GLOBAL INEQUALITY IN LIFE EXPECTANCIES

WHAT ROLE DOES INCOME INEQUALITY PLAY?

Nirajana Banerjee (nirajana004@gmail.com)

Bachelor's student, Economics Department, Presidency University

Ritojeet Basu (ritojeetbasu@gmail.com)

Bachelor's student, Economics Department, Presidency University

Ananya De (ananya.sbh@gmail.com)

Bachelor's student, Economics Department, Presidency University

Monalisa Poali (monalisapoali@gmail.com)

Bachelor's student, Economics Department, Presidency University

Abstract

Context: Life expectancy best helps to capture the health and well-being of a population. However, wide differences are seen in life expectancies across different countries and at different points of time.

Research objective: This study examines trends in global inequalities in Life Expectancy at birth between countries from 1960 to 2017, and studies how income inequality affects inequality in life expectancy.

Data and methodology: Life expectancy at birth is the main variable under study. We have also used data on GDP per capita at purchasing power parity, Health Expenditure per capita and Government Expenditure on health per capita. Six measures of inequality have been used, primary being the Gini coefficient. We have divided the countries into four groups according to the World Bank Classification Scheme, and have used graphs, choropleth maps and Moran's I for statistical analysis. The causal relationship between inequality in Life Expectancy and per capita GDP is examined using reduced form regression models.

Findings: The life expectancy at birth shows a steadily rising trend. The choropleth maps indicate considerable spatial variations in life expectancy, which are stable over time. There is a decline in the global inequality in life expectancy over time. Moreover, the decomposition analysis for the Gini coefficient shows that in any year, the between groups inequality is more important than the within groups inequality. This is corroborated by the decomposition of Gini over time. We infer from the devised regression models that there is a positive association between per capita income inequality and inequality in life expectancy. The regressions point out evidence that inequality in lifespan and inequality in per capita income can be explained in terms of inequality in health expenditure.

Key words: Life expectancy at birth, GDP per capita, Gini coefficient, Decomposition of inequality.

Introduction

The objective of this paper is to analyse global inequalities in life expectancy at birth (referred to as life expectancy hereafter) between countries from 1960 to 2017, and its relationship to income inequality. We have classified the countries into four per capita income groups, based on the World Bank Classification Scheme. We also propose to examine inequalities in countries using the income criteria. The study will also attempt to examine the association between inequality in life expectancy and inequality in GDP per capita among countries.

Life expectancy is by far the best measure of health. Among all the major human development indicators, life expectancy plays a crucial role in analysing the constantly changing trends in the average life-span of a person during the period under study. As per the United Nations Development Programme (UNDP), Sustainable Development Goals (SDG) (UNDP, 2015) for 2030, “Good health and well being” has been considered as the third main objective for achieving health related SDGs. In order to attain this goal, the World Health Organization (WHO) has considered monitoring life expectancy. Thus, the increasing pattern in the life expectancy indicator is a crucial measure to “Ensure healthy lives and promote well-being for all at all ages” (SDG3).

However, life expectancy is seen to exhibit disparities between different groups of individuals, and between different countries. This is a matter of concern. For instance, in 2017, the average Japanese was expected to live for around 84 years, whereas, the average individual from Sierra Leone was expected to live for only about 58 years, a glaring difference of 26 years in average expected survival. In this project, we aim to understand these trends of inequality and the possible reasons behind them.

Historically, there have been two approaches to studying inequality. One focuses on inequality *within* countries, while the other analyses inequality *between* countries. While there has been a considerable body of work on inequality in life expectancy, the literature in this field has largely concentrated on within-country inequality, focussing on identifying vulnerable groups. In contrast, global trends in inequality in life expectancy between countries have been a relatively neglected area of research (Permanyer and Scholl, 2019).

However, a study of variations in life expectancy *between* countries is important because of many reasons. Firstly, of course, is the availability of long run data permitting us to study this issue. Secondly, a situation where the lifespan of an individual depends upon where he/she is born is not acceptable because of ethical reasons. Thirdly, we can examine the impact of fundamental changes in technology, economy and society upon between-country variations in lifespans. In particular, the effect of globalisation on living conditions and life expectancy is an important issue. Finally, exploring how life span changes over time between countries enables us to understand human mortality and improve the quality of public health policies.

Studies of life span at the global level indicate increasing life expectancy over time. This is often attributed to economic growth, reflected in increase in per capita national income. With rising national income, life expectancy of a country is also seen to increase (Preston, 1975); it is graphically represented as the Preston Curve. However, increasing national income was not the only reason for the change in life expectancy; exogenous factors accounted for around 75-90 percent of the growth in life expectancy, for both developed and developing nations (Preston, 1975).

The increasing trend in life span, however, hides considerable variations in life span between countries. Smits and Monden (2009) found that there was substantial variation in the length of adult life across nations and time periods. They found most of the life inequality to be the within-country inequality. They suggested that the life inequality should be adjusted for the life expectancy to have a better measure of life differentials across populations. They tested two hypotheses: the *forerunner hypothesis* and the *diffusion hypothesis*. The former hypothesized that those nations that reached a certain level of life expectancy ahead of others faced a higher level of life inequality; it was found to be true in their analysis. The diffusion hypothesis tested whether countries that lagged behind in life expectancy had a lower level of life inequality, given that they could adopt the better health practices in the forerunner countries. White (2002) concluded that the rate at which life expectancy increased for the developing nations was faster than the rate at which life expectancy increased for the developed nations of the world.

In a study of the Human Mortality Database (2009), Edwards (2011) found out that the inequality in the expectation of life at birth had clearly declined during the 30 years after 1970 for the world as a whole and for advanced countries with high-quality data. In addition,

Edwards also found that the share of within-country inequality, out of total inequality, declined over the years and the proportion of between-country inequality rose over the years. The work done by Anand and Ravallion (1993) in this field show that the commonly observed positive relationship between life expectancy and affluence across countries ceases to hold once provisions are made to account for incidence of poverty and the public expenditure on health. Finally, Moser, Shkolnikov and Leon (2005) studied life expectancy in a more composite context as they highlighted factors such as infant mortality, fertility, the global outbreak of contagious diseases such as HIV/AIDS in Sub-Saharan Africa and human development into their scope of analysis. They observed a convergence in mortality up to 1980 and a divergence thereafter.

Due to the dearth of literature on cross-country or between-country inequality, it was imperative to be involved in such a study. For our analysis, we have taken the life expectancy at birth figures for 218 countries; they are divided into 4 groups based on levels of per capita income. While we have primarily used the Gini coefficient, other inequality measures like the Theil Index, standard deviation, coefficient of variation, standard deviation and the mean logarithmic deviation have also been used. We have also used spatial statistics. Finally, we have used reduced form regression models to examine the role of income inequality in explaining inequalities in life expectancy.

2. Research objective and hypotheses

The study seeks to examine trends in inequality in life expectancy over time at the global level, and decompose this inequality at specific years, and its change over time, into sub-groups of countries classified in terms of per capita income levels. We will also analyse the role of income inequality underlying inequalities in life spans. Although we will start with several measures of inequality, we will ultimately focus on the Gini coefficient in this study. The analysis will be undertaken using, not only the Gini coefficient, but also its two components—a *within* country measure of inequality, and a *between* country measure of inequality. The concept of within and between inequality and how the Gini may be decomposed into these two components has been discussed in a classic paper by Lambert and Aaronson (1993), and is discussed in the methodology section. We will also use the Mookherjee-Shorrocks decomposition (Mookherjee and Shorrocks, 1982) to decompose the change in inequality over the period of study.

The study rests on the following research hypotheses:

H₁: *Inequality in life expectancy is decreasing over time.*

H₂: *The inequality in life expectancy may be attributed to (i) differences in mean income of countries classified by income levels (between), and (ii) income differences between countries in each sub-group (within). The weight of the former (between) will be greater.*

H₃: *The change in global inequality in life spans over the period of study may similarly be attributed to these two causes, with between group changes playing a major role.*

H₄: *The inequality in life expectancy may be attributed to inequality in per capita income.*

3. Data and Methodology

3.1 Variables

The main variable of the study is Life expectancy at birth, measured in years. It is estimated for total population, male, and female. Life expectancy, defined as life expectancy at birth, indicates the number of years a newborn infant would live if prevailing patterns of mortality at the time of its birth were to stay the same throughout its life.¹ Data on life expectancy at birth is obtained from The World Bank *World Development Indicators* site, and is based on estimates provided by the United Nations Population Division, in *World Population Prospects: 2019 Revision*. In addition, we have used data on Gross Domestic Product (GDP) per capita at purchasing power parity (PPP), Health expenditure per capita, and Government expenditure on health per capita. These data have also been downloaded from the WDI website (<https://databank.worldbank.org/reports.aspx?source=world-development-indicators>).

In Table 1, we sum up the statistical characteristics of the study variable, life expectancy at birth. Statistics are given for all countries as well as for countries divided into groups on the basis of per capita income.

¹ Given in The World Bank Development Indicators data set, <https://bit.ly/2YXM3sr>, accessed on 13 May 2020.

Table 1: Descriptive statistics of Life expectancy at birth—World, and Countries divided into Income groups

Group	N	Maximum	Minimum	Mean	Median	Variance
Low Income	3696	85.41708	42.672	73.39772	73.71464	31.49335
Low Middle Income	1798	74.43	26.172	49.79139	49.0225	85.01185
High Middle Income	2696	76.31	18.907	57.7985	58.154	85.07006
High Income	3139	79.914	36.535	66.40062	68.17	56.34059
World	11329	85.41708	18.907	64.00028	67.134	131.2127

3.2 Measures of Inequality

For measuring global inequalities in life expectancy within countries, we have used six measures:

- (a) *Relative Mean Deviation*: The Relative Mean Deviation (RMD) shows the degree of deviation of individual observations (y_i), in a population of size n from the mean ($\bar{y}$). The formula for the calculation is given by:

$$I_{RMD} = \frac{1}{2\bar{y}n} \sum_{i=1}^n |y_i - \bar{y}|$$

- (b) *Coefficient of Variation*: The Coefficient of Variation (COV) is one of the relative measures of dispersion. It is given as the ratio of the standard deviation to mean. This measure is more suitable than using the variance because variance has the property that inequality varies with the level of mean and this property turns out to be problematic while conducting the analysis. Hence to avoid this problem of mean-dependence, we use COV, whose formula can be given by:

$$I_{CV} = \sqrt{I_{VAR} / \bar{y}}$$

where,

$$I_{VAR} = \frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2$$

- (c) *Standard Deviation of logarithms*: The Standard Deviation of Logarithms (SDL) is another quite useful measure in measuring inequality. It is the sum of the squares of

deviation of the logarithms of individual observations (y_i) from the log of the geometric mean (given by $\bar{y}_G$). The formula is stated as:

$$I_{SDL} = \frac{1}{n} \sum_{i=1}^n (\log y_i - \log \bar{y}_G)^2$$

(d) *Gini Coefficient*: The Gini Coefficient is a widely used measure in economic researches on inequality. It varies from 0 to 1, with higher numbers indicating greater inequality. It is the mean of the absolute value of the inter-individual differences in age at death, divided by the life expectancy:

$$I_{GINI} = \frac{2}{n^2 \bar{y}} \sum_{i=1}^n i(y_i - \bar{y})$$

(e) *Mean Logarithmic deviation (Generalised Entropy of Parameter 0) and Theil Index*: The Theil Index and the Mean Logarithmic Deviation are entropy measures taken from the Information Theory by Theil (1969, 1972). The two entropy measures are:

$$I_{THEIL} = \frac{1}{n} \sum_{i=1}^n \frac{y_i}{\bar{y}} \log \frac{y_i}{\bar{y}}$$

$$I_{MLD} = \frac{1}{n} \sum_{i=1}^n \log \frac{\bar{y}}{y_i}$$

One of the most striking features about these measures is that they are all highly correlated and hence behave quite similarly across countries and time.

3.3 Statistical analysis

The temporal trends in life expectancy and inequality in life expectancy are analysed using graphs. We have also examined similar trends for groups of countries. For our purpose, we have divided the countries into four groups based on the World Bank classification of countries by per capita income.² Four groups have been formed: High Income Countries (HIC), High Middle Income Countries (HMIC), Low Middle Income Countries (LMIC), and Low Income Countries (LIC). Choropleth maps have also been used to examine the spatial patterns in life expectancy. These are maps which use differences in shading or colouring within predefined areas to indicate the average values of a particular quantity (here life expectancy) in those areas. To test whether the data has a spatial pattern, or whether it is

² Given in The World Bank Development Indicators data set, <https://bit.ly/2YXM3sr>, accessed on 13 May 2020.

random we have also estimated Global Moran's I. It measures how a variable in a particular region is similar to the values of the same variable in the surrounding regions. Moran's I is given by:

$$I = \frac{N \sum_i \sum_j W_{ij} (x_i - \bar{x})(x_j - \bar{x})}{W \sum_i (x_i - \bar{x})^2}$$

when N is the number of spatial units indexed by i and j; x is the variable of interest; W_{ij} is a matrix of spatial weights with zeroes on the diagonal, and W is the sum of all w_{ij} .

3.4 Decomposition of inequality measures³

A question often encountered in studies of inequality relates to the extent to which (say) income inequality in the total population is a consequence of income differences between population subgroups classified by characteristics such as age, gender, race, educational level or area of residence (Deutsch and Silber, 1999). This is a relevant question in studies of inequality in life spans also.

One of the major problems encountered on decomposing the Gini coefficient into between groups and within groups contributions, is that we are left with a residual term, that is cumbersome and difficult to interpret, in the case when the subgroup ranges overlap. In such a situation, a graphical analysis becomes a useful tool in understanding the residual term as a sub-area of the Lorenz diagram. In the second part of our analysis, we decompose the inequality measures into within and between components. Lambert and Aronson (1993) have shown that the Gini coefficient may also be decomposed into these two (and a residual component). This is explained below.

Consider an income vector $X_1, X_2, \dots, X_N$ in the population. We call it Ω . The mean income is μ , and the incomes are divided into n subgroups with income vectors given by Ω_k , each having a size of N_k . The means are μ_k ($k = 1, 2, \dots, k$). Suppose that there is no inequality, so that all income $N\mu = \sum N_k \mu_k$ is distributed equally among the population so that each individual has an income of μ . In this situation, the bottom 100p% of incomes in the income distribution would get 100p% of total income under this scenario. When we introduce inequality, the income share of these individuals reduces, to $L(p)$ say; where $L(p)$ is the Lorenz Curve. The transformation:

³ This section largely follows Lambert and Aronson (1993), 1221-1223.

$$P \rightarrow L(p) \quad (2)$$

represents a downwards shift, and in terms of area, is measured by the Gini Coefficient:

$$G = 2 \int_0^1 [p - L(p)] dp. \quad (3)$$

We now try to introduce inequality in three stages. We shall then obtain an area decomposition of total inequality on the Lorenz diagram. In describing the stages, we start with the situation of perfect equality, where everyone gets μ .

1. First, we introduce the ***between-group inequality***: We bring about a redistribution of the total $N\mu = \sum N_k \mu_k$ between the subgroups in such a manner so that each person in Ω_k gets μ_k ($1 \leq k \leq n$). The Lorenz curve for this notional income distribution may be denoted as $L_B(q)$. Here, q denotes the rank of an individual in the income distribution. We assume without loss of generalisation that the Ω_k are so labelled such that we have $\mu_1 \leq \mu_2 \leq \dots \leq \mu_n$; that is, we consider all the people in Ω_1 , then consider all in Ω_2 , etc. How people get ranked within a particular subgroup Ω_k , does not matter for the definition of $L_B(q)$.
2. Next, we incorporate ***inequality within the groups***, by taking the actual incomes of people, but maintaining the same order by subgroups as done before, considering all the people in Ω_1 first, then considering all in Ω_2 , etc. Within subgroups, we now arrange people in ascending order of income. Hence, the richest individual in Ω_1 will stand next to the poorest individual in Ω_2 , etc., and the last individual of all is the richest in Ω_n . Let $C(q)$ denote the concentration curve for income with respect to this 'lexicographic income parade'. In that case, $C(q)$ measures the proportion of total income received by the first Nq individuals in this income parade.
3. Finally, we consider the ***overlapping effect*** which occurs between subgroup income distributions: Here, we re-rank individuals as necessary, to form the true income distribution, from the overall poorest to the overall richest. Then, we calculate the income shares in the distribution of re-ranked income units.

The transformation of income shares starting from an initial situation of perfect equality, $q \rightarrow L_B(q)$, which may be written as:

$$q \rightarrow L_B(q) \rightarrow C(q) \rightarrow L(q) \quad (4)$$

summarises this three-step procedure to describe overall inequality. Thus, we first took into account between-group inequality, then the within-group inequalities, ignoring the overlaps;

and finally, we considered the overlaps by calculating the income shares relative to the true ranking (non-lexicographic). At each stage, there is a downwards shift, as illustrated in Fig. 1).

The area measures of these downward shifts are called as follows:

$$A_B = 2 \int_0^1 [q - L_B(q)] dq, \quad (5)$$

$$A_W = 2 \int_1^0 [L_B(q) - C(q)] dq, \quad (6)$$

and,

$$A_O = 2 \int_0^1 [C(q) - L(q)] dq \quad (7)$$

respectively. The Gini coefficient is thus:

$$G \equiv A_B + A_W + A_O \quad (8)$$

Figure 1: Introducing inequality in stages

Source: Lambert and Aronson (1993), 1223.

3.5 Decomposition of change in inequality over time

The above decomposition analysis studies the inequality at a single point of time. But, in studies over time, changes in economic, and socio-demographic variables occur; such changes will affect the change in inequality over time. Mookherjee and Shorrocks (1982)

argue that the “potential importance of these structural changes suggests that their impact on the income distribution needs to be examined systematically before any attempt can be made to interpret the observed trend in overall income inequality adequately” (886).

The method usually adopted for this purpose

“attempts to create a comparable time series of distributions based on the structural characteristics pertaining in one particular year. For example, if we wished to investigate the effect of shifts in the age structure of the population, this approach would disaggregate the distribution in one year, t_1 , and estimate what the distribution would have been like if the age structure had been the same as that observed in some other year t_0 , on the assumption that other distributional characteristics (such as age-income differentials and inequality within age groups) remain invariant to this demographic change. The comparison between the actual distribution at t_1 and the constructed distribution at t_1 with t_0 's age structure is taken to indicate the impact of changes in the age structure over the period concerned” (Mookherjee and Shorrocks, 1982: 886).

In our case, we start with an initial distribution of life span and income at a starting point. At the end point, both distributions will have changed. Standard measures (Semple, 1975; Love and Wolfson, 1976; Dinwiddy and Reed, 1977; Royal Commission, 1977; Rowley and Henderson, 1978; MacLeod and Horner, 1980) propose to estimate what the distribution of life expectancy would have been if the income distribution had remained same. The difference between observed inequality in life expectancy and the estimated inequality in life expectancy is attributed to the change in distribution of income.

Mookherjee and Shorrocks (1982) propose an alternative method for decomposition over time based on:

“use of additively decomposable inequality measures, and involves disaggregating total inequality in each year into various components and examining the time paths of these individual contributions. The conventional type of cross section decomposition in a single year associates two main contributions (the 'within group' and 'between group' terms) with each disaggregation characteristic. These contributions are in turn completely specified by information about the subgroup mean incomes, population shares and inequality values. *In the decomposition of the trend in inequality, we identify three main components corresponding to the influence of changes in*

subgroup inequality values, shifts in the subgroup population shares and relative variations in the subgroup mean incomes” (Mookherjee and Shorrocks, 1982: 887; italics, our).

Two decomposition results are obtained: one an exact decomposition, described in equation (12) of the paper, and the other, an approximation, described in equation (13). The approximate decomposition yields four components—labelled A, B, C, and D, which correspond to equations (14a), (14b), (14c), and (14d) in the original paper, respectively. This is a slight approximation of the exact decomposition, but offers a more practical interpretation than the exact decomposition.

A: Effect due to changes in within subgroup inequality,

B: Effect due to changes in population shares of within component,

C: Effect due to changes in population shares of between component, and

D: Effect due to relative changes in subgroup means.

The exact decomposition, provided in equation (12) of the original paper, unfortunately does not provide the same neatly-identified conceptual components; however, they can be presented in four distinct terms, of which the first two corresponds to A and B of the approximate decomposition, while the remaining two cannot be interpreted.

3.6 Econometric methodology

The causal relation between inequality in life expectancy and per capita GDP is examined using a reduced form regression of inequality in life expectancy on inequality in per capita income:

$$LE_G = \alpha_1 + \beta_1 GDP_G. \quad [9]$$

Similar regressions are also estimated for within and between measures:

$$LE_B = \alpha_2 + \beta_2 GDP_B. \quad [9a]$$

$$LE_W = \alpha_3 + \beta_3 GDP_W. \quad [9b]$$

Our hypothesis is that the regression coefficients (β_1 , β_2 , and β_3) should all be positive. The reduced form enables us to focus on the relationship between two variables; however, particularly in the case of time series data, the presence of confounding variables may lead to the observation of spurious or misleading relationships.⁴

⁴ Suppose x_1 and x_2 are not related with each other. But both are positively related to y . If y has gone up, then x_1 and x_2 will also increase. Scatter plots of x_1 and x_2 will indicate a positive relationship between the two, even though they are not actually related

In particular, given that we are using time series data, time may be a confounding variable. To eliminate the role of time, we regress LE_G and GDP_G upon time, estimate the residuals and then regress residual of LE_G upon residual of GDP_G :

$$LE_{RG} = \delta_1 + \gamma_1 GDP_{RG}. \quad [10]$$

This is repeated for LE_B and LE_W :

$$LE_{RW} = \delta_2 + \gamma_2 GDP_{RW}. \quad [10a]$$

$$LE_{RB} = \delta_3 + \gamma_3 GDP_{RB}. \quad [10b]$$

4. Findings

An analysis of global life expectancy at birth over time indicates a steadily increasing trend (Fig. 2). In the period 1960 to 2017, it has increased by about 25 years from 50 years to 75 years. This is also observed for the four groups of countries, HIC, HMIC, LMIC and LIC. Further, life expectancy at birth has been uniformly higher for richer countries, indicating a positive association between per capita income and life expectancy.

Figure 2: Life expectancy at birth—World, Countries classified by Income

Source: Own estimate

4.1 Choropleth maps

Spatial variations in life expectancy at birth are also pronounced, and relatively stable over time (Fig. 3A-3D). Life expectancy is low in Sub-Saharan Africa, parts of Latin America, and Asia (except for China, Japan, Malaysia, and Singapore). While life expectancy was initially high in former USSR and the Soviet bloc countries (Fig. 3A and 3B), the disintegration of Communism in Europe in the 1990s has led to lifespan declining in East Europe and Russia (Fig. 3C and 3D). In contrast, North America, Western Europe and Australia have had consistently high lifespan over the period of study. It may be noted that these countries are also relatively more affluent than other countries, reiterating the association between per capita income and life expectancy.

Figure 3a: Geographical distribution of life expectancy - 1960

Figure 3b: Geographical distribution of life expectancy – 1980

Figure 3c: Geographical distribution of life expectancy - 2000

Figure 3d: Geographical distribution of life expectancy - 2017

Source: Own estimate

Moran's Global I is also high: 0.760 (1960), 0.645 (1980), 0.806 (2000), and 0.763 (2017), indicating the presence of spatial spill-overs (Appendix Fig. A1).

Figure 4a: LISA map for life expectancy at birth: 1960

Figure 4b: LISA map for life expectancy at birth: 1980

Figure 4c: LISA map for life expectancy at birth: 2000

Figure 4d: LISA map for life expectancy at birth: 2017

Source: Own estimate

In particular, we find positive spatial spill-overs in Europe (1960 and 1980), which decreases to Western Europe in subsequent periods (Fig. 4A-D). Negative spill-overs are observed in Sub-Saharan Africa, which increases over time to cover the southern parts of the continent over time.

4.2 Inequality in life expectancy

In Fig. 5 we have plotted the Gini coefficient of life expectancy. Global inequality in life spans has decreased steadily, except for the first decade in this millennium. Per capita income appears to be associated with, not only life spans, but inequality in life expectancy. While all country groups display a decreasing trend in inequality in life spans in general, the level of inequality is lower in HICs and HMICs.

Figure 5: Gini of life expectancy at birth—World, and countries by income groups

Source: Own estimate

Inequality has also been estimated using other measures (Fig. 6). It may be seen that the findings reflect the global decline in inequality in life spans between countries. Correlation coefficient between different measures (Appendix Table A1) is also very high (about 0.99 between each measure).

Figure 6: Inequality in life expectancy at birth—selected years

Source: Own estimate

4.3 Results of decomposition

In the next step we examine the summary of results of the decomposition analysis for the Gini coefficient (Table 2). Detailed results are provided in Appendix, along with decomposition for other inequality measures.

It may be seen that inequality in life expectancy has halved over the study period from 0.13 to 0.06. Further, the decompositions indicate, generally, that in any year the between groups inequality is more important than the within groups inequality. In case of the Gini coefficient, the weight of the between groups inequality is about 80 percent, while the within groups inequality counts only for about 12-14 percent of total inequality. This is true for other decomposable measures also (Appendix Table A2. Although, the relative weights of these two components, as well as that of the overlapping element, vary from one decomposition to the other, the correlation between each total, within group and between group inequality is, however, close to unity (about 0.99, see Appendix Table A3). Since the country groups have been formed based on levels of per capita income, it implies that the difference in mean life spans *between* HI, HMI, LMI and LI countries is greater than the inequality in life expectancy *within* each of these groups.

Table 2: Results of decomposition analysis for Gini coefficient

Year	Within (%)	Between (%)	Overlap (%)	Gini
1960	14.64	78.59	6.93	0.1298
1970	12.76	81.11	5.47	0.1097
1980	12.70	81.52	5.29	0.0945
1990	12.01	84.06	3.6	0.0833
2000	13.24	80.63	6.02	0.0831
2010	14.34	80.31	5.74	0.0697
2017	13.53	81.19	5.07	0.0591

Source: Own estimate

The above decomposition is for specific years. But what happens to the change in distribution of life expectancies over the entire period of study? The answer to this question is provided by the Mookherjee-Shorrocks decomposition results (Table 2). It may be seen that inequality decreases marginally by 0.02 between 1960 and 2017. Again, the between effects component has greater weight. Approximately, about two-thirds of this decline may be attributed to relative changes in means of each group. In case of the exact decomposition, only about 35 percent may be attributed to relative changes in within subgroup distribution and changes in population shares within each group. This underlines the lower weight of the within effect components in the decomposition analysis.

Table 3: Results of Shorrocks-Mukherjee decomposition analysis over 1960-2017

Components of Inequality change	Value	Percentage
Approximation decomposition components:		
A. Effect due to changes in within subgroup inequality	-0.007516	35.01
B. Effect due to changes in population shares of within component	-0.000075	0.35
C. Effect due to changes in population shares of between component	-0.000056	0.26
D. Effect due to relative changes in subgroup means	-0.013824	64.39
Sum of components	-0.021470	

Exact decomposition components:		
A. Effect due to changes in within subgroup inequality	-0.007516	34.95
B. Effect due to changes in population shares of within component	-0.000075	0.35
C.	-0.005539	25.76
D.	-0.008376	38.95
Sum of components	-0.021507	
Exact difference in I(0)	-0.021507	

Source: Own estimate

4.4 Explaining Inequality in LE

The decomposition analysis indicates that change in inequality between the country groups formed on the basis of per capita income levels is crucial in understanding the decline in inequality in life expectancy at the global level. This hints that income inequality may be an underlying factor explaining the global change in life span inequality. Scatter plots (see Appendix Fig. A2) indicate the presence of a relation, but the relation is not positive as we hypothesised. To check this, we estimated reduced form regression models of Gini, within and between of life expectancy upon Gini, within and between of per capita income, respectively (Table 4). These equations correspond to [9, 9a, 9b].

Table 4: Results for regression of Inequality in LE upon Inequality in GDP

Life expectancy	Gini	Between	Within
$\beta_{\text{GDP per capita}}$	-1.83	0.14	-0.03
t-value	-4.96	45.47	-0.74
Prob	0.00	0.00	0.46
Intercept	1.61	0.01	0.02
t-value	5.22	16.30	1.84
Prob	0.00	0.00	0.08
N	28	28	28
F	24.61	2067.53	0.55
Prob	0.00	0.00	0.46
R ²	0.47	0.99	0.01

Source: Own estimate

From Table 4, we see that inequality in Life expectancy is significantly related with inequality in GDP per capita—for the overall and between components. However, although we would expect a positive sign for the coefficient of GDP per capita, in the “Gini” and “within” models, they are negative.

A possible reason is that time plays a confounding role. To eliminate the role of time, we have regressed each variable upon time, and estimated residuals—corresponding to the equation set [9, 9a, 9b]. In Table 5, we present results of a regression model for residuals. The coefficients are statistically significant in each model—Gini, between and within models. This time, we get the hypothesised positive sign—indicating that higher income inequalities are reflected in higher inequalities in life expectancies.

Table 5: Results for regression of Residuals of LE upon Residuals of GDP per capita—eliminating time effect

R(Life expectancy)	Gini	Between	Within
$\beta_{R(\text{GDP per capita})}$	1.08	0.20	0.17
t-value	4.60	6.34	9.51
Prob	0.00	0.00	0.00
Intercept	0.00	0.00	0.00
t-value	0.00	0.00	0.00
Prob	1.00	1.00	1.00
N	28	28	28
F	21.19	40.19	90.38
Prob	0.00	0.00	0.00
R^2	0.29	0.48	0.79

Source: Own estimate

What is the explanation for this positive association between inequality in per capita income and life expectancy? A possible explanation runs in terms of health expenditure per capita. Expenditure on health is a form of investment in human capital; it will increase productivity, and hence increase GDP per capita. On the other, increased expenditure on health is expected to improve health and life expectancy. So, the role of inequality in health expenditure in our research context needs to be examined. We take two indicators for health spending:

Government expenditure on health per capita, and Expenditure on health per capita (which includes private out of pocket expenditure as well as Government expenditure).

Table 6: Results for regression of Residuals of LE upon Residuals of GDP per capita—eliminating time and health expenditure effects

R(Life expectancy)	Government exp on health			Expenditure on health		
	Gini	Between	Within	Gini	Between	Within
$\beta_{R(\text{GDP per capita})}$	-0.25	0.03	-0.23	-0.32	0.04	-0.20
t-value	-2.14	1.31	-3.76	-2.14	1.54	-2.63
Prob	0.05	0.21	0.00	0.05	0.14	0.02
Intercept	0.00	0.00	0.00	0.00	0.00	0.00
t-value	0.00	0.00	0.00	0.00	0.00	0.00
Prob	1.00	1.00	1.00	1.00	1.00	1.00
N	17	17	17	17	17	17
F	4.56	1.72	14.12	4.56	2.37	6.90
Prob	0.05	0.21	0.00	0.05	0.14	0.02
R ²	0.16	0.09	0.47	0.19	0.13	0.43

Source: Own estimate

In the last step of our analysis, we regress Inequality in Life expectancy and Inequality in GDP per capita on time and health expenditure; residuals are estimated and regressed upon each other.

$$I_{LE} = \alpha + \beta_1 \text{Time} + \beta_1 \text{HealthExp (for Gini, Within and Between)} \Rightarrow R_{LE} \quad [11]$$

$$I_{\text{GDP per capita}} = \alpha + \beta_1 \text{Time} + \beta_1 \text{HealthExp (for Gini, Within and Between)} \Rightarrow R_{\text{GDP per capita}} \quad [12]$$

$$R_{LE} = \alpha + \beta R_{\text{GDP per capita}} \text{ (for Gini, Within and Between)} \quad [13]$$

when

I: Inequality (Gini, Within and Between)

R: Residual

Results are similar for [13] whether we take Government expenditure per capita, or Total expenditure on health per capita (Table 6). The coefficient of residual of GDP per capita is significant at 5% level only in the Within model. The sign, too, is negative—which is not expected. However, as we have seen earlier, the Within effect component comprises a relatively small proportion of the decomposed Gini coefficient. In contrast, the coefficient of residual of GDP per capita in the models for Gini and the Between component is insignificant at 5% level. This result may be interpreted as evidence that the positive relation between

inequality in life span and inequality in per capita income may be explained in terms of inequality in health expenditure.

5. Conclusion

The study proposes four hypotheses:

H₁: *Inequality in life expectancy is decreasing over time.*

H₂: *The inequality in life expectancy may be attributed to (i) differences in mean income of countries classified by income levels (between), and (ii) income differences between countries in each sub-group (within). The weight of the former (between) will be greater.*

H₃: *The change in global inequality in life spans over the period of study may similarly be attributed to these two causes, with between group changes playing a major role.*

H₄: *The inequality in life expectancy may be attributed to inequality in per capita income.*

Our analysis provides evidence in support of all our hypotheses. Inequality in life expectancy is decreasing over time (H₁). The decomposition of inequality measures reveals that both between and within components decline over time; however, the between effects dominate over the within effects (H₂). It implies that differences in the mean life expectancy in the four country groups (HIC, HMIC, LMIC, and LIC) are the dominant force reducing inequality in life span (H₃). Finally, we find evidence of an association between inequality in life span and inequality in per capita income; it is observed for the overall, between, and within components (H₄). This result is explained in terms of inequality in health expenditure. Our analysis, therefore, underlines the importance of increased spending on health expenditure. It will not only lead to human capital formation, but also reduce inequalities within countries.

Acknowledgement

We would like to thank Dr Zakir Husain, for his expertise and guidance. He has taught us so much along the way and without his support this paper would not have been possible. We are also grateful to Dr Diganta Mukherjee, Sampling & Official Statistics Unit, Indian Statistical Institute, Kolkata for his suggestions regarding the data analysis. The responsibility for any remaining errors lies with the authors.

REFERENCES

- Anand, S. & Ravallion, M. (1993). Human Development in Poor Countries: On the Role of Private Income and Public Services. *Journal of Economic Perspectives* [online], 7(1). 133-150. [Viewed 24 May 2020]. Available from: doi: <https://bit.ly/2TAbf51>
- Deutsch J., and Silber J. (1999) Inequality Decomposition by Population Subgroups and the Analysis of Interdistributional Inequality. In: Silber J. (eds) *Handbook of Income Inequality Measurement*. Recent Economic Thought Series, vol 71. Springer, Dordrecht: 363-397.
- Dinwiddy, R. & Reed, D. 1977. *The Effects of Certain Social and Demographic Changes on Income Distribution*. Background Paper No. 3 to the Royal Commission on the Distribution of Income and Wealth. London: HMSO.
- Edwards, Ryan D. (2011). Changes in World Inequality in Length of Life. *Population and Development Review* [online]. 37(3), 498-528. [Viewed 24 May 2020]. Available from: doi: <https://bit.ly/3cZPiUs>
- Lambert, Peter J. & Aronson, J. Richard (1993). Inequality Decomposition Analysis and the Gini Coefficient Revisited. *Economic Journal*. 103(420), 1221-1227, [Viewed 24 May 2020]. Available from: doi: <https://bit.ly/3d2gCRZ>
- Love, R. & Wolfson, M. C. 1976. *Income Inequality: Statistical Methodology and Canadian Illustrations*. Ottawa: Statistics Canada.
- MacLeod, N. and Horner, K. (1980). 'Analysing postwar changes in Canadian income distribution.' In *Reflections on Canadian Incomes* (Economic Council of Canada). Ottawa: Canadian Government Publishing Centre.
- Mookherjee, D., & Shorrocks, A. (1982). A Decomposition Analysis of the Trend in UK Income Inequality. *The Economic Journal*. 92(368), 886-902. doi: 10.2307/2232673
- Moser, K., Shkolnikov, V., Leon, David A. (2005). World mortality 1950-2000: divergence replace convergence from the late 1980s. *Bulletin of the World Health Organisation*(2005) [online]. 83(3), 202-209. [Viewed 24 May2020]. Available from: doi: <https://bit.ly/2yzGfe4>
- Permanyer, I., & Scholl, N. (2019). Global trends in lifespan inequality: 1950-2015. *PLoS ONE* [online]. 14(5): e0215742. [Viewed 24 May 2020]. Available from: doi: <https://bit.ly/2ywxDVr>

Preston, Samuel H. (1975). The Changing Relation between Mortality and Level of Economic Development. *Population Studies* [online]. **29**(2), 231-248. [Viewed 24 May 2020]. Available from: doi: 10.2307/2173509

Pyatt, G. (1976). On the Interpretation and Disaggregation of Gini Coefficients. *The Economic Journal*. **86**(June), 243-255. [Viewed 24 May 2020]. Available from: doi: <https://bit.ly/2LUIHP9>

Rowley, J. C. R. & Henderson, D. W. (1978). 'Partial standardisation of Canadian family incomes for changes in composition and other characteristics.' Economic Council of Canada.

Royal Commission on the Distribution of Income and Wealth (1977). *Report No. 5: Third Report on the Standing Reference*. London: HMSO.

Semple, M. 1975. The effect of changes in household composition on the distribution of income 1961-73. *Economic Trends*. **266**, 99-105.

Smits, J., & Monden, C. (2009). Length of life inequality around the globe. *Social science & medicine* (1982) [online]. **68**(6), 1114–1123. [Viewed 24 May 2020]. Available from: doi: <https://bit.ly/2ZzGeSw>

UNDP, (2015). Sustainable Development Goals [online]. *UNDP in India*. [Viewed 24 May 2020]. Available from: <https://bit.ly/2XmM8Uo>

White, Kevin M. 2002. Longevity advances in high-income countries, 1955-96. *Population and Development Review*. **28**(1), 59-76. [Viewed 24 May 2020]. Available from: doi: <https://bit.ly/3d2gQsj>

SOFTWARE MODULES

Kerm, P. (2007). INEQUAL7: module to estimate series of standard inequality measures for income distributions. Stata Technical Bulletin, 23; 02/12/2007.

Aliaga, R. and Montoya, S. (1999). GINIDESC: module to compute Gini index with within- and between-group inequality decomposition, 21/03/1999.

Silva, A. (2016). MSDECO: module to calculate the Mookherjee & Shorrocks (1982) over-time inequality decomposition by subgroup, 16/07/2016.

APPENDIX

Appendix Table A1: Correlation coefficient between measures of inequality

	<i>GE(-1)</i>	<i>GE(0)</i>	<i>GE(1)</i>	<i>GE(2)</i>	<i>Gini</i>	<i>RMD</i>	<i>CoV</i>	<i>SD of Logs</i>
GE(-1)	1							
GE(0)	0.9997	1						
GE(1)	0.9990	0.9998	1					
GE(2)	0.9979	0.9991	0.9998	1				
Gini	0.9992	0.9990	0.9981	0.9968	1			
RMD	0.9989	0.9985	0.9976	0.9961	0.9999	1		
CoV	0.9978	0.9969	0.9953	0.9931	0.9990	0.9991	1	
SD of Logs	0.9925	0.9900	0.9870	0.9835	0.9941	0.9945	0.9977	1

Source: Own estimate

Appendix Table A2: Decomposition of inequality in global life expectancy

Absolute	1960	1970	1980	1990	2000	2010	2017
Within							
GE(-1)	0.01063	0.0072	0.00662	0.00453	0.00526	0.00375	0.00221
GE(0)	0.00963	0.00643	0.00557	0.00391	0.00463	0.00336	0.00204
GE(1)	0.00901	0.00589	0.00484	0.00346	0.00414	0.00305	0.0019
GE(2)	0.0087	0.00553	0.00432	0.00312	0.00375	0.0028	0.00179
Gini	0.019	0.014	0.012	0.01	0.011	0.01	0.008
Between							
GE(-1)	0.01854	0.01456	0.01093	0.00919	0.00819	0.00529	0.00387
GE(0)	0.01774	0.01391	0.01052	0.0088	0.00793	0.0052	0.00382
GE(1)	0.01715	0.01341	0.01019	0.00847	0.00771	0.00512	0.00378
GE(2)	0.01677	0.01304	0.00993	0.00821	0.00753	0.00506	0.00374
Gini	0.102	0.089	0.077	0.07	0.067	0.056	0.048
Overlap							
Gini	0.009	0.006	0.005	0.003	0.005	0.004	0.003
Total							
GE(-1)	0.02917	0.02176	0.01756	0.01373	0.01346	0.00904	0.00608
GE(0)	0.02737	0.02034	0.01609	0.01271	0.01256	0.00856	0.00586
GE(1)	0.02617	0.0193	0.01502	0.01193	0.01185	0.00817	0.00568
GE(2)	0.02547	0.01856	0.01425	0.01133	0.01129	0.00786	0.00553
Gini	0.12978	0.10973	0.09446	0.08327	0.0831	0.06973	0.05912
Percentage share of within and between components							
Within							
GE(-1)	36.44	33.09	37.70	32.99	39.08	41.48	36.35
GE(0)	35.18	31.61	34.62	30.76	36.86	39.25	34.81
GE(1)	34.43	30.52	32.22	29.00	34.94	37.33	33.45
GE(2)	34.16	29.80	30.32	27.54	33.22	35.62	32.37
Gini	14.64	12.76	12.70	12.01	13.24	14.34	13.53
Between							
GE(-1)	63.56	66.91	62.24	66.93	60.85	58.52	63.65
GE(0)	64.82	68.39	65.38	69.24	63.14	60.75	65.19

GE(1)	65.53	69.48	67.84	71.00	65.06	62.67	66.55
GE(2)	65.84	70.26	69.68	72.46	66.70	64.38	67.63
Gini	78.59	81.11	81.52	84.06	80.63	80.31	81.19
Overlap							
Gini	6.93	5.47	5.29	3.60	6.02	5.74	5.07

Source: Own estimate

Appendix Table A3: Correlation coefficient of between and within components for inequality measures

Between component					
	<i>GE(-1)</i>	<i>GE(0)</i>	<i>GE(1)</i>	<i>GE(2)</i>	<i>Gini</i>
GE(-1)	1				
GE(0)	0.99801	1			
GE(1)	0.99307	0.99848	1		
GE(2)	0.98594	0.99439	0.99869	1	
Gini	0.98224	0.99168	0.99682	0.99886	1
Within component					
	<i>GE(-1)</i>	<i>GE(0)</i>	<i>GE(1)</i>	<i>GE(2)</i>	<i>Gini</i>
GE(-1)	1				
GE(0)	0.99998	1			
GE(1)	0.99989	0.99997	1		
GE(2)	0.99971	0.99985	0.99995	1	
Gini	0.99755	0.99751	0.99723	0.99672	1

Source: Own estimate

Appendix Figure A1: Moran's Global I in life expectancy for all countries

Source: Own estimate

Appendix Figure A2: Scatter plots of inequality in income per capita and inequality in life expectancy for all countries

Source: Own estimate