SSRG International Journal of Economics and Management Studies
ISSN: 2393-9125/ https://doi.org/10.14445/23939125/IJEMS-V9I6P103
Volume 9 Issue 6, 13-18, June 2022


*Original Article*

# Determinants of Human Development Index: An Econometric Analysis of Indian States

Mintu Kalita

*Department of Economics, Gauhati University Assam, India*

Received: 24 April 2022 Revised: 04 June 2022 Accepted: 09 June 2022 Published: 30 June 2022

***Abstract*** *- Instead of mere economic growth, the recent focus of economists and policymakers has been on human development. The human development index (HDI) captures the status of human development in a region/country. This paper attempts to study the variation of the HDI values of the Indian states/UTs. By applying logistic regression to data on independent variables such as road length, the number of educational institutes, and the number of hospital beds for 36 Indian states/UTs for the year 2015, we find that the variables kilometers of roads per square kilometers of area, and the number of hospital beds per 10,000 of the population have a statistically significant positive impact on HDI. Thus, states/UTs with comparatively lower HDI scores should focus on creating additional kilometres of roads per square kilometres of area, and making more hospital beds per 10,000 people available to foster human development.*

***Keywords*** *- Human development, Limited dependent variable, Logistic regression, The impact of infrastructure, Sub-national HDI.*

## 1. Introduction

The Human Development Index (HDI) was developed by United Nations Development Programme (UNDP) to measure and compare the status of human development among countries and regions of the world. The index is a composite index consisting of three primary dimensions of human development: income, health, and education.

The present study attempts to explain the variation in the status of human development, as measured by HDI, among Indian states and figure out the determinants of such variation. One of the important determinants of the level of income of an economy or region is the status of the region's physical infrastructure. Improved infrastructures such as roads, ports, etc., reduce inefficiency in movements of factors and goods, set in motion a conducive environment for private investment to grow, and enhance the quality of life of the people. Thus infrastructural scenario of a region can be linked to its level of income and hence further to the status of human development of that region.

Education is indispensable to human development. An educated workforce is more productive, skilled, and better aware of their health, rights, and social duties. The availability of educational institutes among the population of a region, apart from other factors, determines the accessibility to education. It also enables them to earn a few more years of schooling. Thus, the number of educational institutes per unit of a region's population is an indicator of the region's level of human development.

The third dimension of human development is health. Deteriorated health incapacitates workers from participating in the labour market, reducing their income and triggering a vicious circle of poverty. Accessibility to health facilities is a major issue in developing regions for most people. The availability of more hospital beds per unit of the population can ensure the delivery of medical care to a larger chunk of the population. Thus, the number of hospital beds per lakh of a region's population may affect the region's human development via the health dimension.

The paper is organised into 6 sections. Section 2 states the objectives. Section 3 briefly reviews the literature on determinants of the Human Development Index across various countries/regions of the world. Section 4 describes the data and methodology adopted for the present study. Section 5 discusses the results obtained from the econometric analysis. Finally, section 6 concludes the study and suggests possible policy measures.

## 2. Objectives

The present study aims to formulate an econometric model for explaining the variation in the HDI across Indian states in the year 2015. The aim is to study the variation in HDI values among Indian states using cross-sectional data.

## 3. Literature Review

There are several existing studies on the determinants of HDI. However, studies about the factors causing variation in HDI in Indian states are limited. Shah(2019), doing a cross-country empirical analysis, finds GDP per capita, literacy rate, life expectancy at birth, Gini index, fertility rate, and $CO_2$ emission as significant factors in affecting the Human Development Index of a country. Khan et al. (2019), considering data on the Human Development Index(HDI), information and communication technology(ICT), economic growth, urbanization, and Foreign Direct Investment(FDI) for the period from 1990 to 2014, find that ICT and economic growth have a positive and significant effect on human development and urbanization, FDI and trade harm human development in case of Pakistan. A nation's institutional quality is also supposed to impact the Human Development Index. Binder & Georgios (2011), considering panel data of 87 countries from 1970 to 2005, find that macroeconomic policies such as stimulation in investment in physical capital and government spending impact HDI with a longer delay and less strongly than Gross Domestic Product. Pardede, Irsad, and Rujiman(2021), analyzing the factors that affect human development in North Sumatra, find economic growth and poverty have significant negative effects on HDI, and direct expenditure per capita has significant positive effects on effect on HDI.

## 4. Data and Methodology

The present study attempts to explain the variations in HDI values among the Indian states and Union Territories using secondary data collected from various sources. The data on HDI values of Indian states are collected from a website, 'Global DataLab' (https://globaldatalab.org/shdi/shdi/ accessed on 29/04/2022). The data on road lengths, number of educational institutes, and number of hospital beds are collected from various reports (Govt. of India, 2015, 2016, 2017) and Economic Survey reports (Govt. of India, 2017) published by various ministries of the government of India.

The present study's list of variables for the 36 states/UTs is presented in Table 1.

Rationales for selecting the dependent and independent variables of the study are already explained in the preceding paragraphs. The rationales for the hypothesised sign of the coefficients of the independent variables are as follows-

1. The more the number of educational institutes per unit of population in a region, the more the average year of schooling that the region's people can avail themselves. Hence a positive relationship between the number of educational institutes per 10,000 population and the HDI value is assumed.

2. More availability of hospital beds per unit of population enhances accessibility to basic health facilities for masses, especially women. Thus more hospital beds per unit of the population can be linked to enhancement of the life expectancy of the people and hence higher HDI. Therefore, a positive association between hospital beds per 10,000 populations and HDI is assumed.

**Table 1. List of Variables**

| Type of the variable | Variable (name in parentheses) | Hypothesized sign |
|---|---|---|
| **Dependent** | The value of the HDI of the observation state/UT (***hdi***) | |
| **Independent** | Number of education institutes per 10,000 of the population (***eduinst***) | + |
| | Number of hospital beds per 10,000 population(***bed***) | + |
| | Kilometers of roads per square kilometres of area (***road***) | + |

3. Increased roads per square kilometres of the area are an indicator of the basic infrastructural development of the region. More kilometres of roads per square kilometres of an area provide not only direct employment in the construction sector, thereby boosting income, but also increasing overall efficiency in the economy, leading to an increase in productivity of factors and hence per capita income. This leads to an improvement in the HDI value. Therefore the variable kilometres of roads per square kilometres of area are hypothesised to be positively linked to HDI.

Therefore, the present study considers the following model-

$$hdi_i = f\,(road_i\,, eduinst_i\,, bed_i\,) \qquad (1)$$

where the functional form(f) is the logistic functional form which is given by-

$$hdi_i = \frac{1}{1+\exp\left(-(\alpha+\beta_1 * road_i+\beta_2 * eduinst_i+\beta_3 * bed_i)\right)} \qquad (2)$$

This equation transforms into the following form-

$$\ln\left(\frac{hdi_i}{1-hdi_i}\right) = \alpha + \beta_1 \times road_i + \beta_2 \times eduinst_i + \beta_3 \times bed_i \qquad (3)$$

or, $hdi\_tans_i = \alpha + \beta_1 \times road_i + \beta_2 \times eduinst_i + \beta_3 \times bed_i$

where, $hdi\_trans_i = \ln(\frac{hdi_i}{1-hdi_i})$

Therefore the Sample Regression Function is –

$$hdi\_tans_i = \alpha + \beta_1 \times road_i + \beta_2 \times eduinst_i + \beta_3 \times bed_i + u_i \qquad (4)$$

where the $u_i$ the term is the stochastic error term

Thus the transformed dependent variable (*hdi_trans*) is linearly related to the regressors. Thus we can apply OLS.

The logistic functional form is chosen because the dependent variable is limited to dependent variables whose values are restricted to the interval (0, 1). This functional form will make the predicted values of the original dependent variable, i.e., *hdi*, lie between (0, 1), however large the values of the independent variables are.

## 5. Results and Discussion

The descriptive statistics of the variables are presented in Table 2.

**Table 2. Descriptive statistics of the variables**

| Variable | Mean | Standard Deviation | Minimum | Maximum |
|---|---|---|---|---|
| ***hdi_trans*** | 0.683 | 0.245 | 0.217 | 1.136 |
| ***hdi*** | 0.662 | 0.055 | 0.554 | 0.757 |
| ***road*** | 3.220 | 5.338 | 0.164 | 25.746 |
| ***eduinst*** | 15.402 | 9.494 | 2.236 | 49.240 |
| ***bed*** | 10.501 | 9.198 | 1.101 | 46.531 |

*Source: Author's calculations*

*NB: The maximum value of hdi_trans being above 1 is permissible because it is the transformed logistic value of the variable hdi.*

The coefficients of the SRF are estimated using the Ordinary Least Square (OLS) linear regression method. The results of the regression analysis are shown in Table 3.

**Table 3. Regression results**

| Variable | Estimated coefficient | Standard error | t value | p-value |
|---|---|---|---|---|
| ***road*** | 0.0134 * | 0.0067 | 1.99 | 0.055 |
| ***eduinst*** | -0.0025 | 0.0038 | -0.07 | 0.508 |
| ***bed*** | 0.0145*** | 0.0035 | 4.17 | 0.000 |
| ***constant*** | 0.5252*** | 0.0837 | 6.27 | 0.000 |
| R-squared: 0.4661 | | No. of observations: 36 | | |
| F(3,32): 9.31 | | p(F)= 0.0001 | | |

*Source: Author's calculations*

The value of $R^{2}$, which measures the proportion of explained variation to the total variation, has come out to be 0.4661. That is, 46.61% of the total variation in the dependent variable ( *hdi_trans*) has been explained by the regressors included in the model. At first look, this value of $R^2$ is not quite high. However, looking at the heterogeneity among the Indian states/UTs, this value of R2 is reasonable. To support the point, the value and significance of the F statistic can be stated. The F statistic indicates the overall significance of all the explanatory variables. Under the null hypothesis that all individual estimated coefficients are zero, the value of F statistics has come out to be 9.31 with a p-value of 0.0001. Thus the null hypothesis stands rejected, implying that the model has overall significance.

### *5.1. Results of Diagnostic tests*:

*5.1.1. MultiCollinearity*

To test whether there is high multicollinerity among the independent variables, the Variance Inflation Factors ( VIF) are calculated. As a rule of thumb, a variable with a VIF score above 10 (or sometimes 5) indicates multicollinarity. The VIF scores in the case of the present model are shown in Table 4.

From table 4, it is evident that all the VIF values are below 5. Therefore the possibility of multicollinerity is ruled out.

**Table 4. VIFs of the independent variables**

| Variable | VIF | 1/VIF |
|---|---|---|
| ***road*** | 1.29 | 0.7775 |
| ***eduinst*** | 1.26 | 0.7910 |
| ***bed*** | 1.02 | 0.9800 |
| **Mean VIF** | 1.19 | |

*Source: Author's calculations*

*5.1.2. Heteroscedasticity:*

In the present case, the residuals from model 4 are plotted against the predicted dependent variable (*hdi_trans*) to get an Insight of the heteroscedasticity problem. This is shown in figure 1.

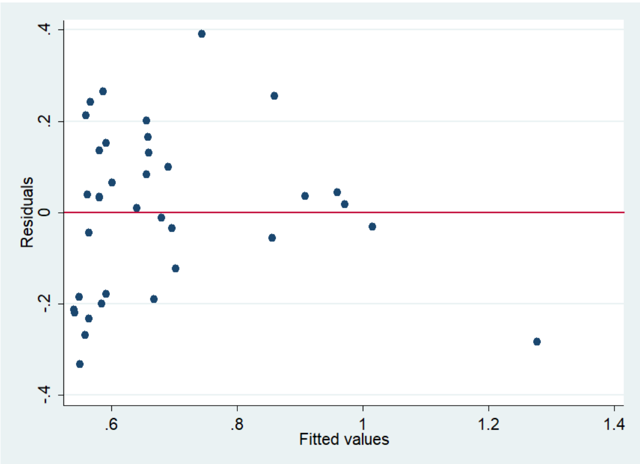


**Fig. 1 Residuals vs. fitted values of hdi_trans**

*NB: A few fitted values of hdi_trans may cross 1 because it is the transformed logistic value of the variable hdi.However, when the fitted hdi_trans are converted back to hdi, all fitted hdi will lie in (0,1).*

One may suspect the presence of heteroscedasticty by looking at figure 1, as the residuals are not clustered around zero across the predicted value range. However, a definite pattern of the residuals cannot be discerned.

Several formal tests are performed to confirm the presence or absence of heteroscedasticy.

One may suspect the presence of heteroscedasticty by looking at figure 1, as the residuals are not clustered around zero across the predicted value range. However, a definite pattern of the residuals cannot be discerned. Several formal tests are performed to confirm the presence or absence of heteroscedasticy.

*5.1.3. Breusch -Pagan Test of heteroscedasticity*

The requirement to conduct the BP test of heteroscedasticity is that the residual from the original model must be normally distributed. The Chi-square test of normality reveals that the residuals obtained from model 4 are normally distributed (Under the null hypothesis that residuals are normally distributed, the Chi-square value at 2 degrees of freedoms (d.f.) is 0.394807 with a p-value of 0.820859).

When the Breusch-Pagan test is conducted by taking the same set of independent variables as in model 4 as the variables under the null hypothesis of homoscedascity, the Chi-square values at 3 degrees of freedom have come out to be 1.72 with a p-value of 0.6323. Thus we fail to reject the null hypothesis, and the error term is homosedastic. However, another popular test, White's test, is performed to reassure the presence or absence of heteroscedasticty.

*5.1.4. White's test of heteroscedasticty*

It has been argued that when White's test avoids the cross-product terms, it becomes a pure test of hetertoscedastcity rather than a test of both heteroscedasticty and specification (Gujarati & Porter, 2005:388). Therefore, the independent variables of model 4 and their squared terms have been included as independent variables in the test procedure. Under the null hypothesis of homoscedasticity, the value of the test statistic, i.e., Chi-square at 6 d.f.s, is found as 10.8221 with a p-value of 0.0940. Thus, the null hypothesis stands rejected, although at a higher p-value, indicating heteroscedasticity. As the two tests have produced contradictory results, the hetereoscedasticty corrected standard errors or robust standard error of the coefficients of model 4 is presented in table 5.

**Table 5. Heteroscedasticity corrected regression results**

| variable | coefficient | Robust standard error | t value | p-value |
|---|---|---|---|---|
| ***road*** | 0.0134*** | 0.0028 | 4.85 | 0.000 |
| ***eduinst*** | - 0.0025 | 0.0030 | -0.84 | 0.406 |
| ***bed*** | 0.0145*** | 0.0043 | 3.38 | 0.002 |
| ***constant*** | 0.5252*** | 0.0843 | 6.23 | 0.000 |
| F( 3, 32) : 19.04 | with p(F): | 0.0000 | | |

*Source: Author's calculations*

Thus the variables *road* and *bed* are found to have a significant positive impact on the dependent variable *hdi_trans* and hence *hdi*. Construction of an additional kilometer of road per square kilometers of area and making an additional hospital bed per 10,000 people available can improve the HDI value of the state/UT. The variables *bed* and *road* are significant at all the conventional levels of significance. However, before considering robust standard error, the road was significant only at a 10% significance level. The exact interpretation of the slope coefficients of these variables is difficult as the partial derivative of the equation (3) concerning, say, road yields $\frac{\partial \widehat{hdi_i}}{\partial road_i} = \beta_1 \times \widehat{hdi}_i \times (1 - \widehat{hdi}_i)$ which value depends on $\widehat{hdi}_i$ Hence non-constant. The variable *eduinst* has appeared to be insignificant and with unexpected signs. This might be because when educational institutes' availability was scaled on population, population figures had overshadowed, giving sparsely populated and hilly states an advantage over populous states/UTs.

## 6. Conclusion and Policy Suggestion

The present study considers 36 Indian states and UTs for the year 2015 and tries to establish an econometric relationship between the HDI values of the states/UTs and the variables that may have a potential impact on HDI, viz., kilometers of roads per square kilometers of area, number of hospital beds per 10,000 of population, and number of educational institutes per 10,000 of the population to explain variances of the HDI values among the states. Through the regression analysis, it has been found that the variables kilometers of roads per square kilometers of area and the number of hospital beds per 10,000 population have a statistically significant positive impact on HDI. Thus, states/UTs with comparatively lower HDI scores should focus on creating additional kilometres of roads per square kilometres of area, and making more hospital beds per 10,000 people available to foster human development.

## Conflicts of Interest

The author declares that there is no conflict of interest regarding the publication of the present paper.

## Funding Statement

This is a work of the author's own. No funding from any agency is received for carrying out this work.

## References


[1] M. Binder and G. Georgiadis, "Determinants of Human Development: Capturing the Role of Institutions," CESIFO working paper no. 3397, Munich: Center for Economic Studies and ifo Institute (CESifo), 2011.

[2] D. N. Gujarati and D. C. Porter, *Basic Econometrics*, 5th ed., New York: McGraw Hill Irwin, pp. 388, 2005.

[3] N. H. Khan, Y. Ju, and S. T. Hassan, "Investigating the Determinants of Human Development Index in Pakistan: An Empirical Analysis," *Environmental Science and Pollution Research*, vol. 26, pp. 19294–19304, 2019.

[4] S. Shah, "Determinants of Human Development Index: A Cross-Country Empirical Analysis," *International Journal of Economics and Management Studies*, vol. 3, no. 7, pp. 40-43, 2016.

[5] N. W. Pardede, Irsad, and Rujiman, "Analysis of Factors That Affecting to Human Development Index in North Sumatra," *International Journal of Research and Review*, vol. 8, no. 4, pp. 453-467, 2021.

[6] GlobalDataLab website. [Online]. Available: https://globaldatalab.org/shdi/shdi/

[7] *Basic Roads Statistics of India 2013-14 & 2014-15*, Ministry of Transport and Highways, Govt. of India, pp. 7, 2016. Available: https://morth.gov.in/sites/default/files/Basic_Road_Statistics_of_India_2013_14_and_2014_15_0.pdf

[8] *Economic Survey of India, 2016-17*, Ministry of Finance, Govt. of India, pp.A148, 2017.

[9] *National Health Profiles 2015*, Central Bureau of Health Intelligence, Govt. of India, pp. 243, 2015. Available: www.cbhidghs.nic.in/index1.php?lang=1&level=2&sublinkid=85&lid=74

[10] *National Health Profiles 2016*, Central Bureau of Health Intelligence, Govt. of India, pp. 244, 2016. Available: www.cbhidghs.nic.in/index1.php?lang=1&level=2&sublinkid=86&lid=1136

[11] *National Health Profiles 2017*, Central Bureau of Health Intelligence, Govt. of India, pp. 272, 2017. Available: www.cbhidghs.nic.in/index1.php?lang=1&level=2&sublinkid=87&lid=1137


## Appendix

**Table 6. The data used in the model**

| States/UTs | HDI | Kilometers of roads per sq. kilometers of area | Number of educational institutes per 10,000 population | Number of (govt.) hospital beds per 10,000 population |
|---|---|---|---|---|
| Andaman and Nicobar Islands | 0.72 | 0.163899 | 11.167137 | 28.24629 |
| Andhra Pradesh | 0.627 | 1.117456 | 13.328129 | 4.003461 |
| Arunachal Pradesh | 0.66 | 0.302855 | 29.601215 | 16.71572 |
| Assam | 0.595 | 4.162676 | 22.972817 | 4.288016 |
| Bihar | 0.554 | 2.187802 | 8.1808308 | 1.109708 |
| Chandigarh | 0.732 | 25.74561 | 2.2360131 | 6.632242 |
| Chhattisgarh | 0.59 | 0.721448 | 21.37114 | 4.702254 |
| Dadra and Nagar Haveli | 0.661 | 1.761711 | 10.386693 | 10.82311 |
| Daman and Diu | 0.688 | 4.702703 | 6.3721238 | 8.222095 |
| Delhi | 0.728 | 21.62306 | 3.6341562 | 14.52412 |
| Goa | 0.753 | 3.950297 | 11.127528 | 21.37747 |
| Gujarat | 0.649 | 0.928879 | 9.0670548 | 4.620804 |
| Haryana | 0.684 | 1.046933 | 9.3785518 | 3.0231 |
| Himachal Pradesh | 0.702 | 0.998563 | 26.957135 | 12.78443 |
| Jammu and Kashmir | 0.672 | 0.175921 | 23.526265 | 7.782286 |
| Jharkhand | 0.58 | 0.535714 | 14.8417 | 3.269054 |
| Karnataka | 0.657 | 1.67791 | 13.220821 | 8.678573 |
| Kerala | 0.757 | 5.015289 | 5.7031567 | 11.49492 |
| Lakshadweep | 0.73 | 7.133333 | 6.9796659 | 46.53111 |
| Madhya Pradesh | 0.581 | 0.937321 | 21.139026 | 3.881074 |

| | | | | |
|---|---|---|---|---|
| Maharashtra | 0.678 | 1.976322 | 10.210428 | 4.578092 |
| Manipur | 0.692 | 1.085995 | 17.858431 | 4.996859 |
| Meghalaya | 0.646 | 0.596192 | 49.24013 | 10.53966 |
| Mizoram | 0.695 | 0.466344 | 35.289636 | 14.85592 |
| Nagaland | 0.677 | 2.242355 | 14.68788 | 9.502138 |
| Orissa | 0.582 | 1.82196 | 17.105024 | 3.974583 |
| Puducherry | 0.729 | 6.269388 | 7.0194951 | 26.21092 |
| Punjab | 0.701 | 2.092212 | 10.987503 | 4.254715 |
| Rajasthan | 0.602 | 0.725096 | 16.358214 | 6.808179 |
| Sikkim | 0.69 | 1.049887 | 21.405982 | 25.5496 |
| Tamil Nadu | 0.688 | 2.007535 | 8.4580058 | 8.904455 |
| Telangana | 0.649 | 0.894483 | 13.053201 | 5.282017 |
| Tripura | 0.641 | 3.565134 | 13.364483 | 11.26046 |
| Uttar Pradesh | 0.572 | 1.724096 | 13.185572 | 3.000065 |
| Uttarakhand | 0.661 | 1.176916 | 24.416307 | 8.439177 |
| West Bengal | 0.617 | 3.335102 | 10.643201 | 8.607509 |

***NB (A)*** *Populations figures of Census 2011 are used to scale the variables. Figures for Andhra Pradesh and Telangana are extracted from their official government websites.* ***(B)****Educational institutes and hospital beds are scaled on 10,000 populations instead of 1,00,000 because the least populous UT Lakshadweep had a population of 64,473 as per 2011 Census* ***(C)*** *Area figures of Census 2011 scale the variable road. Figures for Andhra Pradesh and Telangana are extracted from their official government websites.* ***(D)*** *Sources of data are listed in the list of references.*