International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878, Volume-8 Issue-3S3, November 2019

Evaluation of Various Economic Factors of Indian States & Union Territories through Mathematical Modeling using Factor Theory

Renuka Devi, A. K. Agrawal, Joydip Dhar, Piyush Kumar Tripathi

Abstract: ***The total value of goods produced and services rendered within a country during a year is its Gross Domestic Products." GDP is the growth measurement technique. In this research paper, we analyze India's gross state domestic products (GSDP) data, using the factor analysis model to find out the impact of GSDP contributing factors. There are several methods to test the adequacy of the factor model. Here researcher has used Principal Component factor analysis approaches by varimax rotation method. Data has been collected by the Ministry of statistics & program implementation of year 2017-2018. Basically Indian economy is segmented into three major sectors as primary sector, secondary sector & tertiary sector. Primary sector means "Agriculture, forestry & fishing", Secondary sector means "Manufacturing, construction & electricity, gas water & other utilities services" and tertiary sectors includes "Trade, Transport, Hotels & financial services" so on. India has 28 states and 9 union territories. But in this paper, we have examined 28 states & 5 union territories on the basis of 15 variables. Factor analysis is the factor redemption technique so here, we have reduced these 15 variables into three common factors.***

Key words: Principal Component Method (PCM), Gross Domestic Product (GDP), MOSPI(Ministry of Statistics & Program Implementation), UT (Union Territories), LPG (Liberalization Privatization Globalization), Gross State Domestic Products (GSDP).

I. INTRODUCTION

Gross Domestic Products (GDP) is an important tool to understand the actual economic health of a nation. Post-independence GDP analysis has been significant part of econometrics of India. Since 1947 to 1990 India was under the influence of Soviet Union and thus has practiced socialistic economic structure. After the economic reform of 1991 India has taken a shift from socialism to capitalism by incorporating LPG model. Importance of GDP numbers has significantly increase in this new capitalistic structure. For this paper, researcher has used 15 variables which are affecting the economic growth of India and also are the components of GDP calculation. The secondary data has been taken from Ministry of Statistics and Program Implementation (MOSPI) of the year 2017-18. The data includes the numbers from 28 states and 5 union Territories of republic of India. Researcher has used factor analysis as a mathematical tool for interpretation and analysis of the data.

Revised Manuscript Received on December 22, 2018.

***** Correspondence Author
Renuka Devi, Amity University Uttar Pradesh, Lucknow, INDIA
A. K. Agrawal, Amity University Uttar Pradesh, Lucknow, INDIA
Joydip Dhar, IIIT & M Gwalior, INDIA
Piyush Kumar Tripathi, Amity University Uttar Pradesh, Lucknow, INDIA

Using factor analysis we classify 15 variables into three factors and the new name of these factors are "Agriculture & Manufacturing", "Forestry & Mining" & "Trade & Aquaculture". At last we compare the performance of the states on the basis of these three factors and analyze contribution of "Agriculture, Industry & Service" factors in state/union territories and Indian economy.

Factor analysis is a co-relational method used to find and describe the underlying factors deriving data values of a large set of variables. It is statistical approach which is used to analyze interrelationship among a large number of variables.

II. LITERATURE REVIEW

India is second populated country but also fastest growing economy in the world. The picture of Indian economy is very wide. So to understand the depth of the Indian economy it's necessary to refer privous researches. As Kaushik Bashu & Annemie Maertens (2007) explained the challenges faced by Indian economy in 1990. This paper highlights the problems like infrastructure, corruption in the government bureaucracy, low growth pattern in agriculture, inequalities between states and so on, were the major problems. Sanjay Kumar (2015) explained the role of financial institution in Indian growth. And established a relationship between development and financial intuitions in India. For a deep knowledge its essential to analyse internal and external both approaches. Anita Bai (2015) defined another dimension, where researcher highlight the performance of 20 countries and select the 15 variables.

III. DESCRIPTION OF THE METHOD

Let the observation random vector X, with 'p' components, has mean vector μ and covariance matrix $\sum$. The Factor analysis model suggest that X is linearly dependent upon a few unobservable random variables $F_1, F_2, F_3, \ldots\ldots, F_m$, called common factors, and 'p' additional source of variation $\delta_1, \delta_2, \ldots\ldots, \delta_p$, called error or sometimes specific factors. Then the factor analysis model is

Regression Line Equation

$$
\begin{aligned}
X_1 &= \mu_1 + \lambda_{11}F_1 + \lambda_{12}F_2 \cdots\cdots + \lambda_{1m}F_m + \delta_1 \\
X_2 &= \mu_2 + \lambda_{21}F_1 + \lambda_{22}F_2 \cdots\cdots + \lambda_{1m}F_m + \delta_2 \\
&\vdots \qquad\qquad\qquad \vdots \qquad\qquad\qquad \vdots \\
&\vdots \qquad\qquad\qquad \vdots \qquad\qquad\qquad \vdots \\
X_j &= \mu_j + \lambda_{j1}F_1 + \lambda_{j2}F_2 \cdots\cdots + \lambda_{jm}F_m + \delta_j \\
&\vdots \qquad\qquad\qquad \vdots \qquad\qquad\qquad \vdots \\
&\vdots \qquad\qquad\qquad \vdots \qquad\qquad\qquad \vdots \\
X_p &= \mu_p + \lambda_{p1}F_1 + \lambda_{p2}F_2 \cdots\cdots + \lambda_{pm}F_m + \delta_p
\end{aligned}
$$

In Form of Matrix Model

$$
\begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_j \\ \vdots \\ X_p \end{bmatrix}_{p\times 1} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_3 \\ \vdots \\ \mu_p \end{bmatrix}_{p\times 1} + \begin{bmatrix} \lambda_{11} & \lambda_{12} & \cdots & \lambda_{1m} \\ \lambda_{21} & \lambda_{22} & \cdots & \lambda_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ \lambda_{j1} & \lambda_{j2} & \cdots & \lambda_{jm} \\ \vdots & \vdots & \vdots & \vdots \\ \lambda_{p1} & \lambda_{p2} & \cdots & \lambda_{pm} \end{bmatrix}_{p\times m} \begin{bmatrix} F_1 \\ F_2 \\ \vdots \\ F_2 \\ \vdots \\ F_m \end{bmatrix}_{m\times 1} + \begin{bmatrix} \delta_1 \\ \delta_2 \\ \vdots \\ \delta_j \\ \vdots \\ \delta_p \end{bmatrix}_{p\times 1}
$$

$$
[X]_{p\times 1} = [\mu]_{p\times 1} + [\lambda]_{p\times m}[F]_{m\times 1} + [\delta]_{p\times 1}
$$

In general form we can write this equation

$$
X - \mu = \lambda F + \delta
$$

This equation is known as factor model. If this model fulfill all assumptions so this model is orthogonal factor model because $F_1, F_2, F_3, \ldots\ldots, F_m$ have no correlation. They are independent to each other.

IV. ASSUMPTIONS OF THE FACTOR ANALYSIS MODEL

1. Measurement errors are homogenous and an average value of 0.

2. There is no association between the factor and the measurement's error

$E(F, \delta) = 0$

3. There are no associations between errors.

$Cov(\delta_i, \delta_j) = 0$

4. $Cov(F) = E(FF^T) = I$

5. $Cov(\delta) = \varphi = \begin{pmatrix} \varphi_{11} & 0 & \cdots & 0 \\ 0 & \varphi_{22} & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & & \varphi_{pp} \end{pmatrix}$

V. DATA ANALYSIS & INTERPRETATION

Indian GDP data is based on Primary, Secondary & Territory Sectors. Here the data is based on 15 variables in 33 States/Union Territories. These 15 variables are as follows:

X_1 = Crop
X_2 = Livestocks
X_3 =Forestry & Logging
X_4 = Fishing & Aquaculture
X_5 = Mining & Quarrying
X_6 = Manufacturing
X_7 = Electricity, Gas, Water Supply & Other Utilities Services
X_8 = Construction
X_9 = Trade & Repair Services
X_{10} = Hotels & Restaurant
X_{11} = Financial Services
X_{12} = Real Estate, Ownership of Dwelling & Professional Services
X_{13} = Public Administration
X_{14} = Transport, Storage, Communication & Services related to broadcasting
X_{15} = Other Services

On the basis of these 15 variables first we calculate the correlation matrix, as follows:

Correlation Matrix

Crop	1.00														
Livestocks	.841	1.00													
Forestry	.720	.674	1.00												
Fishing	.391	.460	.095	1.00											
Mining	.550	.541	.811	.200	1.00										
Manufacturing	.633	.627	.733	.210	.543	1.00									
Electricity	.755	.756	.804	.265	.694	.868	1.00								
Construction	.763	.803	.686	.351	.458	.863	.814	1.00							
Trade	.718	.742	.698	.336	.501	.861	.811	.918	1.00						
Hotels	.633	.645	.597	.287	.395	.892	.755	.898	.906	1.00					
Financial Services	.517	.475	.681	.170	.554	.886	.848	.705	.809	.796	1.00				
Real Estate	.552	.525	.617	.158	.439	.902	.781	.774	.852	.927	.887	1.00			

Retrieval Number: C10641183S319/2019©BEIESP
DOI: 10.35940/ijrte.C1064.1183S319

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

Public Administration	.857	.785	.696	.355	.540	.772	.822	.879	.849	.780	.734	.716	1.00		
Transport	.735	.758	.639	.409	.517	.815	.862	.881	.907	.879	.836	.841	.908	1.00	
Other Services	.685	.745	.733	.403	.561	.883	.852	.905	.968	.915	.862	.859	.843	.925	1.00

Kaiser-Mayer-Olkin (KMO) And Bartlett's Test

KMO and Bartlett's test are two different techniques to identify that factor analysis is useful for your data or not. KMO measure the sampling adequacy and Bartlett's test is for sphericity. In table below it shows that the value of KMO measure of sampling is 0.745 which is greater than 0.50 , so the result of factor analysis is useful.

KMO and Bartlett's Test

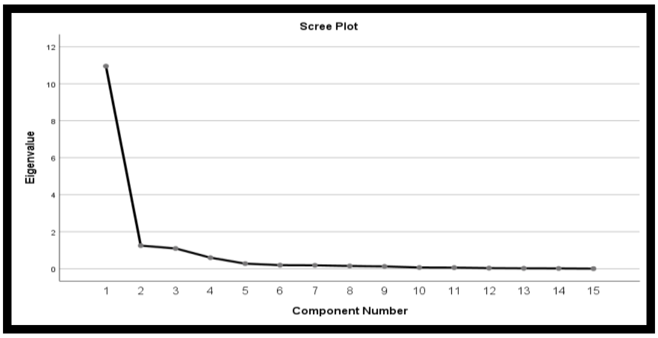
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.745
Bartlett's Test of Sphericity	Approx. Chi-Square	737.748
	df	105
	Sig.	.000

Following table shows the analysis of Initial Eigen values and factor analysis of gross domestic products by extraction & rotation sum of square loadings:

Table- 3: Represents Eigen Values, Percentage of variance and Cumulative percentage of variance Extraction Method: Principal Component Analysis.

Components	Initial Eigen Values			Extraction Sum Of Square Loadings			Rotation Sum Of Square Loadings		
	Total	% Of Variance	Cumulative %	Total	% Of Variance	Cumulative %	Total	% Of Variance	Cumulative %
1.	10.950	72.997	72.997	10.950	72.997	72.997	7.271	48.473	48.473
2.	1.248	8.320	81.317	1.248	8.320	81.317	3.596	23.970	72.443
3.	1.094	7.293	88.610	1.094	7.293	88.610	2.425	16.167	88.610
4.	.597	3.980	92.591						
5.	.275	1.831	94.422						
6.	.192	1.277	95.699						
7.	.182	1.210	96.909						
8.	.148	.985	97.895						
9.	.123	.817	98.712						
10.	.066	.437	99.149						
11.	.058	.388	99.537						
12.	.032	.217	99.754						
13.	.020	.134	99.888						
14.	.013	.087	99.976						
15.	.004	.024	100.000						

From the above table we can see that first component total Eigen value is 10.950 which is greater than 1. The first component represents the 72% of total variance. The second component total Eigen value is 1.248 Which is also greater than 1. And second component represents the 8% of total variance. On the basis of this analysis only two components representing the total 81% of variance. By extracting three components, the total Eigen value of third component is 1.094 which is also greater than 1. And the third component represents only 7% of total variance. So here are three factors whose Initial Eigen values are greater than 1. According to the analysis we consider the three factor loading model. In varimax rotation three factor model the extraction sum of square loading and rotation sum of square loading is also representing the same result.



This graph representing the relationship between the Eigen value and 15 components. In this graph it is clear that only three components whose value is greater than 1. And 4 to 15 components whose value is less than 1 and near to 0. So on the basis of table & graph we can say that our model is based on three factors only.

Component Matrix : Principal Component Method With Varimax Rotation

In the this table three factor loading model is calculated by Principal Component Method with Varimax rotation. The results are given below:

Table 4 : Extraction Method: Principal Component Analysis

Variables	Three Factor Solution						
	Estimated Factor Loadings			Estimated Rotated Factor Loadings			Communalities
	F_1	F_2	F_3	F_1 *	F_2 *	F_3 *	h_i^2
Crop	.811	.383	.145	.398	.600	.554	.825
Livestocks	.811	.438	.065	.400	.542	.632	.854
Forestry	.806	-.015	.519	.454	.841	.068	.919
Fishing	.364	.699	-.389	.084	-.020	.875	.773
Mining	.642	.115	.642	.216	.886	.072	.837
Manufacturing	.919	-.257	-.032	.864	.389	.117	.911
Electricity	.926	-.030	.184	.679	.616	.228	.893
Construction	.931	.052	-.155	.766	.347	.432	.893
Trade	.946	-.055	-.154	.830	.336	.346	.922
Hotels	.905	-.193	-.282	.912	.182	.265	.936
Financial Services	.862	-.364	-.005	.862	.364	-.005	.876
Real Estate	.871	-.377	-.168	.935	.228	.052	.929
Public Administration	.914	.144	-.022	.658	.467	.453	.856
Transport	.945	.019	-.169	.798	.336	.415	.922
Other Services	.963	-.047	-.128	.830	.368	.349	.946
Cumulative proportion of total variance	72.997	81.317	88.610	42.473	72.443	88.610	

Above table represent the factor loadings of 15 variables on the basis of three factors. And factor three having the highest (88.61%) value. Which is greater than first factor total variance percentage (72.997%). As we can see that 14 variables achieve the high factor value in first factor(F_1). And 72% total variance is covered by first factor(F_1). Only Fishing factor load very low value. But in second factor (F_2) there are only three factors whose value is high (or more than 0.3). And in the third factor (F_3) there are ' Forestry and Mining' variables which score the highest value (more than 0.5).

In the varimax rotation matrix picture is different. Varimax rotation method is extracting the components and define the variables clearly. There are 10 variables whose loading value is high (more than 0.5) in factor one (F_1 *) . In rotation matrix, factor two (F_2 *) has 5 variables whose loading value is high (more than 0.5). And factor three (F_3 *) has only 3 variables whose value is high (more than 0.5).

On the basis of above analysis we can reduce these 15 variables into three major sectors i.e. First factor -“Agriculture & Service Factor”, Second factor -“Forestry & Mining Factor ” & Third factor - “Trade & Aquaculture” respectively.

Sorted Factor Score Table: Regression Method

Factor score table is representing the factor values of states and union territories on the basis of these three factors. Results are shown below in the table:

Table 5

Factor-1		Factor-2		Factor-3	
State	Factor Score	State	Factor Score	State	Factor Score
Maharashtra	3.08249	Rajasthan	3.08566	Andhra Pradesh	3.91687
Tamil Nadu	2.41772	Maharashtra	2.5449	West Bengal	2.20737
Karnataka	2.209	Madhya Pradesh	2.04725	Uttar Pradesh	1.62209
Delhi	1.29832	Odisha	1.13828	Tamil Nadu	0.81936
Kerala	0.8172	Uttar Pradesh	1.11326	Bihar	0.68677
Uttar Pradesh	0.78919	Chhatisgarh	0.64498	Madhya Pradesh	0.54865
Telengana	0.5833	Assam	0.45778	Kerala	0.51119
West Bengal	0.57889	Jharkhand	0.44652	Punjab	0.10849
Haryana	0.19133	Punjab	0.25821	Telengana	0.08529
Bihar	0.03569	Andhra Pradesh	-0.08951	Haryana	-0.00349

Retrieval Number: C10641183S319/2019©BEIESP
DOI: 10.35940/ijrte.C1064.1183S319

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

In the above table there are 10 major states/union territories. As we can see that in the first factor the "Maharashtra" load the highest value (3.08249). And the Tamil Nadu (2.41772), Karnataka (2.209) & Delhi (1.29832) is loading more than 1 value respectively. In factor two the highest value score by "Rajasthan" i.e. 3.08566. And four other states/ union territories are scoring more than 1 value. Like Maharashtra (2.5449), Madhya Pradesh (2.04725), Odisha (1.13828) & Uttar Pradesh (1.1326) are scoring more than 1 value. And the last factor , in factor three Andhra Pradesh achieving the highest score i.e. 3.91687 and other two states/union territories scoring the more than 1 value. West Bengal (2.20737) & Uttar Pradesh (1.62209) are the other hieghest scoring state/union territories in factor three.

VI. CONCLUSIONS

This paper was based on the idea to investigate the role of GDP in Indian development. This writing involves the use of the Principal component tools for its investigation. In this paper we explored 15 different variables which are there in Indian GDP. The main emphasis of this paper was to reduce the 15 variables into three common factors. The nature of Indian GDP is homogeneous at macro level & heterogeneous at micro level. The most important outcome of this analysis is Maharashtra, Rajasthan & Andhra Pradesh scoring the highest value in this factor analysis. And service and industry are contributing the more percentage in Indian GDP.

REFERENCES

1. AC. Rencher, Methods of Multivariate Analysis, Vol 492, John Wiley & Sons, (2003)
2. Anwar Ali Shah, Effects of Macroeconomic Variables On Gross Domestic Products (GDP) In Pakistan, Procedia Economic And Finance ,(2013)
3. Anita Bai, An Application of Factor Analysis In The Evaluation of Country Economic Rank, Procedia Economic And Finance ,(2015)
4. Sanjay Kumar LENKA, Measuring Finance Development In India : A PCA Approach, Vol XXII, Theoretical and Applied Economics (2015)
5. McDonald R, (1985), Factor Analysis and Related Methods Hills dale , NJ: Erlbaum
6. Mandeep Singh and Dr. Kuldip Kaur, India's Service Sector and It's determinants: An Empirical Investigation, Journal of Economic and development studies, Vol.2 (2014)