A NOTE ON POLIO COUNT: SOME EMPIRICAL EVIDENCE FROM INDIA

MUKHERJEE, Debasri*

Abstract: Although the world wide global polio eradication program aimed at exterminating the disease by the year 2003, seven countries including India are still considered to be polio endemic. India witnessed a polio explosion during the year 2002. In recent years there is definite evidence in favor of importations of polio from these seven countries to other countries as well. Certified polio free nations also continue to face the risk of transmission until the disease is completely eradicated. Using a negative binomial model which corrects for overdispersion problem in count data, this paper examines various factors affecting the polio counts in India. Expenditure on the polio eradication program which includes cost of vaccination and the other related costs has a significantly negative impact on polio counts as expected. The paper calculates the expenditure elasticity of polio count which is important from the point of view of policy making.

Keywords: Poisson, negative binomial, polio, immunization expenditure
JEL Classification No. C1, I12, O1

1. Introduction.

Polio is a highly infectious disease which affects children usually under the age of three. Once infected, the child can suffer total paralysis within hours. The disease has no cure, only prevention through vaccination can be achieved, and once immunized, life long prevention is guaranteed. In 1988, the forty-first World Health Assembly launched a global initiative program to completely eradicate polio by the end of the year 2003. International organizations adopted a polio eradication strategy primarily by making vaccines available world-wide with an aim of freeing the world from polio. Overall, in the 15 years since the Global Polio Eradication Initiative was launched, the number of victims has fallen considerably and the number of polio-endemic countries has reduced from 125 to 7. These seven remaining polio endemic countries include India, Nigeria, Pakistan, Egypt, Afghanistan, Niger and Somalia. Unfortunately, the world faced an increase in global cases in 2002 over 2001 (1918 cases as opposed to 483 cases) and this increase can be attributed to an epidemic in India (1600 cases as opposed to 268 cases[2]). According to WHO [3] "*As long as a single child remains infected with poliovirus, children in all countries are at risk of contracting the disease. In January 2003, a child was paralyzed by polio in Lebanon, the first case seen in that country in nearly ten years. Genetic sequencing confirmed this case was an importation from India. The poliovirus can easily be imported into a polio-free country and can spread rapidly amongst unimmunized populations*". Even if the transmission of the disease is geographically restricted, there are evidences of polio attacks in so-called polio free areas as well. "*Between 2000 to 2003, a total of 12 such importations (of the virus) were detected with*

* Debasri Mukherjee, Economics Department, Western Michigan University
Kalamazoo, MI 49008, Email: debasri.mukherjee@wmich.edu

[2] See Indiastat.com and World Health Organization Report:
http://www.who.int/immunization_monitoring/en/diseases/poliomyelitis/case_count.cfm
[3] See World Health Organization Report on polio: http://www.who.int/topics/poliomyelitis/en/

over 70 children paralyzed in Africa, Asia, Europe and the Western Pacific". See Mascarenas *et al* (2005). Some of the northern states of India, especially Uttar Pradesh are still badly affected containing more than sixty percent of the cases found all over India.

Although actual immunization is a task of the government at the grass root level, the material cost of vaccination is not high and is provided by international organizations. The overall costs are associated with material and preservation costs of the vaccine, cost of technical assistance as well as that of campaigns in favor of immunization. Various factors account for the endemic although poverty, lack of education and poor social and political system are the prime ones. Fogel (1984) emphasizes the role of economic poverty while Preston (1976) emphasizes the role of poor public health measures in this context. Public health measures reduce the possibility of being exposed to the disease whereas decrease in poverty level is associated with increase in resistance power. Surprisingly, in some states of India (Kerala, for example) we find substantial evidence of reduction in child mortality rate without substantial increase in income per capita. See Gauri and Khaleghian (2002). They attribute it to relatively high degrees of female liberation, improved health care system, inclination towards education and lack of a rigid class structure. All these combine to a positive attitude towards education as well. Failure to immunize is caused by many factors including fear of negative effects, social and religious stigma, and especially poor functioning of the local government. Lack of adequate storage opportunities needed to preserve the vaccine in the heat, use of dirty syringes, unequipped staff, and use of expired vaccines also lead to immunization failure. The Media center (2003) of World Health Organization[4] states, *"Eighty-three per cent of all new polio cases are now found in India. This country and Uttar Pradesh in particular, are the number one priorities for stopping transmission of the polio virus around the world"*. Although economists have discussed the factors leading to the continuation of the disease in various countries including India, no serious econometric attempt has been made to look into the issue in India, a country that still contains a considerable number of victims of poliovirus.

This paper provides a regression analysis of factors affecting polio counts in India. We also attempt to quantify the expenditure elasticity of polio count in particular. The plan of the paper is as follows. In section 2 we present the econometric methodology and the description of the data set used. We analyze the results in section 3 and section 4 concludes.

2. Methodology and Data.

The data on polio victims are discrete in nature containing large number of zeros. Due to such integer property of the dependent variable (number of polio victims) any continuous regression specification will not suffice. Poisson regression is a natural starting point for modeling count data. Poisson model assumes that the conditional mean and the variance of the dependent variable must be the same. However, quite often, data show overdispersion and hence mean-variance equality restriction is rejected. Negative binomial is a standard modeling strategy for that. In our case we use both Poisson and negative binomial models although the standard test shows that negative binomial

[4] http://www.who.int/mediacentre/news/releases/2003/pr30/en/

distribution is more appropriate for modeling polio count. Poisson and negative binomial models have been widely used in economics especially in health economics and other related areas. Cameron *et al* (1986) have used Poisson and negative binomial regression frameworks to discuss the factors affecting number of doctors' visit in a particular time interval for Australian household survey data. Cameron *et al* (1988) have also used aforementioned regression models to analyze data on health care utilization in the United States. Using data from Ethiopian household survey, Muhe *et al* (1996) have analyzed the factors leading to child mortality (count data on child morbidity), while Samandari *et al* (2004) have quantified the effect of immunization coverage rate on the incidence of hepatitis A (number of victims) for different states in the United States. Both the papers have used Poisson regression framework.

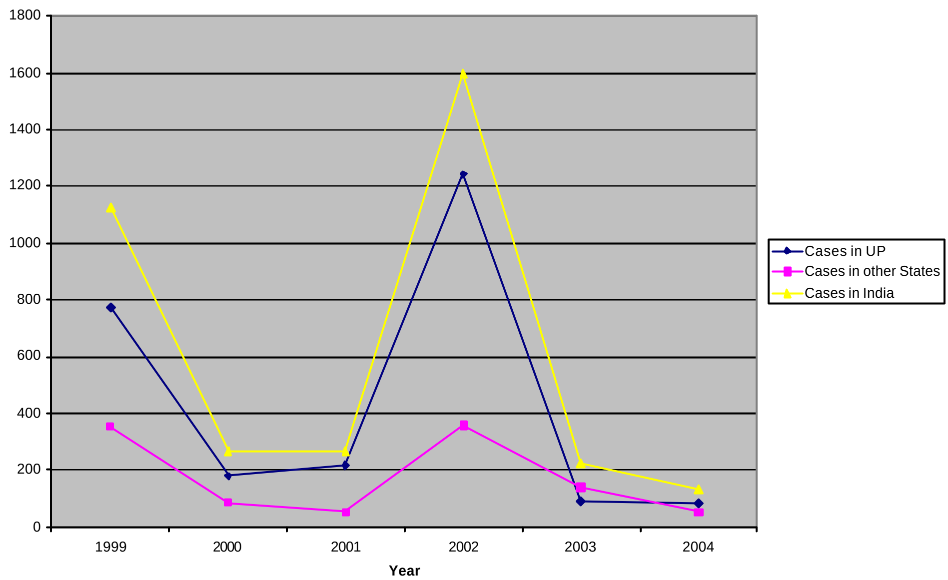
In this paper, using Indian state level data for 6 years, we plan to pin point major factors leading to the prolongation of polio. As pointed out by Fogel (1984) and Gauri and Khaleghian (2002), the two important determinants are poverty and education (literacy rate). Desai and Alva (1998) emphasized the role of maternal education in the Indian context. Both lower poverty and higher literacy rate are expected to reduce the count. Infant mortality is considered to be a widely used indicator for poverty. It is more important in this context because only children are affected by the polio. Infant mortality variable not only acts as a proxy for poverty level, but it also indicates the level of child care and child health condition (nutrition and resistance) in particular. High infant mortality rate, being an indicator of poor child care and high poverty level, is expected to increase the count.[5] Population itself may lead to an increase in the count. In addition to that, population density may also lead to an increase in the polio incidence simply because the chances of spread are much higher in a highly dense area. Expenditure on polio eradication is expected to lower the incidence.

Greenough (1995) and Nitcher (1995) have emphasized the role of general awareness and positive public attitude towards vaccination. Public resistance to immunization has also been noticed in some underdeveloped countries. This emerges from the belief that the vaccines are strategically meant to impede population growth among the minority population. The resistance also stems from the conviction that vaccination is an experimentation of the developed countries on the poorer countries and the results of such experiments are yet to be known. Mass campaign in favor of vaccination and support from the political leaders in this respect are important. See for example, Hull and Aylward (2001). In fact, the expenditure (cost) variable captures the coverage rate as well which also embeds political will and overall attitude towards vaccination. One would expect a significantly negative impact of this variable on polio incidence. We use a dummy variable for Uttar Pradesh which is considered to be the polio epicenter in India (dummy equals 1 if the observation comes from this state, 0 otherwise). We also add a dummy variable to capture the year 2002 effect (dummy equals 1 if the observation belongs to this year, 0 otherwise). The expected signs of these two dummy variables are positive. Figure 1 presents a raw plot for polio cases from 1999-2004.

---

[5] Note that although polio is not really a life threatening disease, we have taken a one period lag in the infant mortality variable, just to avoid any possible simultaneity bias.

Figure 1: Polio Cases



It is clear from the plot that the number of polio cases went sky-high during the year 2002. It also appears that polio cases in Uttar Pradesh account for a substantial number of polio cases in the country. We consider the number of polio cases (y_{it}) in a state i during the year t as the dependent variable. We also use various continuous and discrete regressors. Our explanatory variables include expenditure (measured in rupees) on polio eradication in state i at time t, child population in thousands (0-6 years) in state i at time t, population density in state i at time t (total population in the state per square mile radius), infant mortality rate (as an indicator of poverty and child care)[6] in state i at time t, female literacy rate in state i at time t and the two dummy variables. We have taken natural log of the above mentioned continuous variables while using them in the regressions. Our expenditure variable includes overall operational expenses as well.[7]

A natural methodological choice for modeling polio count data would be Poisson regression. Note that although our data set is a pooled one with small time period (state level data for 6 years) we haven't used any panel fixed effect heterogeneity parameter simply because our purpose is to capture explicitly the effects of Uttar Pradesh dummy and the year 2002 dummy[8]. Hence, essentially the analysis is similar to the cross section analysis. The probability function of Poisson model, $p(Y_{it})$ of Y_{it} can be stated as

[6] Number of death per 1000 infants.
[7] However, the detailed information on the break-ups is not available.
[8] Since fixed effect model does not allow us to add any special time invariant or individual invariant dummy variable in the regression, we will not be able to explore explicitly the effects of the aforementioned state and the year dummies in such framework. Hence, we stick to the pooled model with no heterogeneity term included, which effectively makes the analysis similar to a cross section analysis although we still get the benefit of a large (pooled) sample.

$$p(Y_{it} = y_{it} \mid x_{it}) = \frac{e^{-m_{it}} m_{it}^{\;y_{it}}}{y_{it}!} \quad , \quad y_{it} = 0,1,,2,\ldots \text{ with}$$

$$\ln m_{it} = x_{it} b \text{ , or } E(y_{it} \mid x_{it}) = m_{it} = e^{x_{it} b}$$

Where y_{it} is the dependent variable (polio count), x_{it} is the vector of regressors (including a constant term) as mentioned above and β represents the vector of parameters associated with the regressors.

Alternatively, following Cameron and Trivedi (1998), the standard probability function of y_{it} for negative binomial model can be stated as

$$p(Y_{it} = y_{it} \mid x_{it}) = \frac{\Gamma(y_{it} + a^{-1})}{\Gamma(y_{it} + 1)\Gamma(a^{-1})} \left(\frac{m_{it}}{a^{-1} + m_{it}} \right)^{y_{it}} \left(\frac{a^{-1}}{a^{-1} + m_{it}} \right)^{a^{-1}}$$

where a is called the dispersion parameter. The conditional variance of the dependent variable is given by $V(y_{it} \mid x_{it}) = m_{it} + a m_{it}^2$ and the conditional mean of the dependent variable is given by $E(y_{it} \mid x_{it}) = m_{it}$. One may test the null hypothesis $H_0 : a = 0$ against the alternative hypothesis $H_1 : a > 0$. The rejection of the null hypothesis provides evidence for the lack of equality between conditional mean and variance and hence justifies the use of negative binomial model as opposed to Poisson model. See Cameron and Trivedi (1998) for details.

We have used state level data for 27 Indian states and union territories for 6 years (1999-2004). The list is given in Table 1A. The data source is INDIASTAT database[9]. Thus our sample size is 162. For each state, six years average values (mean) for the dependent variable as well as for the independent variables are reported in Table 1A. In order to provide an overview, we also report six years average values for the same variables for the country as a whole in Table 1A (last row). A brief outline of the regional classifications is also presented in the same table.

[9] www.indiastat.com Only the data for density is constructed from data on state area which is collected from STATOIDS (http://www.statoids.com/uin.html). Data on all other variables are collected from INDIASTAT. We had to omit few small states and union territories due to lack of data availability. Data on infant mortality rate are missing for some years for Nagaland. We have used the average trend based on the other available years.

Table 1: State-wise and overall country-wide six years average values for the variables considered

	States & union territories	Polio Cases	Expen diture (M)	Child Population (thousand)	Pop. density (sq.m)	Literacy Rate (female)	Inco me (B)
N.	Bihar	63	153.7	20124	2682	33.57	371.9
	Chandigarh	0.67	1.00	112	20992	76.65	26.6
	Delhi	18	24.3	205	25654	75	396.1
	Haryana	11.67	33.44	3131	1186	56.31	285.6
	Himachal Pradesh	0	9.77	874	321	68.08	785.1
	Uttar Pradesh	430	37.53	32340	1876	42.98	985.1
S.	Andhra Pradesh	7.16	59.16	9806	722	51.17	786.1
	Goa	0	27.30	175.09	1158	75.51	45.4
	Karnataka	11	40.32	6866	716	57.45	614.2
	Kerala	0.17	21.52	3747	2176	87.86	351.5
	Pondicherry	0	1.14	134	6076	74.13	252.3
	Tamil Nadu	1.67	470.4	6859	1244	64.55	789.6
E.	Orissa	1.17	35.41	5123	603	50.97	211.7
	West Bengal	18	89.81	11171	2349	60.22	833.6
W.	Gujarat	6.5	33.78	6690	651	58.60	692.0
	Maharashtra	6.83	68.13	12601	778	67.51	1382.9
C.	Madhya Pradesh	8.5	110.0	14383	688	50.28	629.0
N. E.	Arunachal Pradesh	0	6.57	227	38	44.24	110.3
	Manipur	0	3.67	342	303	59.70	195.2
	Meghalaya	0	4.94	502	292	60.41	23.9
	Nagaland	0	3.54	247	274	61.92	21.8
	Sikkim	0	1.92	83	211	61.46	6.8
	Tripura	0	4.23	526	970	65.41	44.3
N. W.	Jammu & Kashmir	0.17	11.91	1451	119	41.82	77.4
	Punjab	2	24.17	3006	1228	63.55	360.4
	Rajasthan	10.5	93.61	10179	416	44.34	459.6
S E. I.	Andaman Nicobar	0	1.18	50	126	75.29	6.5
	India	603	1325	156799	803	54.16	10833

Expenditure in Million of Rupees Income in Billion of Rupees (1Bn=1000 Million). Regions: N. North, S. South, E. East, W. West, C. Central, N.E. North East, N.W. North West, S.E. I. South-East Island. Pop Density is total per square mile. The corresponding value per square kilometer is the result of multiplying this value by

3. Results, goodness of fit tests and sensitivity analysis:

A. Results: Regression results for Poisson model are reported in Table 2A.

First we include all the variables. While the coefficients of expenditure, population, population density, and the two dummy variables have the expected signs (they are also significant at 1% level), the literacy and the infant mortality rate variables are significant only at the 5% level with their expected signs. Suspecting some presence of collinearity between the literacy rate and the poverty (infant mortality rate)[10], we drop the literacy variable. Then the coefficient of the infant mortality variable increases although it continues to remain significant at the 5% level.

Alternatively, if we drop infant mortality variable, the value of the literacy variable goes up (and the significance level also changes from 5% to 1%). We have also tried income variable (income per thousand child population) as an alternative to infant mortality variable, and we find no significant impact of this variable when literacy variable is present in the regression. Only after dropping the literacy variable, we find significant impact of income variable (with the expected negative sign) pointing toward some collinearity.[11]

Inclusion of income variable instead of infant mortality variable does not change the sign and the significances of the other variables. The coefficient of the cost variable changes only by a negligible amount as well. Gauri and Khaleghian (2002) argue that it is not the income variable but the female education which plays an important role in polio eradication in India[12].

Our results are consistent with their findings. For infant mortality variable however, we always get the expected sign and it always turns out to be significant. The overall fit also improves when we use this variable instead of income variable. Since the standard R^2 measure does not accurately reflect the goodness of fit of the Poisson model, we have reported pseudo R^2 or LR statistics and the log likelihood. However, in our sample about 57% of the observations are zeros for the dependent variable. Also, the test for overdispersion rejects the null hypothesis of no overdispersion (see Cameron and Trivedi, 1998) and the estimated value of the dispersion parameter α turns out to be significantly positive. We therefore, resort to a negative binomial model. The results from this model are reported in Table 2B.

---

[10] Simple correlation coefficient among these two variables is -0.64.
[11] For brevity, we don't report this result.
[12] We have also tried overall (male and female combined) literacy rate variable but it turns out to be less significant than the female literacy rate variable.

Table 2A: Dependent variable is polio incidence count (Poisson model)

constant	-9.39 ** (3.46)	-17.45*** (2.20	*-7.07** (3.12*
Expenditure	-0.55 ***(0.05)	-0.55 *** (0.05)	*-0.58 ***(0.05)*
Infant Mortality	0.012**(0.006)	0.017 ** (0.007)	-
Child Population	1.33 *** (0.14)	1.62 *** (0.19)	*1.375 *** (0.13)*
Population Density	0.87 *** (0.17)	0.96 *** (0.17)	*0.773 *** (0.16)*
Literacy Rate	-1.13 ** (0.53)	-	*-1.401 ***(0.51)*
2002dummy	1.64 *** (0.16)	1.63 *** (0.15)	*1.691 *** (0.15)*
Uttar Pradesh	1.93 *** (0.19)	1.61 *** (0.25)	*2.164 *** (0.19)*
Log likelihood	-670	-692	*-686*
Pseudo R^2	*0.928*	*0.926*	*0.927*

Table 2B: Dependent variable is polio incidence count (negative binomial model)

constant	-14.00 ***(3.40)	-18.24***(1.38)	*-9.09 **(3.64)*
Expenditure	-0.56 ***(0.16)	-0.54 ***(0.15)	*-0.62 ***()0.16*
Infant Mortality	0.02**(0.01)	0.03***(0.01)	-
Child Population	1.61 ***(0.20)	1.64*** (0.19)	*1.74*** (0.18)*
Population Density	1.00 *** (0.14)	0.99 *** (0.14)	*0.93***(0.13)*
Literacy Rate	-1.00 *(0.62)	-	*-1.86 ** (0.81)*
2002dummy	1.14***(0.34)	1.16***(0.34)	*1.22 ***(0.35*
Uttar Pradesh	1.16*** (0.33)	1.25*** (0.33)	*1.21***(0.36)*
Log likelihood	-305	-306	*-307*
a (dispersion parameter)(t-stat)	*1.64***(4.55)*	*1.67***(4.77)*	*1.67***(4.64)*

Note for Table 2A and Table 2B: Heteroscedasticity corrected robust standard errors are reported below the coefficient estimates. Total number of observations is 162. * implies significance at 10% level, ** indicates significance at 5% level and *** denotes significance at 1% level. We have used log transformations for the continuous variables.

The results are very similar for Poisson and negative binomial models as far as the signs and the significances of the variables are concerned. Here also the significance of the literacy variable increases from 10% to 5% if the infant mortality variable is dropped. Significance of the infant mortality variable also increases from 5% to 1% if we drop the literacy variable. Both the models estimate the expenditure elasticity of polio count to be around negative 0.56.[13] Since the expenditure variable has been used in log form, the estimated coefficient of this variable can be interpreted as a measure of elasticity. However, the log likelihood improves much in case of negative binomial model as expected. All the above specifications show that the two dummy variables (Uttar Pradesh state dummy and 2002 year dummy) are always highly significant (at 1% level).

[13] When all the variables are present in the regression.

B. Overall Goodness of Fit:

Following Cameron and Trivedi (2005), we have performed a chi-square goodness of fit test for the base regression model (full sample with all the variables including Uttar Pradesh dummy). The null hypothesis here is that the theoretical frequencies are equal to the observed frequencies. It tests the extent to which the fitted frequencies differ from the observed frequencies. Rejection of the null, therefore points towards the fact that fitted frequencies are significantly different from the observed ones, implying a poor fit. The test statistic for the negative binomial model obtains a value around 14.20 with a p-value around 0.11., which fails to reject the null, implying a good fit for the negative binomial model.[14] The (Pearson) goodness of fit test for Poisson model obtains a value of 1528 with a p-value of 0.000, which clearly rejects the Poisson model.[15] We have also tabulated the actual and the fitted frequencies. These tabulated values for polio cases up to 9 are presented below. Table 1B clearly shows that Poisson model considerably underpredicts zero counts and the negative binomial predicts them quite well.

Table 1B: Relative Frequency in percentage term (Full Sample)

Cases	Actual	Poisson Fitted	*Negative Binomial Fitted*
0	56.79	38.62	*54.12*
1	6.79	12.24	*9.62*
2	5.56	8.27	*5.20*
3	4.32	6.22	*3.6*
4	3.09	4.59	*2.74*
5	1.23	3.39	*2.16*
6	0.62	2.58	*1.78*
7	1.23	2.05	*1.48*
8	1.23	1.69	*1.26*
9	*0.62*	*1.43*	*1.08*

The table above shows that NB2 is a better parametric model for this data.[16]

C. Sensitivity Analysis for Uttar Pradesh:

For a sensitivity analysis we drop Uttar Pradesh dummy variable in the base regression (with all other variables included) and the log likelihood drops substantially from -670 to -912 in case of Poisson and from -305 to -307, in case of negative binomial, as expected.[17] We then test the null hypothesis that the Uttar Pradesh dummy is not a significant factor. The computed chi-square for the negative binomial is 2*[307-305] = 4

[14] The test is described in Cameron et al (2005) and Cameron et al (1998). The program for the negative binomial goodness of fit test is available at http://cameron.econ.ucdavis.edu/mmabook/mmaprograms.html. the author would like to thank Colin Cameron for his useful suggestions in this regard.

[15] Pearson Goodness of fit statistic is reported by Stata. Also note that the deviance is 1057 with a p-value of 0.000.

[16] See Cameron et al (2005). The remaining actual and fitted frequencies are associated with counts 10 and above.

[17] The details of these results are not reported.

and the chi-square critical value for one degree of freedom is 3.84. We therefore, reject the null, implying that Uttar Pradesh is indeed an important explanatory variable.[18]

We then exclude Uttar Pradesh from our sample and repeat the main analysis. Results for this reduced sample are reported in Tables 2C and 2D. For this new sample, we consider the same set of explanatory variables as before but Uttar Pradesh dummy variable.

Table 2C: Dependent variable is polio incidence count (Poisson model)

constant	-9.18***(3.55)	-17.58***(2.14)	*-6.50**(3.02)*
Expenditure	-0.53***(0.11)	-0.54***(0.12)	*-0.53***(0.11)*
Infant Mortality	0.014*(0.008)	0.02***(0.008)	-
Child Population	1.29***(0.17)	1.58***(0.18)	*1.32***(0.15)*
Population Density	0.89*** (0.19)	1.00*** (0.18)	*0.77***(0.16)*
Literacy Rate	-1.15***(0.43)	-	*-1.44*** (0.39)*
2002dummy	1.06*** (0.21)	1.09*** (0.22)	*1.04*** (0.22)*
Log likelihood	-564	-585	*-575*
Pseudo R^2	*0.68*	*0.67*	*0.678*

Table 2D: Dependent variable is polio incidence count (negative binomial model)

constant	-14.19***(3.74)	-18.54***(1.48)	*-9.03** (3.74)*
Expenditure	-0.58*** (0.19)	-0.56*** (0.18)	*-0.60*** (0.18)*
Infant Mortality	0.025* (0.014)	0.03** (0.013)	-
Child Population	1.62*** 0.22	1.64*** 0.21	*1.73**** *0.20*
Population Density	1.02***(0.14)	1.02***(0.14)	*0.94***(0.13)*
Literacy Rate	-0.92 (0.69)	-	*-1.90** (0.84)*
2002dummy	1.09*** (0.37)	1.11*** (0.37)	*1.15***(0.39)*
Log likelihood	-264	-264.2	*-266*
a (dispersion parameter)(t-stat)	*1.83***(4.69)*	*1.86***(4.77)*	*1.88***(4.82)*

Note for Table 2C and Table 2D: Heteroscedasticity corrected robust standard errors are reported below the coefficient estimates. Total number of observations is 156. We exclude Uttar Pradesh from the sample here. * implies significance at 10% level, ** indicates significance at 5% level and *** denotes significance at 1% level. We have used log transformations for the continuous variables.

We find that the results are similar to those in Tables 2A and 2B as far as the signs and the significances of the coefficients are concerned. This is not surprising because for our full sample, we still control for Uttar Pradesh dummy variable. However, the new sample

[18] This test is based on Greene (2003) (page 746).

reveals some interesting facts.[19] The significances of the infant mortality and the literacy variables fall in this case although each of the variables remains significant at the 5% level, once the other variable is dropped (for the negative binomial model). The dispersion parameter is still positive and highly significant, pointing towards overdispersion. It is interesting to note that the coefficient for the cost variable (cost elasticity) increases in absolute value for the negative binomial model, as expected. Also the log likelihood improves much in both Poisson and negative binomial frameworks.

The negative binomial chi-square goodness of fit statistic (while testing the null that the actual and the fitted frequencies are similar) obtains a value of 8.40 with a p-value of 0.49, implying a good fit. For Poisson, deviance and Pearson goodness of fit statistics yield values of 889 and 1215 respectively with p-values being 0.000 in both the cases, implying a poor fit. Similar story is reflected while comparing tabulated values of the actual and the fitted frequencies of these two models for the reduced sample.[20]

4. Conclusion

This paper focuses on the factors contributing to prolongation of polio in India, still one of the largest polio affected countries in the world. To our knowledge, this is the first econometric attempt to look at various factors causing partial failure of the polio eradication program in India, undertaken by international agencies. We find that Uttar Pradesh is the worst polio affected state. Uttar Pradesh is an economically and socially backward state. Bihar is also one of the backward (in terms of social and economic indicators) northern states in India and it also registers a high number of polio cases. However, on an average, no state comes even close to Uttar Pradesh as far as polio incidence is concerned.[21] Therefore, it is beyond doubt that this state needs some special attention.

Also we find significant evidence of a sudden outbreak in the year 2002. Thus the government and the international agencies can not afford to be sluggish at any stage until the disease is completely exterminated. Even after we control for the year 2002 effect and Uttar Pradesh dummy, we still find that poverty adversely affects the epidemic. This is in conformity with the existing literature. Female literacy rate also has an important role to play and this is partly because people have unknown fear against the vaccines and some basic level of literacy may help reducing it. This indicates that the eradication program should include more vigorous awareness programs for the people. As expected, expenditure on polio has a significantly negative effect (with elasticity around negative 0.56 for the full sample and around 0.58 for the reduced sample) on the polio count. A more aggressive eradication program with a higher coverage rate is therefore, called for.

---

[19] We are thankful to an anonymous referee for drawing our attention towards this.
[20] This table is not reported as it gives a similar conclusion as in Table 1B.
[21] We have also used dummy variables for some of the northern states including Bihar but they are not significant.

References:

Cameron A.C. and P.K. Trivedi (1986), "Econometric Models Based on Count Data: Comparisons and Applications of Some Estimators and Tests", *Journal of Applied Econometrics*, 1,1,29-53.

Cameron A.C., P.K. Trivedi, F. Milne and J. Piggott (1988), "A Microeconometric Model of the Demand for Health Care and Health Insurance in Australia", *Review of Economic Studies*, 55, 1, 85-106.

Cameron A.C. and P.K. Trivedi (1998), *Regression Analysis for Count Data*, Econometric Society Monograph No.30, Cambridge University Press.

Cameron A.C. and P.K. Trivedi (2005), *Microeconometrics: Methods and Applications*, Cambridge University Press.

Desai S. and S. Alva (1998), "Maternal Education and Child Health: Is There a Strong Causal Relationship?", *Demography*, 35(1), 71-81.

Fogel R.W. (1984), "Nutrition and the Decline of Mortality since 1700: Some Preliminary Findings", *NBER Working Paper No. 1402*.

Gauri V. and P. Khaleghian (2002), "Immunization in Developing Countries: It's Political and Organizational Determinants", *World Development*, 30, 12, 2109-2132.

Greene W.H. (2003), *Econometric Analysis*, Fifth edition, Prentice Hall.

Greenough P. (1995), "Global Immunization and Culture: Compliance and resistance in large Scale Public Health campaigns", *Social Science and Medicine*, 41(5), 605-607.

Hull H.F. and R.B. Aylward (2001), "Progress Towards Global Polio Eradication", *Vaccine*, 19, 4378-4384.

Mascarenas A., J. Salinas, A. Tasset-Tisseau, C. Mascarenas and M.M. Khan (2005), "Polio Immunization Policy in Mexico: Economic Assessment of Current Practice and Future Alternatives", *Public Health*, 119, 542-549.

Muhe L., P. Byass, L. Freiz, A. Sandstrom, and S. Wall (1996), "A One Year Community Study of Under-fives in Rural Ethiopia: Health and Behavioral Determinants of Morbidity", *Public Health*, 110, 215-219.

Nitcher M. (1995), "Vaccinations in the Third World: A Consideration of Community Demand", *Social Science and Medicine*, 41(5), 617-632.

Preston S.H. (1976), *Mortality Patterns in National Populations*, New York: Academic Press.

Samandari T., B.P. Bell, G.L. Armstrong (2004), "Quantifying the Impact of Hepatitis A Immunization in the United States, 1995-2001", *Vaccine*, 22, 4342-4350.

WHO, *Polio Case Count*, http://www.who.int/immunization_monitoring/en/diseases/poliomyelitis/case_count.cfm

WHO, *Poliomyelitis*, http://www.who.int/topics/poliomyelitis/en/

NDIASTAT database: www.indiastat.com