PRINCIPAL COMPONENT ANALYSIS OF SOCIOECONOMIC FACTORS AND THEIR ASSOCIATION WITH LIFE EXPECTANCY AT BIRTH IN ASIA

Humera Razzak
Ex-Lecturer, Department of Statistics
Government College University Faisalabad, **PAKISTAN**

Mehboob Ali
Lecturer in Statistics, Government Post Graduate College
Jauharabad District Khushab, **PAKISTAN**
&
Maqsood Ali
Statistical Officer
Punjab Bureau of Statistics, **PAKISTAN**

ABSTRACT

Life expectancy (LE) is considered as one of key health outcome and a major indicator of human development as well. Wide ranges of socioeconomic and demographic factors have major impact on life expectancy rate at birth in various countries. Association of several socioeconomic factors with life expectancy at birth and the influencing factors in forty countries of Asia has been explored in this paper. Less surprisingly the results and discussions obtained in this paper are in agreement with previous research. A close relationship between several socioeconomic variables and life expectancy at birth is found. Principal components analysis (PCA) and backward regression is performed on quantitative secondary data collected from various databases which shows that life expectancy at birth is statistically significant at 5% level of significance and have positive association with four factors extracted from PCA. Strong significant positive correlation is found between life expectancy at birth and health expenditures, gross national income, good governance and healthy life. However crude birth rate, crude death rate and infant mortality rate has negative relationship with life expectancy at birth which shows life expectancy at birth decreases as crude birth rate, crude death rate and infant mortality rate increases. The reference year for this study is 2012.

Keywords: Life Expectancy, Socioeconomic Factors, Principal Components Analysis, Asia.

INTRODUCTION

Enjoying healthy and long life in presence of with better health facilities, good education and governess is a dream of every citizen in every nation (Lei et al., 2009) and these are the promises likely to be forgotten by every government which they makes while their establishment. Since life expectancy at birth is a major indicator of human development which is somehow associated with level of income yet nations with low income but with skillful demonstration of resources and expansion of human capabilities has achieved better developments (Yavari and Mehrnoosh, 2006; Sen, 1999).

Life expectancy rate has been greatly affected by socio economic risk factors universally especially in vulnerable communities living in under developed nations. Although statistical data on socio economic variables including population growth, health expenditures, gross national income, crude birth rate, crude death rate, infant mortality rate, population density, education, gender equality, income distribution, good governance, and healthy life is inconsistently available yet identifying and summarizing the relationship between life expectancy gap and socioeconomic risk factors mentioned above is of paramount importance. Two major hurdles that go hand in analyzing economic aspects of human development are selection of set of explanatory variables that explains variation in other economic phenomenon and specification of functional form. However third major problem which is addressed in this study is statistical analysis of human indicators in unviability of data especially in developing countries where human indicators are not maintained on regular basis.

Our study is aimed to analyze the influence of some human and health indicators in the variation in life expectancy in Asia. Necessary data of 40 countries from Asia is obtained from World Bank open data and Sustainable Society Foundation, Netherlands. Two statistical techniques are used. PC analysis technique will determine the set of human and health indicators relevant in determination of other indicators and backward regression analysis which determines an economic model for explaining and estimating relationship between variation in life expectancy at birth and socioeconomic risk factors.

LITERATURE REVIEW

Studies regarding determinants of life expectancy mostly include identifying and correlating factors influencing LE. Limitations regarding to the unavailability of LE data and lack of strong evidence against biological connection between socioeconomic determinants of life expectancy must be kept in mind before reviewing the studies regarding determinants of life expectancy. Factors related to both medical research and policy making has a great influence on life expectancy as factors like socioeconomic status, smoking cessation and overweight reduction plays a critical role in losing the Indigenous LE gap (Zhao, 2013).

Scotland, Northern Island and England are forming new government policies to reduce life expectancy gap in result of various studies indicating health inequalities which are main determinants of life expectancy and are directly associated with income inequalities and a huge gap is witnessed in life expectancy of people living in deprived areas (Health inequalities and the social determinants of health, 2012).

Numerous empirical studies have been already done in past exploring association of socioeconomic risk factors with life expectancy in European countries (Iacobuta and Cuza, 2012; Wang et al., 2013). Our study aims to explore several characteristics influencing life expectancy rate in Asian countries.

Increased per capita real income and higher expenditure on health has results in to a longer life expectancy according to a study made in Bangladesh (Mahumud et al., 2013) which shows that an average 8 days can be increased in a life expectancy by increasing one unit of per person Health Expenditure Per Capita (HEPC) and an increase of 33 days is expected by one US Dollar (USD) increment in GDP per capita.

MATERIAL AND METHODS

Data Collection

Data are noted at country level for 40 Asian countries. The following countries of Asia are not included in our study due to non-availability of data of these countries: Afghanistan, Brunel, Darussalam, East Timer (Timer Lester), Macau, Maldives, Palestine, Singapore and Taiwan. Data of Life Expectancy at Birth, Population Growth, Health Expenditures, Gross National Income, Crude Birth Rate, Crude Death Rate, Infant Mortality Rate and Population Density were obtained from World Bank website and Education, Gender Equality, Income Distribution, Good Governance and Healthy Life data were obtained from webpage of Sustainable Society Foundation, Netherlands. The reference year for this study is 2012.

Principal Component Analysis

According to (Jolliffe, 2002) explanation about the main idea of the PC transformation, PCA is used to retain few (<p) derived variables preserving most of the information provided by the variance of the p random variables. This linear transform has been widely adopted in data analysis and compression (Banerjee, 2012).

Let X be a vector of p random variables $X' = [X_1, X_2, \ldots, X_p]$ having the covariance matrix $\sum$ with eigenvalues $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_p \geq 0$.
Let the element of X has the following linear combinations
$Y_j = \alpha'_j X = \alpha_{j1}X_1 + \alpha_{j2}X_2 + \ldots + \alpha_{jp}X_p = \sum_{k=1}^{p} \alpha jkXk \ , \ j = 1,2, \ldots, p$
With a vector of p components $\alpha_{j1}, \alpha_{j2}, \ldots, \alpha_{jp}$.
Then $Var(Y_{j,}) = \alpha'_j \sum \alpha_j$ $\quad j = 1,2, \ldots, p$ (1.1)
$Cov(Y_{j,} Y_k) = \alpha'_j \sum \alpha_k$ $\quad j = 1,2, \ldots, p$ (1.2)

The PCs are those uncorrelated linear combinations $Y_1, Y_2, \ldots, Yp$ whose variances in (1.1) are as large as possible (Richard and Dean, 2001).
Emphasis on the variances is given in finding the PCs. First of all we look for a linear combination with maximum variance, such that
$\alpha'_1 X = \alpha_{11}X_1 + \alpha_{12}X_2 + \ldots + \alpha_{1p}X_p = \sum_{k=1}^{p} \alpha 1kXk$

Next, we look for a linear combination α'_2 X uncorrelated with α'_1 X having maximum variance, and so on, at the end we reach at kth stage of linear combination $\alpha'_k X$ having maximum variance and also being uncorrelated with α'_1 X, α'_2 X,…, α'_{k-1} X. The kth PC is kth derived variable α'_k X. Although upto p PCs could be derived but we restrict our findings till the qth stage ($q \leq p$) when most of the variation in X have been accounted for by q PCs.
Given $Var(Y_{j,}) = \alpha'_j \sum \alpha_j = \lambda_j$ $\quad j = 1,2, \ldots, p$
is the variance of PC which is equal to the corresponding eigen value
The total variance of PCs is considered as the total variance in a data set, which is given below

$$\sigma_{11} + \sigma_{22} + \ldots + \sigma_{pp} = \sum_{j=1}^{p} Var(Xj) = \lambda 1 + \lambda 2 + \ldots + \lambda p = \sum_{j=1}^{p} Var(Xj)$$

By standardizing the variables $X' = [X_1, X_2, \ldots, X_p]$ of similar scale with mean zero and unit standard deviation, we have the following corresponding standardized variables

$$Z = [Zj = \frac{(Xj - \mu j)}{\sigma jj}] \quad j = 1,2, \ldots, p$$

In matrix $Z = (V^{1/2})^{-1} (X - \mu)$
where $V^{1/2}$ is the diagonal standard deviation matrix having following properties
$E(Z) = 0$
$Cov(Z) = \rho$

The eigenvectors of the correlation matrix ρ of X will provide the PCs of Z, having all the properties of X by referring Y_j to the jth PC and (λ_j, α_j) to eigenvalue – eigenvector pair.
Now
The jth PC of the standardized variables Z' = [$z_1, z_2, \ldots, z_p$] can be shown as below

$$Y_j = \alpha'_j Z = \alpha'_j (V^{1/2})^{-1} (X - \mu),$$

Such that

$$\sum_{j=1}^{p} \text{Var}(Xj) = \sum_{j=1}^{p} \text{Var}(Zj) = p \qquad j = 1,2,\ldots,p$$

Having the following he eigenvalue- eigenvector pairs for ρ

$$(\lambda_1, \alpha_1), (\lambda_2, \alpha_2), \ldots, (\lambda_j, \alpha_j) \text{ with } \lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_p \geq 0$$

Interpretations of outcomes of Principal Component Analysis

The loading or the eigenvector $\alpha j = \alpha_1, \alpha_2, \ldots, \alpha_p$ shows the importance of the variable for a given PC. The eigenvector with the highest eigenvalue is the most dominant principle component of the dataset (PC_1). It expresses the most significant relationship between the data dimensions (Jeong et al., 2008). The type of crime components can be found by analyzing the positive and negative coefficients in subsequent components (Rencher, 2002). The information about the weights of original variables when calculating each PC can be found in loading matrix which shows association between PC and original variable (Fang, 2011).

The proportion of variance:
The best explanation of the original variables is obtained by the proportion of variance which is given below

$$\Psi q = \frac{\sum_{j=1}^{p} \lambda j}{P} = \frac{\sum_{j=1}^{p} \text{Var (Zj)}}{P}$$

A useful criterion for determining the number of components to be retained in the analysis is called cumulative proportion of explained variance. A good graphical representation of the ability of the PCs to explain the variation in the data is a scree plot (Cattell, 1966).

RESULTS AND DISCUSSION

Table 1: The correlation coefficient between life expectancy at birth and analyzed variables

Sr. No.	Independent Variables	r	Sig.
1	Population Growth (PG)	0.107	0.511
2	Health Expenditures (HE)	0.702	0.000
3	Gross National Income (GNI)	0.629	0.000
4	Crude Birth Rate (CBR)	-0.556	0.000
5	Crude Death Rate (CDR)	-0.220	0.173
6	Infant Mortality Rate (IMR)	-0.756	0.000
7	Population Density (PD)	0.145	0.372
8	Education (E)	0.325	0.041
9	Gender Equality (GE)	0.050	0.761
10	Income Distribution (ID)	0.008	0.961

11	Good Governance (GG)	0.534	0.000
12	Healthy Life (HL)	0.748	0.000

Different levels of associations can be seen in table 1. Bivariate correlation analysis between LE and 12 variables is presented using Pearson correlation coefficients (Jaba, 2002). Student t-test is used for testing statistical significance of Pearson's correlation coefficients. Strong significant positive correlation is found between life expectancy and health expenditures, gross national income, good governance and healthy life. However crude birth rate, crude death rate and infant mortality rate has negative relationship with LE which shows LE decreases as crude birth rate, crude death rate and infant mortality rate increases.

Regression Model

Backward regression is used and health expenditures, infant mortality rate and education are included in model to explain variation in LE. A significant explanation of variation in dependent variable using backward regression is presented below:

Table 2: Regression model summary

Model	**Unstandardized Coefficients**		**Standardized Coefficients**		
	Beta	**Std. Error**	**Beta**	**t**	**Sig.**
Constant	83.932	3.524		23.816	0.000
Health Expenditures	0.002	0.001	0.452	4.941	0.000
Infant Mortality Rate	-0.220	0.034	-0.719	-6.437	0.000
Education	-0.113	0.042	-0.282	-2.716	0.000
Kaiser-Meyer-Olkin Measure of Sampling Adequacy			0.683		
Bartlett's Test of Sphericity	Approx. Chi-Square		265.818		
df			66		
Sig.			0.000		

Variable health expenditures has positive sign with its coefficient indication positive effect of health expenditures on LE whereas remaining two factors has negative effect on LE. About 76% of the total variation (R^2= 0.768) in dependent variable is explained by this model. In order to verify adequacy of data that it is suitable for factor analysis or not, preliminary analysis is performed using SPSS. The null hypothesis of variables in correlation matrix of population are uncorrelated is tested using Bartlett's test of Sphericity. Whereas degree of prediction of each variable with help of other variables is achieved by using Kaiser-Meyer-Olkin measure of sampling adequacy. From table 2 it can also be seen that null hypothesis of variables are uncorrelated is rejected at Bartlett's test of Sphericity using chi-square test statistic, Sig.= 0.00 which is less than conventional value 0.05. Also Kaiser-Meyer-Olkin measure of sampling adequacy (0.683) shows value higher than 0.5 which indicate that solution through factor analysis is acceptable.

Table 3: Rotated component matrix[a]

Sr. No.	Variables	Factor 1	Factor 2	Factor 3	Factor 4	Communality
1	Population Growth	0.091	0.834	0.261	-0.121	0.786
2	Health Expenditures	0.729	0.015	0.236	0.173	0.618
3	Gross National Income	0.613	0.299	0.572	0.122	0.806
4	Crude Birth Rate	-0.648	0.409	-0.322	-0.205	0.732
5	Crude Death Rate	-0.147	-0.850	0.021	-0.080	0.751
6	Infant Mortality Rate	-0.872	0.005	-1.44E-5	0.099	0.771
7	Population Density	0.104	-0.079	-0.170	0.923	0.898
8	Education	0.723	-0.170	-0.428	-0.381	0.879
9	Gender Equality	0.080	-0.704	0.044	0.037	0.506
10	Income Distribution	0.009	0.006	0.840	-0.215	0.752
11	Good Governance	0.848	0.207	0.029	0.052	0.766
12	Healthy Life	0.934	0.076	0.076	0.097	0.890
Eigen Values		4.249	2.254	1.479	1.175	
% of Variance		35.408	18.780	12.324	9.789	
Cumulative %		35.408	54.188	66.512	76.301	

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation converged in 6 iterations.

Table-3 displays eigenvalues, percent and cumulative percent of explained variance which helps us to decide how many factors (or components) are being retained. As rule of thumb factors having eigenvalues greater than one are sufficient to be retained (Malhotra and Dash, 2011). However by considering scree plot in figure 1, it is reasonable to retain first four components as fourth eigenvalue is 1.175. Thus by retaining first four PCs up to 76.30% of the variability in the total data set can reasonably be explained.

We can see that factor-1 is highly negatively associated to IMR and have high positive association with HL and GG followed by HE and E. Factor-2 is highly positively correlated with PG and high negative association is found between factor-2 and CDR followed by GE, E and PD also confirmed by regression and correlation analysis. Factor-3 is highly positively correlated to ID. A very strong positive association is detected between factor-4 and PD. Discussion above reveals that overall HL, GG, ID, GNI, HE and PG has positive impact on economic developments in Asia. By inspecting squared multiple correlation coefficient (or R^2) of each variable with all other variables in communality column, we can see only one small value of R^2 for variable gender inequality which indicates that this variable is unique and not correlated with other variables. All remaining variables are moderately correlated with each other.

ISSN 2309-3218

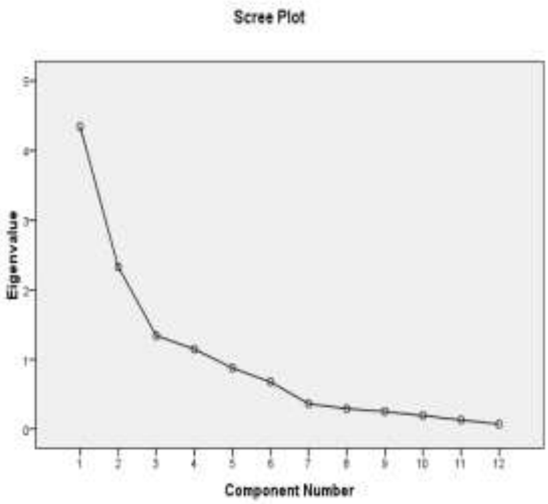

Fig. 1

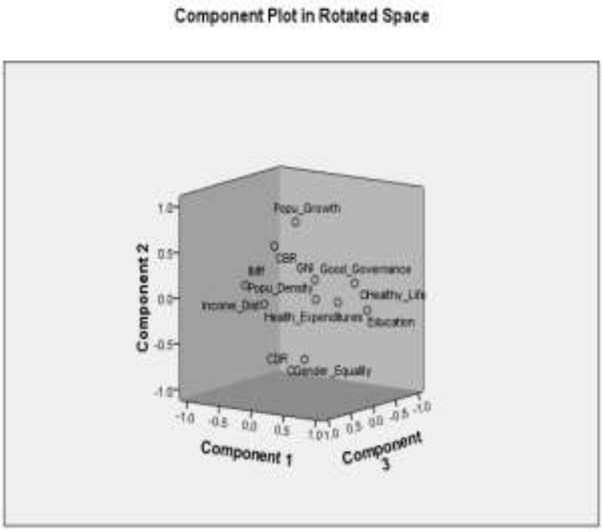

Fig. 2

Three factor axis graphical view is presented in figure 2, which provides us graphical representation of correlation among various variables (Gagea and Iacobuta, 2010). CDR and CBR are located in negative quadrant of 1st factorial axis whereas HL and GG are in positive quadrant. CDR and PG explains 2nd factorial axis in negative and positive quadrant respectively. Whereas 3rd factorial axis is explained by E and ID in negative and positive quadrant respectively.

Table 4: Countries wise factor scores and ranks of countries in Asia

Countries	Factor Score 1	Rank	Factor Score 2	Rank	Factor Score 3	Rank	Factor Score 4	Rank
Armenia	0.31016	16	-1.12663	35	-0.63064	32	-0.32273	23
Azerbaijan	-0.38664	27	-0.27103	25	-0.14960	20	0.13311	14
Bangladesh	-1.00770	33	-0.08505	21	-0.59088	30	4.58862	1
Bhutan	-0.48419	28	0.20430	14	0.67208	6	-0.67874	33
Cambodia	-1.05131	34	0.48164	11	0.10633	14	-0.33433	25
China	0.37732	13	-0.96228	33	0.38759	11	-0.05102	16
Georgia	0.36576	14	-1.68041	39	0.72556	5	-0.64871	31
India	-0.72450	30	0.05129	20	-0.44233	27	1.12643	4
Indonesia	-0.05959	23	-0.37274	27	-0.50076	28	-0.25755	20
Iran	-0.00991	22	0.21213	13	-0.29582	26	-0.71431	35
Iraq	-1.11712	36	1.01034	7	-0.71559	33	-0.31324	21
Israel	1.83439	2	0.11545	17	-0.07743	18	0.31137	12
Japan	2.96318	1	-1.23988	38	0.01144	16	1.35214	2
Jordan	0.12154	18	1.12385	6	-0.93622	37	-0.65169	32
Kazakhstan	0.24868	17	-0.59933	29	-0.96301	38	-1.25318	40
Korea, North	-0.21204	25	-1.15317	36	0.94309	4	0.84395	6
Korea, South	1.65476	3	-0.08732	22	-1.92056	40	0.56800	10
Kuwait	1.00155	6	1.35469	4	-1.11267	39	-0.23534	19
Laos	-1.27431	37	0.06707	18	-0.16593	22	-0.31546	22
Lebanon	0.54274	9	0.16643	15	-0.61523	31	0.84774	5
Malaysia	0.52929	10	-0.12844	23	0.99078	3	-0.60767	30
Mongolia	0.05067	20	-0.30998	26	-0.88567	36	-1.18065	38

Myanmar	-1.51741	39	-1.22347	37	2.10735	2	-0.40674	27
Nepal	-1.00380	32	0.06690	19	0.66665	7	0.19243	13
Oman	0.73385	7	2.62471	1	-0.00795	17	-0.81054	37
Pakistan	-1.86174	40	0.51919	10	0.02101	15	1.23599	3
Philippines	-0.06737	24	-0.68723	31	-0.10858	19	0.02975	15
Qatar	1.27440	4	2.12904	2	4.30274	1	0.59859	9
Russia	0.50179	11	-2.08047	40	0.50777	10	-1.22222	39
Saudi Arabia	0.62088	8	0.88125	8	-0.50852	29	-0.75143	36
Sri Lanka	0.09475	19	-0.74669	32	-0.19071	23	0.65095	8
Syria	-0.23003	26	0.88072	9	-0.83660	35	-0.18469	18
Tajikistan	-0.98250	31	0.26704	12	-0.82212	34	-0.71065	34
Thailand	0.34371	15	-1.11834	34	0.36305	12	-0.16749	17
Turkey	0.48635	12	0.15622	16	0.51218	9	-0.51338	28
Turkmenistan	-1.07554	35	-0.65514	30	0.52618	8	-0.58180	29
U. Arab Emirates	1.14597	5	1.57615	3	0.29019	13	0.45829	11
Uzbekistan	-0.64734	29	-0.15128	24	-0.21296	24	-0.36039	26
Vietnam	0.02288	21	-0.38917	28	-0.28532	25	0.66677	7
Yemen	-1.51159	38	1.17960	5	-0.15889	21	-0.33016	24

Countries wise factor scores and ranks of countries in Asia based on principal components 1, 2, 3 and 4 are presented in table 4. Differences between LE at birth and indicators variables in Asian countries according to their ranks in factors framework reveals following information. In order to find out which country was on the top of hierarchy in terms of the variables, the highest value of the score for a particular country highly loaded on the factor is inspected (Singariya, 2013). It is observed that value of first factor score is varied from (-1.86) to (2.96). On the basis of standardized variables countries having positive scores on the first component can be interoperated as above 'average' level of the countries in Asia or we can say that these countries are better as compared to other countries of Asia. Out of forty countries only nineteen countries are considered above average and remaining are considered as below average. Major influencing indicators are GG and HL. Less surprisingly Japan stood at top with a score of (2.96) followed by Israel with score of (1.83) while Pakistan (-1.86) preceded by Myanmar (-1.52) were at the lowest rank of first factor. Surprisingly wide gap between Japan and Pakistan is inspected in terms of IMR and CDR.

Inspecting 2nd factor score which varies from (2.62) to (-2.08) shows Oman and Qatar stood at highest rank and, Russia and Georgia at lowest rank. Out of 40 countries nineteen countries are found to have above average in PG and remaining countries falling in category of below average level countries of Asia due to CDR and GE. Inspection of 3rd factor score reveals variation of factor scores (4.30) to (-1.92) showing Qatar at highest rank of above average countries of Asia in terms of PD and South Korea at lowest rank in below average countries in ID. Study of factor 4 scores shows variation of factor scores (4.58) to (-1.92). Bangladesh stood at highest rank of above average countries in terms of high PD and Health facilities while Kazakhstan at the lowest rank in below average countries category mainly due to low ID and Education facilities.

CONCLUSIONS

Association of several socioeconomic factors with life expectancy at birth and the influencing factors in forty countries of Asia has been explored in this paper. Less surprisingly the results and discussions obtained in this paper are in agreement with previous research. A close relationship between several socioeconomic variables and LE at birth is found. The conclusions and suggestions are as follow:

Strong significant positive correlation is found between life expectancy and health expenditures, gross national income, healthy life and good governance. However crude birth rate, crude death rate and infant mortality rate has negative relationship with LE which shows LE decreases as crude birth rate, crude death rate and infant mortality rate increases.

Wide disparities among countries in health care infrastructural facilities, income distribution and education facilities have been witnessed in above analysis. Pakistan is turned out to be an outlier in first factor score. Countries like Pakistan and Myanmar are required to have more emphasis on strategies to reduce infant mortality rate and crude death rate in order to increase longevity and health issues must be properly addressed (Kambiz et al., 2011). Whereas countries like South Korea and Kazakhstan required strengthen in policy making for proper income distribution and education opportunities as well (Mahfuz, 2008; Kambiz et al., 2011).

REFERENCES

Banerjee, A. (2012). Impact of Principal Component Analysis in the Application of Image Processing. *International Journal of Advanced Research in Computer Science and Software Engineering,* 2(1): http://www.ijarcsse.com/docs/papers/january2012/V2I1052.pdf visited on 10/01/2015.

Cattell, R.B. (1966). The Scree Test for the number of Factor. *Multivariate Behavioral Research,* 1: 245-276.

Fang, Y. (2011). Multivariate Methods Analysis of Crime Data in Los Angeles Communities. *University of California,* Los Angeles.

Gagea, M. & Iacobuta, A.O. (2010). The Role of individual Values in Personal Development, in Analele Ştiinţifice ale Univ. Al.I.Cuza Iaşi, Secţiunea Ştiinţe Economice, Tom LVII, 451-463.

Health inequalities and the social determinants of health, (2012). *Royal College of Nursing, Policy and international department 20 Cavendish Square London,* Policy Briefing No. 01/12.

Iacobuta, A.O. & Cuza, A.I. (2012). Socio-economic determinants of life expectancy: a cross-country analysis. *Journal of Academy of Business and Economics*, 12(2): 76-82.

Jaba, E. (2002). Statistica, Editia a treia, Editura Economica, Bucuresti.

Jeong, D.H., Ziemkiewicz, C., Ribarsky, W. & Chang, R. (2008). Understanding Principal Component Analysis Using a Visual Analytics Tool. *Charlotte Visualization Center,* UNC Charlotte.

Jolliffe, I.T. (2002). Principal Component Analysis 2nd edn. New York: Springer-verlag.

Kambiz, P., Gollu, R.B, Gashti, H.P. & Shahrivar, R.B. (2011). Studying the relationship between health and economic growth in OIC member states. *Interdisciplinary Journal of Contemporary Research in Business,* 3(8): 1041-1054.

Lei, H., Li, L., Liu, X. & Mao, A. (2009). Quantitative study on socioeconomic determinants of life expectancy in Beijing, China. *Journal of Evidence- Based Medicine,* 2: 92-98.

Mahfuz, K. (2008). Determinants of Life Expectancy in Developing Countries. *The Journal of Developing Areas*, 41(2): 185-204.

Mahumud, R.A., Rawal, L.B., Hossain, G., Hossain, R. & Isam, N. (2013). Impact of Life

Expectancy on Economics Growth and Health care expenditures: A Case of Bangladesh. *Universal Journal of Public Health*, 1(4): 180-186.

Malhotra, N. & Dash S. (2011). Marketing Research-An applied orientation, Sixth Edition, *Pearson Publication.*

Rencher, A.C. (2002). Methods Multivariate Analysis, 2nd edn. New York: *Wiley*.

Richard, A.J. & Dean, W.W. (2001). Applied Multivariate Statistical Analysis, 3rd edn. New Delhi: *Prentice Hall.*

Sen, A. (1999). Development as Freedom. Alfred A. Knopf, New York.

Singariya, M.R. (2013). Principal Component Analysis of socioeconomic factors and their association with life expectancy in India. *The Journal of Economics,* Photon 113: 163-172.

Wang, H., Schumacher, A.E., Levitz, C.E., Mokdad, A.H. & Murray, C.J.L. (2013). Left behind: widening disparities for males and females in US country life expectancy, 1985-2010. *Population Health Matrices,* 11(8): 1-15.

Yavari, K. & Mehrnoosh, M. (2006). Determinants Life Expectancy: A Cross-Country Analysis. *Iranian Economic Review,* 11(15): 131-142.

Zhao, Y.**,** Wright**,** J., Begg, S. & Guthridge, S. (2013)**.** Decomposing Indigenous life expectancy gap by risk factors: A life table analysis. *Population Health Metrics*. 11(1): http://www.pophealthmetrics.com/content/11/1/1 visited on 10/01/2015.