TESTING CONVERGENCE IN LIFE EXPECTANCIES: COUNT REGRESSION MODELS ON PANEL DATA

Stefano MAINARDI *

Abstract:
Long-term growth convergence has extensively been investigated based on economic variables. Indicators of social development and health status are generally focused on their contribution to growth or on assessing national health care systems. Yet, as a general yardstick of well-being, life expectancy should be regarded as a criterion to measure cross-country development patterns over long periods. Following a review of two approaches to estimating convergence, hypotheses and findings of recent studies on public health and growth are examined. Reformulating the analytical framework of both strands of research, discrete choice and parametric and semi-parametric Poisson regressions are applied to a three-decade panel of 132 countries. Determinants of achievements tend to impact differently across countries, with this distinction occurring particularly between negative and positive counts. Indications of convergence are tempered by results accounting for possible non-linear relationships, which further highlight the discrepancy between country groups with average life expectancy losses and gains.

Keywords: long-term growth, convergence, life expectancy, public health, count regression

JEL Classification: C23, C24, C25, I12, O11

1. Introduction

Average health conditions and economic performance are linked by synergic relationships. Improvements in health status contribute to human capital development and as such are a basic determinant of growth. Conversely, by generally fostering higher standards in health infrastructure and technology, nutrition, sanitation and safety, sustained economic growth can enhance the scope for improved health by the population. While studies on long-term growth have typically focused on the former direction of causality, Barro (2001) has recently examined the influence of the social and economic environment on basic health and demographic indicators. However, the latter analysis is aimed at assessing the quality dimension of econo-

*) Department of Informatics and Econometrics, Card. S. Wyszynski University, Dewajtis 5, 018 15 Warsaw, Poland (e-mail: smainardi@interia.pl).

mic development, not at testing a long-term convergence hypothesis on this dimension, following an approach similar to applications testing convergence in economic growth terms.

Life expectancy has long been regarded as a general yardstick related to several indicators of health and well-being (see Hicks, Streeten, 1979). While output-capital schedule convexity represents a controversial assumption for growth theories, biological constraints necessarily imply a convex health-related production function. This type of function has been adopted by studies in health economics, with a view to evaluate what the health care system is able to achieve across different countries above an imputed minimum situation of full inefficiency, and relative to an estimated maximum potential with available resources. Rather than considering issues of long-term convergence, the objective is to assess and compare national health care systems in a particular period, in terms of intrinsic social goals (see Evans et al., 2001a, Wibulpolprasert, Tangcharoensathien, 2001). Hence, the analysis is still based on functional forms relating *levels* of population health with inputs used to produce health.

The following analysis is aimed at testing long-term convergence in life expectancy across countries at different levels of development. While partly sharing problems encountered in economic growth convergence tests, the use of life expectancy difference-counts as a dependent variable requires the choice of specific functional forms, regression models and control variables. With a view to highlighting similarities and dissimilarities relative to economic growth convergence studies, problems arising from theory and empirical testing are outlined in section 2. Empirical evidence and hypotheses brought forward by previous applied work on public health and growth are discussed in section 3. Drawing on the analytical framework of contributions in both strands, count data models are formulated and applied to 1970 – 1999 cross-country panels. Econometric results are presented in section 4, preceded by a brief discussion on the choice of functional form, variables, and count regression models. Section 5 concludes.

2. Convergence Tests

Assuming decreasing returns to capital, the neoclassical growth model predicts that, under steady state, capital-scarce economies would grow faster than capital-rich ones. By contrast, the additional assumption of heterogeneous agents allows for multiple steady states even in a neoclassical framework of diminishing marginal productivity of capital and constant returns to scale (see Galor, 1996). Given a representative sample period, statistical tests can be distinguished between the Baumol-type regression relating long-term growth with initial levels of income, and the Bernard-Durlauf-Quah (BDQ) approach of testing for the presence of deterministic or stochastic trend components in per capita income gaps. In analytical terms, define T a fixed time horizon, $g_{i,T}$ the average compound rate of growth per capita over a period 0-T, and $y_{i,t}$ the PPP-adjusted per capita income of a sample country. The variable z_t represents per capita income disparity at time t between two countries, or of a country versus, for instance, a sample mean. The regression equations corresponding to the two approaches can be specified as follows, respectively (with level and differenced variables generally expressed in logarithmic form; t, ε_i, and η_t represent a linear trend variable and residual terms, respectively):[1)]

1) Regression (1) is suited to test absolute convergence, and is augmented with the inclusion of control variables to test for conditional convergence. Taking the growth and income per capita variables in lo-

$$g_{i,T} = \alpha + \beta y_{i,0} + \varepsilon_i \quad (1)$$

$$\Delta z_t = \mu + \alpha t + \delta z_{t-1} + \eta_t \quad (2)$$

Baumol-type regressions are typically applicable to cross-section or panel data. Since the estimated β parameter is mostly found to lie within the range (-1, 0), the null hypothesis of no convergence tends to be rejected in favour of convergence in absolute or conditional terms, depending on features of country sample. By contrast, the BDQ procedure, which relies on time series, and can be extended to panel data, has turned out to generally yield opposite results. In the Baumol approach, the use of interim years in panel data allows avoidance of excessive sensitiveness to the choice of initial and final years.

Bernard and Durlauf (1996) demonstrate that the two approaches are incompatible, since they arise from different premises. The first approach is geared to analyze a *weak* concept of convergence, meant as a catch up process of countries in relatively backward position, with key variables being in transition towards their limiting distributions. The second approach checks the equality of long-term forecasts of these variables, subject to the hypothesis that they are near their limiting distributions (specifically, sample means are good proxies of asymptotic means). As such, this approach corresponds to a *strong* kind of convergence, as first defined by Chatterji (1992), namely equalisation of real per capita income to a steady state level. The significance of results from each of the two approaches has been questioned on different grounds, depending on sample characteristics, functional specification and variables used.

The Baumol approach is unable to identify different convergence clusters, thus leading to spurious results if data are generated by multiple long-run equilibria. In heterogeneous country samples, the β parameter can be biased in favour of the null hypothesis due to omission of (steady state-related) control variables, except in the theoretical case of orthogonality between included and omitted regressors. On the other hand, a negative bias in this coefficient (against the no convergence null) can be introduced by the use of some of these control variables, such as typically the rate of domestic investment and population growth. Given their lack of exogeneity to economic growth, a simultaneity bias affects the β parameter (see Cho, 1996). Other explanatory variables may represent fixed effects, which cannot be modified. Alternative cross-country regression specifications, in terms of functional forms of the equation and specification of the variables, have been suggested, so as to allow for dual convergence equilibrium and to account for strong convergence, respectively (see Mainardi, 1995).

In a pooled regression with large *T* and small *N*, a further source of pro-convergence bias may be derived from the coexistence of stationary and non-stationary variables, or level variables characterised by high persistence (see Chumacero, 2001). By contrast, panels with small *T* may lead to the opposite error, namely an underestimation of convergence, given country-specific institutional and technological features. Higher rates of conditional convergence obtained from fixed effects

---

garithmic form, a β estimate lower than –1 would indicate reversal of positions. Analyzed in this form, weak convergence does not guarantee a reduction of absolute gaps. The BDQ test (regression (2)) is illustratively expressed as a Dickey-Fuller (DF) regression. In this case (or its augmented DF version), the no convergence null hypothesis implies at least one of the following: $\mu \neq 0$, $\alpha \neq 0$, $\delta = 0$ (as opposed to the stationarity condition $\delta < 0$). If z_t represents deviations from sample mean values, this model can be respecified as follows (or in its ADF version with no intercept and trend, as in Ben-David, 1996, with critical values nearly *t*-distributed for panel data): $z_t = \delta z_{t-1} + \theta_t$, with the convergence condition becoming $\delta < 1$.

5-year interval 1960 – 1985 panel estimates, compared to single cross-country or OLS pooled regression results, are explained by Islam (1995) in terms of a downward bias due to omitted unobservable individual country effects not accounted for in the latter models. In this respect, OLS pooled regressions typically suffer from temporal and spatial residual correlation and heteroscedasticity (see Beck, Katz, 1995).

In turn, the BDQ approach may spuriously fail to reject the no convergence null in many cases, if data are largely driven by transition dynamics or reflect a mixture of economies close to and far from unique steady state equilibrium. The time series approach is also subject to the partly arbitrary preliminary filtering of cycles out of long-term trends, even if the use of PPP-adjusted data already dampens the variability due to exchange rate fluctuations. An additional drawback of both types of test derives from the fact that the steady states towards which countries are heading may themselves change over time, due to policy changes and technological progress, and results can be sensitive to the choice of the initial and final year. Furthermore, both approaches rely on properties of first (and eventually also second) moments of the distribution of the variables concerned. The tests do not allow to identify changes in the shape of the cross-country distribution of the income variable, or to account for the relevance of different national population sizes and high skewness in the income distribution of some countries (see Quah, 1996; Jones, 1997).

3. Health Indicators and Growth

Over the last three decades, epidemiological conditions in developing countries have changed due to technological advances in cost-effective primary health care, such as the diffusion of oral rehydratation therapy and vaccines with easier storage and delivery. In view of these developments, standard indicators as life expectancy and infant mortality are regarded as less useful to measure the health status of a population. Efforts have been made to estimate a new set of health-related indicators, which isolate periods of forced inactiveness due to disease (disability-free or active life expectancy), or account for different degrees of health impairment during a lifespan (disability-adjusted life expectancy, henceforth DALE). Estimates for these indicators tend not to cover more than 15-20 years for a number of developed countries, while for most developing economies they have started to be computed only for recent years (see Mathers et al., 2000). Cross-country comparisons are hindered by differences in criteria guiding national health surveys. Among industrial countries, a general slowdown in the rate of progression of chronic diseases appears to be accompanied in some cases by persistence of moderate disabilities (see Robine et al., 1999). In broad terms, there seems to be no trade-off between quantity and quality of life, with the two sets of measures being positively associated in most country cases.[2)]

In the absence of sufficiently long and reliable time series on health-related life expectancy, registered life expectancies have been used for this analysis. Despite remarkable achievements in primary health care coverage, at the global level and

2) Except for Niger, which performs relatively more poorly in terms of DALE, and Burundi, for which the opposite appears to occur, the bottom ten countries in life expectancy and DALE in 1999 are the same (all in sub-Saharan Africa). On the opposite extreme, all ten countries with the highest DALE estimated for that year belong to the top twenty of the actual life expectancy scale, which are clustered within the (78, 81) range. According to both indicators, Sierra Leone and Japan occupy the last and first position, respectively (see Mathers et al., 2000; World Bank, 2001).

within developing countries severe imbalances persist in terms of standard health indicators. The insurgence and diffusion of chronic debilitating diseases in several developing economies seem to have even reversed trends towards a narrowing of the gap *vis-à-vis* industrial countries, as likely to occur until the early 1990s. Recent studies in public health economics address the following causal relationships: health policy (+ control variables) → (healthy) life expectancy (+ other health status indicators) → growth. In this analysis, in order to test cross-country long-term convergence in life expectancies, a re-specification of regression (1) has been applied to a partly reversed causal link: initial level of life expectancy (+ control variables including health policy and growth) → changes in life expectancy. However, some of the hypotheses and empirical findings of these studies are functional to the analysis.

As observed for the role of government in economic growth and for other sectors, the kind and quality of government interventions appear to matter more than its relative size in the domestic economy in upgrading the population health status. However, there is no agreement regarding the main determinants of an effective public health policy. After controlling for social and economic conditions, shifting public health expenditures towards primary or preventive care is found by Gupta et al. (1999) to have positive effects on infant and child mortality of fifty developing and transition economies in 1993 –1994. Based on a similar country sample, the same authors (Gupta et al., 2001) find that the amount of public health spending per capita has an impact on the health status of poor households, especially in low-income countries. In both studies, multicollinearity among variables may partly impinge upon the robustness of the results. For the former study, this concerns percentage share of total health care spending in GDP versus two control variables, namely immunisation coverage and adult literacy rate; for the latter, per capita consumption in PPP terms versus private and public per capita health expenses and primary school enrolment ratio (ibid.). Other control variables used in these studies include population per cent share with access to improved sanitation and urbanisation rate. By potentially ensuring easier access to health care facilities, urban areas can be expected to be associated with lower child mortality rates (and higher average life expectancy). However, among others, the lack of affordable services for low-income urban households may cause the opposite impact. In Gupta et al. (1999), the associated parameter turns out to be statistically insignificant, while no results are provided in this respect by Gupta et al. (2001). Rather than the allocation of public spending, for other authors the main problem lies in its efficacy, reflected in uniform and sufficient utilisation of available public resources and a balanced mix of public and private facilities (see Filmer et al., 1997).

Given the strong association between child mortality and life expectancy, cross-country studies focused on the latter variable tend to use similar explanatory variables, and face similar empirical constraints. Possibly to a greater extent than for child mortality, a variable accounting for the relative weight of primary health care may not capture other important differences in national health budget allocations. A programme strongly oriented to fighting fatal, rather than non-lethal, diseases may lead to gains in life expectancy, although this may be accompanied by an increased burden of (non-lethal) morbidity, and vice versa (see Robine et al., 1999). Besides urbanisation and education proxy variables, Barro (2001) models life expectancy as a function of PPP-adjusted per capita income and income inequality, with 5-year interval panel data over the period 1970 – 1998: estimated parameters have the expected sign and are statistically significant. Evans et al. (2001b) choose PPP-adjusted health expenditure per capita and average years of schooling in adult population as broad measures of respectively health and non-health inputs in a health

production equation with DALE as dependent variable, analyzed with 1993 – 1997 panel data. Income per capita is not included, due to its high correlation with both regressors and its only indirect effects on health, through impact on variables as food intake, housing and education. A fixed-effects frontier production model yields estimates of health system efficiency for each country. Countries devoting relatively lesser resources to health (below PPP-adjusted 1997 USD 60 per capita) suffer a higher degree of health system inefficiency, particularly if they have been affected by a history of civil conflict and a high HIV infection rate.

The use of life expectancy as a proxy for living standards has been criticised on the base of its asymptotic upper limits and the more than proportional efforts required by incremental improvements at increasing levels of achieved longevity (see Kakwani, 1993). In recent recession periods these improvements are found to persist, possibly due to the effects of long-term trends and adjustments in household and government expenditures. Based on an achievement-adjusted measure of longevity, cross-country estimates indicate a decreasing elasticity of this index to higher levels of PPP income per capita, and a positive impact of the latter variable and its changes on improvements in living standards (ibid.).[3] Similarly but in reverse terms, in a 1965 – 1990 general cross-country panel, life expectancy is found to affect economic growth in an asymmetric way (see Bhargava et al., 2001). While being strong at low per capita income levels, its influence becomes negligible for upper middle- and high-income countries, thus implying that other health- and human capital-related control variables assume greater relevance in explaining growth.

4. Model Specification and Estimation

4.1 *Testing for Strong Convergence in Life Expectancies*

Hypotheses and empirical evidence reviewed in the preceding section can provide insights for testing convergence in life expectancy rates across countries. Unlike economic growth theories, convexity in a life expectancy production function is an uncontroversial hypothesis, even if there is no consensus in the identification of an upper limit. However, alternative convergence/divergence paths may emerge, depending on whether unique or multiple equilibrium solutions are predominant. Hence, within a sufficiently long period, one should test for absolute or conditional convergence of a *strong* type. Moreover, after controlling for institutional and socio-economic characteristics, transition dynamic determinants of morbidity and mortality affect individual country patterns in different ways. Alternative theoretical approaches can be followed to explain these differences.

In view of the long persistence, frequent near uni-directionality, and, for most country estimates, non-continuous nature of life expectancy data, a preliminary log-transformation of this variable is not necessary, in contrast with empirical analyses on economic variables (see Charemza, Deadman, 1992). With discrete panel data, a suitable model for long-term convergence may be formulated as follows ($\varphi_{i,t}$ is the

3) Kakwani interprets the positive sign of the wealth level variable in the improvement index equations in terms of increasing returns of cumulated infrastructure and health sector investment at higher income levels. This is the conceptual analogue in health and social economics of the divergence pattern hypothesised by some endogenous growth models (such as Romer, 1986). However, the achievement function produces heavy skewness. For instance, a 5-year improvement starting from 35 years of life expectancy is worth only 5.5 % of an equivalent increase reached from 70. Moreover, as in economic growth studies, results may be distorted by the omission of control variables.

residual term, and *m* represents the time span used to measure life expectancy changes for each country *i*, e.g. ten-year periods):

$$(le_{i,t} - le_{i,t-m}) = dle_{i,t} = \nu + \gamma le_{i,t-m} + \varphi_{i,t} \quad (3)$$

The analogy with equation (1) is only apparent. Given the simple linear form, a negative γ parameter in (3) implies *strong* convergence. If variables are measured in deviation form from the respective means, the thus transformed equation becomes the panel version of equation (2) above, except for the zero intercept instead of the terms $\mu + \alpha t$. This does not hold true if means are replaced by median values in this transformation. In terms of conditional convergence, equation (3) can be generalised as follows (the subscript *j* refers to variables within two broad categories, as explained below, and the respective parameters δ_j and ζ_j):

$$dle_{i,t} = \nu + \gamma le_{i,t-m} + \sum_j \delta_j h_{i,t-m,j} + \sum_j \zeta_j o_{i,t-m,j} + \varphi_{i,t} \quad (4)$$

Equation (4) includes indicators of health h_j and other (non-health) control variables o_j (as those indicated below), influencing changes in life expectancy. Relative to health *input* indicators, long-term improvements can be fostered by health systems achieving high levels of population coverage and efficiency (as defined in Evans et al., 2001b), and targeted as a priority in public budget allocation decisions. The relevance of public health expenditures appears to be a controversial issue in studies examining life expectancy in level terms, and may be better assessed once the relative importance of the private health sector is accounted for. Complementary to these health sector characteristics, other potential enhancements derive from basic infrastructure development, as proxied by access to improved water sources, and, in dynamic terms, scope for improvements in nutrition, leisure, housing conditions and education, allowed by high growth rates in income per capita.

The potential positive effects of favourable socioeconomic conditions are severely tempered in the presence of widespread infectious diseases, which have an impact on life expectancies. To this purpose, specific health *output* indicators should also be accounted for, since they act as inputs in equation (4). Apart from the direct impact on mortality rates of infectious diseases, countries where individuals spend on average relatively longer periods of their lives in poor health are in a worse position to increment the overall expected lifespan over time. Substantial differences can be observed according to leading causes of mortality. In all regions except Africa, ischaemic heart, cerebrovascular and pulmonary diseases and respiratory infections tend to rank among the most relevant causes of death (coupled with trachea/lung cancer in Europe, and diarrhoeal diseases in the Middle East and South and South East Asia). By contrast, in Africa HIV/AIDS accounts for almost one fifth of deaths (see WHO, 1999). In view of this heterogeneity and the heavy incidence of HIV/AIDS as a determinant of premature deaths in developing countries in the 1990s, a two-step approach can be adopted, instead of a one-step estimation of (4).

The variable $dle_{i,t}$ can be seen as a discretisation of a continuous time latent variable (which is most often estimated in terms of integer figures), with the bulk of the observations lying within a narrow interval. As in the case of event counts, estimation models should be suited to the restricted and discrete nature of the dependent variable, and should account for possible lack of equidispersion and problems of *apparent* and *true* contagion.

Specification and testing in count regression models. In count model applications, the equidispersion assumption of standard Poisson models is often rejected, due to the presence of long/heavy right tails and consequent disparity between conditional mean and variance values,

even after accounting for various regressors. In this case, ML Poisson estimates are inefficient and tend to yield inflated t statistics (see Cameron, Trivedi, 1998). Often connected although not necessarily associated with overdispersion (extra-Poisson variation) (see Mullahy, 1997), are problems of population mixing and occurrence dependence between events. Population mixing, also defined as *apparent* contagion, arises from cross-section unobserved heterogeneity and misspecified variance functions. Zero-inflated Poisson (ZIP) models account for a special case of unobserved or neglected heterogeneity (and in some cases also of occurrence dependence), with a spurious swelling of zero values caused by a mixture of structural and chance reasons. Occurrence dependence, or *true* contagion, refers to positive or negative dependence between two successive events for an individual, and can result in misspecified mean functions.

Non-homogeneous Poisson models and Poisson-gamma mixtures, such as negative binomial models, applied to overdispersed cross-section data provide more efficient estimates than the standard Poisson. However, inconsistent estimates are obtained if the conditional mean is misspecified, and no distinction between the two possible sources of overdispersion and excess number of zeros is generally allowed by these models (see Winkelmann, 1995; Gurmu, 1997). Empirical evidence by Gurmu and Trivedi (1996) seems to suggest that a negative binomial regression, despite being theoretically superior to the Poisson in tracing overdispersed data, may over-adjust the low prediction power of the latter model for high counts in the presence of heavy tails. Hurdle and finite mixture (including zero-inflated) models provide a more flexible approach, accounting for changes in both mean and variance, and allowing for both over- and underdispersion (see Mullahy, 1986; Cameron, Trivedi, 1998). Other estimators formulated to this purpose have theoretical limitations, given by probability approximations (double Poisson: Efron, 1986) or restrictive assumptions on the possible range of the random variable (generalised Poisson: Wang, Famoye, 1997). Alternatively, absence of equidispersion and excess/scarcity of zeros can be modeled through a semi-parametric ML estimator based on a p^{th}-order power series expansion around a prior baseline density (see Cameron, Trivedi, 1998).

Within the latter approach, Cameron and Johansson (1997) formulate a Poisson polynomial model of order p (PPp), which turns out to be particularly suited to underdispersed data (another application is given by Vredin, Johansson, 1998). With $p = 1$, given a random variable y with baseline Poisson density $f(y \mid \mu)$ and a vector of conditional mean μ, the model (PP1) rests on the following first-order expansion-augmented Poisson density:

$$h_1(y \mid \mu,\alpha) = [(e^{-\mu}\, \mu^y)/y!]\, (1 + \alpha y)^2/\eta(\alpha, \mu) \tag{5}$$

The constant $\eta(\alpha, \mu) = 1 + 2\alpha\mu_i + \alpha^2(\mu_i + \mu_i^2)$ normalises this probability density to a unit sum, and μ and $(\mu + \mu^2)$ are respectively the first and second moment of the baseline density $f(y \mid \mu)$. The first order polynomial term $(1 + \alpha y)$ is squared to ensure non-negativity of the density. Analytical proofs are derived in Cameron, Johansson (1997). PP1 estimates in Table 3 are obtained with NLS applied to relevant explanatory variables, following ML estimation of conditional mean values μ_i on the same restricted model (PH3).

To test for overdispersion due to apparent contagion in large T and large N panels, Ruser (1991) proposes an auxiliary cross-section regression relating estimated variances of residuals to mean dependent variable values, for each individual. For true contagion, the same approach can be applied in a time series framework, with variances and mean values calculated for each period. A more common regression-based test for overdispersion, similarly based on the relationship between Poisson and negative binomial distributions and not allowing a distinction between apparent and true contagion, is given by (Greene, 1998; Cameron, Trivedi, 1998):

$$[(y_{i,t} - \mu_{i,t})^2/\, \mu_{i,t}] = 1 + \alpha(\mu_{i,t})^{\beta} + \eta_{i,t} \tag{6}$$

where $y_{i,t}$ represents observed panel frequencies, $\mu_{i,t}$ is the conditional mean of $y_{i,t}$, and $\eta_{i,t}$ is an error term. Overdispersion is implied by $\alpha > 0$ (against the null of hypothesis $\alpha = 0$ of equidispersion), with the parameter β being generally assumed to be equal to zero (NB1 model) or one (NB2 model). In this analysis, equation (6) is estimated with NLS, so as to simultaneously test for values of both α and β parameters. Given a multivariate equation such as (4), hurdle models are likely to cause or strengthen underdispersion in the dependent variable, due to the inclusion of control variables and censoring/truncation features of these models.

Regarding true and apparent contagion in this application, the evidence of consistent life expectancy improvements over long periods can reflect a genuine increased likelihood of such occurrences for countries with a successful past performance, or be a statistical artefact of aggregation, within this category, of countries with substantial differences in experiencing these changes. Panel data typically help in better distinguishing individual behaviour from unobserved heterogeneity.

Differently from studies on microeconomic data with a zero-bounded discrete dependent variable (see Hall et al., 1986; Deb, Trivedi, 1997; Gurmu, Trivedi, 1996; Gurmu, 1997), cross-country unobserved heterogeneity in life expectancy changes cannot be explained strictly in terms of a principal agent framework, and may concern negative (besides zero and positive) counts. However, this approach may be used for testing similar hypotheses at a macro level. To this purpose, two-step models can be applied, with control variables having a different impact on different country subsets. Among health indicators, the rate of HIV infection is regarded as the main determinant of slowdowns or reversals in life expectancy trends over the last two decades in some countries. A hurdle model implies a clear dichotomy separating countries unable to maintain early achievements within the period analyzed from others, with the latter eventually including countries tracing an upward convergence. As in Ben-David (1997), this is given by a narrowing of cross-country variation due to progressive convergence to the top of the scale. An alternative interpretation can suggest a finite mixture model, thus implying a smooth transition in the changing impact of the variables across the two country sub-populations. The presence of negative counts makes the latter model unfeasible in this case, unless all counts are re-scaled to non-negative values. The re-scaled DLE would be subject to problems of high underdispersion, with the mean exceeding the standard deviation more than sixfold.

4. 2 *Empirical Analysis*

4. 2. 1 *Sample Features and Absolute Convergence*

For most of the variables (see Table 1), the analysis has relied on World Bank and WHO time series (devdata.worldbank.org/hnpstats, www.who.int/whosis, World Bank, 2001 and previous years, WHO, 1999). Additional information has been drawn from supplementary sources, relative to an efficiency index of national health systems (see Evans et al., 2001b), percent of life expectancy spent in poor health (see Mathers et al., 2000) and Summers-Heston PPP-adjusted real per capita income estimates (www.bized.ac.uk:8080/dataserv). The database concerns 132 countries (world except few countries with insufficient available information, including Angola, Yemen and Cambodia, and small Caribbean and Pacific Ocean nations) and three ten-year spans. Average data for the 1990s for former Czechoslovakia, Ethiopia (including Eritrea), Yugoslavia and USSR were constructed based on population proportions of new countries, while the reverse weighting procedure was applied to Germany, for the former Federal and Democratic Republics as separate states.

Due to year-to-year instability of the figures reported for some social indicators in several countries, four variables are in the form of dummies instead of period averages (HEXP, LHC, SDW1/2). Relative to the variable HEXP, the chosen threshold criterion identifies slightly less than one fourth of the country sample with national budget allocations particularly favourable to the health sector. As discussed in section 3, assessments of national health system performance and cross-country estimates of lifespan affected by disability are only available for the last decade. Unlike for public sector per cent share in total health expenditures, whose

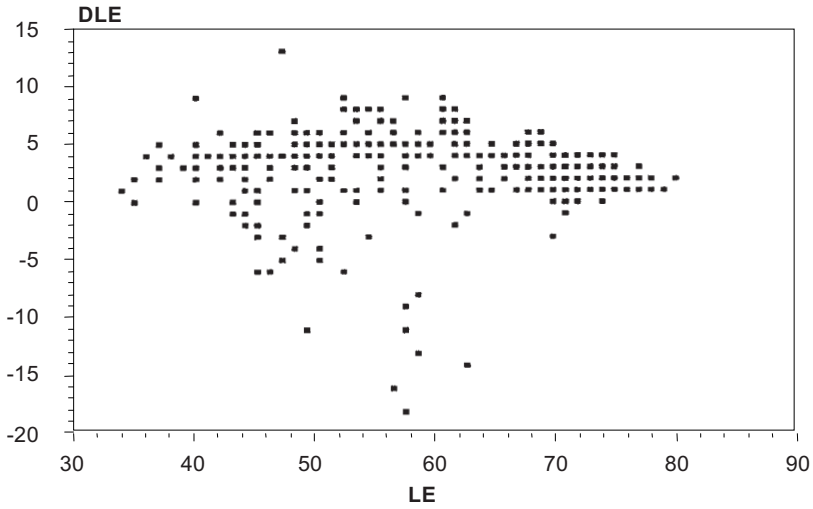
Table 1
List of Variables and Descriptive Statistics

Variable	Definition	Mean	Standard deviation	Skewness	Kurtosis
Dependent variable					
DLE	Difference counts in life expectancy (1970-80, 1980-90, 1990-99)	2.56	3.27	-2.22	13.20
DLECENS	DLE censored from below (0 for DLE ≤ 0) and above (9 for DLE ≥ 9)	2.97	2.05	0.53	2.94
Independent variables (dummies from HEFF1 onwards)					
AGPC	Average compound growth rate in real income per capita (ten-year lag) (PPP-adjusted 1985 USD GDP chain index; 1960-70, 1970-80, 1980-90)	1.79	2.93	-0.56	4.16
HIV	Prevalence of HIV (percentage of adults)	1.19	3.74	4.92	32.90
LE	(initial sub-period) Life expectancy	60.30	11.70	-0.31	1.84
LEFMR	(initial sub-period) Female to male population life expectancy ratio	1.07	0.03	-0.14	6.10
LNPUBH	Public sector per cent share of national health expenditure (around 1995, in natural logarithms; in square brackets: 1990 – 1998 average)	4.01 [3.91]	0.38 [0.45]	-0.66 [-0.99]	2.75 [3.62]
LNURB	Urbanization rate (percentage of population, in natural logarithms)	3.72	0.67	-0.99	3.61
HEFF1	Medium and high efficiency in national health system (1 if efficiency index ≥ 0.6 % in 1993 – 1997; 0 otherwise)	0.68	0.47	-0.78	1.60
HEFF2	High efficiency in national health system (1 if efficiency index ≥ 0.8 % in 1993 – 1997; 0 otherwise)	0.38	0.48	0.50	1.25
HEXP	Relative importance of national health expenditure in GNP/GDP (1 if ≥ 7 % on average over 1980 – 1999; 0 otherwise)	0.24	0.43	1.18	2.40
LHC	High per cent share of population with access to local health care (1 if ≥ 70 % on average over 1980 – 1999; 0 otherwise)	0.70	0.46	-0.90	1.81
SDW1	Substantial *increase* in per cent share of population with access to safe drinking water and improved water source (1 if ≥ 20 % over 1980 – 1999; 0 otherwise)	0.39	0.49	0.46	1.21
SDW2	High *level* of per cent share of population with access to safe drinking water and improved water source (1 if sub-period average ≥ 70 %; 0 otherwise)	0.61	0.49	-0.45	1.20
YPH1	Medium and high percentage of lifespan spent in poor health (1 if ≥ 10 % in 1999; 0 otherwise)	0.73	0.44	-1.06	2.13
YPH2	High percentage of lifespan spent in poor health (1 if ≥ 15 % in 1999; 0 otherwise)	0.32	0.47	0.78	1.60

estimates are also limited to the 1990s, international disparities according to these criteria can be expected to persist over long periods, at least in terms of broad categories. Hence, the respective estimates have been relied on to construct dummy variables for the whole period (HEFF1/2, YPH1/2). In both cases, the chosen thresholds subdivide the sample in three batches of comparable sizes, with the in-between batches covering nearly 31 % and 42 % of the observations, for HEFF and YPH respectively.[4)]

The count and continuous variables appear to stray away from normal distribution properties in terms of skewness and kurtosis, with the dependent variable being leftward-skewed and platykurtic (see Table 1). Long tails and marked skewness especially characterise the variable HIV, which is also the one presenting the highest dispersion in terms of coefficient of variation. To test for absolute convergence, equation (3) was first estimated with maximum likelihood (ML) accounting for multiplicative heteroscedasticity, with the latter being modelled in log-linear form as a function of the variable HIV. Results, based on all 396 observations, confirm the heteroscedastic pattern (visually observable in Figure 1) and point to the presence

Figure 1
Scattergram of Difference Counts versus Initial Sub-period Life Expectancies



4) If used simultaneously, parameters for HEFF1/2 and YPH1/2 should be interpreted in terms of the partly overlapping group nature of these dummies (HEFF1 vs. constant, and HEFF2 vs. HEFF1), thus yielding analogous regression results as separate group dummies. National health expenditures are measured as a percentage of GNP (WHO) or GDP (World Bank), with the latter estimates tending to be slightly higher in some developing countries. Statistical information for this variable is often of poor quality, with no national health accounts officially reported and reliance on national and provincial budgets and household surveys (see World Bank, 2000). National surveys on access to safe drinking water were first conducted in 1980 based on the WHO definition of treated surface water, or untreated water from protected springs and wells, public standpipes or boreholes within reasonable distance from a dwelling unit (200 metres in urban areas, unspecified for rural areas). As from 1990, reported estimates for this variable refer to availability of at least 20 litres a person a day from an improved source, as defined above, within one kilometre of the dwelling (see World Bank, 2001).

of strong convergence for the sample as a whole, with γ equal to -0.066 (1% significance level). The hypothesis of homoscedasticity is rejected at the 1% significance level by Breusch-Pagan Lagrange multiplier test and a Wald test. The log-linear residual variance function is statistically significant in both intercept and (HIV-related) slope parameter, with the respective estimates being 2 and 0.22. By comparison, a simple OLS regression, with heteroscedastic (even if serially uncorrelated according to the DW test) errors, suggests a slower speed of convergence, with γ being -0.025 (10% significance level).

Poisson regressions on non-negative counts lead to similar results, with slow convergence revealed by values of γ approximately equal to -0.02, -0.03 and –0.058, for the simple model and models with fixed and random effects, respectively (statistically significant at the 1% level). The Hausman test rejects the null of no correlation of individual effects with the regressor, thus favouring the fixed to the random effects model as a consistent estimator ($\chi^2(1) = 384.2$). The fixed effects model is appropriate in this case also given the almost complete population coverage by the sample (see Kennedy, 1998). A regression-based test of overdispersion, based on the application of equation (6) on results of the simple Poisson estimates for non-negative counts, fails to reject the null hypothesis of equidispersion.

4. 2. 2 *Conditional Convergence: Ordered Probit and Logit*

To test for conditional convergence, two alternative approaches have been followed. First, one-step discrete choice models accounting for the lack of precise observability of the dependent variable in its continuum, namely ordered probit and logit, are applied to a censored transformation of difference life expectancy counts (DLECENS). Second, two-step count regression models on the original variable (DLE) are tested. The latter approach hypothesises a hurdle-type process underlying different patterns across countries, with a first step based on negative and zero-censored counts (distinguishing among counts, unlike the dichotomous model in standard hurdles), and a second step on non-negative counts truncated at –1 or 0. This assumes that the hurdle is crossed when long-term life expectancy changes are positive, or at least non-negative. The variable YPH1 is dropped from the equations because of multicollinearity problems.

Results of ordered probit and logit are presented in Table 2. Estimated threshold parameters $\zeta(i)$ mark the maximum likelihood probability limits for each unobserved continuous time change in life expectancy in the ith range, corresponding to the ith observed count occurrence (with $i = 0, 1, \ldots m\text{-}1$). As the maximum value in DLECENS, m is equal to 9 after aggregating a higher exceptional sub-period count recorded for Oman in the 1970s. Since information on variable AGPC is missing for this country in the 1970s, in practice DLECENS is right-censored (besides left-censored) only in model OP3. Not surprisingly, Oman ranks first in Evans et al. (2001b)'s health sector efficiency index, and is regarded as a case of successful reduction of child mortality in the last three decades. The threshold $\zeta(0)$ is normalised to zero, and separates in this case non-positive counts from other observations. A positive estimated parameter implies a shift in probabilities from low- to high-count categories with higher values of the associated explanatory variable, and vice versa for a negative parameter (see Gerdtham, Johannesson, 1997). Due to some missing values, regarding especially the variable AGPC, the size of estimation samples falls short of the 396-data panel. Since the variable LNPUBH is only related to the last decade, its inclusion limits cross-country estimates to the 1990s. Although this is not directly relevant for long-term convergence, it helps detect differences between long and medium term effects.

Table 2
Long-term Changes in Life Expectancy: Ordered Probit and Logit Estimates

Variable	Ordered probit [OP1]	Ordered probit [OP2]	Ordered logit [OL]	Ordered probit [OP3]	Ordered probit [OP4] (1990 –1999)	Ordered probit [OP5] (1990 – 1999)
Constant	10.50 (5.20)	10.70 (5.32)	21.20 (5.86)	10.30 (5.17)	19.80 (2.97)	16.90 (3.19)
LE	-0.12 (-10.05)	-0.12 (-8.99)	-0.24 (-10.10)	-0.14 (-11.80)	-0.16 (-2.63)	-0.14 (-3.47)
LEFMR	-2.60 (-1.40)"	-3.22 (-1.73)*	-6.47 (-1.90)*	-1.95 (-1.06)"	-10.50 (-1.66)*	-9.83 (-2.00)
AGPC	0.05 (2.40)	0.02 (1.18)"	0.07 (2.06)		0.01 (0.21)"	0.05 (1.34)"
SDW1	0.61 (3.71)	0.64 (3.82)	0.93 (3.33)	0.54 (3.52)	0.31 (0.61)"	0.79 (2.46)
SDW2	0.22 (1.04)"	0.23 (1.14)"	0.37 (1.06)"	0.20 (0.95)"	0.35 (0.28)"	0.34 (0.79)"
HEXP	-0.07 (-0.40)"	-0.12 (-0.74)"	-0.03 (-0.10)"	-0.11 (-0.72)"	0.14 (0.36)"	0.32 (0.91)"
LHC	0.45 (2.33)	0.42 (2.16)	0.57 (1.75)*	0.59 (3.27)	-0.94 (-1.30)"	-0.22 (-0.57)*
HIV	-0.40 (-8.52)	-0.39 (-7.98)	-0.74 (-9.17)	-0.39 (-9.01)	-0.67 (-3.05)	-0.68 (-3.77)
YPH2	-0.93 (-3.13)	-0.89 (-3.03)	-1.71 (-3.50)	-1.14 (-4.06)	-2.20 (-1.85)*	-0.69 (-1.01)"
HEFF1	0.50 (1.69)*	0.47 (1.63)*	1.02 (2.12)	0.47 (1.64)*	0.53 (0.41)"	0.66 (1.22)"
HEFF2	0.62 (3.73)	0.59 (3.54)	1.07 (3.77)	0.72 (4.75)	0.28 (0.74)"	0.36 (1.09)"
LNPUBH					0.34 (0.68)"	
LNURB	0.27 (1.59)"	0.31 (1.79)*	0.67 (2.43)	0.35 (2.11)	0.76 (1.21)"	0.94 (1.95)
*T*90		-0.34 (-2.07)				
ζ(1)	0.88 (8.01)	0.89 (8.02)	1.72 (7.36)	0.83 (8.26)	1.27 (3.27)	1.06 (4.74)
ζ(2)	1.69 (13.02)	1.70 (13.02)	3.13 (11.64)	1.60 (13.50)	2.59 (6.72)	2.18 (8.97)
ζ(3)	2.43 (17.08)	2.46 (17.23)	4.44 (14.98)	2.33 (18.04)	3.26 (7.05)	2.81 (9.70)
ζ(4)	3.24 (21.33)	3.26 (21.35)	5.90 (18.27)	3.06 (22.00)	4.81 (7.62)	4.05 (10.42)
ζ(5)	3.83 (23.44)	3.86 (23.52)	7.01 (19.53)	3.70 (24.21)		5.05 (7.68)
ζ(6)	4.50 (21.70)	4.51 (21.84)	8.21 (18.43)	4.33 (22.89)		
ζ(7)	4.95 (17.80)	4.98 (17.92)	9.10 (15.71)	4.76 (19.94)		
ζ(8)	5.25 (16.05)	5.28 (16.23)	9.72 (14.03)	5.19 (16.50)		
r(1)	0.24 (4.54)	0.2 (3.9)	0.26 (5.12)			
r(3)	0.16 (3.02)	0.17 (3.34)	0.15 (2.86)			
LnL	-564.7	-562.1	-557.6	-626.8	-97.4	-146.8
χ^2	301.9 (12)	307.0 (13)	316.1 (12)	321.3 (11)	102.7 (13)	128.1 (12)
Pseudo R^2	0.21	0.21	0.22	0.20	0.34	0.30
% correct	0.39		0.4	0.33	0.52	0.46
N	354	354	354	384	90	124

Notes for Tables 2 and 3. Dependent variable: (Table 2) DLECENS; (Table 3) DLE. *T*-statistics in parentheses (5% significance level, except *5-10 %, "more than 10 %). *r*(1), *r*(3) simple autocorrelation coefficients of estimated residuals, lags 1 and 3 (with *t*-statistics). LnL fitted maximum log-likelihood. χ^2 Lagrange ratio test of fitted versus intercept-only log-likelihood (degrees of freedom in parentheses). Pseudo R^2 (likelihood ratio index) = 1 – [LnL/LnL$_0$]. R^2_{dev} deviance R^2 (see Cameron, Trivedi, 1998). *T*90 dummy for last decade (1 for 1990 – 1999; 0 otherwise). ζ(*i*) threshold parameter. *N* sample size.

Ordered probit and ordered logit on panel data yield equivalent results, in terms of sign and relative size of estimated parameters (Table 2: OP1 and OL). This holds true also for diagnostic statistics, reported in Table 2. Both models tend to underpredict the zero/below-zero and unit counts, and fail to trace correctly the long right tail in the frequency distribution. Compared with the actual pattern, an excessive clumping of fitted values can be noticed in the count range (2, 4). Model OP1 ap-

pears to be better suited to distinguish countries and periods with no or negative changes in life expectancy from others. Both regressions are affected by some degree of serial and spatial correlation in the residuals, revealed by statistically significant autocorrelation coefficients of order 1 and 3, respectively.

Within-country residual correlation may be handled with fixed/random effects models or, alternatively, dynamic panel models (see Ruser, 1991). While fixed effects are not applicable in this case, random effects ML fails to converge, which is likely to be due to over-parameterisation. A dynamic model is unsuitable with a 3-period panel and, relative to the dependent variable, in view of equation (4). Spatial (contemporaneous) residual correlation would in turn require accounting for time-related fixed effects, through sub-period dummies or Zellner estimation technique, or the search for omitted variables.[5] Sub-period dummies in model OP1 provide no significant parameter results for the 1980s and highlights a downward shift from 1990, apparently coupled with loss of statistical significance of the economic growth variable. However, the 1990s dummy also fails to remove spatial residual correlation (see Table 2: OP2).

Once a number of control variables is accounted for, convergence is found to occur over the period analyzed. The scope for life expectancy improvements is constrained by levels of diffusion of HIV and other causes of morbidity (YPH2). Among health sector control variables, access to local health care and system efficiency are found to positively influence this scope, while neither favourable budget allocations nor the ownership structure (OP4: LNPUBH) appear to have a significant role.[6] Moreover, countries undertaking substantial improvements in population coverage in access to safe drinking water and adequate water supply sources and undergoing relatively higher growth in the preceding years are more likely to reap benefits as reflected by positive changes in average life expectancy. Once the above factors are accounted for, widespread access to safe water (in level terms), urbanisation, and gender disparity in life expectancy do not appear to have a substantial influence (except according to the logit parameter estimate for urbanisation).

Removing the variable AGPC from the regression does not remarkably alter parameter estimates (OP3). Ordered probit models based on the same regressors, with or without inclusion of an additional variable (LNPUBH), and limited to the 1990 – 1999 period for the dependent variable, show a higher overall explanatory and predictive power. This result is obtained even if some parameters lose statistical significance or become statistically insignificant (with the exception of LE and HIV), and may be partly explained by the smaller range of the dependent variable counts in the sub-sample.

---

5) The introduction of slope dummies for each explanatory variable, besides the intercept dummies, is equivalent to running separate cross-section regressions. The apparent low explanatory power of the models, indicated by the pseudo R^2, should be weighted against the downward bias often associated with the likelihood ratio index in discrete densities (see Cameron, Trivedi, 1998). The statistical package used for this research would allow Zellner estimation of linear and non-linear equation systems, but not specifically for discrete choice models (see Greene, 1998). In any case, this option is not applicable when T is very small. The small T feature should also be considered when interpreting first order residual autocorrelation estimates (see Tables 2 and 3).

6) Regression OP4 was re-run by replacing the 1995 public sector per cent share estimate in health expenditure (from WHO 1999) with the corresponding 1990 – 1998 average figure (see World Bank, 2001). Summary statistics are reported in Table 1, in square brackets. Results (not reported for space reasons), based on a sample of 110 observations, are similar in terms of diagnostic tests and most parameter estimates, with the exception of SDW1 (which turns out to be significant, as in OP5) and LE (with a parameter value of -0.115).

4. 2. 3. *Conditional Convergence: Poisson Hurdle*

Table 3 presents results of Poisson hurdle models. In the first stage, to avoid multicollinearity (added to statistical insignificance for HEFF2), the variables YPH1 and HEFF2 are left out of the regression. First-step results are obtained from a standard Poisson and a zero-inflated Poisson (ZIP). In the second stage, Poisson regressions are applied to non-negative counts. The test for overdispersion based on regression (6) yields statistically insignificant parameter estimates for negative zero-censored counts (PH1), and for the β parameter in non-negative counts (PH2). In the latter case, α is nearly –0.5 (with an associated *t*-statistic of –3.68), thus indicating some degree of underdispersion. These results are substantiated by unrealistically high estimates obtained for upper extreme predicted values, in both stages of the model, if a negative binomial model is applied. Modeling unobserved heterogeneity through a hurdle approach does not remove the mismatch between actual and integer-rounded fitted values around the mode and in the right tail of the distribution (see Table 4 and Note 7). Negative counts are predicted more accurately, although with an upward bias for unit values.[7)]

Table 3
Long-term Changes in Life Expectancy: Parametric and Semi-parametric Poisson Hurdle Estimates

Variable	(-)DLE (negatives, zero-censored: note 7)				DLE (zeros and positives) [PH2]		DLE (zero and positives) [PH3]		DLE (zero and positives) [PP1] (NLS)		DLE (positives only) [PH4]	
	Standard Poisson [PH1]		ZIP									
Constant	-7.25	(-1.45)"	-1.15	(-0.27)"	3.90	(3.83)	3.74	(12.90)	3.70	(33.00)	3.69	(3.65)
LE	0.13	(4.77)	0.17	(4.17)	-0.05	(-8.20)	-0.06	(-11.20)	-0.05	(-27.70)	-0.05	(-8.18)
LEFMR	2.05	(0.40)"	-3.01	(-0.68)"	-0.53	(-0.51)"					-0.17	(-0.16)"
AGPC	-0.19	(-4.63)	-0.22	(-4.51)	0.01	(0.98)"					0.01	(1.00)"
SDW1	0.33	(1.51)"	0.09	(0.42)"	0.18	(2.21)	0.19	(2.67)	0.19	(6.90)	0.12	(1.43)"
SDW2	-1.12	(-3.14)	-0.70	(-1.73)*	0.06	(0.56)"					0.01	(0.13)"
HEXP	-0.28	(-0.9)"	-0.28	(-0.76)"	-0.12	(-1.35)"					-0.12	(-1.30)"
LHC	0.34	(1.22)"	0.28	(0.78)"	0.16	(1.76)*	0.26	(3.13)	0.24	(8.44)	0.14	(1.56)"
HIV	0.14	(9.53)	0.10	(6.22)	-0.19	(-4.94)	-0.20	(-5.27)	-0.19	(-8.70)	-0.13	(-3.17)
YPH2	-0.69	(-1.76)*	-0.72	(-1.07)"	-0.40	(-2.98)	-0.61	(-6.18)	-0.59	(-15.99)	-0.47	(-3.41)
HEFF1	-3.91	(-6.27)	-4.65	(-4.82)	0.16	(1.17)"					0.17	(1.28)"
HEFF2					0.21	(2.59)	0.25	(3.43)	0.23	(8.54)	0.21	(2.58)
LNURB	-0.63	(-2.89)	-1.05	(-2.35)	0.15	(1.84)*	0.18	(2.53)	0.17	(6.30)	0.13	(1.59)"
τ(ZIP)			-0.22	(-0.94)"								
α(PP1)									0.28	(13.94)		
r(1)	-0.09	(1.67)*	0.16	(2.99)			0.2	(3.86)				
r(3)	0.03	(0.56)"	0.11	(1.99)			0.05	(0.97)"				
LnL	-163.1		-133.0		-556.8		-619.9		-427.6		-530.8	
χ^2	583.3	(11)	(Vuong stat.= 4.39)		175.1	(12)	208.8	(7)	(adj. R^2 = 0.84)		150.0	(12)
Pseudo R^2	0.64		0.71		0.14		0.14		0.44		0.12	
R^2_{dev}	0.71				0.47		0.47				0.51	
% correct	0.85		0.84		0.32		0.29		0.91		0.35	
N	354		354		325		355		355		313	

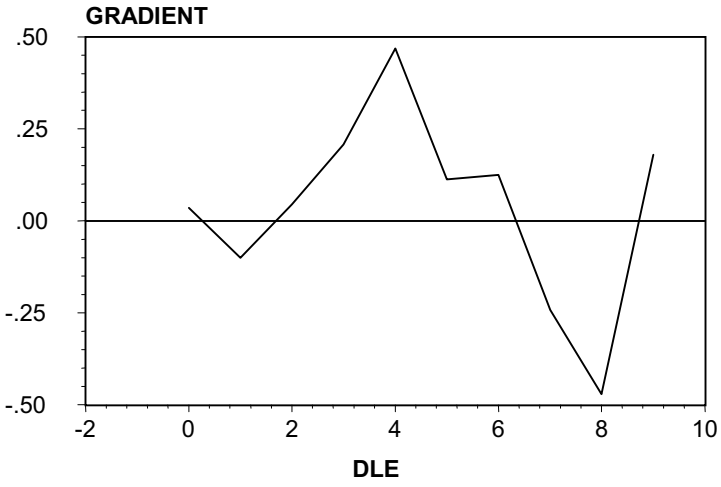
7) As a consequence of this bias, the sum of predicted counts in the two stages of the hurdle model slightly exceeds the total number of counts. Moreover, whereas in ordered probit/logit predicted frequencies are integer numbers, in Table 4 Poisson fitted values are rounded down to respective integers, thus

Table 4
Actual and Fitted Frequency Distributions

Count (*N*=354)	-18 to -15	-14 to -11	-10 to -7	-6	-5	-4	-3	-2	-1	0 (≥ 0 325)	
Actual	2	3	1	3	2	2	4	4	8	12 [≤ 0 41]	
OP1					[35]						
OL					[29]						
PH1-2	[1<-18]	0	3	0	1	0	0	5	22	(322)	6
ZIP-PH3	1	0	3	1	2	1	2	7	21	(316)	6
PP1											14

Count (*N*=354)	0 (≥ 0 325)		1	2	3	4	5	6	7	8	9
Actual	12 [≤ 0 41]		52	69	65	59	32	21	8	3	4
OP1	[35]		22	104	82	85	7	17	0	0	2
OL	[29]		32	99	79	77	16	22	0	0	0
PH1-2	(322)	6	46	134	67	36	22	9	5	0	0
ZIP-PH3	(316)	6	40	144	61	36	22	8	7	1	0
PP1		14	77	106	60	27	14	14	8	2	[3 ≤ 10]

Figure 2
Directional Gradients for Non-negative Counts: OP1 against PH2



spuriously swelling low non-zero counts (see Greene, 1998). Hence, Table 3 should, strictly speaking, be used to compare models within each type, while comparison in terms of overall predictive power should focus on the percentage of correctly predicted counts (based on centred rounding), and directional gradients (see Figure 2). To interpret the estimated parameters in the first step of the hurdle model, one should notice that the dependent variable has reverse sign, so as to allow for count data regressions.

The ZIP model is preferable to the standard Poisson according to the LR-based Vuong statistic, testing non-nested competing families of conditional distributions (see Vuong, 1989). Nonetheless, the parameter τ of the zero-state probability, modelled as a logistic function of this parameter and a vector of the explanatory variables and respective estimated coefficients, is not statistically significant (see Table 4 and Greene, 1998). Moreover, residual autocorrelation coefficients suggest the presence of serial and spatial correlation in the ZIP model, and serial correlation in model PH3 (see Table 3). Among the three regressions compared according to this criterion, only one (PH1) does not appear to violate the conditional independence assumption of the standard Poisson distribution. As in ordered probit and logit models, no ML convergence is obtained if random effects are introduced, while a fixed effects model applied to the restricted specification in PH3 yields insignificant results in all parameters except for LE and, at a nearly 10% significance level, HIV.

On the whole, the hurdle model has more explanatory power in the first step zero-censored sample than in second step sub-samples, according to goodness-of-fit statistics (see Table 3). In conditional terms, convergence occurs more rapidly in the first than the second stage (PH1 and ZIP versus PH2-3-4). In both cases, the diffusion of HIV severely limits the capability of life expectancy improvements, while no influence is associated with relative size of health expenditures and gender disparity. The percentage of lifespan spent in poor health seems to be more relevant for non-negative counts. Relative to other control variables, remarkable differences can be observed between the two stages. A comparison of parameter estimates in terms of statistical significance is limited by the over-estimated standard errors in model PH2, with consequent downward-biased *t*-tests and higher probability of type II error induced by underdispersion.

Access to local health care and high efficiency in the national health system play a relevant role mainly in countries "crossing the hurdle". Before this occurrence, countries with low levels of health system performance (not included in HEFF1) are at high risk. Among non-health indicators, high standards in water and similar utilities, high levels of urbanization and preceding experiences of sustained economic growth put a country in a better position to contain or avoid reversals in life expectancy achievements. By contrast, a successful history of widening of safe water coverage contributes to ensure the ability of countries to raise life expectancy changes beyond the zero-threshold (along with urbanisation) (PH2-3).

Aside from possible measurement errors, the partial failure to correctly trace actual occurrences can be due to misspecifications in the conditional mean and/or the conditional variance functions. The frequency distribution of non-negative fitted counts is not of easy tractability, since it presents features of clustering around the mean, typical of underdispersion, coupled with long rightward skewness, often associated with overdispersion. Among these counts, based on predicted versus observed values of the dependent variable (DLE) in PH2-3-4, overpredictions exceeding three years of difference concern countries suffering from internal conflicts (Gambia and Guinea Bissau in the 1970s, and Sierra Leone in the 1970s and 1980s). Equivalent levels of underprediction are found for a heterogenous group of other countries, mainly limited to the 1970s (Botswana, Costa Rica, El Salvador, Lesotho, Madagascar, Oman, South Korea and Saudi Arabia).

A quadratic term can be introduced for the variable LE in model PH1, to account for possible non-linearities in the conditional mean function and further analyze the convergence process. This specification does not substantially alter the above results on signs and relative sizes of estimated parameters, nor does it produce a noticeable improvement in the regression fit. However, the negative sign of the (statistically significant at the 1% level) parameter of squared initial life expectancy in-

dicates a slowdown of the convergence process in the upper part of the distribution. DLE (with inverted sign) reaches its maximum in terms of first-order conditions at a life expectancy value of slightly more than 65 years. This would imply a lower equilibrium conditional convergence for countries below this expectancy level, with a higher rate of convergence than in PH1, and a process of divergence of these countries from the upper life expectancy group. Using the same non-linear transformation for non-negative counts in model PH2 also yields a parameter with 1% significance level, but removes statistical significance from all regression coefficients except for LE, HIV, HEFF2 and, to some extent, YPH2. The turning point is much lower than in the first step (in a reverse hyperbola, once the inverse sign is accounted for), with very few country/period cases (including Afghanistan, Burkina Faso, Ethiopia, Guinea and Mali) below 45 years of life expectancy, that is the identified turning point, being theoretically cut off from the convergence process. The estimated parameters, for LE and LE^2 respectively, are 0.71 and –0.005 in PH1, and 0.17 and –0.002 in PH2. An additional squared term for HIV has a statistically insignificant estimated parameter.

Regarding possible misspecifications in the conditional variance and limited to non-negative counts (given the above dispersion test results), a Poisson first-order polynomial model is applied (see equation (5)). This model is found to outperform standard Poisson equations in terms of regression fit and prediction power, especially for zero and high counts (PP1 vs. PH2 and PH3: Table 3 and 4). The polynomial coefficient α is statistically significant at the 1% level, and its value falls within the simulated region of underdispersion (see Cameron, Johansson, 1997). By contrast, the use of non-linear least squares (NLS) to estimate the restricted standard Poisson model, including only variables with parameters significant in the unrestricted ML estimation (PH2), yields analogous, though less significant, results as ML estimates of model PH3.

5. Conclusion

If assumed to reflect exclusively current well-being and be highly liable to changes due to, for instance, institutional decay and resource scarcity, life expectancy can be regarded as a *flow* variable, as per capita income (Dasgupta, Mäler, 2001, p. 5: "…Life expectancy at birth this year says nothing about life expectancy at birth next year"). This would contrast with *stock* variables, which can be used as proxies for inter-temporal well-being, such as adult literacy rate. Hence, the rationale for composite welfare indices including both types of indicators would be fundamentally flawed. In this study, the implicit assumption has been that life expectancy is characterised by both flow and stock components. Being itself partly a result of, partly conditioning a continuous upgrade and investment in health, education and infrastructure development, life expectancy can be seen as a constituent of wellbeing through time, especially if disability-adjusted measures are used. Once sufficiently long and comparable time series become available, future research should examine cross-country convergence/divergence patterns in terms of DALE.

This analysis builds on recent literature developments on economic growth convergence and health care system assessment. By the end of last century, most countries have slowed down the process of marked decline in mortality rates, fostered by modern advances in primary health care. In terms of demographic transition dynamics, in the thirty-year period analyzed relatively poorer economies can be located between stage 2 and stage 3, or somewhere within stage 3. In this framework, *weak* convergence is bound to occur in the long run. Alternative models have been applied here to regression equations aimed at testing *strong* convergence in life

expectancy and trying to avoid some of the problems associated with two common approaches to convergence testing, including imbalanced regression, temporal aggregation bias and individual country heterogeneity. Among unaddressed problems, the study could be deepened by accounting for different population sizes and life expectancy inequalities within individual countries. In prospective terms, reciprocal causality lead-lag links between the dependent variable and economic growth could also be considered, including possible negative fiscal effects of gains in longevity along with other demographic changes.

Statistical results apparently suggest that on the whole countries have tended to converge, in absolute (albeit slowly) and conditional terms. The convergence pace does not seem to have substantially accelerated over time, if long-term panel and medium-term 1990 – 1999 period probit estimates are compared. However, relative to absolute convergence over 1970 – 1999, the speed of this process clearly increases if country-specific effects are introduced, through panel models or corrections for heteroscedasticity to a large extent related to the rate of HIV infection. Similarly, conditional convergence speed parameters sharply decline when moving from negative, zero-censored to non-negative truncated counts in Poisson hurdle models, and are relatively high if threshold parameters are used, in ordered probit/logit regressions (where these parameters largely account for fixed effects). While applied to a different growth proxy variable, these results are consistent with Islam's (1995) finding of a downward bias in convergence estimates in a cross-country approach, due to partly unobservable individual country effects. Paradoxically, this leads to underline the importance of effective national health policy interventions.

Discrete choice regressions on zero-censored positive counts produce parameter results which reflect to some extent both sides of the alternative hurdle approach. Relative to 325 non-negative counts, the directional gradient function of average fitted-cell probabilities of models OP1 versus PH2 is plotted in Figure 2. The Poisson hurdle appears to shape better the actual distribution for unit counts and upper counts except for extreme right-tail observations, fare worse for mid-counts (3 to 6), and be nearly equivalent for the remaining low counts. Therefore, evidence is mixed, especially if account is taken of the different probability distributions and the lack of robustness of upper count results due to few observations per count (see Lindsay, Roeder, 1992). However, theoretical reasons for treating negative counts differently from zero and positive counts receive support in Poisson hurdle estimates, with health and non-health indicators having a different impact on life expectancy changes in the two country/sub-period samples. In partial revision of the above results, a diverging pattern of dual convergence arises if account is taken of possible non-linear relationships between life expectancy changes and the respective initial sub-period level variable in both sides of the hurdle. Countries experiencing negative reversals in life expectancy, as is particularly the case in Southern Africa, should be targeted with *ad hoc* strategies.

References

Barro, R. (2001), "Quantity and Quality of Economic Growth," mimeo. Fifth Annual Conference of the Central Bank of Chile, November (www.bcentral.cl/Estudios/Conferencias).

Beck, N., Katz, J. (1995), "What to Do (and not to Do) with Time-Series Cross-Section Data." *American Political Science Review*, 89(3), pp. 634-647.

Ben-David, D. (1996), "Trade and Convergence Among Countries." *Journal of International Economics*, 40(2), pp. 279-298.

_____ (1997), "Convergence Clubs and Diverging Economies." mimeo. Berglas School of Economics (spirit.tau.ac.il/public).

Bernard, A. B., Durlauf, S. N. (1996), "Interpreting Tests of the Convergence Hypothesis." *Journal of Econometrics*, 71, pp. 161-173.
Bhargava, A., Jamison, D., Lau, L., Murray, C. (2001), "Modelling the Effects of Health on Economic Growth." Geneva, WHO, GPE Discussion Paper Series 33.
Cameron, A. C., Johansson, P. (1997), "Count Data Regression Using Series Expansions: with Applications." *Journal of Applied Econometrics*, 12(3), pp. 203-224.
Cameron, A. C., Trivedi, P. K. (1998), *Regression Analysis of Count Data*. London: Cambridge University Press.
Charemza, W., Deadman, D. (1992), *New Directions in Econometric Practice*. Aldershot: Edward Elgar.
Chatterji, M. (1992), "Convergence Clubs and Endogenous Growth." *Oxford Review of Economic Policy*, 8(4), pp. 57-69.
Cho, D. (1996), "An Alternative Interpretation of Conditional Convergence Results." *Journal of Money, Credit and Banking*, 28(4), pp. 669-681.
Chumacero, R. (2001), "Absolute Convergence, Period," mimeo. Fifth Annual Conference of the Central Bank of Chile, November (www.bcentral.cl/Estudios/Conferencias).
Dasgupta, P., Mäler K. G. (2001), "Wealth as a Criterion for Sustainable Development," mimeo. Beijer Institute (www.beijer.kva.se/publications).
Deb, P., Trivedi, P. K. (1997), "Demand for Medical Care by the Elderly: A Finite Mixture Approach." *Journal of Applied Econometrics*, 12, pp. 313-336.
Efron, B. (1986), "Double Exponential Families and Their Use in Generalized Linear Regression." *Journal of the American Statistical Association*, 81(3), pp. 709-721.
Evans, D., Tandon, A., Murray, C., Lauer, J. (2001), "Comparative Efficiency of National Health Systems: Cross-National Econometric Analysis." *British Medical Journal* (323), pp. 307-310(a).
_____ (2001), "The Comparative Efficiency of National Health Systems in Producing Health: An Analysis of 191 Countries." Geneva, WHO, GPE Discussion Paper Series 29(b).
Filmer, D., Hammer, J., Pritchett, L. (1997), "Health Policy in Poor Countries: Weak Links in the Chain," mimeo (econ.worldbank.org/docs).
Galor, O. (1996), "Convergence? Inferences from Theoretical Models." *The Economic Journal*, 106(3), pp. 1056-1069.
Gerdtham, U. G., Johannesson, M. (1997), "The Relationship between Happiness, Health and Socio-Economic Factors: Results Based on Swedish Micro Data." Stockholm School of Economics, Working Paper Series in Economics and Finance, 207 (netec.wustl.edu/wopec/data/papers).
Greene, W. H. (1998), *LIMDEP Version 7.0. User's Manual*. Econometric Software Inc. New York.
Gupta, S., Verhoeven, M., Tiongson, E. (1999), "Does Higher Government Spending Buy Better Results in Education and Health Care?" Washington, IMF, *IMF Working Papers*, WP/99/21.
_____ (2001), "Public Spending on Health Care and the Poor." Washington, IMF, *IMF Working Papers*, WP/01/127.
Gurmu, S. (1997), "Semi-Parametric Estimation of Hurdle Regression Models with an Application to Medicaid Utilisation." *Journal of Applied Econometrics*, 12, pp. 225-242.
Gurmu, S., Trivedi, P. K. (1996), "Excess Zeros in Count Models for Recreational Trips." *Journal of Business and Economic Statistics*, 14(4), pp. 469-477.
Hall, B., Griliches, Z., Hausman, J. (1986), "Patents and R&D: Is There a Lag?" *International Economic Review*, 27, pp. 265-283.
Hicks, N., Streeten, P. (1979), "Indicators of Development: The Search for a Basic Needs Yardstick." *World Development*, 7, pp. 567-580.
Islam, N. (1995), "Growth Empirics: A Panel Data Approach." *Quarterly Journal of Economics*, 110(4), pp. 1127-1170.
Jones, C. I. (1997), "On the Evolution of the World Income Distribution." *Journal of Economic Perspectives*, 11(3), pp. 19-36.
Kakwani, N. (1993), "Performance in Living Standards. An International Comparison." *Journal of Development Economics*, 41, pp. 307-336.
Kennedy, P. (1998), *A Guide to Econometrics*. Oxford: Blackwell.
Lindsay, B., Roeder, K. (1992), "Residual Diagnostics for Mixture Models." *Journal of the American Statistical Association*, 87(3), pp. 785-794.

Mainardi, S. (1995), "Mineral Resources and Growth. Towards a Long-term Convergence?" *Resources Policy*, 21(3), pp. 155-168.
Mathers, C., Sadana, R., Salomon, J., Lopez, A. (2000), "Estimates of DALE for 191 Countries: Methods and Results." Geneva, WHO, GPE Discussion Paper Series, 16.
Mullahy, J. (1986), "Specification and Testing of Some Modified Count Data Models." *Journal of Econometrics*, 33, pp. 341-365.
_____ (1997), "Heterogeneity, Excess Zeros, and the Structure of Count Data Models." *Journal of Applied Econometrics*, 12, pp. 337-350.
Quah, D. T. (1996), "Two Peaks: Growth and Convergence Models of Distribution Dynamics." *The Economic Journal*, 106(3), pp. 1045-1055.
Robine, J. M., Romieu, I., Cambois, E. (1999), "Healthy Expectancy Indicators." *Bulletin of the World Health Organization*, 77(2), pp. 181-185.
Romer, P. M. (1986), "Increasing Returns and Long Run Growth." *Journal of Political Economy*, 94, pp. 1002-1037.
Ruser, J. W. (1991), "Workers' Compensation and Occupational Injuries and Illnesses." *Journal of Labour Economics*, 9(4), pp. 325-350.
Vredin-Johansson, M. (1998), "What Determines the Number of Alternative Travel Modes for Work Trips?" *Umeå Economic Studies* (www.econ.umu.se/ues).
Vuong, Q. H. (1989), "Likelihood Ratio Tests for Model Selection and Non-nested Hypotheses." *Econometrica*, 57(2), pp. 307-333.
Wang, W., Famoye, F. (1997), "Modelling Household Fertility Decisions with Generalized Poisson Regression." *Journal of Population Economics*, 10(3), pp. 273-283.
Wibulpolprasert, S., Tangcharoensathien, V. (2001), "Health System Performance – What's Next?" *Bulletin of the World Health Organization*, 79(6), pp. 489.
Winkelmann, R. (1995), "Duration Dependence and Dispersion in Count-Data Models." *Journal of Business and Economic Statistics*, 13(4), pp. 467-474.
WB (2000), *World Development Report 2000/2001*. Washington DC: World Bank.
___ (2001), *World Development Indicators*. Washington DC: World Bank.
WHO (1999), *The World Health Report 1999*. Geneva: World Health Organization.