An Econometric Investigation into Variations in Life Expectancy at Birth across Countries

Unmilan Kalita

M. Phil Scholar, Department of Economics, Gauhati University, Assam, India

Abstract: ***Life expectancy at birth has garnered prominent attention among policymakers and researchers worldwide due to its pivotal role associated with the degree of social and economic development among the countries worldwide. World Bank and United Nations have time and again asserted its significance and as such, economists have, over the years, studied the impact of different socio-economic factors on life expectancy using various econometric modelling techniques. These studies bear importance in the sense that they work as guidelines for setting out policies which have bearings on the overall well-being of an economy. The present study envisages to examine the impact of economy, human development, demography, health, education and environment on the life expectancy at birth across 40 nations using a linear and a double-log regression model. The observations indicate several expected relationships between the variables, with a few unexpected results. The results, in general, traces out empirical observations that have been experienced in this field of study and solidifies their stance vis-à-vis international policymaking efforts.***

Keywords: life expectancy at birth, OLS, economy, fertility, population, environment

1. Introduction

It has been more than 500 years since Juan Ponce de Leon set out on his quest to find *FonsJuventutis* and until this day, the general public as well as researchers in many disciplines aspire to understand the intricacies of prolonging life. Such a quest continues even today wherein advances in medicine and diet routines have enhanced the health and nutrition status of people. It is undeniable that nutritious food in the right amount, supplemented by medicinal support in the form of vaccinations, prevention of disease outbreaks and robust policymaking, can substantially impact the life expectancy of individuals while social unrest, malnutrition and epidemics can drastically reduce their lifespans.

Life expectancy at birth (LEB) is a pivotal indicator of the nature of development of a nation, that successfully reiterates the health status of its population while acting as a suitable benchmark against which development status of nations can be compared and derived upon. The level of development can be gauged upon vis-à-vis investments in social wellbeing, sanitation, education, ecological stability, environmental improvement and sustainable development. An increased life expectancy additionally indicates the improved levels of per capita income of a nation that can have substantial impacts upon poverty alleviation and livelihood status of a nation's people. Globally, life expectancy at birth has been calculated on an average as 71.5 years with 68 years for males and 72 years for females (UN, 2015). Interestingly, this estimate varies widely across the world with Monaco having the highest life expectancy at 89.52 years and Chad being the lowest with just 49.81 years (World Factbook, 2015). A significant feature of LEB trends is the existence low rates in Sub-Saharan nations which indicate their otherwise prominent features of malnutrition, poverty and social unrest. OECD (2017) states that gains in LEB may be brought about by better education and improved standards of living of people, along with a web of social and economic factors that can invariably uplift the health status and overall economic position of a nation.

The present paper envisages to examine the impact of economy, human development, demography, health, education and environment on the life expectancy at birth across 40 nations. The study involves an econometric cross-sectional analysis and the data has been collected from secondary sources that have been standardised to fit the needs under our study. Multiple regression models have been used in drawing conclusions about the significance of incorporated explanatory variables in the analysis. Section 2 includes an explanation of the material and methods used in the study. This is followed by section 3 which pertains to the results and discussions, while section 4 concludes the study.

2. Literature Review

While conducting a review of research done in the context of life expectancy, it is noticeable that the area of literature is indeed vast. However, such literature pertains to numerous scientific and social fields, and given our scope of study, constricting the review to relevant fields in necessary. To being with, the relationship between income and life expectancy was shown by Rogers (1979) who found it to be directly related. But previously Grossman (1972) had noted that inflation affects life expectancy negatively along with impacting household welfare. Besides this, Kuhn (1992) found that investing in social security and health schemes had favourable impact on the mortality rate of people. Sala-i-Martin (1995), in a famous study, studied the relationship between economic growth and life expectancy and had found it favourable as well. Ehrlich (1991) had also earlier observed that investing in human capital increased life expectancy. Regarding determinants of life expectancy, Rogers (1989) had found that urbanisation, agriculture, literacy, sanitation, food intake and health facilities played significant role in supporting LEB across 95 nations. This was reiterated by Mahfuz (2008) and Posnet (1992).

World Bank in a prominent study noted the positive relationship between LEB and per capita income with respect to developing countries (WB, 1997). It was done based on a cross-sectional study across most nations. A similar study was done by Hussain (2002) for 91 countries using fertility rate, GDP per capita, literacy and food intake. He observed significant relationships among the variables. Ohemeng (2015) did the study on socio-economic determinants of LEB for Nigeria using VAR and VCEM, and observed that per capita income, education and government spending significantly affected life expectancy. Delavari et al (2016) conducted a study of LEB and socio-economic determinants for Iran with respect to GDP per capita, availability of doctors, urbanisation, food security, carbon emissions, total fertility rate, inflation and literacy status, from 1985-2013, using OLS model. They observed that GDP per capita, number of doctors, literacy rate and food security had a positive and significant impact on LEB while fertility rate, urbanisation, carbon emission and inflation did not have any significant impact. Bayati et al (2013) in a prominent study on determinants of life expectancy in the Eastern Mediterranean Region, observed that per capita income, education index, food security, urbanisation and employment ratio had significant impact on LEB. They concluded that to improve health status in such nations policymakers had to focus on building policies that target factors external to the healthcare system. Besides this, Monsef et al (2015) in his study of determinants of life expectancy using a panel data approach observed that unemployment and inflation severely impacted LEB along with capital formation and gross national income.Urbanisation was found to be a major socio-environmental cause impacting LEB.

Therefore, given the brief literature review¸ it has been understood that a study determining the impact of factors as has been taken in our study has not been done. Since it is the target of a researcher to fill in necessary gaps in research, it becomes imperative that a study is done covering the variables and their impact on LEB. Moreover, the ordinary least squares (OLS) technique has been predominantly used in computing the impacts and as such, given its simplistic nature of usage, we have considered to employ the same under our study.

3. Conceptual framework

Life expectancy at birth: LEBis the average number of years that a newborn could expect to live if he or she were to pass through life subject to the age-specific mortality rates of a given period (UN 1994). For our study, this has been chosen as the dependent variable. The data for 40 countries has been collected. LEB is regressed with the following explanatory (independent) variables:

1) **Economy**: For economy, we have considered GNI and human development estimates as proxy variables.
   a) **Gross National Income (GNI)** per capita at purchasing power parity (PPP $) has been taken so as to compare economic statistics across countries. GNI per capita is defined as the gross domestic product (GDP) plus net receipts of primary income (employee compensation and investment income) from abroad, divided by mid-year population. GNI PPP is gross national income per capita converted to international dollars using purchasing power parity rates (World Bank).
   b) **Human Development**: Dummy variables pertaining to the state of human development of a nation were taken under consideration which is an important indicator of life expectancy. The categorical variables as given below were taken in accordance with the categories of nations specified in the Human Development Report (UNDP, 2016). The intention for placing such a variable into our analysis was motivated by the intuition that higher the human development of a nation, the longer their citizens will live.
   Very High Human Development: base category variable
   High Human Development = 1, otherwise = 0
   Medium Human Development = 1, otherwise = 0
   Low Human Development = 1, otherwise = 0

2) **Demography:** Annual Population Growth rate (%) as an explanatory variable was considered relevant because increasing population can lead to shortage of resources and decreasing prosperity as a nation's wealth must be spread more among its individuals. It is defined as theaverage annual rate of change of population size during a specified period (say, a year).
3) **Health:** Health includes availability of physicians & the total fertility rate as proxy variables.
   a) **Physicians (per 10,000 people):** Greater the number of physicians implies greater chances of receiving effective treatment when in need. It is defined as theNumber of medical doctors (physicians), both generalists and specialists, expressed per 10,000 people.
   b) **Total Fertility Rate (births per woman):**It is defined as the number of children who would be born to a woman if she were to live to the end of her child-bearing years and bear children at each age in accordance with prevailing age-specific fertility rates.
4) **Education:** The more knowledge an individual has, the more he or she can make informed life decisions, and improve his/her quality of life. Population with at least some basic education (%) was considered as an indicator for educational attainment.
5) **Environment:** Environment soundness was measured in the form of access to safe drinking water (% of total population) and total forest cover (% of total land area).
   a) **Access to safe drinking water (%):** It is measured by the proportion of population with access to an adequate amount of safe drinking water located within a convenient distance from the user's dwelling.
   b) **Forest cover (%):**It is defined as the amount of forest area tracked over time (say, a year).

Objective
To examine the impact of economy, human development, demography, health, education and environment on life expectancy at birth.

4. Limitations of the Study

The study is entirely based on secondary data and hence, is susceptible to inconsistencies in the tabulated data. However, it must be noted that the data used is internationally standardised and collected from prominent agencies. The domain of study has been constricted to 40 nations as data for all the nations is not available. Additionally, we have used the OLS model for the linear regression analysis. This forms one of the basic econometric models and hence, is unable to factor in several effects that are otherwise factored in by panel data models. Given the scope of our study, this model has been found suitable and it has been expected the results will be free from inconsistencies in estimation.

5. Materials and Methods

5.1 Assumptions on the regression model:

1) It is assumed that our independent variables employa significant influence on LEB of the nations. The relationship between LEB and the explanatory variables is assumed to be linear and subject to random error.
2) Since, data for all the nations was not readily available, we have assumed that our sample of 40 countries is a good reflection of the overall world population, and that variables significant in our model will also apply to other nations as well.
3) Since the data is entirely secondary, inconsistencies in datasets is viable. However, for our study, we assume that the data is consistent. Moreover, it is also assumed that socio-economic situations across nations arent extremely different.

5.2 Specification of the models:

The following two linear regression models were fitted to the data:

a) Linear model:

$$LE = \alpha + \beta_1(GNI) + \beta_2(POP) + \beta_3(PHY) + \beta_4(FER) + \beta_5(EDU) + \beta_6(WAT) + \beta_7(FOR) + \delta_2(HHD) + \delta_3(MHD) + \delta_4(LHD) + u \quad \text{eq (1)}$$

where, α is an intercept, $\beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6, \beta_7$ are the regression coefficients of the explanatory variables, $\delta_2, \delta_3, \delta_4$ are the coefficients of the dummy variables and u is the random error term. The variable abbreviations are elaborated in table-1.

b) Cobb-Douglas form (double-log model):

Given our variable specifications (table-1), the model can be expressed as -

$$LE = a\,(GNI)^{\beta_1}(POP)^{\beta_2}(PHY)^{\beta_3}(FER)^{\beta_4}(EDU)^{\beta_5}(WAT)^{\beta_6}(FOR)^{\beta_7}e^{\delta_2(HHD)}e^{\delta_3(MHD)}e^{\delta_4(LHD)}e^{u} \quad \text{eq(ii)}$$

Since the model (ii) is non-linear in the parameters, we log-transform it to obtain a linear regression model –

$$\ln(LE) = \ln a + \beta_1 \ln(GNI) + \beta_2 \ln(POP) + \beta_3 \ln(PHY) + \beta_4 \ln(FER) + \beta_5 \ln(EDU) + \beta_6 \ln(WAT) + \beta_7 \ln(FOR) + \delta_2(HHD) + \delta_3(MHD) + \delta_4(LHD) + u$$

$$= \alpha + \beta_1 \ln(GNI) + \beta_2 \ln(POP) + \beta_3 \ln(PHY) + \beta_4 \ln(FER) + \beta_5 \ln(EDU) + \beta_6 \ln(WAT) + \beta_7 \ln(FOR) + \delta_2(HHD) + \delta_3(MHD) + \delta_4(LHD) + u \quad \text{(iv)}$$

where, $\alpha = \ln a$.

Thus, the model (iv) is linear in the parameters $\alpha, \beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6, \beta_7, \delta_2, \delta_3, \delta_4$. This model is also known as a log-log or double-log model.

Table 1: Description of variables and expected signs of coefficients

	Variable name	Variable description	Expected impact on life expectancy
1	LE	Life Expectancy at Birth (in years)	----
2	GNI	Gross National Income (PPP $)	Positive
3	POP	Annual Population Growth Rate (%)	Negative
4	PHY	Physicians per 10,000 people	Positive
5	FER	Fertility Rate (births per woman)	Negative
6	EDU	Population with at least some basic education	Positive
7	WAT	Access to safe drinking water	Positive
8	FOR	Forest cover (% of total land area)	Positive
9	HHD	High Human Development = 1, otherwise = 0	----
10	MHD	Medium Human Development = 1, otherwise = 0	----
11	LHD	Low Human Development = 1, otherwise = 0	----

Table 2: Descriptive statistics of the variables

Sl. No.	Variable	Maximum	Minimum	Mean	Standard Deviation
1	LE	83.30	53.6	68.7534	14.227
2	GNI	62471	1098	14249.90	14943.711
3	POP	5.90	-2.50	1.7275	1.450
4	PHY	67.2	0.4	16.033	16.638
5	FER	5.70	1.20	3.125	1.409
6	EDU	100	5.20	55.6350	26.675
7	WAT	100	39	83.825	17.548
8	FOR	70.10	0.00	28.425	21.443

5.3 Materials

- Secondary research & data collection: The models discussed above were estimated using international cross-sectional dataof 40 countries. The basic criterion for choice of the countries was that the relevant data was readily available. The secondary data has been obtained from the World Bank, UNDP (HDR) and United Nations websites, the links to which have been mentioned in the bibliography.
- Model: As stated in the previous section, two linear models were separately used for the analysis - a linear regression model (i) and a double-log regression model (iv).
- Software: To perform the analysis, statistical software IBM SPSS v23.0& Microsoft Excel, 2013 were used.

6. Results and Discussion

The significance of variables in the present analysis has been examined through both linear and double-log regressions. The results are reported in Table-3.[iii] The estimated coefficients of GNI, PHY, EDU, FOR in linear model and coefficients of PHY, EDU, WAT in the double-log model have expected signs as given in table-1. The estimated GNI, FER & the dummy variables HHD, MHD and LHD turned out to be statistically significant in both the models. The coefficient of POP turned out to be highly significant in the linear model.

Surprisingly, coefficients of PHY and WAT turned out to be statistically insignificant in both the models.

The **estimated linear model** can explain 79.8% variation in life expectancy at birth in our study of 40 countries. Although the goodness of fit is moderately high, only 6 variables out of 10 exert great significance with respect to life expectancy at birth due to discrepancies existing within the data set. Some of the important inferences drawn from the analysis of the linear model are mentioned below:

- The insignificance of EDU is surprising, as its p-value is near one. It seems counterintuitive that basiceducation does not influence life expectancy.
- The sign of estimated POP came out to be positive which indicates a positive relationship between life expectancy at birth and population growth rate.
- The sign of estimated WAT indicatesa negative relationship with life expectancy at birth which is contrary to prevalent theory. This necessitates further explanation.
- Also surprising is that estimated HHD and MHD dummy variables have negative signs because once again, it would makesense that a country that is more developed would be able to provide a better standard ofliving than an underdeveloped country.

From the estimates, it can be seen that the variables Gross National Income (PPP $), annual population growth rate (%), fertility rate (births per woman), dummy variables high human development, medium human development and low human development exert high influence on life expectancy at birth.

GNI (PPP $) has a positive correlation with life expectancy at birth which approves of their usual relationship that individuals born in wealthier countries, on average, can expect to live longer than those born in poor countries (Preston, 1975). This is evident from the scatter-plot below.

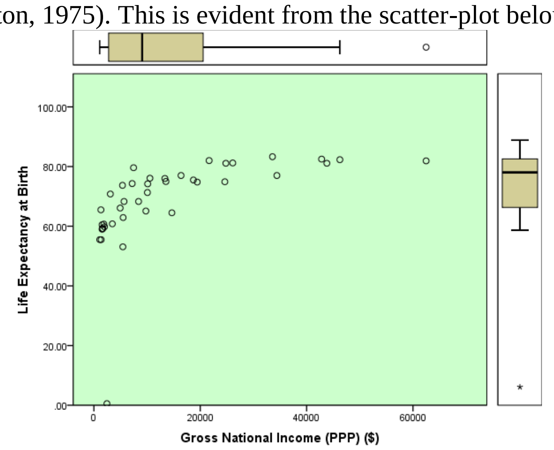

Figure 1: LEB and GNI (PPP)

Fertility and life expectancy at birth were found to be negatively related which resonates the concept that increased life expectancy may mean lower fertility (Mace, 2008).

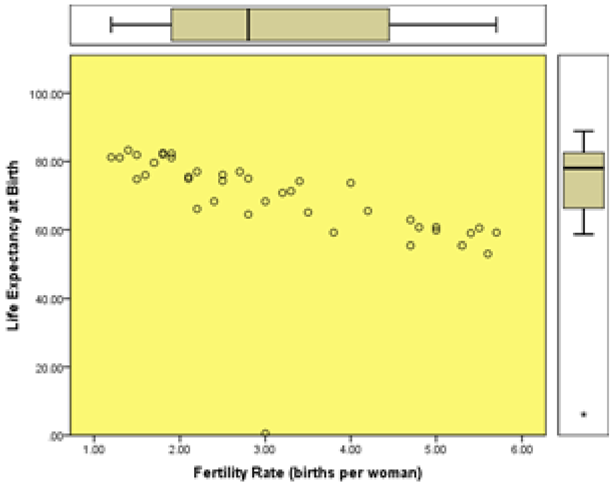

Figure 2: LEB and Fertility rate

The categorical variable Low Human Development has a negative relationship with life expectancy which resounds the concept that more the human development of a nation, higher is its life expectancy (Livi-Bacci, 2001).

Table 3: Results of regression analysis of life expectancy at birth using a linear & a double-log model for the year 2015

SL. No.	Variables (Linear model)	Estimated Coefficients (standard error) {p value}	Variables (Double Log model)	Estimated coefficients (standard error) {p value}
1.	Constant	112.967 (16.592) {9.567}	Constant	4.756 (0.295) {8.240}
2.	GNI	0.000** (0.000) {0.021}	lnGNI	-0.47** (0.019) {0.021}
3.	POP	7.011*** (1.603) {0.000}	lnPOP	0.013 (0.020) {0.515}
4.	PHY	0.102 (0.115) {0.381}	ln PHY	0.010 (0.014) {0.483}
5.	FER	-5.120** (2.394) {0.041}	ln FER	-0.124** (0.045) {0.011}
6.	EDU	0.001 (0.076) {0.990}	ln EDU	0.008 (0.017) {0.625}
7.	WAT	-0.164 (0.156) {0.301}	lnWAT	0.038 (0.057) {0.513}
8.	FOR	0.077 (0.062) {0.222}	ln FOR	-0.008 (0.007) {0.251}
9.	HHD	-18.906*** (5.912) {0.003}	HHD	-0.093*** (0.031) {0.006}
10.	MHD	-28.667*** (6.969) {0.000}	MHD	-0.197*** (0.039) {0.000}
11.	LHD	-42.095*** (8.862) {0.000}	LHD	-0.302*** (0.062) {0.000}
	R^2 F	0.798 11.468***	R^2 F	0.938 34.648***

Note:

- A total of 40 countries' information have been used to carry out the study.
- Figures within ( ) and { } indicate standard error (SE) and p (sig) values respectively.
- ***, ** and * indicate that the estimated coefficients and test statistic are statistically significant at 1, 5 and 10 percent levels respectively.

The **estimated double-logmodel** shows a strong linear fit with an R^2 of 0.938. This means that 93.8 % of the variance has been accounted for in the model. Therefore, we can assume that our data set is sufficient to create a regression model for life expectancy at birth.

A discussion of the observations has been given below:

- Estimated GNI is statistically significant but with a negative sign. This suggests there is a negative relationship between life expectancy at birth and GNI, and the latter asserts significance influence on the former. This is contrary to prevalent theory and is in contrast with the result derived from the linear model. This issue needs further elaboration in a systematic manner.
- Estimated POP has turned out insignificant in this model which is contrary to the result derived from the linear model.
- Resonating with the results derived from the linear model, estimated HHD and MHD dummy variables have negative signs once again. This is indicative of a renewed research in this field.

Results based on the model show that variables GNI (PPP $), fertility rate (births per woman), dummy variables high human development, medium human development and low human development exert high influence on life expectancy at birth.

Fertility and the categorical variable Low Human Development were found to be negatively related with life expectancy at birth similar to the linear model.

Taken together, the results of our analysis have some relevant policy implications. Increase in GNI (PPP $) and higher expenditure on population planning are undoubtedly important for increasing life expectancy at birth, although these may not necessarily do so in these nations. Forest cover (%) has shown no significant effect on the dependant variable, but as the p-value indicates, provided appropriate data and in-depth study, it is possible to show that the area covered with forests and woodlands might have a positive relationship with life expectancy at birth. This has an important implication in today's world where forest cover is dwindling at a rapid rate in the form of preventing deforestation so as to achieve sustainable development. The impact of geographical location and climate on the human health should also be taken into account to achieve a sound result. It must be noted that analyses of panel data for a considerable period of time for selected countries would demonstrate different results, but unavailability of long-term data for relevant indicators is a problem that must be handled skilfully.

7. Conclusion

The present study tries to explore some of the determinants of life expectancy at birth pertaining to 40 countries with different levels of human development using a linear and a Cobb-Douglas (double-log) model for the year 2015 using cross-section data. Ten widely used variables have been used as regressors to examine their significance in determining life expectancy of these countries. The results reveal that some of the variables traditionally considered to be influential turned out to be insignificant in contrary to the findings of previous studies while some turned out to have a different relation with the dependent variable than is traditionally considered.

References

[1] Barro, R. J., and Sala-I-Martin, X. I. (1995) *Economic Growth*. McGraw-Hill: New York.
[2] Bayati, M., et al. (2013) "Determinants of Life Expectancy in Eastern Mediterranean Region: A health production function" *International Journal of Health Policy and Management.*
[3] Beaglehole R, Bonita R, Kjellstrom T (1993)*Basic Epidemiology*;WHO, Geneva.
[4] Delavari, S. et al. (2016) "Life expectancy and its socioeconomic determinants in Iran. *Electronic Physician*, 8(10): 3061-3068.
[5] Ehrlich, I., and Lui, F. T. (1991). "Intergenerational trade, longevity, and economic growth".*Journal of Political Economy*, 99, 1029-1059
[6] Gulis, G. (2010)"Life Expectancy as an Indicator of Environmental Health",*European Journal of Epidemiology*, 16(2): 161-165; Springer.
[7] Husain, A. R. (2002). "Life expectancy in developing countries: a cross-section analysis". *The Bangladesh Development Studies Journal*, 28(1/2), 161-178
[8] Jabaa, E. et al (2014) "The relationship between life expectancy at birth and health expenditures estimated by a cross-country and time-series analysis";College of Business, Tennessee State University.
[9] Kabir, M.; "Determinants of Life Expectancy in Developing Countries*".The Journal of Developing Areas*, 41(2): 185-204, 2008.
[10] Kuhn, P., et al (1992). "Social security, longevity, and moral hazard" *Journal of Public Economics*, 49(1), 91-106
[11] Livi-Bacci,(2001). *Concise History of World Population;* Blackwell Publishers. 82-109.
[12] Monsef, A., et al (2015). Determinants of Life Expectancy: A Panel Data Approach.*Asian Economic and Financial Review*. 5(11): 1251-1257.
[13] Ohemeng, W., et al (2015). "Socio-economic determinants of life expectancy in Nigeria (1980-2011)" *Health Economics Review*,5:2.
[14] Rodgers, G. (1979) *"Income and inequality as determinants of mortality: an international cross-section analysis. Population studies*" 33(2), 343-351.

Websites

[15] https://data.worldbank.org/indicator/NY.GNP.PCAP.PP.CD?end=2015&start=1990&year_high_desc=true (2015)

[16] https://data.worldbank.org/indicator/SP.POP.GROW?end=2015&start=1960 (2015)
[17] www.un.org
[18] http://hdr.undp.org/en/2016-report
[19] https://data.unicef.org/topic/water-and-sanitation/drinking-water/