

Global Trends in Life Expectancy: A Club Approach

ULLA LEHMIJOKI, Ph.D., University Lecturer
Department of Economics, University of Helsinki

Abstract

This paper discusses the post-war trends in life expectancy worldwide. Even though applying the specification test to a sample of 125 countries suggests that some life expectancy clubs exist, their number and borderlines are not properly distinguished by mechanical splits of the sample. Hence, the clubs are discovered by regression tree analysis. The potential threshold variables are initial per capita income, literacy, fertility change, and the HIV prevalence rate in 2005. Four clubs appear, characterized as High Literacy, Low Literacy, Medium Literacy, and High AIDS, between which considerable life expectancy differentials appear.

Excluding the HIV prevalence rate from the threshold candidates re-allocates a considerable number of the members of the High AIDS club, indicating that incomes, literacy, and fertility are unable to predict AIDS completely. The similarity of economic and demographic conditions in the Low Literacy and High AIDS clubs, however, raises concerns about life expectancy convergence in the future.

Keywords: Life expectancy convergence, Regression tree, Life expectancy clubs

Introduction

Since World War II, many countries have experienced enormous gains in life expectancy and, in the most spectacular cases, have nearly doubled it. Growth has typically been greater the lower was the initial life expectancy, leading to considerable convergence across countries. This convergence, however, has not been equivocal, as some countries seem to be caught up in a low-life expectancy trap and, in the most unfortunate cases, life expectancy has even decreased.¹

Contradictory findings on life expectancy convergence are currently the subject of heated debate. The optimists claim that, in spite of some exceptional cases, the trend is unique and most low-life expectancy countries will catch the high-life expectancy ones. The reason for this is the widespread diffusion of health technologies, driven by market forces and promoted by international health programs. The optimists maintain that the basic health technology is quite easy to adopt and implement, and thus will lead to a rapid decrease in infections and infant mortality in developing countries, whereas

¹ For empirical results, see Ram (1998), Neumayer (2003), Becker et al. (2005) and Riley (2005), to mention only some.

gains are less easily available in more developed countries, where life expectancy gains mostly come through the decrease in cardiovascular diseases (Omran 1971; Deaton 2003; Vallin and Meslè 2004; Edwards and Tuljapurkar 2005; Cutler et al. 2006).

Some recent findings show, however, that optimism may be premature as life expectancy convergence has stuttered (McMichael et al. 2004). The convergence club approach has thus gained acceptance (Mayer-Foulkes 2003; Edwards and Tuljapurkar 2005). This approach assumes that life expectancy exhibits multiple trends, thus producing several groups or clusters, which – to emphasize the mutual proximity of members – are called *clubs* (Mayer-Foulkes 2003).² There are several reasons for the appearance of such clubs. First, since World War II, world incomes have exhibited divergence rather than convergence as poor countries have lagged far behind the rich. Because income is an important determinant of life expectancy (Fogel 1994; 2004; Preston 1975), one expects that it should follow the same trend. Second, human capital shows a strong intergenerational heritage, as less educated people have less capacity as parents and are less able to take care of their children's health. They also educate their children less, which produces higher mortality among their grand-children.

High fertility may also have a negative impact on life expectancy because high-fertility families have more limited resources per child and a short period between births may decrease breast-feeding and endanger the nutrition of infants (Cigno 1998). Because fertility habits are slow to change, a persistent vicious cycle may follow (Becker et al. 1990). Thus, in spite of the diffusion of health technology, unfavorable initial conditions are difficult to overcome, and some countries may be locked into low life expectancy traps. A new problem is the outbreak of the *AIDS* epidemic in poor countries. In many cases, the epidemic is now so serious that life expectancy is decreasing because of *AIDS* deaths alone (Neumayer 2004); the situation is the most critical in Sub-Saharan Africa and alarming in many countries outside this area as well (UNAIDS and WHO 2007). Recently, Caldwell (2000) has connected the high *AIDS* numbers to fertility-friendly cultures which, combined with a long postpartum sexual abstinence among women, may be one of the reasons for the extramarital sexual activities among men typical of many *AIDS* countries. High fertility rates may thus indicate high *AIDS* risk as well.

Consider the post-war trends in life expectancy worldwide. The largest gains have typically been seen in those countries where life expectancy was initially low, which is in line with the convergence hypothesis. Nevertheless, the convergence of life expectancies has not been universal. In the present paper, I ask whether the life expectancy optimism is evidenced by the data. Since an alternative hypothesis is that life expectancy trends can be illustrated and understood more accurately if they are explored in clusters or clubs, I investigate the need to partition the data to see whether multiple trends exist and what causes may produce such multiplicity. An important goal is to

² Analogously, economics applies the concept of convergence club to a group of countries which exhibit convergence of income.

identify the life expectancy laggards and the reasons for falling behind to understand the policies which could promote their life-expectancy growth.

The role of socio-economic conditions as determinants of life expectancy has been widely discussed since Preston (1975). Whether the epidemic *AIDS* should be seen as a primary reason for increased mortality or rather as a consequence of poor socio-economic conditions has recently been debated (Caldwell 2000). In this paper, I try to shed some light on this question by distinguishing the clubs in two steps, first with *AIDS* among the potential threshold variables and, second, without *AIDS*. The differences in the clubs generated by these steps should then give information about the cause-consequence problem in terms of *AIDS*.

The outline of the paper is the following: Section 2 explains the data and the methods, Section 3 performs the specification test and partitions the data into life-expectancy clubs, and Section 4 analyzes the life expectancy trends and the differences between the clubs. Section 5 performs the classification without *AIDS* and discusses some caveats and interesting exceptions in the data. Section 6 summarizes the results and Section 7 closes the paper.

Data and methods

This research explores life expectancy trends from 1960 to 2001. In understanding the need for clubs, one has to suggest variables which might explain why life expectancy has increased in some countries but decreased in others. Given that no unique rule exists for choosing such variables, one can rely on the theories above, which suggest that potential threshold variables are income and human capital. I take these variables as candidates for threshold variables here. As a measure of income, I take the real GDP per capita and the literacy rate as a measure of human capital. To keep them exogenous from life-expectancy change, I measure them at the beginning of the research period. Caldwell (2000) maintains that fertility-friendly cultures promote extramarital sexual relationships, thus increasing the risk of *AIDS*. At the beginning of the research period, fertility in developing countries was universally high, but more variation appeared in fertility change from 1950 to 1960, and I take this as a proxy for the degree to which the culture is fertility-friendly. Finally, the *HIV* prevalence rate in 2005 is taken as one of the threshold candidates. To summarize, the variables and their data sources are:

- Life expectancy* at birth, both sexes. United Nations (2007).
- GDP per capita*, in constant dollars (base year 2000). Maddison (2001).
- Literacy rate*, % of people aged 15 and above. World Bank (2001).
- Fertility change* from 1950 to 1960, children per woman. United Nations (2007).
- AIDS*, measured as *HIV* prevalence rate in 2005. United Nations (2007).

All these variables are presumably exogenous to life expectancy changes from 1960 to 2001. Data for these variables for these years is available for 125 countries. Several interesting countries are thus excluded from the sample because of missing data, the most important being the former Soviet Union countries (see McMichael et al. 2004); unfortunately, these countries provide the literacy and GDP data only from 1970 and 1990 onwards. Figure 1 illustrates the data by showing the change in life expectancy from 1960 to 2001 as a function of life expectancy in 1960.



Figure 1. The change in life expectancy from 1960 to 2001 regressed against life expectancy in 1960.

This paper aims to classify the sample into sub-samples or clubs which exhibit life-expectancy trends as similar as possible. The number and boundaries of these clubs are determined by the regression tree method, which is suggested by Breiman et al. (1984) and Durlauf and Johnson (1995). Regression tree analysis is a data-sorting method which partitions the sample into sub-samples to find the best piecewise linear model. Its main advantage is that it chooses variables which most efficiently discriminate among the data, thus giving important information about the clubs. Technically, its algorithm chooses the threshold variables (and their critical values) to minimize *SSR*, the sum of squared residuals of the model, which is calculated as the sum of *SSRs* in the sub-samples. The algorithm calculates all possible splits for all threshold candidates to find the variable and critical value which minimize the *SSR* of the model in the new clubs.

Only one-step look ahead and binary splits are used. Successive splits grow into a tree starting from the root (the full sample) to leaves (clubs). The algorithm continues the partitioning until it is unable to find further splits to reduce the model's *SSR* or when 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 *SSR* of the model always decreases, every split makes the interpretation of the tree more difficult, so that an excessive number of clubs is not desirable. Several criteria for choosing the best number are available, but I apply Mallow's C_p in this paper.³ A detailed description of the regression tree method is available in Breiman et al. (1984) and Durlauf and Johnson (1995).

Regression Tree Analysis

To see whether there are sub-samples in the data, we partition the sample into two sub-samples according to the means of the potential threshold variables.⁴ Let $E_{i,t}$ stand for the life expectancy in country i at time t and T for the period length (41 years). We estimate the standard convergence equation

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

by least squares for the full sample and sub-samples separately. The null hypothesis is that all 125 countries obey a common linear model. The Wald-test for the similarity of the coefficients α and β in the sub-samples yields F-statistics with a value of 20.74, which, with (4, 121) degrees of freedom, gives a p -value of 0.00. Hence the null is rejected in favor of separate sub-samples.

Even though the specification test above suggests that some sub-samples exist in the data, their number and borderlines are not properly revealed by mechanical splits. We therefore determine the life-expectancy clubs by using the regression tree analysis.

Figure 2 reports the result of the regression tree analysis. The algorithm chooses the first split in terms of the initial literacy rate, indicating that initial literacy most efficiently discriminates between life expectancy gains in countries. The split value is 90.45%. Hence, countries with the initial literacy rate above this value constitute the first life-expectancy club (26 countries). This club is denoted the "*High Literacy*" club (symbol h). The second split comes in terms of *AIDS*; countries with a *HIV* prevalence rate in 2005 above 2.54% constitute the second life-expectancy club, which contains 29 such countries and is called the "*High AIDS*" club (symbol a). The next split, again in terms of literacy with the split value of 58.15%, partitions the rest of the sample into "*Medium Literacy*" and "*Low Literacy*" clubs with 26 and 44 members (symbols m and l) respectively.

³ In the Mallow's C_p , the "punishment" associated with each split is $2 \cdot \min SSR \cdot (K-1)$, in which the $\min SSR$ is the *SSR* of the model in the largest possible tree and K is the number of clubs. The number of the clubs is then chosen to minimize $SSR + \text{punishment}$.

⁴ The first (second) sub-sample consists of countries with a higher (lower) than mean per capita GDP and literacy and a lower (higher) than mean *HIV* prevalence rate and a faster (slower) decrease in the total fertility rate. For the values of these means, see the Appendix.

Figure 3 illustrates how the *SSR* of the model decreases with successive splits. The full sample has an *SSR* of 242.73. After the first split this decreases to 153.51, but the figure reported by the algorithm remains flat, thus indicating that this decrease is relatively small. Only after the second split (the number of clubs becomes three) does the *SSR* take a considerable jump, falling to 69.89. The next split decreases it to 44.56. After this, the *SSR* decreases at a slowing rate, Mallow's *Cp* indicating that the best club number is four, as is illustrated in Figure 2.

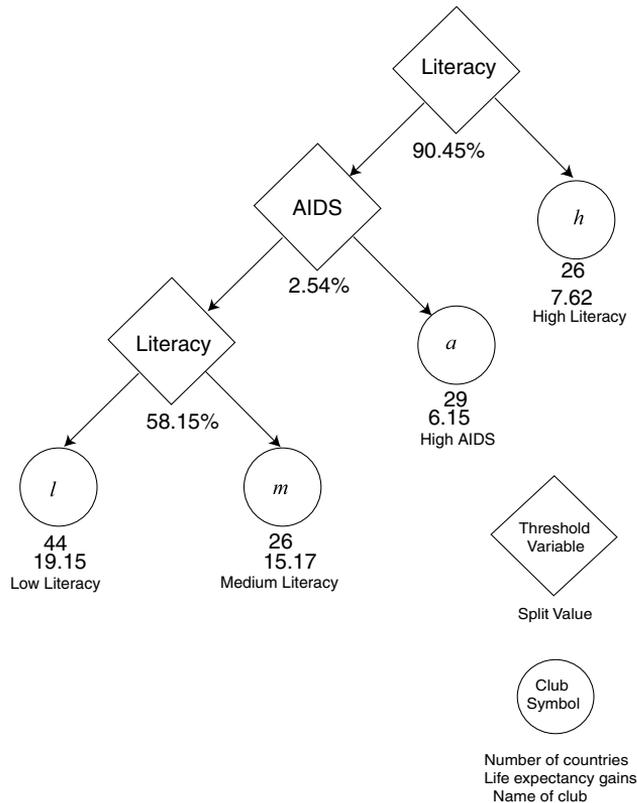


Figure 2. The regression tree. The left arrow indicates the observations for which the threshold variable $<$ Split Value. The right arrow indicates the observations for which the threshold variable \geq Split Value.

The order by which the variables enter to partition the data provides interesting information about their importance. Given that two splits come in terms of the literacy rate, the analysis suggests that literacy is the variable which efficiently discriminates between the clubs. Naturally, it is not the only one. Table 1 shows that the clubs have different initial incomes as well, so that not all life-expectancy disparities can be explained by literacy alone. However, its role is the most decisive, supporting the idea that the effects of the initial human

capital and literacy 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 appears as the critical condition for this growth.

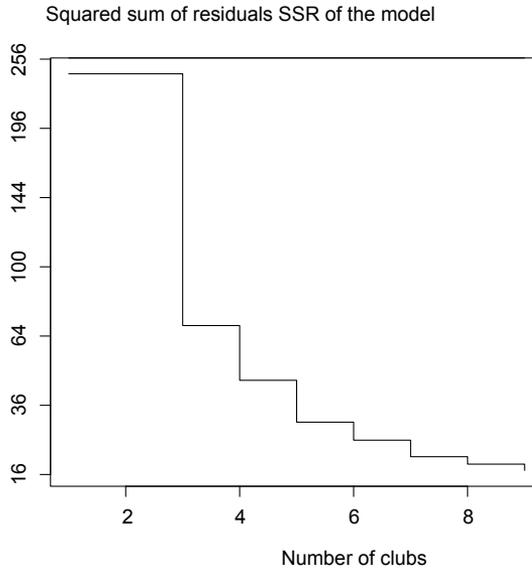


Figure 3. The SSR of the model as a function of the number of the clubs

Table 1. Life expectancy statistics.

Club	LIF60	LIF01	DLIF	LIT60	GDP60	AIDS05	GDP01
High Literacy	69.72	77.34	7.62	98.23	6611.09	0.26	17263.76
Med. Literacy	55.16	71.06	15.17	76.71	2539.62	0.51	7099.89
Low Literacy	43.28	62.43	19.15	29.63	1256.79	0.71	2243.71
High AIDS	42.94	49.09	6.15	34.80	1284.24	9.72	2253.68

Table 1 reports the main statistics in life expectancy clubs. The variables are the club-specific life-expectancy averages in 1960 (*LIF60*), life expectancy in 2001 (*LIF01*), the change in life expectancy between 1960 and 2001 (*DLIF*), the literacy rate in 1960 (*LIT60*), per capita GDP in 1960 (*GDP60*), the *HIV* prevalence rate in 2005 (*AIDS05*), and per capita GDP in 2001 (*GDP01*).

Life-expectancy clubs

Figure 4 illustrates how the full sample is partitioned into four clubs. The most prominent feature is how clearly the countries are classified in terms of their initial life expectancy, even if this was not used as a threshold variable. Figure 4 shows that the High Literacy club consists of high initial life-expectancy countries, whereas Medium and Low Literacy clubs consist of medium and low initial life-expectancy countries respectively. The High AIDS club consists mainly of low initial life-expectancy countries.

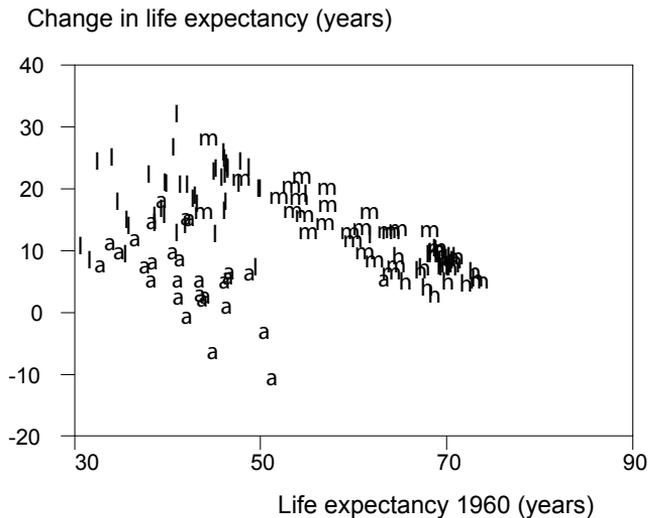


Figure 4. The Life-expectancy clubs.

Figure 5 (panel a) presents the club-specific averages for beginning and end-of-period life expectancies. Compare the High and Low Literacy clubs, for example. In 1960, the average life expectancies in these clubs were 69.72 and 43.28 years. Since life expectancy increased by 19.15 years in the latter club but only 7.62 years in the former, the disparity decreased and the life expectancies were 77.34 and 62.43 years respectively in 2001. Hence, convergence appeared between these clubs, and an analogous observation holds for the Medium Literacy club, which reached the High Literacy club, even though at a slower rate.

Nevertheless, the progress in the Low Literacy club was not satisfactory in the light of its initial life expectancy, as a comparison with the Medium Literacy club clearly shows (see Table 1). In 1960, the gap between High and Low Literacy clubs was more than 26 years, providing considerable potential for catching up to the latter. On the other hand, the gap between High and Medium Literacy clubs was only some six years, implying that potential for improvement in the Medium Literacy club was much smaller. Its life expectancy, however, increased more than twice the gap, whereas the increase in the Low Literacy club was less than its gap. In relative terms, therefore, the Medium

Literacy club proceeded faster. Given that the regression tree analysis above emphasizes the importance of health diffusion, it is likely that the low (initial) literacy in the former club decreased diffusion of new knowledge in the Low Literacy club, thus preventing the high potential for life expectancy growth from being fully realized.

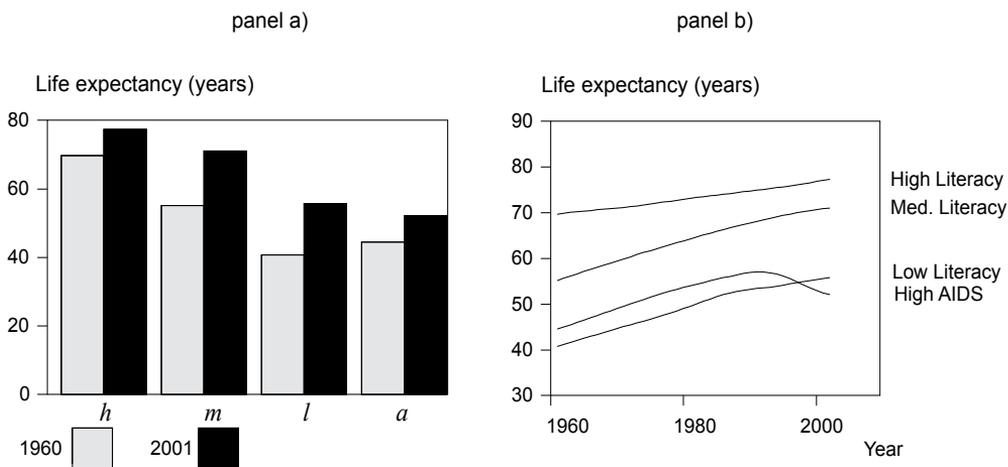


Figure 5. Beginning and end-of-period average life expectancies in clubs (panel a) and trends in life expectancy (panel b).

Soares (2007) provides further insight into the role of literacy by maintaining that the effects of some health measures, such as large-scale immunization and improvement in public health infrastructure, are independent of the personal characteristics of the agents. But since some gains related to preparation of food, water treatment, and nursing infants, for example, derive from personal practices, their effects depend more on absorption of knowledge on the agent's side (Soares 2007). Thus, the observed life-expectancy growth in the Low Literacy club was probably derived from the former types of measure alone, while the Medium Literacy club was able to derive an advantage from both types.

Even though several objections are seen in life-expectancy convergence among literacy based clubs, the High AIDS club is still much more problematic. Starting from similar life-expectancy numbers than the Low Literacy club, it fell behind even in absolute terms, since its average life expectancy increase was only 6.15 years. Figure 5 (panel b) reports the evolution of life expectancy in each club, showing how a promising convergence trend was interrupted by the *AIDS* epidemic in the High AIDS club. The most likely reason for the slow-down in life expectancy growth was this epidemic, since the *AIDS* prevalence number in 2005, reported in Table 1, confirms this. As compared with the Low Literacy club, for example, one sees that the initial literacy rate in the High AIDS club was markedly higher; indicating that slow diffusion of health technology is not the primary reason for the life-expectancy disaster in this club.

Classification without AIDS

The analysis above comes with some caveats. The most critical point is whether the *HIV* prevalence rates in 2005 can be considered as an exogenous factor. This is not self-evident. Some authors actually suggest that *AIDS* is an economic and social rather than medical problem. Even if *AIDS* only appears after contamination by the human immunodeficiency virus *HIV*, the probability of such contamination is determined by social and economic conditions, the roots of which can be found in poverty, low literacy, and fertility-friendly cultures. From this perspective, *AIDS* is a consequence rather than cause (Caldwell 2000). If this is true, the analysis above is misleading in the sense that it has used an endogenous factor as one of the explanatory variables. In this case, even though the analysis above illustrates the past well, its predictive power is not satisfactory.

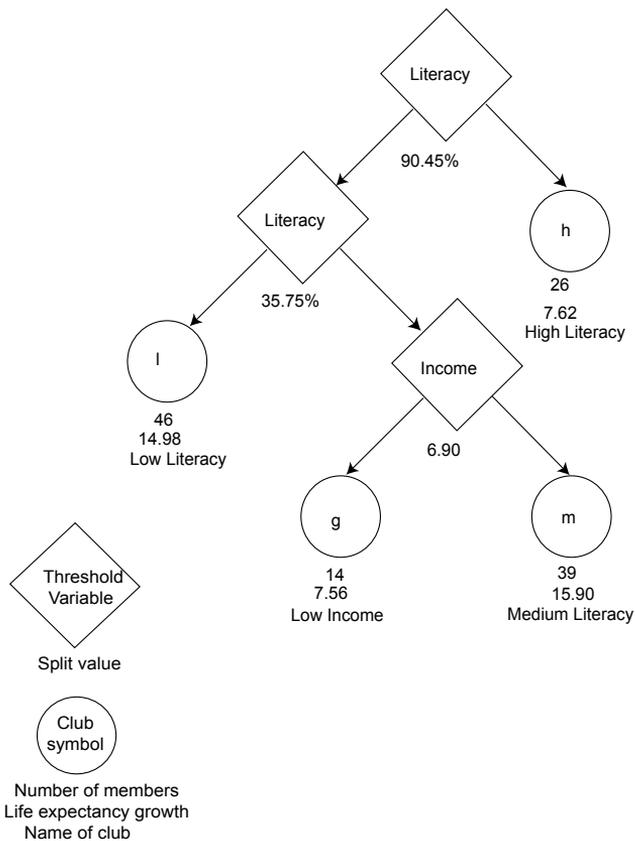


Figure 6. The second regression tree.

Therefore, we re-run the regression tree analysis leaving *AIDS* out of the threshold candidates. The result depicted in Figure 6 shows that the first split is precisely the same as in the first run (Figure 2), but the second split comes in terms of initial literacy and the fourth in terms of GDP per capita. The algorithm identifies High, Medium, and

Low Literacy clubs and the Low Income club (symbol *g*), in which the initial GDP per capita is exceptionally low. Figure 7 (panel a) illustrates how these clubs are allocated in terms of their initial life expectancies. Comparison with Figure 4 shows that, in spite of the changed splitting order, the literacy-based clubs remain relatively similar. Most re-allocations take place in the former High AIDS club, as approximately half of its members (15) are allocated to the Low Literacy club and another half remains in the newly constructed Low Income club, which also gains some members from other clubs. Hence, low initial income does not predict *AIDS* completely. Given that almost all High AIDS countries were in Sub-Saharan Africa, it seems that the virus epidemic nature of *AIDS* is of great importance.

Nevertheless, the average *HIV* prevalence rate in the newly constructed Low Income club (11.14%) was actually higher than the average *HIV* prevalence rate in the former High AIDS club (9.72%). This is intuitively appealing, since it is likely that in case of the circulating virus extreme poverty both increases the risk of contamination and decreases opportunity to treat infected people adequately. In many cases, one can find historical reasons for post-war poverty. The Sub-Saharan countries are good examples. Being seriously damaged by their colonial roots, they had fallen into deep poverty. When the *HI* virus then started to circulate, they were in a very vulnerable state and the life-expectancy toll was dramatic. History explains why African countries were the first to face the exacerbated effects of epidemic *AIDS*. However, if one takes the role of low income seriously, one sees poverty as a risk for this epidemic in the future. Therefore, the income numbers for the year 2001 at Table 1 (GDP01), being remarkably low in the Low Literacy club as well, raises concern over the escalation of *AIDS* outside Sub-Saharan Africa.

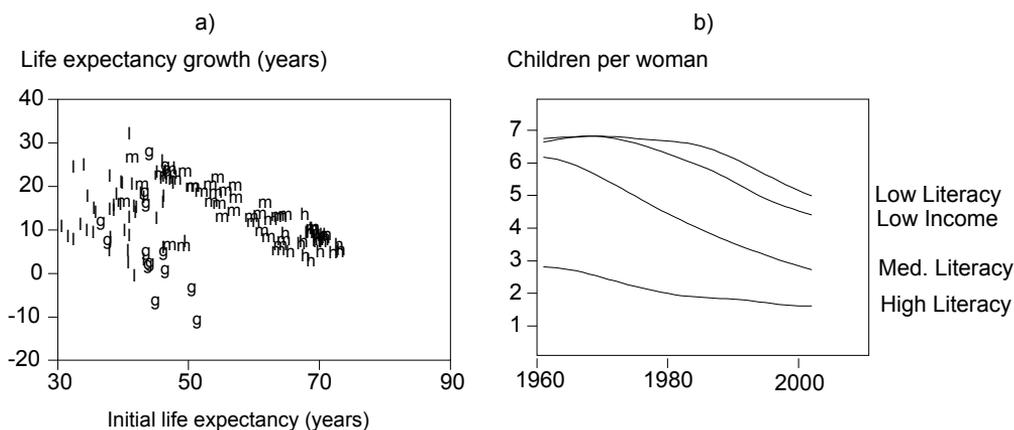


Figure 7. The change in life expectancy from 1960 to 2001 regressed against life expectancy in 1960 (panel a) and trends in fertility in the clubs (panel b).

Another puzzling feature is that the "fertility-friendly" culture, so convincingly connected *AIDS* by Caldwell (2000), plays no role as a threshold variable in the analysis above. A possible explanation is that the - fertility change before the research period, applied as the proxy for "fertility-friendly" culture, does not measure such a feature at the time of the outbreak of *HIV*, which occurred much later. Figure 7 (panel b) illustrates the evolution of the total fertility rate in the life-expectancy clubs during the research period, showing that fertility has however remained high in the High *AIDS* club, indicating that the analysis here is not able to falsify Caldwell's hypothesis.⁵ The worrying fact is that fertility has remained even higher in the Low Literacy club. In 2001, the average total fertility in this club was still 4.99 children per woman and there still were 16 countries with fertility above five children per woman. These same countries were also among the poorest in 2001. Therefore, if Caldwell is right, several countries in the Low Literacy club face a risk of the *AIDS* epidemic escalating in the future.

Results

This paper discusses the post-war trends in life expectancy worldwide. The specification test suggests that some life expectancy clubs exist in a sample of 125 countries, and their number and boundaries are revealed by regression tree analysis.

It turns out that four clubs appear, characterized as High Literacy, Low Literacy, Medium Literacy, and High *AIDS* clubs respectively, between which there are marked longevity differentials. The club with the highest literacy rate proceeds only at a modest rate and the other literacy clubs approach it. This is in line with the convergence hypothesis: the former club has already reached the highest stage of health transition but the latter can adopt and implement new knowledge relatively easily.

The convergence is far from universal, however. The fastest growth was seen in the Low Literacy club, in which life expectancy increased by almost twenty years from 1960 to 2001. In spite of this remarkable achievement, this life-expectancy growth was sluggish relative to the low value of initial life expectancy and the Medium Literacy club actually proceeded faster in relative terms. Since the initial gap between the High and Low Literacy club was considerable, indicating that large potential for improvement exists in the latter, it seems likely that that low literacy has decelerated the diffusion of information, thus preventing the full utilization of the life-expectancy gap.

The increase in life expectancy was slowest in the High *AIDS* club, which fell behind the other clubs since high *AIDS* rates frustrated the longevity achievements derived before the outburst of the epidemic. However, there is some question over whether *AIDS* is a consequence or cause. A second regression tree analysis, performed without *AIDS*, suggests that extreme poverty may explain a great deal of the mortality from

⁵ Since the fertility during the period is endogenous to life expectancy through maternal mortality and replacement fertility, it is not a suitable candidate as a threshold variable in regression tree analysis.

AIDS since most of the original High *AIDS* countries are now allocated to the Low Income club. On the other hand, Caldwell's high fertility hypothesis, the alternative socio-economic explanation for *AIDS*, derives no support in the analysis here. A possible explanation is that this analysis is unable to distinguish and measure the cultural features emphasized by Caldwell (2000).

Conclusions

In comparison with earlier results (see Ram 1998, Neumayer 2003, and Becker et al. 2005, to mention some), the club approach here emphasizes the multiplicity of life-expectancy trends, and the unique rate of convergence, giving rise to overall optimism in terms of future life expectancies, is shown to be less firmly rooted in the data. Classifying the sample indicates that optimism is early since poverty and low literacy seem to stall life expectancy growth in many countries. The finding that poverty explains a great deal of the mortality from *AIDS* also provokes considerable concern about the future life expectancy growth in the poorer countries. Furthermore, because low literacy still seems to limit efficient absorption of health information, continuing efforts are necessary to remove this constraint.

The slow progress in the High Literacy club also needs further research, since some authors have suggested that privatization of health and increasing complexity of health information are its main causes (Soares 2007). We should thus ask whether the slowdown in life expectancy growth in these countries is determined biologically or caused by man-made obstacles (Wallin and Mesle 2004). It is possible that we have accepted this slowdown as an intrinsic element of convergence too easily, and one of the future challenges is to investigate whether it could be alleviated by well-constructed health programs.

The results of this paper are subject to some caveats since several interesting countries are missing from the data. The life-expectancy laggards are now found in poor developing countries, whereas slow growth – and even a decrease – has also taken place in the former Soviet Union countries and its satellites (McMichael et al. 2004). The reasons for these disasters are different from those discussed here as these countries have mainly suffered from heavy drinking, smoking, and violence, as well as from corruption of social structures, and would thus have probably constituted an additional life-expectancy club.

References

- Becker G, K. Murphy and R. Tamura. 1990. Human Capital, Fertility, and Economic Growth. *Journal of Political Economy* 98(5): 12–37.
- Becker G.S., T.J. Phillipson and R.R. Soares. 2005. The Quantity and Quality of Life and the Evolution of World Inequality. *American Economic Review* 95: 277–291.
- Breiman L, J.L. Friedman, R.A. Olshen and C.J. Stone. 1985. *Classification and Regression Trees*. Belmont. CA: Wadsworth.

- Caldwell, J. (2000). Rethinking the African AIDS Epidemic. *Population and Development Review* 26(1): 117–135.
- Cigno, A. (1998). Fertility Decisions when Infant Survival is Endogenous. *Journal of Population Economics* 11(1): 21–28.
- Cutler, D., A. Deaton and A. Lljeras-Muney. 2006. The Determinants of Mortality. *Journal of Economic Perspectives* 20(3): 97–120.
- Deaton, A. 2003. Health, Inequality, and Economic Development. *Journal of Economic Literature* 49(1), 113–158.
- Durlauf, S and P. Johnson. 1995. Multiple Regimes and Cross-country Behaviour. *Journal of Applied Econometrics* 10: 365–384.
- Edwards, R.D. and S. Tuljapurkar. 2005. Inequality in Life-Spans and a New Perspective on Mortality Convergence across Industrialized Countries. *Population and Development Review* 31(4): 645–674.
- Fogel, R.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, R.W. 2004. *The Escape from Hunger and Premature Death, 1700-2100 – Europe, America, and the Third World*. Cambridge, Cambridge University Press.
- Maddison, A. 2001. *The World Economy: Historical Statistics*. Paris: OECD, CD-Rom.
- Mayer-Foulkes, D. 2003. Convergence Clubs in Cross-Country Life Expectancy Dynamics. In: *Perspectives on Growth and Poverty*, edited by R. Van der Hoeven and A. Shorrocks. UNU/WIDER, Helsinki, Finland.
- McMichael, A., M. McKee, V. Scholnikov and T. Valkonen. 2004. Mortality Trends and Setbacks. Global Convergence or Divergence. *Lancet* 363, April 3.
- Neumayer, E. 2003. Beyond Income: Convergence of Living Standards, Big Time. *Structural Change and Economic Dynamics* 14: 275–296.
- Neumayer, E. 2004. HIV/AIDS and Cross-National Convergence in Life Expectancy. *Population and Development Review* 30(4): 727–742.
- Omran, A. 1971. The Epidemiological Transition: A Theory of the Epidemiology of Population Change. *Milbank Memorial Fund Quarterly* 49(4): 509–538.
- Preston, S.H. 1975. The Changing Relation between Mortality and Level of Economic Development. *Population Studies* 29(2): 231–248.
- Ram, R. 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, J.R. 2005. The Timing and Pace of Health Transition around the World. *Population and Development Review* 31(4): 741–764.
- Soares, R.R. 2007. On the Determinants of Mortality Reductions in the Developing World. *NBER Working Paper* No. W12837.
- Vallin, J. and E. Meslé. 2004. Convergences and Divergences in Mortality: A New Approach to Health Transition. *Demographic Research*, Special Collection 2, Article 2, pp. 11–44.
- World Bank. 2001. The World Development Indicators. 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.

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: Brazil, Burma, Sri Lanka, Chile, Colombia, Costa Rica, Cuba, The Dominican Republic, Ecuador, Greece, Israel, Jamaica, South Korea, Lebanon, Malaysia, Mauritius, Mexico, Panama, Paraguay, Peru, The Philippines, Portugal, Singapore, Vietnam, Thailand, Venezuela.

Low Literacy club: Afghanistan, Albania, Algeria, Bahrain, Bangladesh, Bolivia,

Cambodia, China, The Comoros, Benin, El Salvador, Eritrea, Gambia, Ghana, Guatemala, Guinea, Honduras, India, Indonesia, Iran, Iraq, Jordan, Laos, Libya, Madagascar, Mali, Mauritania, Mongolia, Morocco, Oman, Nepal, Nicaragua, Niger, Pakistan, Senegal, Sierra Leone, Somalia, Sudan, The Syrian Arab Republic, Tunisia, Turkey, Egypt, Burkina Faso, Yemen.

High Aids club: Angola, Botswana, Burundi, Cameroon, The Central African Republic, Chad, Congo, Zaire, Equatorial Guinea, Djibouti, Gabon, Haiti, Côte d'Ivoire, Kenya, Lesotho, Malawi, Mozambique, Namibia, Nigeria, Guinea-Bissau, Rwanda, South Africa, Zimbabwe, Swaziland, Togo, Trinidad and Tobago, Uganda, Tanzania, Zambia.

