Administrative Development: A Journal of HIPA, Shimla. Volume VII (2), 2020.

AN ASSESSMENT OF HEALTH STATUS ACROSS INDIAN STATES: APPLICATION OF LOGISTIC REGRESSION AND PRINCIPAL COMPONENT ANALYSIS

Sumit Dey [*] & Sondeep Dey[†]

Abstract

Linkage between health status of population and economic growth of a country is a well established fact. In this context, this paper is an attempt to assess the health status of Indian population through a self developed health status index (HSI). The HSI has been developed by using principal component analysis. Further to identify the determinants of health status a logistic regression has been estimated. The HSI has been calculated only for 19 states of India and for the year 2016.The findings of the study shows that health status of Indian population is not at all satisfactory.

Keywords: *Health Status, Economic growth, Principal component Analysis, Logistic Regression.*

1.1 Introduction

Human Capital has emerged to be an important sector of Investment for the countries all over the world because it is a key driver of growth for an economy. People with longer life are expected to save more than individuals with poor health. As a result, higher savings will contribute more in the national output leading to more investment prospects (Ullah , Malik & Hassan,2019) In other words good health status ensures the macro-economic equilibrium at higher level of income and employment. Weil (2001) discovered in his study that a one percentage point increase in adult

[*] *Assistant Professor, Department of Economics, Bholanath College, e-mail :dey.sumit421@gmail.com*

[†] *Research Scholar, Department of Commerce, Dibrugarh University, e-mail: sondeepdey79@gmail.com*

survival rates translates into a 1.68 percent increase in labour productivity. Again loss in labour productivity due to accidents and health problems was estimated to a figure equivalent to nearly 4.2% of the Gross Domestic Product of Spain in 2005(Moreno, 2011). A decline in life expectancy and ill health will retard economic growth by lessening productivity of labour. (Ullah et al.,2019). Countries with the weakest conditions of health have a much harder time in achieving sustained growth than do countries with better conditions of health (Subramanian, Belli & Kawachi, 2002). All these findings clearly exhibit the positive relationship between good health status and economic growth, where both reinforce each other.

While it comes to health status, inequitable distribution can be noticed across the countries. A child born in 1999 in one of the 24 healthiest countries of the world can expect to live for more than 70 healthy years. By contrast, a child born in one of the 51 least healthy countries can expect to live less than 50 years (Bloom e& Canning 2001).

India, the world's fifth largest economy is the home to 1.3 billion people consisting a huge labour force. In the last couple of decades India has achieved tremendous progress in various sectors, but the bitter truth is that the overall performance in the social sector and health in particular is far from satisfactory. If we notice the health status prevailing in the country, some major health indicators show a disheartening result. Though the Maternal mortality rate(MMR)in India has gradually decreased to 145(2017) from 370 in the year 2000, it is quite high in comparison to its neighbouring countries such as Sri Lanka, China and Thailand with contributory performance in MMR rate which are 36,29 and 37 respectively in 2017(World Bank,2017). Again if we look at the TB cases prevailing in the country, India has been ranked the first position with 2.8 million new cases in the year 2015 alone (Tbfacts.org.,2015). Though India shares 20% of the global burden of diseases, its contribution to global healthcare infrastructure is highly inadequate (KPMG – FICCI, 2015) Therefore, it is of utmost importance that health sector of India should be encouraged to grow more rapidly and efficiently in order to achieve a sustainedrate of growth.

In India, the problem of health becomes more complex due to the intra-state and regional differences in health outcome and infrastructure. For instance, under 5 mortality in Chhattisgarh or Uttar Pradesh is 6–7 times more than in Kerala, where the proportion of educated mothers is almost double (Salve & Yadavar, 2017).

Health status of the population of a country is reflected by a number of indicators such as infant mortality rate, total fertility rate, institutional delivery, sanitisation coverage etc. Thus, the nature of the indicator may be positive or negative depending on its characteristics. For positive indicators (e.g. Sanitisation coverage) higher value denotes better performance whereas in case of negative indicators (e.g. Infant mortality rate) lower the value, better the performance. Thus, to assess the health status of a population all these indicators are needed to be considered simultaneously. Accordingly, in this paper an attempt has been made to develop a composite health index by incorporating the positive and negative indicators. The composite index has been developed by using principal component method. As the necessary data was available for only19major states of India and up to year 2016, the index has been prepared accordingly, considering the reference year 2016. Further, to identify the statistically significant determinants of health status through a logistic regression has been estimated.

A Composite Health Index has already been prepared by NITI Aayog, Government of India in the year 2018 incorporating 23 indicators in the domains of Health Outcomes, Governance and Information, and Key Inputs/Processes where utmost weightage has been given to Health Outcome. But in this paper a health status index has been developed, considering the dimensions of health outcome only. Health Outcome has been considered because health outcome plays a significant role in determining the level of disparity among the population base and provides guidance to clinicians on taking action as well as it reveals the areas in which interventions could improve care.

1.2 Rationale of the Study

Though the problem of poor health and its consequences is an international problem but the problem is more acute for developing and less developed countries. Poor health conditions in these countries are actually creating a vicious cycle of economic problems, as represented by the following chart-

Chart-1.1: Vicious Cycle of Unhealthy Population

Source: Author's Self formation

A country having a healthy population base will ensure that the labour force working within the country are also in good health condition which will in turn increase the level of productivity prevailing among them whereas an opposite picture is noticed in countries consisting unhealthy population. Weak and inefficient labour force means greater amount of absenteeism in their respective field of work which will drag down their income drastically. Again, it is quite obvious that lower income will lead to lower volume of savings and investment and hence, the rate of economic growth. Lower rate of economic growth, in turn results in poverty and unemployment.

Again, the association between poor health and poverty is well established by literature. Thus, from the above chart it can be concluded that the unhealthy population traps the developing and less developed countries in a cycle of critical economic problems and hence, low rate of economic growth. Hence to break this vicious cycle, the health sector of an economy should be prioritised.

From the above discussion, it becomes quite rationale to assess the health status of Indian population as well as to identify the determinants of health status.

1.3 Objectives of the study

1. To assess the health status of Indian population in terms of a self-developed health status index.
2. To identify the factors influencing the health status of Indian Population.

1.4 Materials & Methods

Data

For the purpose of carrying out the study, secondary data have been used. The required data have been collected from NITI Aayog's report entitled "Healthy States, Progressive India", Central Bureau of Health Intelligence's report entitled " National Health Profile,2018", official websites of Press Information Bureau, Government of India, Reserve Bank Of India, Open Government Data Platform, India and World Health Organisation. Certain other information have been collected from various books, literatures, e-journals, articles etc.

Sample size

The study consists of data set from 19 major states out of the 29 states in the country as all the required data were available for only these states whereas for the rest of the states and UTs required data were not available.

Year of reference

2016 has been considered the reference year while constructing the health status index as the necessary data was available till 2016 only.

1.5 Construction of Health Status Index: Application of Principal Component Analysis

The main objective of this paper is to construct a health status index for 19 states of India. Accordingly, the first step involves selection of variables. However, it is worth mentioning that health status of a population is reflected by both positive and negative indicators. Positive indicators refer to those indicators for which higher values are always preferred (e.g Sanitisation coverage). Negative indicators, on the other hand refer to those indicators for which lower values are always preferred (e.g. Infant Mortality Rate).Thus the proposed health outcome index is a composite index of positive as well as negative indicator. List of the indicators considered in the study is given in Appendix-I.

The scaled value (S_i) for the i^{th} positive indicator, with data value as X_i, is calculated as follows.

$$S_i = \frac{X_i - Minimum\ Value}{Maximum Value - Minimum\ Value}$$

Similarly the scaled value (S_i) for the i^{th} Negative indicator, with data value Xi, is calculated as follows -

$$S_i = \frac{Maximum\ Value - X_i}{Maximum\ Value - Minimum Value}$$

The above two equations ensure the value of the dimension indices to remain in the range of 0 to 1.

After scaling the dimension indicators, the next step is to select proper weights for them. To assign proper weights to the dimension indices, principal component analysis has been extensively used in literature.

Principal component method is essentially a dimension reduction technique that creates a new set of orthogonal or uncorrelated variables, known as principal component from a given set of possibly correlated variables. In mathematical terms, from an initial set of n correlated variables, PCA creates uncorrelated indices or components, where each component is a linear weighted combination of the initial variables **(Vyas and Kumarayake, 2006)**. One of the important features of PCA method is that the first principal component (P1) has the maximum variance and hence captures more information than the second principal component and so on. Hence, the factor scores from the first principal component have been used to construct different indices by many **researchers (Córdova Abby.2009, Vyass and Kumaranayake 2006)**

In our case the first principal component explains 42.32% of the total variation of the data, hence factor scores from the first principal component has been used as weights to create a dependent variable (Yi), representing the health outcome index for each state which has a mean equal to zero(0) and standard deviation equal to one (1). However, a positive factor score associated with a variable indicates higher value of Yi and vice –versa. Finally, the composite index derived through PCA can be expressed as under.

$$Y_i = \alpha_1 X_1 + \alpha_2 X_2 + \ldots\ldots\ldots\ldots + \alpha_k X_k \quad \textit{Eq-I}$$

Where, Yi represents the Health outcome index value for 19 states of India, α represents the weights for k^{th}sub indices (11 dimension indices) derived from first principal component, such that sum of the squares of weights is equal to one ($\alpha_1^2 + \alpha_2^2 + \ldots\ldots\ldots\alpha_k^2 = 1$). In table -1, a statistical summary of the calculated Health status Index has been presented.

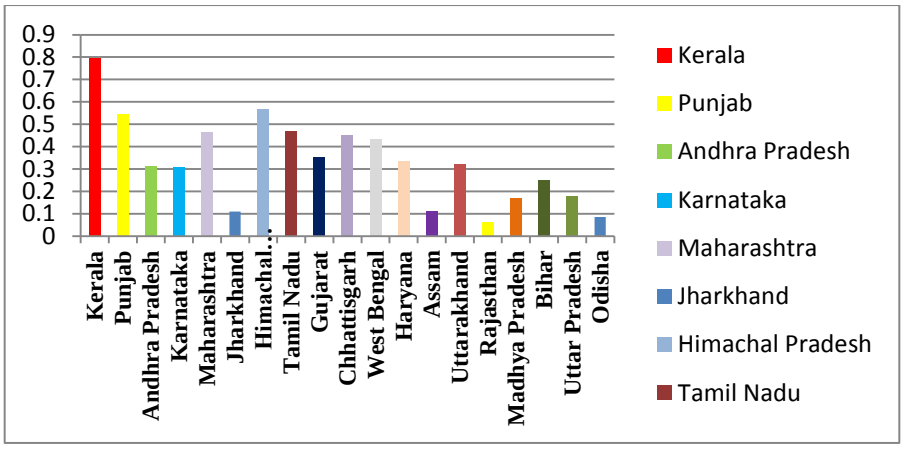
Table-1: Statistical Summary of the Health Status Index

Health Outcome Index	Value
Mean	0.339
Standard Deviation	0.041
Maximum	0.796
Minimum	0.109

Source: Author's self- calculation based on secondary data survey

The above table depicts a statistical summary of the calculated health outcome index value for 19 states of India. The mean index value being very less at 0.339, indicates the health outcome status across the sample states of India, is far from satisfactory. The standard deviation value of the index being very less implies that health status of Indian states is more or less the same. The health status index value for all the 19 states is given in appendix II and plotted in chart-2.

Chart-2: Health Status of the Sample Indian States on the basis of HSI



Source: Self Calculation

1.6 Identification of the Determinants of Health Status of Indian States: Application of Logit Model

OLS regression technique is often inadequate in the study of bounded dependent variables and may produce predicted values that lie outside the

unit interval and the effects of the explanatory variable tend to be non-linear **(Gallani et. al.,2015).** Since the health Status index(HSI), the dependent variable of the regression model is bounded within the range of 0 to 1, therefore a nonlinear formation of the regression model sounds more appealing **(Maity. et al., 2014).**

Therefore to avoid the unboundness problem **(Ramanathan, 2008),** a logistic regression model has been considered to identify the determinants of health status. The regression model is specified as under

$$HSI_i = \frac{1}{1+e^{-(\alpha+\sum_{j=1}^{6}\beta_j X_j+U_i)}} \qquad (iv)$$

After some mathematical manipulation, equation (iv) can be rewritten as-

$$HSI_i = \frac{e^{(\alpha+\sum_{j=1}^{6}\beta_j X_j+U_i}}{1+e^{(\alpha+\sum_{j=1}^{6}\beta_j X_j+U_i)}}$$

Or, $$\frac{HSI_i}{1-HSI_i} = e^{\alpha+\sum_{j=1}^{6}\beta_j X_j+U_i}$$

Or, $$\log\left(\frac{HSI_i}{1-HSI_i}\right) = \alpha+\sum_{j=1}^{6}\beta_j X_j+U_i$$

Or $$Y_i = \alpha+\sum_{j=1}^{6}\beta_j X_j+U_i \qquad \left[Y_i = \log\left(\frac{HSI_i}{1-HSI_i}\right)\right]$$

Incorporating the independent variable the regression model can be finally expressed as under

$$\mathbf{Y_i = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + U_i} \qquad ....(V)$$

Where Yi= Log odd ratio of health status index of the i^{th} state.

α= Constant term.
β_s = Coefficients of the explanatory variables
Ui= The error term

The description and justification of the explanatory variables, incorporated in equation V is described in table -2.

Table-2: Summary of Hypothesis, Justification and Expected effect of Explanatory Variables

Independent Variables	Hypothesis	Justification
Density of population (X_{1i})	*H_{01}: There is no statistically significant relationship between density of population and health status.*	Higher population implies lesser per capita availability of health services and hence low health status.
Literacy Rate (X_{2i})	*H_{02}: There is no statistically significant relationship between literacy rate and health status.*	Health and Education are interrelated component. Education is supposed to make people realize the importance of good health and hygiene
Poverty Rate (X_{3i})	*H_{03}: There is no statistically significant relationship between poverty and health status.*	Poor people could not take proper care of their health and hygiene, hence poverty may be a negative determinant of health status
Health Expenditure as a percentage of SGDP(X_{4i})	*H_{04}: There is no statistically significant relationship between govt. health expenditure and health status.*	Higher level of health expenditure is expected to improve the health status of people.
Per capita income(X_{5i})	*H_{05}: There is no statistically significant relationship between per-capita income and health status*	At higher level of income people become more conscious about their health status and it also enhances the access to better medical facilities. Therefore higher level of per capita income may have a positive impact on health status.
Percentage of rural population (X_{6i})	*H_{05}: There is no statistically significant relationship between percentage of rural population and health status*	In India, adequate medical facilities have not been extended to rural areas yet. Therefore it is expected that higher the percentage of rural population lower will be the health status.

Table-3: Determinants of Health Status: Estimated Result of Logit Model

Variables	Coefficient	Marginal Effect	*p*
Intercept (α)	-1.229	--	0.307
Density of population *(X_{1i})*	-0.0021	-0.89	0.094
Literacy Rate *(X_{2i})*	0.015**	0.140	0.021
Poverty Rate *(X_{3i})*	0.00294	0.0047	0.787
Health Expenditure as a percentage of SGDP *(X_{4i})*	0.0347**	0.112	0.035
Per capita income *(X_{5i})*	0.0347**	0.439	0.021
Percentage of rural population *(X_{6i})*	-0.0371	-0.000260	0.956
R^2(McFadden) 0.32			

Source: Author's own calculation based on secondary data

Note: CensReg package of R statistical software has been applied to estimate the result ** significant at 5% level.

The estimated result of the regression model identifies the statistically significant determinants of health status. Out of the six expected determinants, three determinants have been found statistically significant. Literacy rates, health expenditure as a percentage of state gross domestic product and per capita income have been found positively affecting the health status. The linkage between these variables and health status is already justified in table-2.

From the marginal effect column of the above table, it appears that 1 % increase in literacy rate and health expenditure will increase the health status index value by 0.14 and 0.11 units respectively. Similarly, the health index value will increase by 0.43 units with every unit increase in income

1.7 Conclusion

Considering the fact that health is one of the most important determinants of economic growth, authors in this paper have attempted to estimate the health status of people living across 19 major states of India, in terms of a

self-developed health status index. The mean index value being very less implies that health status of Indian states is not at all satisfactory and hence, utmost care is needed to improve the scenario. The second objective of the study requires identifying the determinants of health status. Accordingly, a logistic regression model has been estimated. The estimated result shows that literacy rate, health expenditure and per capita income of states are the significant determinants of health status. Hence it can be concluded that to improve the health status, govt should take proper measures to increase the literacy rate, per-capita income and health expenditure as a percentage of state gross domestic product.

References

Books, Journals & Articles:

1. Beaglehole. R, and Bonita. R.1998. Public Health at the Crossroads: which way forward? The Lancet, Volume 351.
2. Bloom. DE and Canning. D.2001.A new health opportunity. Development 44(1):36–43
3. Córdova. Abby. 2009. Methodological Note: Measuring Relative Wealth using Household Asset Indicators. Americas Barometer Insights: 2009 (No.6).
4. Gallani Susana. et al. 2015.Application of fractional Response Model to the study of Bounded Dependent Variables in Accounting Research.SSRN Electronic Journal
5. Kumar. A. and Das. K.C.2014. Drinking water and sanitation facility in India and its linkages with Diarrhea among children under five: Evidences from recent data.International journal of Humanities and Social Science Invention, volume 3, issue 4, pp 50-60
6. Maity et al. 2014.Socioeconomic Status of Kuki Tribal Women: A Case Study Churachandpur District, Manipur. American-Eurasian Journal of Scientific Research,9 (5): 120-128, 2014.

7. Moreno. J O.2012. Loss of labour productivity caused by disease and health problems: what is the magnitude of its effect on Spain's Economy? Eur J Health Econ, Pp 605-614
8. Ramnathan. R. 2008 "Introductory Econometrics", 5th Ed. New Delhi: Cengage development, pp: 258-260.
9. Salve. P. and Yadavar. S.2017. States with More Educated Women have Healthier Children. India Spend. Available from: http://www. indiaspend.com/cover-story/states-with-more-educated-women-havehealthier- children-65307
10. Subramanian. S. V. Belli, Paolo. and Ichiro, Kawachi.2002 The macroeconomic determinants of healthAnnu. Rev. Public Health, pp. 23:287–302
11. Ullah. S. Malik,MN.and Hassan. M.2019. Impact of Health on Labour Productivity:Empirical Evidence from Pakistan.European online journal of natural and socialsciences, volume 8, no. 1,139-147
12. Vyass. Seema. and Lilani,Kumaranayake.2006.Constructing Socioeconomic Status Indexes: How to Use Principal Component Analysis.Health Policy and Planning 21, no. 6 (2006): 459-68.
13. Weil. D. 2007. Accounting for the Effect of Health on Economic Growth.Quality Journal of Economics,1265-1306
14. Zodpey. SP. and Negandhi, PH.2018.Tracking India's Progress in Health Sector after 70 Years of Independence.Indian J Public Health;62:1-3.

Reports

15. Census of India. 2011. ABRIDGED LIFE TABLES- 2003-07 to 2006-10
16. Central Bureau of Health Intelligence.2018. National Health Profile. 13th issue
17. KPMG – FICCI.2015. Healthcare: The Neglected GDP Driver – Need for a Paradigm Shift.
18. NitiAayog.2019. Healthy States Progressive India- reports on the ranks of states and Union Territories
19. Open Government Data Platform India. 2011-12. State-wise Unemployment Rate (per 1000)

20. Press Information Bureau, Government Of India.2014. Cases of Cancer
21. Reserve Bank of India.2012-16. State-wise life expectancy
22. The World Bank.2017. Maternal Mortality Ratio(Modeled estimate per 100000 live births)
23. Tbfacts.org.2015. Reported cases of TB.

Appendix -I: Result of Principal Component Analysis

Variables	Nature of the Indicator	Mean	Standard Deviation	Factor Score Coefficients
Infant Mortality Rate	Negative	0.394	0.271	0.106
Total Fertility Rate	Negative	0.340	0.309	-0.260
Low birth weight	Negative	0.544	0.261	-0.072
No. of Respiratory Disease cases	Negative	0.661	0.253	-0.107
No. of Tuberculosis cases	Negative	0.528	0.250	-0.087
No. of Cancer cases	Negative	0.568	0.308	0.012
Sex Ratio at Birth	Positive	0.546	0.291	0.019
Full Immunization	Positive	0.703	0.253	0.037
Institutional Delivery	Positive	0.629	0.286	0.192
Sanitisation coverage	Positive	0.424	0.280	0.200
Life Expectancy at Birth	Positive	0.437	0.276	0.245

Source: Author's own calculation based on secondary data

Appendix-II: Health Status Index Value of Indian States

States	Health Status Index Value	Rank
Kerala	0.796	1
Himachal Pradesh	0.569	2
Punjab	0.544	3
Tamil Nadu	0.470	4
Maharashtra	0.463	5
Chhattisgarh	0.451	6
West Bengal	0.435	7
Gujarat	0.351	8
Haryana	0.336	9
Uttarakhand	0.323	10
Andhra Pradesh	0.312	11
Karnataka	0.308	12
Bihar	0.250	13
Uttar Pradesh	0.180	14
Madhya Pradesh	0.170	15
Assam	0.159	16
Jharkhand	0.113	17
Odisha	0.111	18
Rajasthan	0.109	19

Source: Author's own calculation based on secondary data