International Journal of Humanities and Social Sciences.
ISSN 2250-3226 Volume 3, Number 5 (2013), pp. 353-357

http://www.ripublication.com

Analysis of Government Expenditure on Education in India

N. Shradha Varma

Assistant Professor, Delhi University, M.A.Economics,
Delhi School of Economics, UGC-NET-JRF

Education expenditure is on a rise in India. With only 5.4 billion being spent in 1980, this figure rose to a dramatic 42.2 billion in 2009. Apart from this, the literacy rate has also been increasing steadily and stood at 71.6% in 2009 whereas it was 42% in 1980. Thus it has been tried to study this positive relationship between government's education expenditure and India's literacy rate in this paper.
For this a regression analysis has been conducted where the impact of education expenditure on literacy rate has been studied. In this analysis, 'Y' denotes literacy rate and 'X' denotes education expenditure such that

$$Y_i = b_1 + b_2X_i + e_i$$

Here b_1 is the autonomous literacy rate.

b_2 is the regression coefficient that explains the impact of education expenditure on literacy rate.

e_i is the error term / non-explanatory variables.

According to this analysis

$$b_1 = 41.31$$

$$b_2 = 0.918$$

This means that literacy rate will be at least 41% irrespective of the government expenditure on education.

The table shown provides the data for last 30 years and the key formulae used are:

$$b_1 = \square - b_2\,\bar{x}$$

$$b_2 = \frac{\sum xy - n\,\bar{x}.}{\sum x_i^2 - n\,\bar{x}^2}$$

$$e_i = Y_i - (b_1 + b_{2.}\,X_i)$$

$$R^2 = 1 - \frac{\sum e_i^2}{\sum y_i^2}$$

From the table it has been proved that

$$\sum e_i^2 = 580.739$$
$$\sum y_i^2 = 969897.07$$
$$R^2 = 0.9994$$
$$R = 0.997$$

So the education expenditure and literacy rate are highly correlated. Education expenditure explains about 99.4% of the variation in the literacy rate.

Now, we come to the errors which model should not have.

A classical regression model should be free from the problems:

1. Multicollinearity : It shows that the relationship between two explanatory variables. There is no multicollinearity in our model as the considered explanatory variable is only one.
2. Heteroscedasticity: A good regression model should be free from it which means unequal variances of the error term. In mathematical terms, it means that variances of e_i^2 is not same for all.
3. Autocorrelation: In a classical regression model, covariance

$$(u_1, u_2) = 0 \text{ for every } u_1 \neq u_2.$$

When this condition is not satisfied, the disturbance term is subject to auto correlation or serial correlation. The consequence of autocorrelation is an inefficient OLS estimator.

To detect autocorrelation, Durbin Watson d test is being used.

Here $d = \sum (e_t - e_{t-1})^2 / \sum e_t^2$

$$= \frac{107.5494}{580.739}$$

$$= 0.185$$

this indicates that there is almost perfect positive autocorrelation.

This autocorrelation can be removed by taking a lag variable.

$$Y_t = b_1 + b_2X_t + e_t$$
$$Y_t = 41.31 + 0.918X_t + e_t \text{......................}(1)$$

This is the equation of our model.

Now we assume that the error terms follow the autocorrelation pattern such that

$$e_t = \rho\, e_{t-1} + v_i \quad (-1 \leq \rho \leq 1)$$

where v_i is the error term.

Now we take lag variable Y_{t-1} and multiply by ρ.

$$\rho Y_{t-1} = \rho b_1 + \rho b_2 X_{t-1} + \rho e_{t-1} \text{......................}(2)$$

Subtracting equation (2) from equation (1)

$$Y_{t-1} - \rho Y_{t-1} = b_1 (1-\rho) + b_2 (X_{t-1} - \rho X_{t-1}) + v_t$$

We can write it as

$$Y^* = b_1^* + b_2 . X^* + v_t$$

The value of ρ according to Theil- Nagar is

$$\rho = \frac{n^2 \left(1-\frac{d}{2}\right)+ k^2}{n^2 - k^2}$$

$$= \frac{900 \ (\ 1- 0.185/2) \ + \ 4}{900-4}$$

$$= 0.1023$$

For the detection of heteroscedasticity we have plotted a graph of e_i^2 - X.

From the graph we can see that there is no relationship between e_i^2 and X, it is random. Thus we can conclude that there is no heteroscedasticity in our model.

Hypothesis testing of b_i:

We have conducted a t-test to find out whether education expenditure has an effect on literacy rate or not. Thus our hypothesis is set up like:

$$H_0 : b_1 = 0$$

$$H_A : b_1 \neq 0$$

$$t = \frac{b_1 - \hat{b}_1}{s.e.(b_1)}$$

$$\text{s.e.}(b_1) = \frac{\sigma u}{\sqrt{\sum x^2}}$$

Since we don't know the value of σ_u, we will calculate it from s_u.

$$s_u = \sum e^2 / \ n\text{-}2$$

$$= 580.734 / 28$$

$$= 20.8$$

We know, $\sum x^2 = 10623$

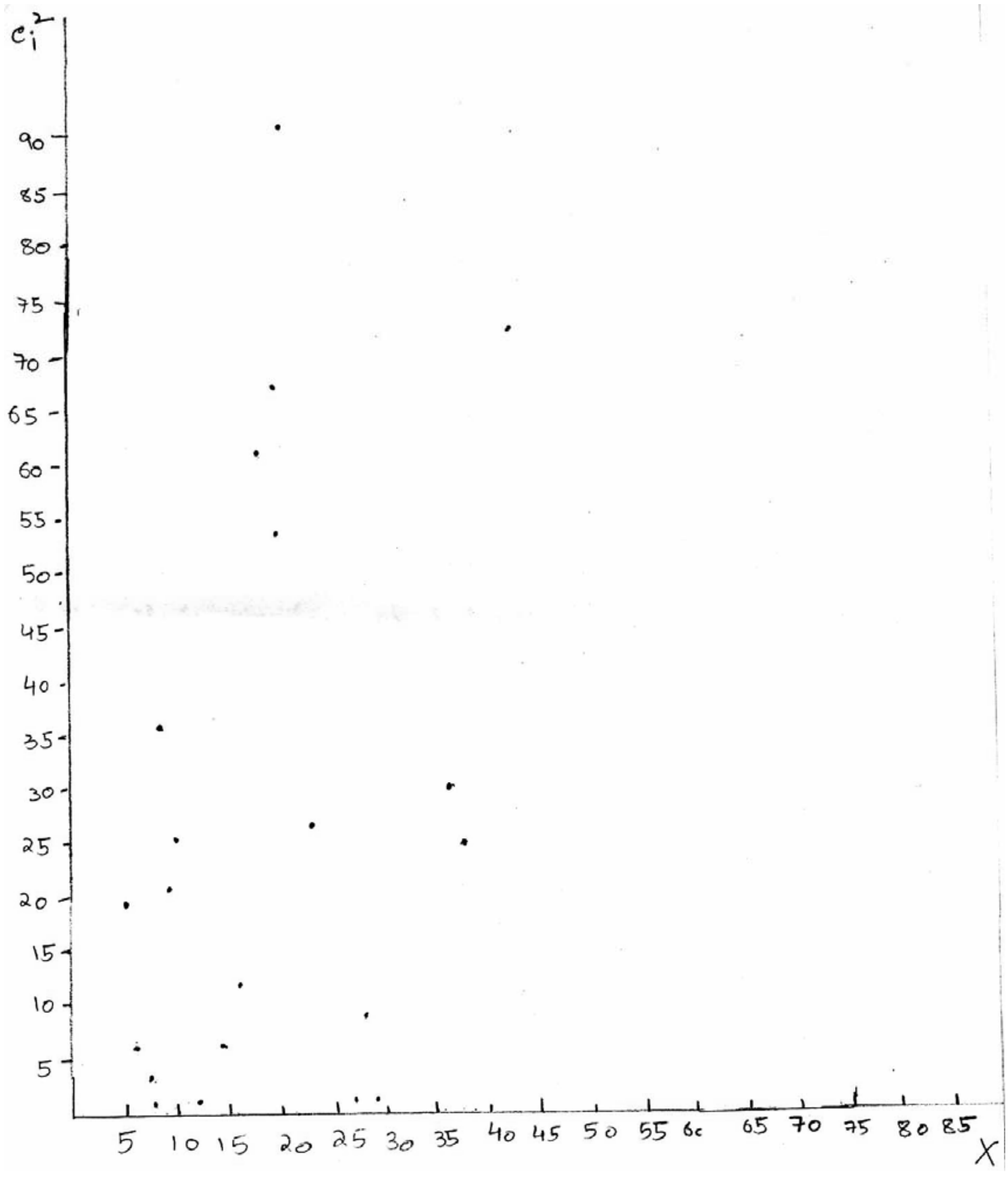
So, $t = \frac{0.918}{20.8/\sqrt{10623}}$

$$= 4.57$$

Critical value of t at 0.05 level of significance and 29 degrees of freedom is 2.045.Hence we reject our hypothesis and so $b_1 \neq 0$.

Thus education expenditure has an effect on the literacy rate. The value of b_1 = 0.918 which implies that if the education expenditure is increased by 1 billion, the literacy rate will rise by 0.918 %.

Years	Education	Literacy R	b2XI	b2Xi + b1	ei = Yi -(b1	ei2	yi2	et - et-1	(et-et-1) sq.
1980	5.4	42	4.9572	46.2692	-4.2692	18.2261	1764		
1981	5.8	43.86	5.3244	46.6364	-2.7764	7.7084	1923	1.4928	2.228
1982	6.1	45	5.5998	46.9118	-1.9118	3.665	2025	0.8646	0.7475
1983	6.7	46.3	6.1506	47.4626	-1.1626	1.3516	2143.69	0.189	0.0357
1984	7.1	48	6.5178	47.829	0.171	0.0292	2304	1.336	1.7784
1985	7.9	48	7.2522	48.5642	-0.5642	0.3183	2304	-0.7352	0.5405
1986	8.3	48.72	7.6194	48.9314	-0.2114	0.447	2373.63	0.3528	0.1244
1987	8.7	49.6	7.9866	49.2986	0.3014	0.0908	2460.16	0.5155	0.2657
1988	10.8	50.5	9.9144	51.2264	-0.7264	0.5277	2550.25	-1.0278	1.0563
1989	11.5	50.9	10.557	51.869	-0.969	0.939	2590.81	0.2426	0.0588
1990	12	51	11.016	52.328	-1.328	1.7636	2601	-0.359	0.1288
1991	9.8	45.3	8.9964	50.3084	-5.0084	25.0841	2052.09	-3.6804	13.545
1992	8.7	42.1	7.9866	49.2986	-7.1986	51.8198	1772.41	-2.1902	-4.796
1993	9.6	45.6	8.8128	50.1248	-4.5248	20.4738	2079.36	2.6738	7.171
1994	11.2	51	10.2816	51.5936	-0.5936	0.3524	2601	3.9312	15.4543
1995	11.5	51.3	10.557	51.869	-0.569	0.3238	2631.69	0.0246	6.0516
1996	12.2	53.2	11.1996	52.5116	0.6884	0.4739	2830.24	1.2574	1.581
1997	14.4	57	13.2192	54.5312	2.4688	6.0949	3249	1.7804	3.1698
1998	15.5	59	14.229	55.541	3.459	11.9645	3481	0.9902	0.9804
1999	19.7	69	18.0846	59.3966	9.6034	92.225	4761	6.1444	37.753
2000	17.2	65	15.7896	57.1016	7.8984	62.3847	4225	-1.705	-2.907
2001	17.9	65.38	16.4322	57.7442	7.6358	58.3054	4274.54	0.2626	0.0689
2002	19	66	17.442	58.754	7.246	52.5045	4356	-0.3898	0.1519
2003	22.4	67	20.5632	61.8752	5.1248	26.2635	4489	-2.1212	4.9994
2004	27	69	24.786	66.098	2.902	8.4216	4761	-2.228	4.9639
2005	26	66.5	23.868	65.18	1.32	1.7424	4422.25	-1.582	2.5027
2006	29	69	26.622	67.934	1.066	1.1363	4761	-0.254	0.0645
2007	38.1	71.3	34.9758	76.2878	-4.9878	24.8781	5083.69	-6.0538	2.8944
2008	37.2	70	34.1496	75.4616	-5.4616	29.829	4900	-0.4738	0.2244
2009	42.2	71.6	38.7396	80.0516	-8.4516	71.4295	5126.56	-2.99	8.9401



e_i^2
90
85
80
75
70
65
60
55
50
45
40
35
30
25
20
15
10
5
5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85
X

THE END