests that health technology diffusion has been the driving force in life-expectancy growth, since the capacity to absorb information seems to be the critical condition for this growth.[4]

[4]Note that the algorithm does not choose fertility to partition the data, even though its association with infant and maternal mortality seems indisputable, and Caldwell (2000) convincingly discusses its role in explaining $AIDS$ contamination rates. A possible explanation for this mismatch is that the fertility change before the research period applied here as a proxy did not adequately predict the fertility rate at the time of the break-out of $AIDS$, which occurred much later and seems to explain most life-expectancy failures, as is discussed in Section 4. An alternative explanation is that literacy explains fertility to such a degree that fertility itself loses its explanatory power.

3 Convergence Within the Clubs

The conventional result is that convergence of life-expectancy appears in broad samples of countries (Ram 1998, Neumayer 2003, Becker et al. 2005, Riley 2005). Table 2 reports that the sample of 125 countries here also exhibits highly significant β-convergence as calculated by Equation (1). The specification test above suggests, however, that the derived estimate may be biased and that one should merely concentrate on clubs. Table 2 shows that statistically significant β-convergence only appeared within the *Medium-Literacy* and *Low-Income* clubs, while less than significant convergence appeared in the *High-Literacy* club. In the *Low-Literacy* club, the association between initial life-expectancy and its growth was negative but not significant at any conventional level. Figure 3 illustrates these results by showing a clear downward sloping trend in the *Medium-Literacy* and *Low-Income* clubs (symbols m and g), whereas no such trend is seen in the *High* and *Low-Literacy* clubs (symbols h and l). Life-expectancy convergence thus seems to be a much more complicated process than is implied by the full-sample analysis alone.

	β	*t-test*	*prob.*
Full Sample	-0.784	-6.814	0.000
High-Literacy	-0.661	-1.352	0.094
Medium-Literacy	-1.450	-7.877	0.000
Low-Literacy	-0.239	-0.607	0.273
Low-Income	-2.723	-1.868	0.043

Table 2: Convergence statistics.

To find an intuitive explanation for the club-specific convergence results above we turn to the role of literacy again as its importance is indicated by the regression analysis. Consider first the *Low-Literacy* club. By construction, the literacy rate in all countries of this club was initially lower than 35.75%, the lowest value being 2.00%. Furthermore, since low (initial) literacy was closely associated with low (initial) life-expectancy (correlation 0.69), the least advanced countries with the highest potential for growth also faced the most serious constraints in realizing this potential, as low literacy seriously limited absorbtion of new knowledge in these countries. Thus, in spite of their good prospects, these countries only proceeded at the average rate, i.e., convergence failed within this club.

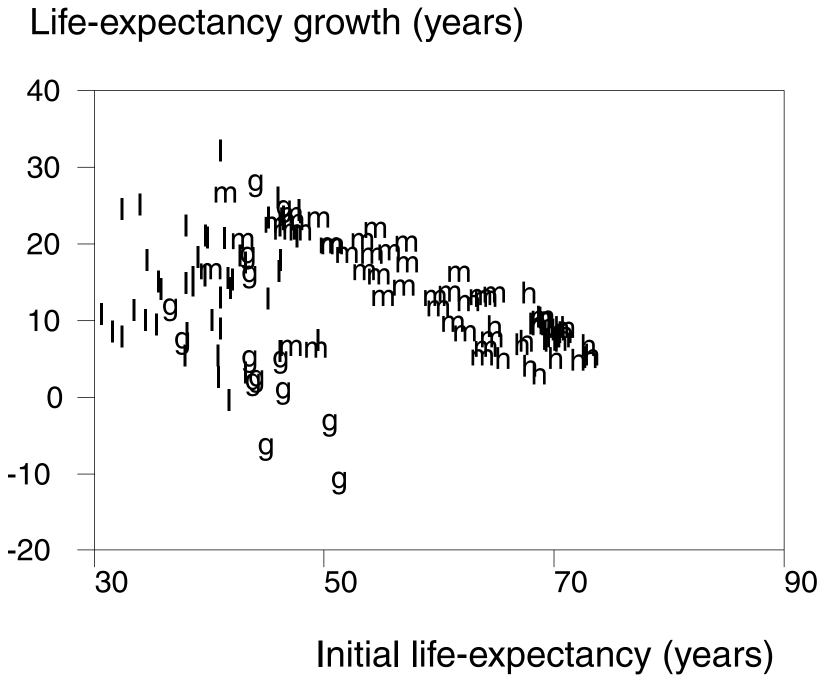

Figure 3: The Longevity Clubs.

Consider then the *Medium-Literacy* club. By construction, its initial literacy rate varied from 35.75% to 90.45%, being closely associated with initial life-expectancy again (correlation 0.68). The least advanced countries should therefore have faced the most serious constraints in this club as well. Despite this, highly significant convergence appeared, so that one wonders why the above reasoning is not valid here. Presumably an explanation can be be found in external effects typically associated with human capital and literacy (Lucas 1988). Externalities arise because literate members of society inform the illiterate ones, either orally or by showing things in practice. Once a critical literacy level is reached, most people achieve new information, and further gains in literacy are less important as they induce little improvement in absorption marginally.[5] To put this differently, one can say that the literacy constraint becomes less binding and the potential for growth dic-

[5]It is possible that the this critical value was close to 35.75%, identified here to distinguish the *Medium-Literacy* club from the *Low-Literacy* club.

tates the life-expectancy growth, thus inducing convergence within this club. This conclusion is confirmed by facts from the *Low-Income* club, in which club, the initial literacy rate was less closely related to initial life-expectancy (correlation only 0.22) so that the worst absorbtion constraints were not so typical of the least advanced countries, and these could utilize their potentials for growth efficiently. Hence most rapid convergence was seen in this club (Figure 3, Table 2).

Table 2 also shows that even though some convergence appeared in the *High-Literacy* club, this convergence was not significant at the 5% level. To comprehend this, note that the *High-Literacy* club had already reached the highest stage of health transition with cardiovascular diseases and aging as the most serious obstacles to life-expectancy growth. Since the causes of these intricate problems are not completely understood thus far, and no simple treatment is available yet, health information tends to be markedly complex in this club. Complicated instructions are given to patients and profit-targeted misinformation is common. Not just literacy but critical information absorbtion becomes important, and it is far from self-evident that adequate health information reaches everyone. Privatization of health is typical, so that positive health externalities are less likely (Vallin and Meslé 2004, Soares 2007). Given that the correlation between initial literacy and life-expectancy is still high (0.60), the most backward countries tend to have the greatest absorbtion constraints in this club. In spite of high literacy rates, the complexity of health information and missing externalities make these constraints binding, thus explaining why no (statistically significant) convergence appeared in this club.

4 Convergence Between the Clubs

The convergence hypothesis maintains that the quickest life-expectancy growth is seen in the initially most backward countries, and one expects this principle to hold for club-specific average life-expectancies as well. Figure 4 illustrates the beginning and end-of-period values. Consider first the *High-Literacy* club where the average life-expectancy was initially as high as 69.72 years and increased further by 7.62 years (see also Table 1). However, as compared with the *Medium-Literacy* club, with an initial value of 55.16 years and an increase of 15.90 years, the *High-Literacy* club proceeded slowly. This is in line with the convergence hypothesis: life-expectancy gains in the *High-Literacy* club were difficult to achieve, whereas the less advanced *Medium-Literacy* club was able to absorb new knowledge easily (Omran 1971, Deaton 2003, Vallin and Meslé 2004, Edwards and Tuljapurkar 2005, Cutler et al. 2006, Soares 2007). As a consequence, the average life-expectancy in the *Medium-Literacy* club approached that in the *High-Literacy* club.

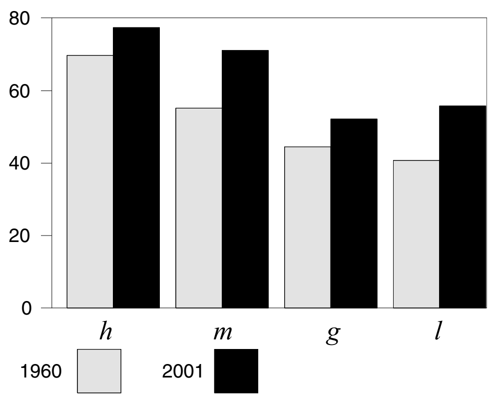

Figure 4: Beginning and end-of-period average life-expectancies.

Consider then the *Low-Literacy* club, the average life-expectancy of which was 40.80 years initially and then increased by 14.98 years (Figure 4, Table 1). At first sight, this seems to be a satisfactory result. But it is not. Comparison with the *Medium-Literacy* club shows this clearly. In 1960, the life-expectancy gap between the *Medium* and *High-Literacy* clubs was 14.56 years, while that between the *Low* and *High-Literacy* clubs was 28.91 years. Hence, the potential for growth was much higher in the *Low-Literacy* club which, however, utilized it only to a limited extent by closing some 52% of the gap, while the *Medium-Literacy* club closed more than 100% of its gap. Presumably, literacy is an explanation since its average initial value was only 20.10% in the *Low-Literacy* club but 68.62% in the *Medium-Literacy* club, implying that these clubs had different opportunities to utilize their growth potential. Health externalities further exacerbated the disparity since they are likely to exist only in the latter club.

Soares (2007) provides further insight into the role of literacy by maintaining that the effects of some health measures, such as large-scale immunizations and improvements in public health infrastructure, are independent of the personal characters of the agents. But since some gains related to preparation of food, treatment of water, and nursing of infants, for example, derive from personal practices, their effects depend more on the absorbtion of knowledge on the agent's side (Soares 2007). Thus the observed life-expectancy growth in the *Low-Literacy* club was probably caused by the former type of measure alone, whereas both types were efficient in the *Medium-Literacy* club.

Now turn to the *Low-Income* club, in which the average life-expectancy increased by 7.57 years, implying that only 30% of its initial life-expectancy gap was closed (Figure 4, Table 1). Given that this club was constructed on the basis of its low (initial) incomes, malnutrition and many income-related diseases must have increased mortality in this club. It has been claimed, however, that these outcomes of poverty have decreased, and the association between life-expectancy and income has become less tight (Soares 2007). It does indeed seem that a great deal of the longevity disaster in the *Low-Income* club was caused by the *AIDS* epidemic. This is clearly illustrated in Figure 5, which shows how the promising convergence trend of this club was interrupted in the 1980s, implying that a great deal of the regression was due to the break-out of *AIDS*. The *HIV* prevalence numbers of 2.67%, 0.26%, 1.48%, 2.45%, and 11.14% in the full sample, *High*, *Medium*, and *Low-Literacy* clubs, and the *Low-Income* club in 2005 support this, even though the association between income and *AIDS* is not one-to one, since the *Low-Income* club consists of four low-*AIDS* countries (Burma, China, The Comorros, and Vietnam), whereas four countries with two-digit *HIV* prevalence rates (Central Africa, Mozambique, Namibia, and South Africa) are classified into other clubs.

Since poverty exacerbates mortality from *AIDS* by imposing serious limits on treatment and prevention (UNAIDS and WHO 2007), it seems that *AIDS* has established the poverty-mortality connection again. In many cases, one can find historical reasons for poverty. The Sub-Saharan countries are good examples; being seriously impaired by their colonial roots, they were markedly vulnerable when the *HI* virus started to spread and the

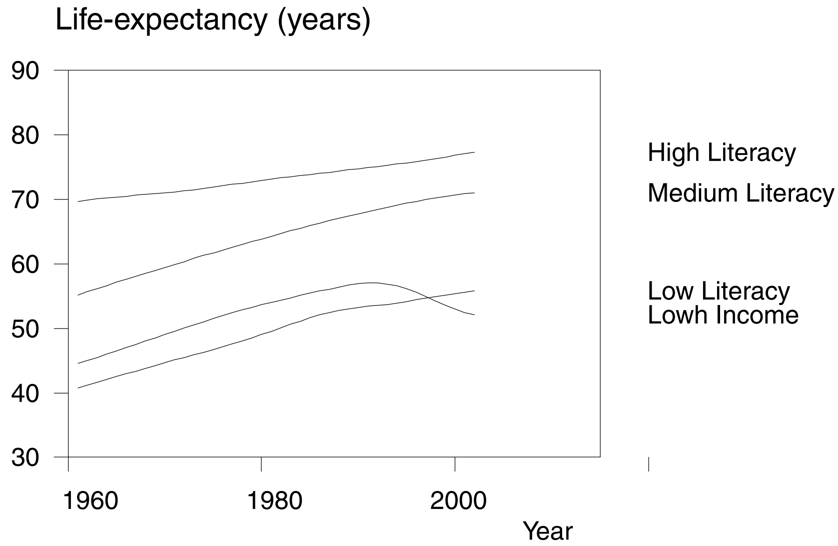

Figure 5: Trends in life-expectancy.

life-expectancy toll was dramatic. History explains why African countries were the first to confront the exacerbated effects of *AIDS*, but if we take the role of poverty seriously we see it as a future risk as well. Table 1 shows that the average income is now the lowest in the *Low-Literacy* club. Therefore, the analysis here raises concerns about the escalation of *AIDS* in this club.

To reiterate, starting from the importance of human capital (literacy) and income indicated by the regression tree analysis above, I have explained the convergence of life-expectancy by these factors alone. Naturally, several other explanations are possible. One is the role of trade because knowledge is often embodied in goods. Another is institutional, as good institutions, legal governments, and political stability are factors promoting diffusion. Cultural factors are important as well, but data availability limits the analysis. Nevertheless, the intergenerational nature of human capital and its critical role in absorbtion of information are intuitively appealing in explaining long-lasting disparities in life-expectancy growth. Only the recent *AIDS* epidemic turns out to be an exception, since its roots seem not to be in in low literacy but in poverty.

5 Conclusions

This paper discusses the post-war convergence of life-expectancy in longevity clubs which are identified in the data by the regression tree method. Four clubs appear. Given that three clubs are constructed by their literacy, this analysis suggests that literacy or, more broadly, human capital, has played an essential role in convergence of life-expectancy.

Convergence of life-expectancy is analyzed both within and between the clubs. The results show that because deficient literacy prevents diffusion of information, the growth in the club with the lowest literacy rate was sluggish. No convergence appeared within this club. By contrast, the club with literacy at the medium level proceeded rapidly and exhibited convergence, presumably the result of high literacy and health externalities rising once critical literacy is achieved. The slow progress in the club with the highest literacy was probably caused by the complexity of health information and privatization of health. No convergence appeared within this club either. Most life-expectancy laggards are members of the *Low-Income* club because the mortality effects of $AIDS$ are exacerbated by poverty.

The results of this paper suggest that continuous efforts are needed to remove the constraints on life-expectancy growth in the less advanced countries. Educational programs are important, but more frank measures are also needed because the health effects of educational programs appear slowly. Profound poverty needs attention because it increases the risk of $AIDS$, and we should also find policies to improve the absorbtion of complex health information in the most advanced countries; measures which increase externalities and decrease privatization could be successful here. But we should also reconsider the roles of biological and man-made obstacles in the slowdown of life-expectancy growth in the most advanced countries, since it is possible that we have accepted it as an intrinsic element of convergence too easily. One of the future challenges thus is to investigate the extent to which this slowdown can be alleviated by well-constituted health policies.

A Appendix: Countries and Clubs

High Literacy Club:
Argentina, Australia, Austria, Belgium, Bulgaria, Canada, The Czech Republic, Denmark, Finland, France, Germany, Hungary, Ireland, Italy, Japan, The Netherlands, New Zealand, Norway, Poland, Romania, Spain, Sweden, Switzerland, The United Kingdom, The United States of America, Uruguay.

Medium Literacy Club:
Albania, Bahrain, Bolivia, Brazil, Chile, Colombia, Costa Rica, Cuba, The Dominican Republic, Ecuador, El Salvador, Greece, Guatemala, Honduras, Indonesia, Israel, Jamaica, Jordan, Lebanon, Madagascar, Malaysia, Mauritius, Mexico, Namibia, Nicaragua, Panama, Paraguay, Peru, The Philippines, Portugal, Singapore, South Africa, South Korea, Sri Lanka, The Syrian Arab Republic, Thailand, Trinidad and Tobago, Turkey, Venezuela.

Low Literacy Club:
Afghanistan, Algeria, Angola, Bangladesh, Benin, Burkina Faso, Burundi, Cambodia, Cameroon, The Central African Republic, Chad, Congo, Côte d'Ivoire, Djibouti, Egypt, Eritrea, Gabon, Gambia, Ghana, Guinea, Guinea-Bissau, Haiti, India, Iran, Iraq, Laos, Libya, Mali, Mauritania, Mongolia, Morocco, Mozambique, Nepal, Niger, Nigeria, Oman, Pakistan, Rwanda, Senegal, Sierra Leone, Somalia, Sudan, Togo, Tunisia, Yemen, Zaire.

Low Income Club:
Botswana, Burma, China, The Comorros, Equatorial Guinea, Kenya, Lesotho, Malawi, Swaziland, Vietnam, Tanzania, Uganda, Zambia, Zimbabwe.

References

Becker Gary S., Kevin Murphy, and Robert Tamura. 1990. "Human capital, fertility, and economic growth," *Journal of Political Economy* 98(5): S12-S37.

Becker Gary S., Tomas J. Phillipson, and Rodrico R. Soares. 2005. "The quantity and quality of life and the evolution of world inequality," *American Economic Review* 95(1): 277–291.

Breiman, Leo, Jerome L. Friedman, R. A. Olshen, and Charles J. Stone. 1984. *Classification and regression trees.* Belmont, CA. Wadsworth.

Caldwell, John. 2000. "Rethinking the African AIDS epidemic," *Population and Development Review* 26(1): 117–135.

Cigno, Alexander. 1998. "Fertility decisions when infant survival is endogenous," *Journal of Population Economics* 11: 21–28.

Cutler, David, Agnus Deaton, and Adriana Lljeras-Muney. 2006. "The determinants of mortality," *Journal of Economic Perspectives* 20(3): 97–120.

Deaton, Agnus. 2003. "Health, inequality, and economic development," *Journal of Economic Literature* 49(1): 113–158.

Durlauf, Steven, and Paul A. Johnson. 1995. "Multiple regimes and cross-country behaviour," *Journal of Applied Econometrics* 10: 365-384.

Edwards, Ryan D., and Shripad Tuljapurkar. 2005. "Inequality in lifespans and a new perspective on mortality convergence across industrialized countries," *Population and Development Review* 31(4): 645–674.

Fogel, Robert W. 1994. "Economic growth, population theory, and physiology: the bearing of long-term processes on the making of economic policy," *American Economic Review* 84(3): 369–395.

Fogel, Robert W. 2004. *The Escape from Hunger and Premature Death, 1700-2100 – Europe, America, and the Third World.* Cambridge, Cambridge University Press.

Lucas, Robert E. 1988. "On the Mechanics of Economic Development," *Journal of Monetary Economics* 22: 3û-42.

Maddison, Angus. 2001. *The World Economy: Historical Statistics.* Paris, OECD, CD-rom.

Mayer-Foulkes, David. 2003. "Convergence clubs in cross-country life-expectancy dynamics," Chapter 5 in: *Perspectives on Growth and Poverty.* Van der Hoeven, Rolph, and Anthony Shorrocks (eds.) Tokyo, United Nations University Press.

Neumayer, Eric. 2003. "Beyond income: convergence in living standards, big time," *Structural Change and Economic Dynamics* 14: 275 A /296.

Omran, Abdel. 1971. "The epidemiological transition: a theory of the epidemiology of population change," *Milbank Memorial Fund Quarterly* 49: 509–538.

Preston, Samuel H. 1975. "The changing relation between mortality and level of economic development," *Population Studies* 29(2): 231–248.

Pritchett, Lant. 1997. "Divergence, big time," *Journal of Economic Perspectives* 11(3): 3–17.

Ram, Rati. 1998. "Forty years of the life span revolution: an exploration of the roles of "convergence," income, and policy," *Economic Development and Cultural Change* 46(4): 849–857.

Riley, James R. 2005. "The timing and pace of health transition around the world," *Population and Development Review* 31(4): 741–764.

Soares, Rodrigo R. 2005. "Mortality reductions, educational attainment, and fertility choice," *American Economic Review* 95(3): 580–601.

Soares, Rodrigo R. 2007. "On the determinants of mortality reductions in the developing world," *NBER Working Paper No. W12837.*

Vallin, Jacques, and France Meslé. 2004. "Convergences and divergences in mortality. a new approach to heath transition," *Demographic Research*, Special Collection 2, Article 2, 11–44.

World Bank. 2000. *World development indicators 2000.* Washington CD. CD-rom.

UNAIDS and WHO. 2007. *AIDS epidemic update.* Geneva, Joint United Nations Programme on HIV/AIDS and World Health.

United Nations. 2007. World population prospects. The 2006 revision. New York .