Impact of Socio-Economic Variables on Life Expectancy: An Empirical Study for 138 Countries

Mahmoud MOURAD

Professor of Econometrics at the Lebanese University

Faculty of economic sciences and business administration-branch V

Abstract

This study has examined the impact of mortality rate of children under five years of age (MORRATE), physicians (PMP), health expenditure per person (HEPP), access to electricity (AELEC) and GDP per capita on life expectancy at birth (LEB) for one hundred and thirty-eight countries taken as cross-sectional data. The MORRATE ranged from 2.4 to 160.2 (per 1,000 people), thus reflecting an inequality in LEB which fluctuates between 44.8 and 82.8. The PMP varies from 0.01 to 7.74, the HEPP between 16.92 and 8264 USD, the AELEC between 4.1% to 100% and finally the GDP per capita oscillates between 326.6 and 102,863 USD. The multiple linear regression model is estimated using the OLS method and several tests for heteroscedasticity are performed. The null hypothesis of homoscedasticity is rejected and therefore the Weighted Generalized Least Squares **(**WGLS**)** method is used to produce unbiased, efficient and consistent estimators. The results showed a negative impact of MORRATE on LEB. A single increase in the number of deceased children leads to a decrease of about 2.12 months in LEB. The HEPP has a positive impact on LEB, so if HEPP rises to 100 USD then the LEB rises by 33 days approximately. When introducing four binary variables characterizing the five continents, and taking Oceania as a reference, the life expectancy in an African country will be about 2.4 years less than the LEB reference. For the other continents, it seems that the values of LEB are very close.

Keywords: multiple linear regression modeling, estimation, heteroscedasticity, tests, life expectancy, binary variables.

1. Introduction

There is no doubt that the life expectancy at birth has evolved a lot since the fifties of the last century. This is due to scientific developments and social awareness, which were positively reflected in the prolongation of human life in all countries of the world. It is certain that the increase in life expectancy is due to the availability of many factors such as the good

governance of social affairs, health care, medical development (Charreaux, 2010), and biomedical research and innovation (Lichtenberg, 2017). Furthermore, combating corruption, improving service delivery, the rule of law, strengthening democratic institutions were relevant suggestions from the United Nations (UN, 2007), (Mourad M., 2018) because they affect significantly the Human Development Index (HDI) which is linked to the indicator of life expectancy.

With time, the availability of education for all individuals, and the emergence of multiple specialties in the medical domain, have produced an increase in the number of physicians, as well as the availability of medical services and awareness of the family health, have caused a tangible limitation in the number of deaths of children under five, and it is in itself one of the most important factors that affect the prolongation of life expectancy. Despite the importance of what we have mentioned, it is true that "medicine has become more powerfully scientific, it has also become increasingly depersonalized" (Miles & Mezzich, 2011), so that it pays excessive attention to the scientific aspect without giving human values the due care. However, a large gap still exists between developed and developing countries.

As stated by (Corvol, 2011), the LEB time series increase with about one year every four years, and over the period 1950-2015[1], the evolution of life expectancy reached a maximum of 31.2 years in East Asia and a minimum of 10.4 years in North America. On the other hand, we are witnessing a significant difference in life expectancy between developed countries and developing countries especially in southern and central Africa which are always distant from other continents as shown in the figure (1). Nevertheless, a decrease in inequality is noticed since 1950, and perhaps this is due to the decrease in child mortality in the world as mentioned by (Jordá & Niño-Zarazúa, 2017). This remarkable improvement in the life expectancy has affected all humans regardless of their social category, remarking that women live longer than the men (Blanpain, 2011). It is evident that medical research in developed countries has played a fundamental role in increasing life expectancy.

---

[1] Source: https://www.inegalites.fr/Les-inegalites-d-esperance-de-vie-dans-le-monde-
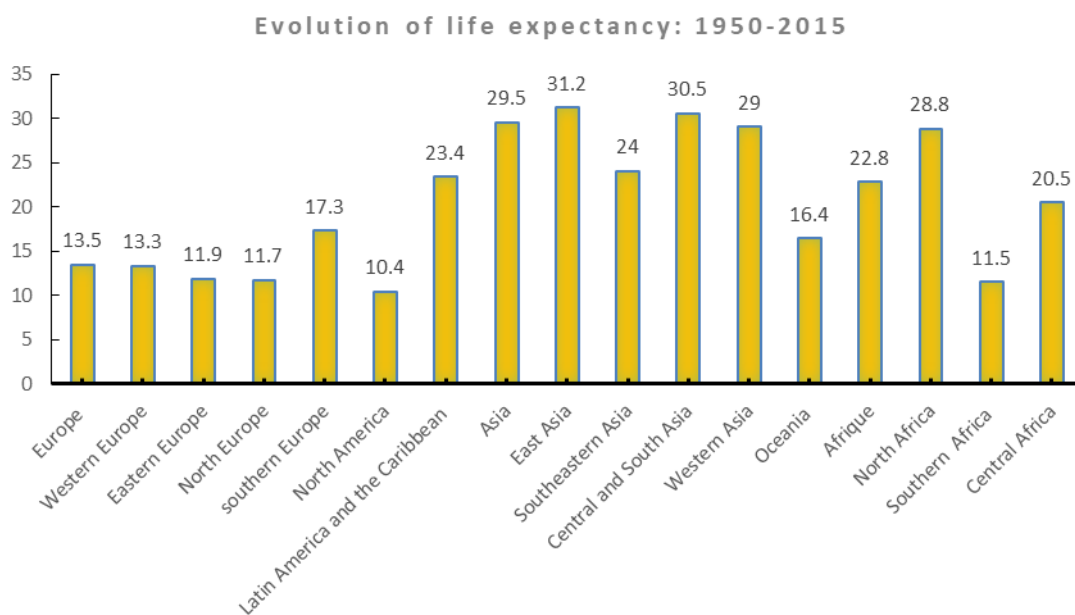
Figure 1

From an economic point of view, there must be different variables that may positively or negatively affect the longevity of a person. From those variables, we will deal with the gross domestic product per capita in U.S. dollars (GDPPCAPITA), that is from a straightforward division of total GDP by the population, the mortality rate under 5 per 1,000 live births (MORRATE), the physicians (per 1,000 people) include generalist and specialist medical practitioners PMP, the health expenditure per person (HEPP) in U.S. dollars, and the access to electricity AELEC (% of population). The data related to one hundred and thirty-eight countries (see Annex) are obtained from the World Bank website after making sure that there were no missing data for all countries, the data was published in 2010, that is before the so-called Arab Spring**.**

This paper focuses on a multiple linear regression model relating first the dependent variable (Y=LEB) to five quantitative explanatory variables that are (X_1=MORRATE), (X_2=PMP), (X_3=HEPP), (X_4=AELEC) and (X_5=GDPPCAPITA). Second by adding four explanatory binary variables to characterize the difference among the five continents. Since the study covers 138 countries (Figure 2) in the year 2010, there is no need to talk about autocorrelation in any model that can be used.

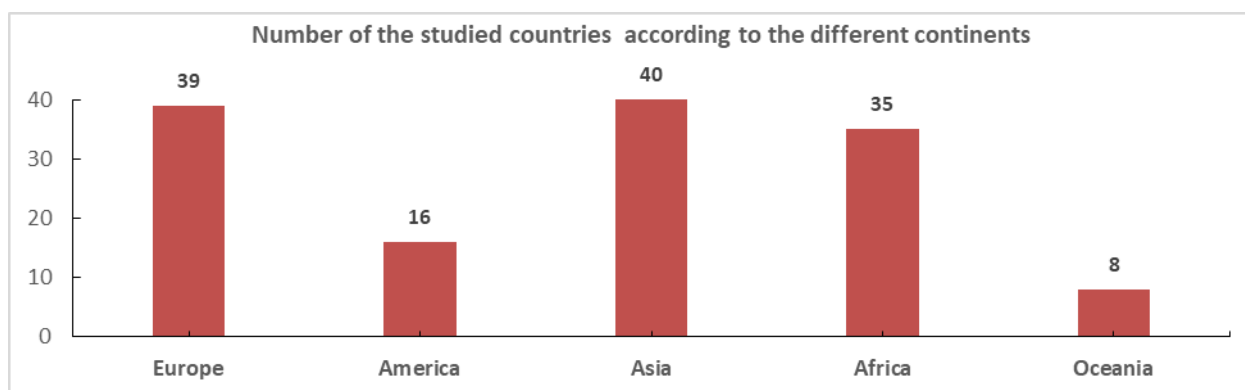Number of the studied countries according to the different continents

Figure 2

Since the study has been based on cross-sectional data, the individual impact of a country will be a considerable effect and thus the issue of the heteroscedasticity of residues will come as a result of the vast differences within countries. The paper is divided into six sections. The first section deals with a general introduction. In the second section, basic statistics and graphical data are shown and a review of literature is done in the third section. In fourth section, an overview of multiple linear regression modeling is performed paying attention to remedies for heteroskedasticity. In the fifth section, the application of MLR model will be carried out using the most known tests for heteroskedasticity. Finally, in the sixth section, a conclusion will be made from the findings.

2. Basic Statistics and Graphical Description of Data

In table (1), for each of the variables above, we calculated the main statistical characteristics as the mean, the standard deviation, the coefficient of variation (CV), the minimum, the maximum and the range. This table allows for a more precise description of the variables.

Table 1.

Basic statistics of the variables						
Variables	AELEC	LEB	MORRATE	PMP	HEPP	GDPPCAPITA
Mean	79.23	70.86	33.70	1.86	1139.42	14082.75
variance	967.24	77.98	1292.70	2.61	3417753.16	368372746.30
Standard deviation	31.10	8.83	35.95	1.62	1848.72	19193.04
coefficient of variation CV	39.25	12.46	106.69	86.89	162.25	136.29
Median	99.00	73.47	17.30	1.73	337.23	4919.06
Maximum	100.00	82.84	160.20	7.74	8264.07	102863.10
Minimum	4.10	44.84	2.40	0.01	16.92	326.60
Range	95.90	38.00	157.80	7.73	8247.15	102536.49

This table shows wide differences between countries in all indicators. For the LEB variable, there is a closeness between the mean (70.86) and the median (73.47), and this reflects the

possibility of a normal distribution of this variable. The CV $\left(\frac{standard\ deviation}{mean} \times 100\right)$, known as relative standard deviation, is an indication of the extent of variability in relation to the mean. A high CV value implies a large dispersion around the mean. For the LEB variable, the value (CV=12.46) indicates the presence of an acceptable homogeneity of the LEB distribution. But there is a considerable inequality in the life expectancy between the countries, in particular, between poor and rich countries (range = 38). For the rest of the variables, the CV values range between 39.25 and 136.29, and this reflects a large farness between observations of each variable, which reduces greatly the importance of using the mean and building on it for important decisions. In fact, the countries are significantly differentiated among themselves in the variables AELEC, MORRATE, PMP, HEPP, and GDPPCAPITA. It is enough to see the big difference between the maximum and minimum limits, and this translates clearly in the range values 95.9, 157.8, 7.73, 8247.15, and 102536.49 respectively.

The graphs are especially important because they give an idea of the fluctuations of the variables, especially about the appearance of signs of heteroskedasticity (no spherical disturbances). In this case, the researcher will be unable to use the least squares method to estimate the parameters because the independent variables will have the particularity of endogeneity, i.e. they have effect on the residuals, and this is expressed by

$E\left(u_i | X_{1,i}, \ldots, X_{k,i}\right) \neq 0.$ It is noticed that the heteroskedasticity does not affect only the features

of biased and consistent estimators, but it causes the loss of efficiency. This, in turn, prevents the achievement of basic statistical distributions, such as Student, Fisher and Chi-square. Thus, other methods should be used, such as the maximum likelihood method, the generalized least squares (GLS) method, the weighted generalized least squares (WGLS) or the feasible generalized least squares (FGLS). For more information about these methods (Mourad 2017, volume I, p. 395-402). In annex (2), for all variables, the graphs are sorted by the MORRATE variable from largest to smallest. The variable MORRATE decreases exponentially from Sierra Leone (MORRATE = 160.2, LEB=44.8) to Luxembourg (MORRATE = 2.4, LEB=80.6). Respecting the sorting by MORRATE, the maximums and minimums are attributed to the countries as shown below:

- MORRATE: Sierra Leone: 160.2; Luxembourg: 2.4
- LEB: Sierra Leone: 44.84; Japan: 82.84
- PMP: Tanzania: 0.008; Qatar: 7.74
- HEPP: Ethiopia: 16.92; United States: 8264.07
- AELEC: Liberia: 4.1, 36.2% of countries have 100% access to electricity
- GDPPCAPITA: Liberia: 326.6 USD, Luxembourg: 102863 USD

Since the LEB variable is taken as dependent, it will be beneficial for this research to present the scatter plots choosing as abscissa each of the other explanatory variables.
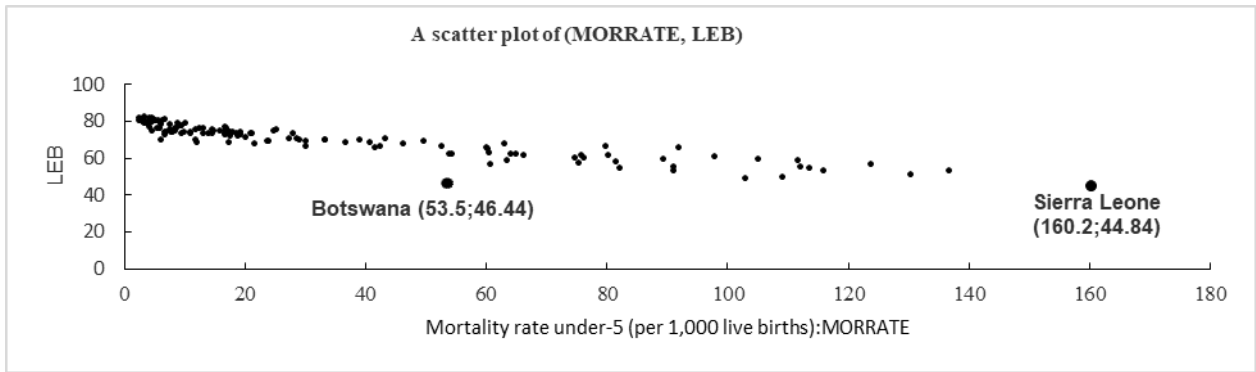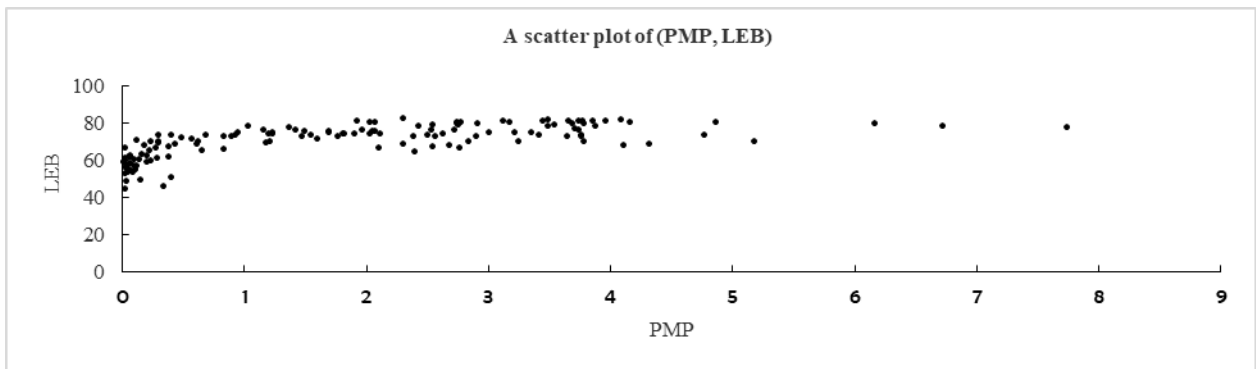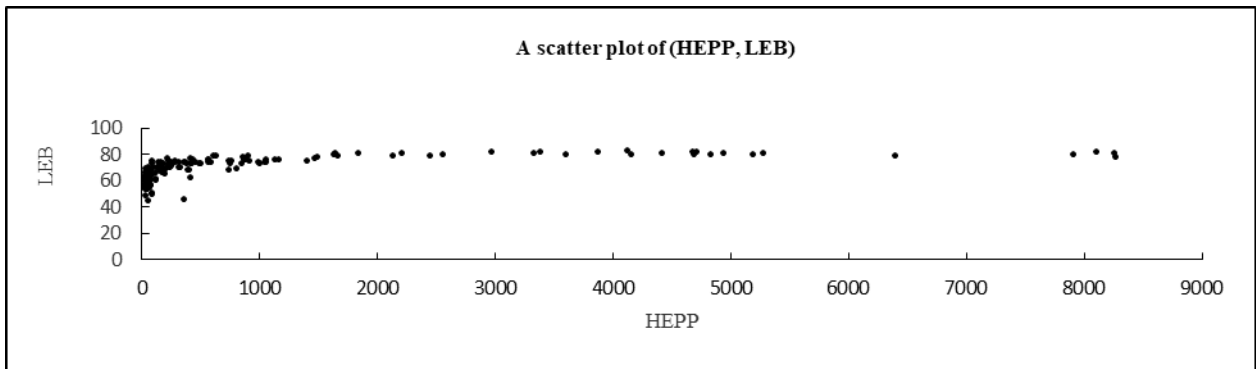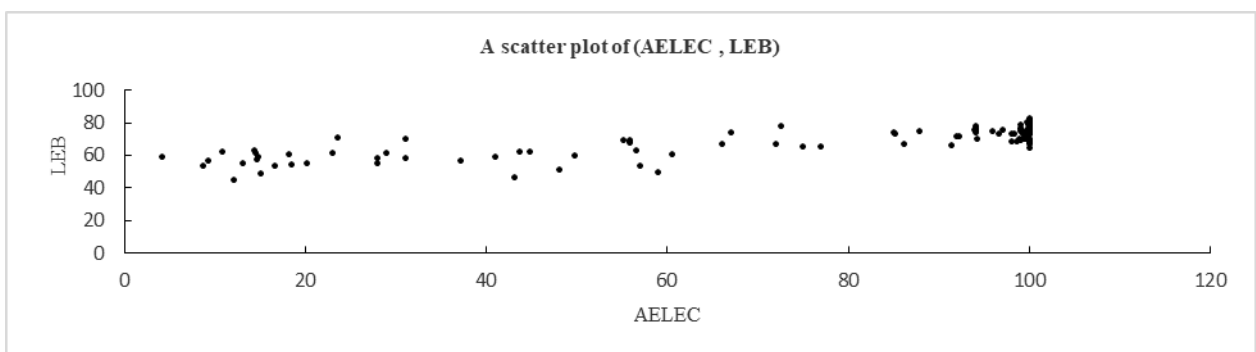
Figure 3

Figure 4

Figure 5

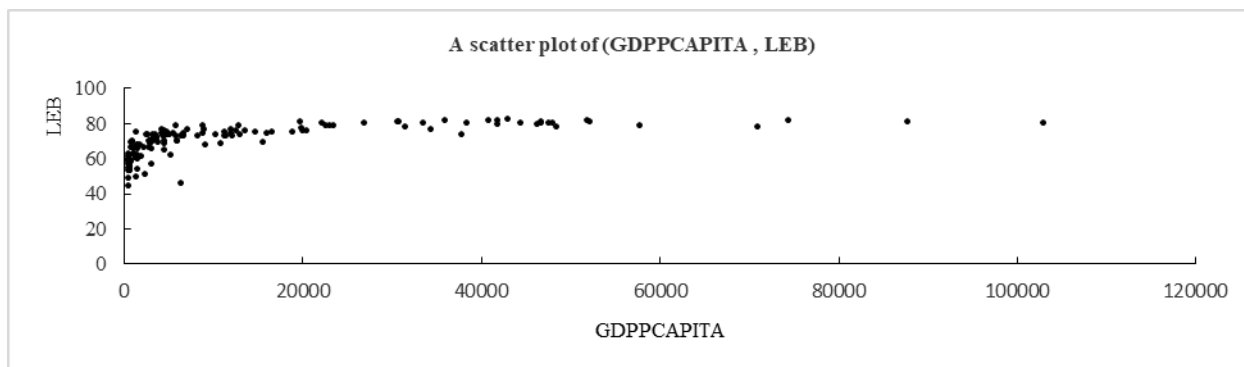Figure 6

A scatter plot of (GDPPCAPITA , LEB)

Figure 7

If we look deeper at these graphics, we notice how a group of points is clustered having roughly the same abscissa. This indicates clearly the existence of a heteroskedasticity multiple linear regression model, which requires the use one of the methods mentioned above to obtain unbiased efficient and consistent estimators.

3. Review of Literature

There are many studies dealing with the issue of life expectancy which is affected by many environmental, health, and scientific factors. In the following, several works will be presented relating in one way or another to our study. For example, the paper by (Ali & Ahmad, 2014) revealed, in short run, a negative and significant impact of CO2 emissions on life expectancy for the Sultanate of Oman, (Ehsan, Farhad, & Farzad, 2020) showed that the highest variability of mortality in Commonwealth of Independent States (CIS), could be explained by CO2 variability using the generalized method of moments (GMM) estimation.

One of the recent literature (Suresh & Vidushi, 2019) that draws great attention is that the health reality of doctors in India threatens to shorten the life expectancy, that is 59 years as compared to 67.9 years of an average person in India. In fact, the category of doctors is expected to have long life because doctors possess sufficient health knowledge to maintain their fitness and health and treat their health symptoms early. This study showed a high probability of doctors suffering from heart diseases and cancer, and the number of deaths is increasing. Moreover, the phenomenon of suicide is undergoing an upward trend, and this is what worries the Indian Medical Association (IMA), which put the finger on the wound through a very important field study that dealt with ten thousand doctors and investigated the causes behind decreased life expectancy of doctors by more than 10 years. After following up the sample for ten years, it was found that 282 doctors died, or 2.82%, divided between 87% males and 13% females. The causes of death were 27% due to heart disease, 25% due to cancer, 2% by infection, and 1% by suicide. While death due to stress was the main cause of premature deaths, followed by a lack of regular medical examinations.

Since our study utilizes the regression technique with heteroskedasticity, the Ordinary Least Square (OLS) method is not valid in this context. For this reason, we focus on the literatures that have dealt with this kind of data. Concerning the use of the Maximum Likelihood (ML),

Weighted Generalized Least squares **(**WGLS), Generalized Least Squares (GLS), Feasible Generalized Least Squares (FGLS) estimation methods, (Mourad M. , 2017) did both theoretical and practical studies, (Chinonso, Oluchukwu, Charity, Nnaemeka, & Chukwunenye, 2020) provided an extension of the ARIMA models when the generated residuals are considered a Fourier series, modeling Malaria Incidence Rates, (Safi S. , 2004) studied the efficiency of the OLS, GLS and estimated GLS (EGLS) estimators when the disturbances reveal a first and second order autoregressive models, (Safi & Abu Saif, 2014) compared the prediction using the GLS method for parameter estimation in the regression models with autocorrelated disturbances considering real data and adopting an ARIMA model for the time series, (Lee & Lund, 2004) proposed the properties of OLS and GLS estimators in a simple linear regression with stationary autocorrelated errors, (Mandy & Fridli, 2001) show under very parsimonious assumptions that FGLS and GLS are asymptotically equivalent when errors follow an invertible MA(1) process.

4. An Overview of Multiple Linear Regression Modeling

The general form of a multiple linear regression model (MLR) is:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_k X_{ki} + u_i \quad \text{(Eq 1)}$$

In matrix form, $(\mathcal{Y} = \mathcal{X}\mathcal{B} + u\ )$, $\mathcal{Y}$ and **u** are $(n \times 1)$- dimensional vectors, $\mathcal{X}$ is a $(n \times (k+1))$ matrix and $\mathcal{B}$ is a $(k+1)$-dimensional vector. If the error variance is not constant across countries $(\sigma_i^2 \neq \sigma_j^2, \sigma_{ij} = 0 \ \forall i \neq j)$ then a pure heteroscedastic MLR model is required. In this case, the variance – covariance matrix of **u** is:

$$\mathbf{V(u)} = \mathbf{\Omega} = \begin{pmatrix} \sigma_1^2 & 0 & 0 & \dots. & 0 \\ 0 & \sigma_2^2 & 0 & \dots & 0 \\ 0 & 0 & \sigma_3^2 & \dots & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \sigma_n^2 \end{pmatrix} \neq \sigma^2 I \quad \text{(Eq 2)}$$

Using the OLS method causes underestimated standard errors of the coefficients, and by consequence, the *t* statistics of the coefficients will be inflated. This, in turn, leads to reject the true null hypothesis. For this, the OLS method must be replaced by an appropriate estimator such that the WGLS method or appropriate transformations for different variables must be used. If $\mathbf{\Omega}$ is known then the GLS estimator of $\mathcal{B}$ has the following expression:

$$\hat{\mathcal{B}}_{GLS} = (\mathcal{X}'\Omega^{-1}\mathcal{X})^{-1}\mathcal{X}'\Omega^{-1}\mathcal{Y} \quad \text{(Eq 3)}$$

This leads to an unbiased consistent and efficient estimator $\hat{\mathcal{B}}_{GLS}$(Mourad, 2017, p.396), (Maddala, 1971), (Amemiya, 1973). We obtain the same result using the transformed equation $\Omega^{\frac{-1}{2}}\mathcal{Y} = \Omega^{\frac{-1}{2}}\mathcal{X}\mathcal{B} + \Omega^{\frac{-1}{2}}u; \ \Omega^{\frac{-1}{2}}\mathbf{\Omega}\Omega^{\frac{-1}{2}} = I$. In time series, it is obvious that $\Omega^{\frac{-1}{2}}$ can have several forms which depend on the model associated to the residuals like AR (1) or MA (1) or others. If $V(u_i|\mathcal{X}) = \omega_i; \ \omega_i > 0, \forall t = 1, \dots, n$ then the Weighted Generalized Least squares (WGLS) consists in the following: $\mathcal{X}_i = (1 \quad X_{1i} \quad \cdots \quad X_{ki})', \ i = 1, \dots, n$

$$\hat{\mathcal{B}}_{WGLS} = \left(\sum_{i=1}^{n}\left(\frac{1}{\omega_i}\right)\mathcal{X}_i\mathcal{X}_i'\right)^{-1} \times \sum_{i=1}^{n}\left(\frac{1}{\omega_i}\right)\mathcal{X}_i Y_i \quad \text{(Eq 4)}$$

If Ω is unknown then it will be necessary to estimate it (let's say $\hat{\Omega}$) to obtain the FGLS estimator (Judge, et al., 1985, chapter 9, p. 351)) :

$$\hat{\mathcal{B}}_{FGLS} = (\mathcal{X}'\hat{\Omega}^{-1}\mathcal{X})^{-1}\mathcal{X}'\hat{\Omega}^{-1}\mathcal{Y} \quad (Eq\ 5)$$

If the ω_i are estimated according to the explanatory variables or other exogenous variables, which are not used in the model, then $(\hat{\mathcal{B}}_{FGLS} = \hat{\mathcal{B}}_{WGLS})$. For the equivalence of GLS and FGLS estimators, (Greene, 2008,p.157) established two conditions:

$$\text{plim}\left[\left(\frac{1}{n}\mathcal{X}'\hat{\Omega}^{-1}\mathcal{X}\right) - \left(\frac{1}{n}\mathcal{X}'\Omega^{-1}\mathcal{X}\right)\right] = 0 \quad (Eq\ 6)$$

$$\text{plim}\left[\left(\frac{1}{\sqrt{n}}\mathcal{X}'\hat{\Omega}^{-1}u\right) - \left(\frac{1}{\sqrt{n}}\mathcal{X}'\Omega^{-1}u\right)\right] = 0 \quad (Eq\ 7)$$

In this context, it is worth noting that (White, 1980) provided a valid solution for parameter estimation suggesting the Heteroskedasticity-consistent standard error (HCSE), and recommending the use of the residual square resulting from the (OLS) estimator so that $\hat{\Omega}$ is a diagonal matrix:

$$\hat{\Omega} = \text{diag}[e_i^2] = \text{diag}(e_1^2, e_2^2, \ldots, e_n^2) \quad (Eq\ 8)$$

The heteroskedasticity-consistent estimator becomes:

$$HC_0 = \hat{\Omega}_{HCSE} = (\mathcal{X}'\mathcal{X})^{-1}\mathcal{X}'\text{diag}[e_i^2]\mathcal{X}(\mathcal{X}'\mathcal{X})^{-1} \quad (Eq\ 9)$$

We obtain $(\mathcal{X}'\text{diag}[e_i^2]\mathcal{X} = \sum_{i=1}^n e_i^2 \mathcal{X}_i \mathcal{X}_i')$ and then we have:

$$HC_0 = (\mathcal{X}'\mathcal{X})^{-1}\left[\sum_{i=1}^n e_i^2 \mathcal{X}_i \mathcal{X}_i'\right](\mathcal{X}'\mathcal{X})^{-1} \quad (Eq\ 10)$$

These are called sandwich estimators. The elements on the main diagonal of HC_0 denote the estimates of variance estimators of the parameters and thus we obtain all standard deviations $HCSE_j, j = 0, 1, \ldots, k$. Thus $t_{\hat{\beta}_j} = \frac{\hat{\beta}_j}{HCSE_j}$ can be calculated and by consequence to test the significance of the parameters, a Student distribution can be carried out. An attention is drawn to the case of small samples because $HCSE_j, j = 0, \ldots, k$ becomes biased (Bera, Suprayitno, & Premaratne, 2002). For more information about this estimator see (Mourad, 2017, p. 438-439), (Hill, Griffiths, & Lim, 2012, p.309-319).

5. Application of MLR Model on the Life Expectancy in 138 Countries

In the first analysis, a MLR model will be estimated. The general equation is formulated as follows:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} + \beta_5 X_{5i} + u_i \quad (Eq\ 11)$$

In the second analysis, four explanatory binary variables will be introduced in (Eq 11) and a WGLS re-estimation will be made. Assuming that the conditions for using the least squares method are available, the parameters will be estimated through the RATS software (version 9). In Table (2), figure the estimate of the model with all the resulting statistics. A priori, the quality of linear regression is very good as indicated by R^2. But it is not advisable to make interpretations before ensuring an accurate estimator.

Table 2.

Estimation by Ordinary Least Squares, dependent Variable Y=LEB, n=138				
$R^2 = 88.19\ \%,\ \ \bar{R}^2 = 87.75\ \%\ \ F(5,132) = 197.23, p - value\ =\ \ 0.000$				
Standard Error of Estimate = 3.091; Sum of Squared Residuals = 1261.172				
Variables	**Coefficients**	**Std Error**	**T-Stat**	**P-value**
Constant	71.10	1.695	41.948	0.000
$X_1 = MORRATE$	-0.164	0.014	-11.496	0.000
$X_2 = PMP$	-0.064	0.248	-0.256	0.798
$X_3 = HEPP$	0.0005	0.0004	1.242	0.217
$X_4 = AELEC$	0.052	0.016	3.146	0.002
$X_5 = GDPPCAPITA$	0.00006	0.00004	1.52	0.131

Before testing the presence of heteroskedastic errors, an overview will be presented to explore it through a graphic regressing the square errors e_i^2 on the adjusted value $\hat{Y}_i$.
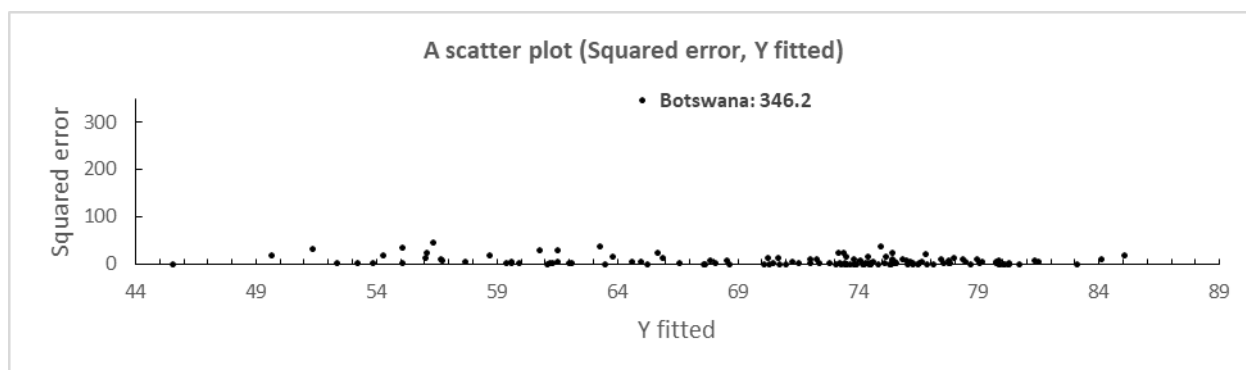

A scatter plot (Squared error, Y fitted)

Figure 8

There is an aberrant observation that concerns Botswana. In 2010, GDP P.C. Annual Growth reached 23%. But in 2009, it was -8.1%. For this reason, this point has been removed and a new graph has been drawn (Figure 9).
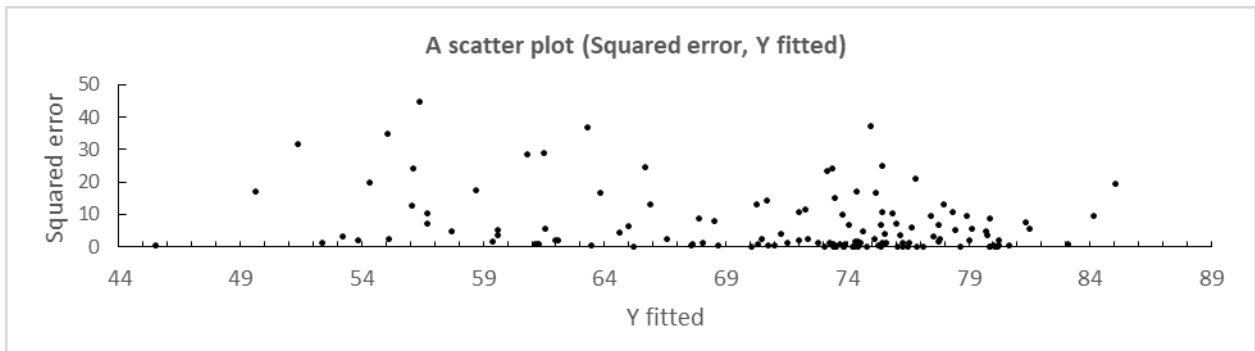
Figure 9

It appears that there is a large dispersion in the squared errors. They change from zero to 45, and this indicates the presence of heteroskedasticity. In the econometric literature, there is a large number of tests that serve to detect the presence of heteroskedasticity. Mourad (2017, p. 407-427) presented nine tests, explaining for each one the necessary steps to reach an adequate decision. Among these tests, we will be limited to five that are: Goldfeld-Quandt test, Spearman's rank correlation test, Harvey-Godfrey test, Breusch–Pagan and White test. For all tests, the null and alternative hypothesis are:

$$H_0 : \text{The model is homoskedastic}$$
$$H_a : \text{The model is heteroskedastic}$$

Table 3

Test for heteroskedasticity		
Tests	Available information	Notes
Goldfeld-Quandt	$GQ = 3.15 \geq F_{0.05;52,52} = \mathbf{1.58}$ Decision: H_0 is rejected	the data are sorted by MORRATE variable from largest to smallest
spearman's rank correlation test	$\hat{\rho}_s = \underbrace{0.209}_{(t=2.49) > t_{0.05;n-2}=1.66}$ Decision: H_0 is rejected	the data are sorted by MORRATE variable from largest to smallest
Harvey-Godfrey	Chi-Squared (5) = 12.37, p-value 0.030 Chi-Squared (5) = 13.76, p-value 0.017 H_0 is rejected	Auxiliary regression(log-log) Auxiliary regression (log - lin)

Breusch–Pagan	Chi-Squared (5) = 4.078, p-value 0.538 Chi-Squared (5) = 8.160, p-value 0.148 H_0 is not rejected	Auxiliary regression(lin-lin) Auxiliary regression(lin-log)
White	Chi-Squared (20) = 33.09, p-value 0.033 H_0 is not rejected	e_i^2 (OLS) is dependent variable in the auxiliary regression

Among these five tests, four tests reject the homoscedasticity of errors.

It seems that the gross domestic product per capita X_5 and the health expenditure per person X_3 are correlated. The cross-correlation for this pair of variables is calculated $(\rho_{35} = 0.918)$ showing a very high value. In fact, if the absolute value of a cross-correlation is greater than 0.8 then the presence of **multicollinearity** cannot be ignored (Mourad, 2017, p.511). The Farrar - Glauber (FG) test can be used to decide about the multicollinearity. This test consists in calculating the statistic:

$$FG = -\left(n - 1 - \frac{1}{6}(2(k+1) + 5)\right)\ln(D); \; D = \begin{vmatrix} 1 & \rho_{35} \\ \rho_{35} & 1 \end{vmatrix} \cong 0.157.$$

Since $FG = 250 > FG \geq \chi^2_{0.05;3} = 7.82$, the null hypothesis of no multicollinearity is rejected at 5% level. For this reason, the variable GDPPCAPITA is removed. In the following, four binary variables $D_i, i = 1, \dots, 4$ are used to characterize the five continents

Binary variables associated to the five continents				
Continents	D_1	D_2	D_3	D_4
Europe	1	0	0	0
America	0	1	0	0
Asia	0	0	1	0
Africa	0	0	0	1
Oceania	0	0	0	0

The final model containing four quantitative explanatory variables and four binary variables is estimated by WGLS, and the results are in Table (4).

Table 4.

Estimation by Weighted Least Squares, dependent Variable Y=LEB, n=138 Standard Error of Estimate $=$ 2.048; Sum of Squared Residuals $=$ 541.195				
Variables	**Coefficients**	**Std Error**	**T-Stat**	**P-value**
Constant	**74.573**	**1.975**	**37.754**	**0.000**
$X_1 = MORRATE$	**-0.177**	**0.016**	**-11.231**	**0.000**
$X_2 = PMP$	0.142	0.199	0.710	0.479
$X_3 = HEPP$	**0.0009**	**0.0001**	**7.044**	**0.000**
$X_4 = AELEC$	0.025	0.019	1.319	0.189
D1	-1.238	0.983	-1.260	0.210
D2	-0.725	1.030	-0.703	0.483
D3	-0.030	0.954	-0.031	0.975
D4	**-2.355**	**1.127**	**-2.090**	**0.039**

The standard error of estimates decreases from 3.091 obtained through the OLS method to 2.048 with the WGLS method[2]. We find that the variable MORRATE (the mortality rate under 5) has a negative effect on life expectancy. If MORRATE increases by one unit then the LEB decreases by 0.177 years, that is, 2.12 months approximately. This very interesting result requires particular concern from rich countries to help poor countries to reduce the mortality rate under 5. For the variables PMP and AELEC, despite the positive value of their estimated parameters, their effects on LEB are not significant (p-values are 48% and 19 % respectively). It was anticipated that the variable HEPP has a positive impact (p-value = 0.000) on the variable LEB. If HEPP rises 100 USD, then the LEB rises 33 days approximately.

Regarding the parameters of binary variables, we cannot use the Student test unilaterally on each of them because the level of significance is expressed by all the variables D_i, $i = 1, \dots, 4$. For this, the null hypothesis $H_0: \beta_i = 0$ for $i = 5, 6, 7, 8$ must be tested versus $H_a: \beta_i \neq 0$ for at least i. Since the Fisher statistic is F(4,129)= 2.56 with p-value 0.042, the H_0 is rejected for 5 % significant level. If $D_i = 0$ for $i = 1, \dots, 4$ then it is a country in Oceania taken as a reference. For fixed values of quantitative variables, if $D_4 = 1$ then it is an

---

[2] The estimation by OLS with Heteroscedasticity-Consistent Standard Errors (HCSE)) was also performed but the results were very close to those in table (2).

African country and the life expectancy will be about 2.4 years less than that for the LEB reference. For the other binary variables, it seems that the values of LEB are close to the LEB in Oceania (all other things being equal).

6. Conclusions and Recommendations

The study revealed a large impact of MORRATE and HEPP variables on the LEB variable. There is a very significant negative impact of MORRATE on LEB. So that every increase in MORRATE is accompanied by a decrease in LEB. If MORRATE increases by one unit then the LEB decreases by 0.177 years, that is, 2.12 months approximately. A positive impact from HEPP on LEB is observed. If HEPP rises by 100 USD, then the LEB rises by 33 days approximately. Whereas the PMP and AELEC variables did not have a significant positive impact on the LEB variable. Since the gross domestic product per capita and the health expenditure per person are highly correlated (**Pearson correlation coefficient**=0.918), the null hypothesis for **multicollinearity** is not rejected and by consequence the variable GDP per capita is removed from the MLR model.

There is a clear difference in the life expectancy between countries, particularly between Africa and the other continents. We find a huge difference between 44.8 years in Sierra Leone and 82.8 years in Japan, so an increase of about 85%. This result has deep humanitarian dimensions, according to which the developed countries, which are distinguished by a very high social well-being, must cooperate intensively with the poor countries, to further push their living standard.

It is no longer acceptable in the 21st century that in Sierra Leone, for example, 160 children die before the age of five while in Luxembourg we are witnessing the death of 2 children in this age category. The HEPP variable also shows a large gap between countries. Its value is about 17 USD in Ethiopia while it reached 8264.07 USD in United States.

For the PMP variable, Tanzania has 0.008 physicians per thousand people, while Qatar has 7.74 physicians. For the AELEC, Liberia has 4.1% versus 36.2% of countries have 100% access to electricity. Finally, the GDPPCAPITA is 326.6 USD in Liberia, but it is 102863 USD in Luxembourg.

Therefore, there should be a proper planning to bridge the gap between continents and make human health a priority for countries. This is up to governments, which give specific attention to scientific research in all fields and in particular the medical field. The health expenditures in the long-term plays an important role in extending longevity and this has been confirmed by many studies, including those we mention (Spillman & Lubitz, 2000) which confirms the statement that «Increases in longevity after the age of 65 years may result in greater spending for long-term care, but the increase in the number of elderly persons has a more important effect on total spending». However, we should not look at this point of view from a financial criterion but rather in the sense that gives man the greatest value in life by exploiting all possible means which produce a positive impact on the life expectancy.

References

Ali, A., & Ahmad, K. (2014). The Impact of Socio-Economic Factors on Life Expectancy for Sultanate of Oman: An Empirical Analysis. *Munich Personal RePEc Archive MPRA*, 1-13. Retrieved from https://mpra.ub.uni-muenchen.de/70871/

Amemiya, T. (1973). Generalized Least Squares with an Estimated Autocovariance Matrix. *Econometrica, 41*(4), 723-732. https://doi.org/10.2307/1914092

Bera, A., Suprayitno, T., & Premaratne, G. (2002). on some heteroskedasticity-robust estimators of variance-covariance matrix of the least-squares estimators. *Journal of Statistical Planning and Inference, 108*, 121-136. https://doi.org/10.1016/S0378-3758(02)00274-4

Blanpain, N. (2011). *L'espérance de vie s'accroît, les inégalités sociales face à la mort demeurent.* Paris: division Enquêtes et études démographiques, N°1372. Institut national de la statistique et des études économiques (INSEE).

Charreaux, G. (2010). *La gouvernance hospitalière : quelques réflexions à partir de la gouvernance d'entreprise. JEL Classification : I190 ; G390.* Retrieved from file:///C:/Users/win10/Downloads/1100705.pdf

Chinonso, M. E., Oluchukwu, C. A., Charity, U. O., Nnaemeka, M. E., & Chukwunenye, I. O. (2020). On the Fourier Residual Modification of Arima Models in Modeling Malaria Incidence Rates among Pregnant Women. *American Journal of Theoretical and Applied Statistics, 9*(1), 1-7.

Corvol, P. (2011). *La recherche médicale et pharmaceutique. Conférence intitulée: LA MONDIALISATION DE LA RECHERCHE.* (G. Fussman, Ed.) Paris: Collège de France. https://doi.org/10.4000/conferences-cdf.311

Ehsan, R., Farhad, T.-H., & Farzad, T.-H. (2020). How Is Mortality Affected by Fossil Fuel Consumption, CO2 Emissions and Economic Factors in CIS Region? *Energies, 13,* 1-13. https://doi.org/10.3390/en13092255

Greene, W. (2008, p.157). *Econometric analysis* (6 ed.). USA: Pearson Prentice Hall.

Hill, R., Griffiths, W., & Lim, G. (2012, p.309-319). *Principles of Econometrics.* (Third, Ed.) John Wiley & Sons (Asia).

Jordá, V., & Niño-Zarazúa, M. (2017). Global inequality in length of life, 1950–2015.WIDER Working Paper 2017/192, JEL classification: D63, I01, I14. United Nations University. UNU-WIDER project 'The Economics and Politics of Taxation and Social Protection'. https://doi.org/10.35188/UNU-WIDER/2017/418-6

Judge, G., Hill, R. C., Griffiths, W. E., E, W., Lütkepohl, H., & Lee, T. (1985, chapter 9, p. 351)). *Introduction to the Theory and Practice of Econometrics* (Second Edition ed.). John Wiley & Sons.

Lee, J., & Lund, R. (2004). Revisiting simple linear regression with autocorrelated errors. *Biometrika, 91*(1), 240-245. https://doi.org/10.1093/biomet/91.1.240

Lichtenberg, F. R. (2017). The impact of biomedical innovation on longevity and health. *Nordic Journal of Health Economics, 5*(1), 45-57. https://doi.org/10.5617/njhe.1290

Maddala, G. S. (1971). Generalized Least Squares with an Estimated Variance Covariance Matrix. *Econometrica, 39*, 23-33. https://doi.org/10.2307/1909137

Mandy, D. M., & Fridli, S. (2001). Exact fgls asymptotics for ma errors. 1-10. Retrieved from https: // www.researchgate.net/publication/23693819

Miles, A., & Mezzich, J. E. (2011). The care of the patient and the soul of the clinic: personcentered medicine as an emergent model of modern clinical. *The International Journal of Person Centered Medicin, 1*(2), 207-222. https://doi.org/10.5750/ijpcm.v1i2.61

Mourad, M. (2017). *Econometrics from Theory to Practice, Volume I* . Beirut, Lebanon: Lebanese University.

Mourad, M. (2018). Management Governance Analysis in the Arab World: MANOVA Approach. *Journal of Public Administration and Governance (JPAG), 8*(4), 1-19. https://doi.org/10.5296/jpag.v8i4.13766

Safi, S. (2004). *The efficiency of OLS in the presence of auto-correlated disturbances in regression models.PhD dissertation, American University Washington, D.C. 20016.*

Safi, S. K., & Abu Saif, E. A. (2014). Using GLS to Generate Forecasts in Regression Models with Auto-correlated Disturbances with with simulation and Palestinian Market Index Data. *American Journal of Theoretical and Applied Statistics, 3*(1), 6-17. https://doi.org/10.11648/j.ajtas.20140301.12

Spillman, B. C., & Lubitz, J. (2000). The Effect of Longevity on Spending for Acute and Long-Term Care. *The New England Journal of Medicine, 342*, 1409-1415. https://doi.org/10.1056/NEJM200005113421906

Suresh, K. P., & Vidushi, S. (2019). Doctor, heal thyself: Addressing the shorter life expectancy of doctors in India. *Indian J Ophthalmol, 67*(7), 1248–1250. https://doi.org/10.4103/ijo.IJO_656_19

UN. (2007). *Good Governance Practices For The Protection Of Human Rights. Office Of The United Nations High Commissioner For Human Rights.* New-York and Geneva: HR/PUB/07/4. UNITED NATIONS PUBLICATION.

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direst test for heteroskedasticity. *Econometrica, 48*, 817-838. https://doi.org/10.2307/1912934

	Annex: The states concerned by the study										
1	Afghanistan	24	Benign	47	Ghana	70	Liberia	93	Oman	116	Swiss
2	Albania	25	Cape Verde	48	Greece	71	Lithuania	94	Uganda	117	Sweden
3	Algeria	26	Cambodia	49	Guinea	72	Luxembourg	95	Uzbekistan	118	Senegal
4	Germany	27	Canada	50	Guinea-Bissau	73	Macedonia	96	Pakistan	119	Tajikistan
5	Saudi Arabia	28	Chile	51	Guyane	74	Madagascar	97	New Guinea	120	Tanzania
6	Argentina	29	China	52	Georgia	75	Malaysia	98	Poland	121	Thailand
7	Armenia	30	Cyprus	53	Hungary	76	Malawi	99	Portugal	122	Togo
8	Australia	31	Colombia	54	India	77	Maldives	100	Peru	123	Tonga
9	Austria	32	Congo	55	Indonesia	78	Mali	101	Qatar	124	Trinidad and Tobago
10	Azerbaijan	33	Korea	56	Iran	79	Malta	102	Romania	125	Tunisia
11	Bahrain	34	Croatia	57	Iraq	80	Morocco	103	United Kingdom	126	Turkmenistan
12	Bangladesh	35	Cuba	58	Ireland	81	Mauritania	104	Rwanda	127	Turkey
13	Barbados	36	Ivory Coast	59	Iceland	82	Mexico	105	Laos	128	Ukraine
14	Belgium	37	Denmark	60	Israel	83	Micronesia	106	Kyrgyztan	129	Uruguay
15	Belize	38	Djibouti	61	Italy	84	Moldova	107	Slovak	130	Vanuatu
16	Bhutan	39	El Salvador	62	Japan	85	Mongolia	108	Czech	131	Viet nam
17	Bosnia-Herzegovina	40	Spain	63	Jordan	86	Montenegro	109	Samoa	132	Yemen
18	Botswana	41	Estonia	64	Kazakhstan	87	Mozambique	110	Serbia	133	Zambia
19	Brunei Darussalam	42	Fiji	65	Kiribati	88	Namibia	111	Sierra Leone	134	Egypt
20	Brazil	43	Finland	66	Kuwait	89	Niger	112	Singapore	135	United Arab Emirates
21	Bulgaria	44	France	67	Latvia	90	Nigeria	113	Slovenia	136	Ecuador
22	Burkina Faso	45	Russia	68	Lebanon	91	Norway	114	Sudan	137	United States
23	Belarus	46	Gambia	69	Libya	92	New Zealand	115	Sri Lanka	138	Ethiopia

Annex 2: Graphics sorted by MORRATE variable from largest to smallest
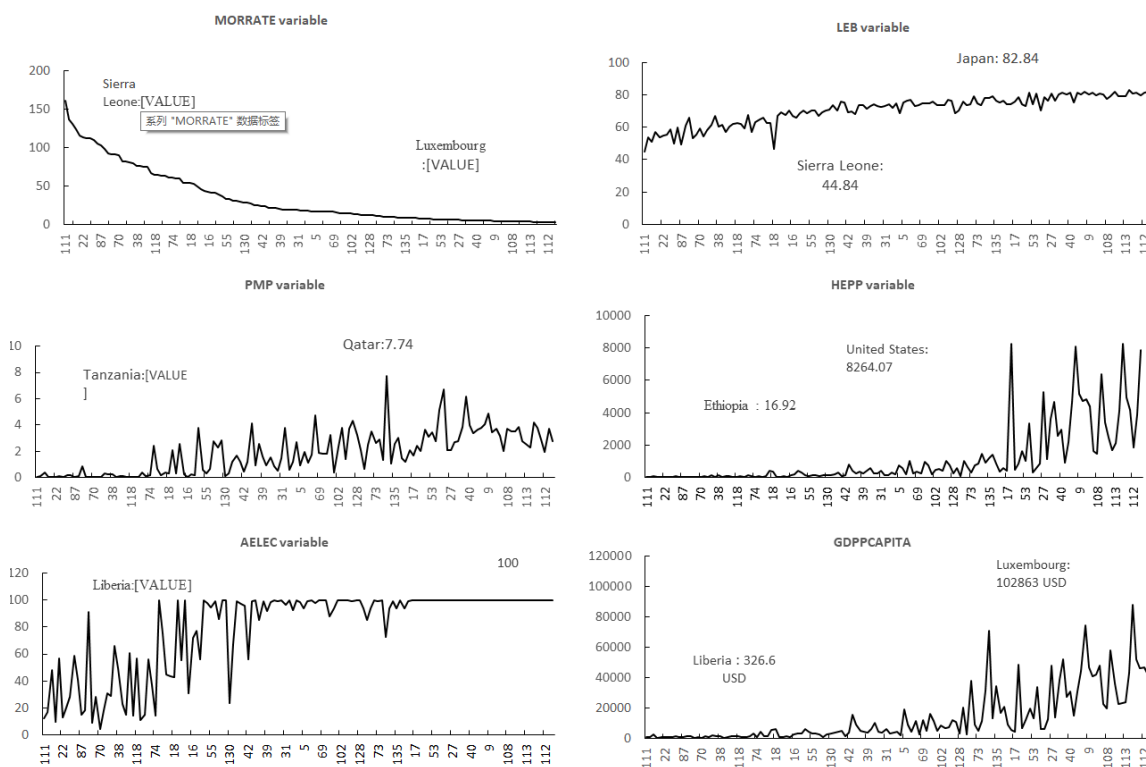
Copyright Disclaimer

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/4.0/).