CH-6.26

RF-CH-6.4 SUDHA

THE DETERMINANTS OF INFANT MORTALITY IN REGIONAL INDIA

By

Michael Beenstock

and

Patricia Sturdy

City University Business School

October 1986

I Introduction

There is a well established empirical tradition in which multiple regression techniques are used to investigate the socio-economic determinants of life expectancy or disaggregated mortality rates, e.g. Da Vanzo (1985), Krishnan (1975) and Beenstock (1980). These studies have typically been conducted using international cross section data although there are exceptions where regional aggregates have been used for a given country. In the latter case, regional variations in mortality rates and various socio-economic variables are studied to determine the factors that systematically account for these variations.

A recent addition to the literature is Jain (1985), who has investigated the data generated by the Survey of Child and Infant Mortality which was carried out in India in 1978. He regressed infant mortality rates on a range of socio-economic variables using state-wise data for rural India. Ruzicka (1984) has used more informal techniques for analysing these data. Here, we present our own analysis of this survey which supplements Jain's efforts in several respects.

First, we show that a richer statistical model of infant mortality can be estimated if the class of causative variables is first converted into factor scores using factor analysis, see Lawley and Maxwell (1971). This approach is useful because many of the causative variables which we investigate are colinear. We therefore find that a broad range of causative variables can be incorporated into the model whereas Jain finds that the range is much narrower.

Secondly, to isolate the effects of e.g. education on infant mortality, we propose a more direct test than the one proposed by Jain. This consists of investigating separately the infant mortality rate among children whose mothers are educated and the infant mortality rates among children whose mothers are not educated. Thirdly, because infant mortality rates are naturally bounded between zero and a thousand, linear regression is not strictly appropriate We therefore experiment with various non-linear transformations of the data on which basis we find that a semi-logistical model provides a superior description of the data. It is well known e.g. (Wyon and Gordon 1971), that in certain Indian states female infant mortality rates are greater than that of their male counterparts. One possible explanation for this is that females respond to the causative variables (e.g. education of the mother)in a quantatively different way to males, i.e. the female model coefficients happen to be different from the male model coefficients. An alternative hypothesis is that there is a genuine preference for male infants and that the causative variables are **not** responsible for differences in male and female infant mortality rates. In Section II we propose a methodology which enables us to discriminate between these competing hypotheses. These and related empirical results are reported in Section III. Methodological issues are described in Section II.

II Methodology

Basic Hypotheses

As in e.g. Jain (1985), a stochastic model is proposed in which the infant mortality rate, at a given point in time, in state i is hypothesised to be determined by a vector of socio-economic variables X, i.e.

$$IMR_i = F(X_i, u_i) \qquad (1)$$

where u_i is a stochastic term, IMR denotes the infant mortality rate and $X_i = (X_{1i}, X_{2i}, \ldots X_{Ki})$.

The constitution of X_i is in the last analysis an empirical matter. However, following earlier research we experiment with the following variables that are included in the Survey of Child and Infant Mortality or are obtained from other sources as described in the appendix.

X_1 = availability of medical facilities, % village with medical facilities greater than 5km distant (+)

X_2 = medical attention at birth, % births attended by trained medical staff (-)

X_3 = nutrition, % population consuming less than 2,100 calories per diem per capita (+)

X_4 = clean drinking water, % population using tap as main source of drinking water (-)

X_5 = poverty, % sample households with per capita monthly household expenditure below 50 Rupees (+)

X_6 = literacy, % of adult female literates (-)

X_7 = vaccination, % female infants given DPT vaccinations (-)

X_8 = Hindu, % population Hindu (?)

X_9 = Muslim, % population Muslim (?)

X_{10} = caste, % population in scheduled caste (?)

X_{11} = tribe, % population in scheduled tribe (?)

X_{12} = overcrowding, % households with one room only (+)

The signs of partial derivations (F_i) have been indicated in parentheses e.g. the larger the proportion of the population not covered by medical facilities, the higher is likely to be the infant mortality rate. In the past, it has proved difficult to estimate equation (1) because of high degrees of colinearity between many of the causative variables. Similar problems beset our own data. Therefore, instead of estimating equation (1) we estimate

$$IMR = G(Z_{ji}, v_i) \qquad j=1,2,\ldots,j< K \qquad (2)$$

where v_i is a random disturbance term and

$$Z_{ji} = \sum_{k=1}^{K} w_{jk} X_{ki}$$

is the factor score of the j'th factor and w_{jk} are the (rotated) factor loadings, see e.g. Maddala (1977)

It follows from this that

$$\frac{\delta IMR}{\delta X_k} = \frac{\delta F}{\delta X_k} = \Sigma_j w_{jk} \frac{\delta G}{\delta Z_j} \qquad (3)$$

is the response of the infant mortality rate to changes in the k'th variable.

Factor analysis of the K variates generates J significant factors (where J < K) which are orthogonal to each other. Thus equation (2) consists of only J independent regressors rather than K correlated regressors as in equation (1). Below we report various estimates of equation (2) and calculate the partial derivatives in terms of equation (3).

Functional Forms

Previous research has on the whole paid little attention to the functional form of equation (1). Jain e.g. estimates equation (1) in terms of a linear model, yet the infant mortality rate is naturally bounded between zero and 1,000. Before these natural limits are reached it seems appropriate that the estimated model should rule out very high or very low infant mortality rates. This is illustrated on fig. 1 where the crosses represent e.g. cross section observations of state-wise infant mortality rates with respect to some positive valued variable, X. Linear regression

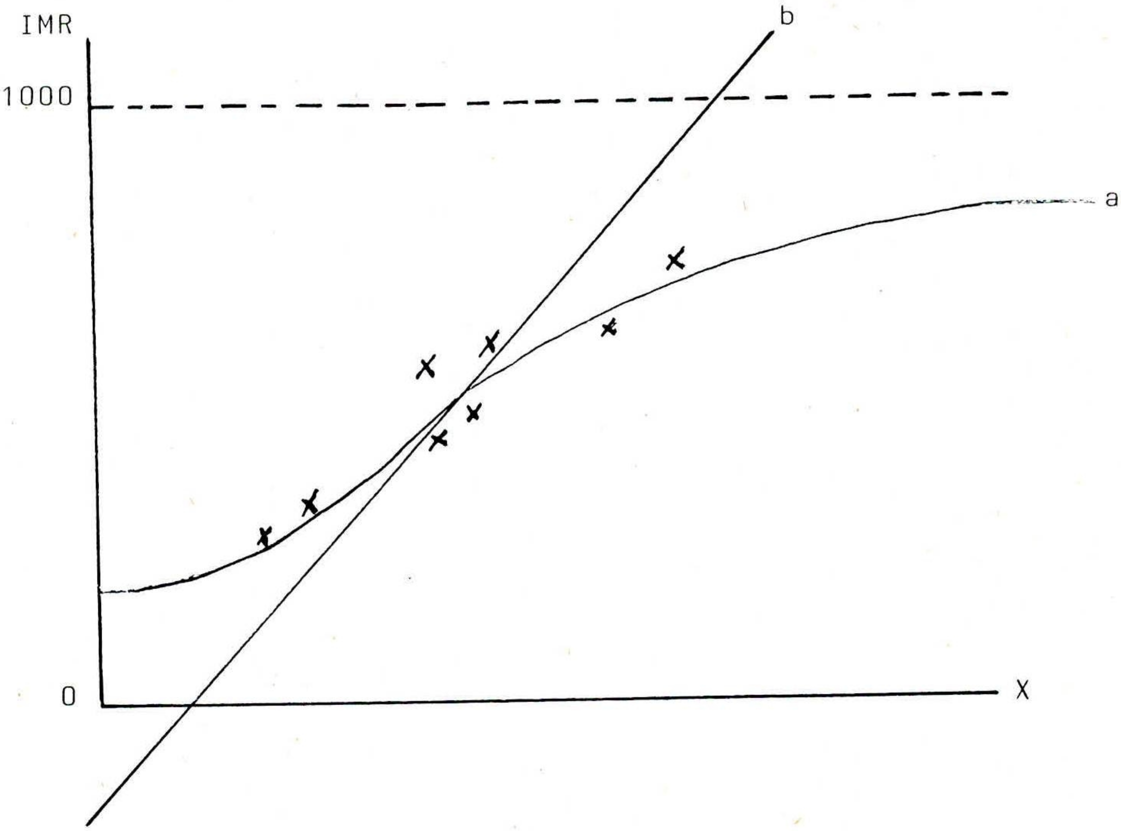

Fig. 1 Choice of Functional Form

would generate a regression line such as b) which implied that the infant mortality rate could either be greater than 1000 or below zero. In contrast, schedule a) implies that the infant mortality rate has unknown upper and lower limits that are below 1000 and greater than zero respectively. Moreover, if schedule a) is indeed the appropriate functional form, the linear model will generate inefficient parameter estimates, because it will not account for the outlying observations.

Below we hypothesise a logistical function of the type:-

$$\ln\left(\frac{IMR/1000}{1-IMR/1000}\right) = \alpha + \beta_k X_k + u \qquad (4)$$

which implies that the infant mortality rate is asymptotically bounded by 1000 $(1+e^{-\alpha})^{-1}$ and 1000. Indeed, we find this fits the data better than the linear model.

Data Control

Suppose we wish to estimate the effect of education on infant mortality using the adult literacy rate as an appropriate proxy. One way of doing this is to estimate equation (1) where one of the X variables is the literacy rate and where IMR is defined for the population as a whole. A second way is to define IMR in terms of the literacy of the parents, to omit the literacy rate as a regressor and then to estimate the following models

$$\ln\left(\frac{IMR_h/1000}{1-IMR_h/1000}\right) = \alpha_h + \beta_{hk}X_k + U_h \qquad h=1,\ 2 \qquad (5)$$

where

IMR_1 = infant mortality rate of illiterate parents

IMR_2 = infant mortality rate of literate parents

If indeed literacy lowers the infant mortality rate, we should find that $\hat{\alpha}_2 < \hat{\alpha}_1$.

Fortunately, the Survey of Child and Infant Mortality controls the data in this way. Indeed, it controls it for the literacy of the mother as well as a range of other variables including source of drinking water, age at marriage and the employment status of mothers. However these controls are not integrated, i.e. we cannot distinguish literacy and drinking water supply simultaneously. We may therefore compare and contrast statistical models of infant mortality for different sub-groups of the population. In this way we are likely to obtain a clearer picture of whether literacy etc. exerts an independent effect on infant mortality.

This methodology lends itself to tests of alternative hypotheses of whether the female infant mortality rate in India is greater than the male one. In certain Indian States, and indeed in our data the female infant mortality rate appears to be relatively high. This may either be due to a pro-male bias or it may reflect the parameters and the variables that are implicit in the causative model for infant mortality. Suppose for illustrative purposes there were two causative variables X_1, and X_2 and the infant mortality models were different for males and females as represented by equations (6) and (7)

$$IMR_f = \alpha_1 X_1 + \alpha_2 X_2 + \alpha_o \qquad (6)$$

$$IMR_m = \beta_1 X_1 + \beta_2 X_2 + \beta_o \qquad (7)$$

IMR_f = female infant mortality rate

IMR_m = male infant mortality rate.

Equations (6) and (7) imply that we can decompose differences in the infant mortality rates as follows:

$$IMR_f - IMR_m = (\alpha_1 - \beta_1)X_1 + (\alpha_2 - \beta_2)X_2 + \alpha_o - \beta_o \qquad (8)$$

Equation (8) indicates that in any particular state differences in the infant mortality rates reflect three possible factors. First insofar as α_1 exceeds β_1 and α_2 exceeds β_2 the female infant mortality rate will be greater than the male infant mortality rate for given values of the causative variables. Secondly whatever the values of the parameters the female infant mortality rate may exceed the male infant mortality rate if the X variables happen to assume particular values. Thirdly if $\alpha_o > \beta_o$ there is an all India pro-boy bias, since even if $\alpha_1 = \beta_1$ and $\alpha_2 = \beta_2$ the female infant mortality rate would exceed its male counterpart.

Below we implement this methodology and test the hypothesis that α_o is significantly greater than β_o. We also estimate independent estimates of $\alpha_1, \alpha_2, \beta_1$, and β_2.

III Empirical Results

Jain's Model

Jain (1985) viewed the problem of infant mortality at three levels: the village level, the household level and the individual level. Using the same variables as Jain, we were essentially able to replicate his results. Table 1 shows how infant mortality may be explained by DPT vaccination, poverty and female literacy. However, it was possible to improve the fit of the model by the inclusion of caste (model 2). It is interesting to note that Jain found a positive relationship between the presence of medical facilities and the infant mortality rate yet when usage of medical facilities was considered it showed a strong negative correlation. Therefore we regard our model 2 in table 1 as a substantial advance on Jain's efforts.

Nevertheless, these results were somewhat disappointing, as we were unable to estimate the possible part played by other independent variables. For example, neither Jain nor ourselves were able to bring the 'clean drinking water' variable into the model. Initial bivariate regressions gave a Pearson coefficient of 0.061. However, through the factor score approach it was possible to estimate the importance of tap water and other variables upon infant mortality.

Table 1

Jain Original Multiple Regression Model with Other Models Based on it

Dependent variable = IMR 1978

Independent variables	Jain's Model β	Model 1 β	Model 1 σ	Model 2 β	Model 2 σ
					SE of regr. coef
High school	-0.079				
DPT vaccination	-0.576* negative slope.	-2.181	(1.817)	-1.992	(0.959)
Poverty	0.437*	0.87	(0.432)	0.924	(0.311)
Adult female literacy	-0.443*	-0.073	(0.656)	-0.56	(0.362)
Birth attendance	-0.008	0.089	(0.765)		
Presence of medical facilities	0.176	0.245	(0.398)		
% sick children seen in medical institutions				-0.331	(0.285)
Caste				1.641	(0.516)
R^2	0.765	0.682		0.847	
$\bar{R}^2$		0.524		0.77	
Standard error of estimate		21.83		15.152	
F	4.34	4.3		11.09	
Constant		109.94		103.46	

* Statistically significant at p = 0.1

β = regression coefficients

σ = standard error of β

Factor analysis - the 4 are the wtd averages of the orignal 12

The Factor Score Models

A four factor model was estimated from eleven or twelve explanatory variables as a data reduction exercise. This choice reflected the degrees of freedom available in the data. However, we have not explored whether a more parsimonious model is more suitable. Table 2 shows the results of this analysis. Factors 1, 2 and 3 contain at least one variable with a high factor loading. For example, adult female literacy and birth attendance in Factor 1, and tap water in Factor 3.

The factor scores for each of 16 states were calculated, and used as regressors in our analysis of infant mortality rates. The model was of the logistical type described in section I. The results are shown in table 3. Models A and B contain 12 and 11 variables respectively, and with the exception of 'poor nutrition', all the signs concur with a priori expectations. Although a better fit was achieved by the reestimation of Jain's equation through the drastic dropping of variables, here we present a model containing many independent variables from which we can estimate the importance of each in relation to infant mortality.

The multipliers in table 3 were computed from equation 3, i.e. the estimated parameters of the regression model were multiplied by the relevant factor loadings. These multipliers measure the effect of the underlying variables on the logarithm of the relative probability of dying to living. For example, in equation A, if there is a

1% increase in children who are vaccinated, the relative probability of dying falls by 23%. These multipliers therefore have a dimension of an elasticity. They estimate the percentage response of the relative probability of dying to the percentage cover of the explanatory variables.

The factor score models confirm the findings of Jain and ourselves insofar as they show that DPT vaccination, female literacy and poverty are important in infant mortality. Model A shows how the lack of medical facilities and attended births play an important part. For instance, for every 1% increase in the births attended by trained medical staff, the relative probability of dying falls by 18%. The factor score approach enabled us to show that clean drinking water could influence infant mortality. For every 1% increase in the population drinking tap water the relative probability of infants dying falls by 10%. Religious and sociological factors also appear to be important determinants. However, it would be difficult to speculate upon the reasons behind these findings. The relationships between caste and infant mortality have been investigated within the control groups. There was a positive relationship between people living in crowded housing conditions and infant mortality. It may be that in the rural situation crowding is less important than in the urban environment. For every 1% of households, living in one room, the relative probability of infants dying falls by only 2 to 6%.

It is noteworthy that the poor nutrition variable was the only one to have an unexpected sign. The nutrition data were collected in 1974 by the National Sample Survey and were included to give a more complete picture of the problem. However, it may be that there was accurate reporting in the educated states, such as Kerala, with the result that 'poor nutrition' would go hand in hand with high female literacy rates. Factor analysis shows that poor nutrition is always grouped with adult female literacy (table 2).

Factor Score Models

Table 2 Factor Analysis of Independent Variables

Variable	Factor Loadings (w_{jk}) Factor 1	Factor 2	Factor 3	Factor 4
Birth attendance	0.90183	-0.0627	0.21902	-0.22269
Lack of medical facilities	-0.14666	0.7805	-0.5066	0.0410
Vaccination	0.5482	-0.1124	0.7604	-0.0241
Adult female literacy	0.8441	-0.36613	0.1448	0.0327
Poverty	0.150	0.3267	-0.618	0.319
Tap water	0.0205	0.2385	0.8341	0.0650
Hindu	-0.238	0.7884	-0.115	0.3088
Muslim	0.1093	-0.6131	-0.0414	0.2829
Caste	-0.0200	-0.009	0.0313	-0.8976
Tribe	-0.308	0.260	-0.430	0.3985
Poor nutrition	0.8102	0.34118	-0.2322	0.2793
Crowding	0.1781	0.6833	0.1273	0.1378

FACTOR	Cumulative percentage variance explained	chi-square
1	48.3	171.71552
2	76.4	120.59752
3	89.4	92.66852
4	100.0	53.40533

Table 3 Logistical Regressions Using Factor Scores

Dependent variable ln(IMR/1000-IMR)

	Model A β	Model A σ	Model B β	Model B σ
Constant	-2.010		-2.004	
S1	-0.164	(0.057)	-0.167	(0.0533)
S2	0.124	(0.057)	-0.189	(0.063)
S3	-0.168	(0.067)	0.130	(0.054)
S4	-0.056	(0.061)	-0.409	(0.057)
$\bar{R}^2$	0.55		0.61	
σ	0.218		0.20	
Constant	-2.011		-2.006	
S1	-0.168	(0.0571)	-0.170	(0.052)
S2	0.122	(0.057)	-0.185	(0.062)
S3	-0.164	(0.066)	0.127	(0.0532)
$\bar{R}^2$	0.56		0.62	
Standard error of estimate	0.217		0.20	

Variables	Multipliers	
Birth attendance	-18	-18
Lack of medical facilities	20	
DPT vaccination	-23	-23
Adult female literacy	-21	-22
Poverty	10	10
Tap water	-11	-13
Hindu	13	12
Muslim	-10	-10
Caste	4	4
Tribe	13	12
Poor nutrition	-6*	-5*
Crowding	2	6

* Denotes unexpected sign

Controlled Data

Table 4 gives the results of factor score models for control groups. To test for the effect of literacy we now implement equation (5) to see whether the constant term is significantly higher for illiterate mothers than for literate mothers, reported in table 4 equations A and B. We find that the constant term for literate mothers is indeed smaller and suggests that the relative probability of infants dying in the case of literate women is 44% smaller than for illiterate women. The model fits better for literate than illiterate mothers. Adding the general female literacy as a regressor made little difference to either model, thus implying that the general literacy of females at the community level does not apear to affect the infant mortality rates of either literate or illiterate mothers.

A comparison was made between mothers who drink tap water and mothers who use other sources of drinking water (Table 4 equations E and F. Again, the constant term in the tap water model was lower indicating that the relative probability of infants dying for mothers drinking "unclean water" was 11% higher than for mothers who drink tap water. However, this difference was not statistically significant.

The infant mortality rate in mothers whose age at marriage was under eighteen was directly compared with the infant mortality rate of mothers whose age at marriage was over twenty one. Table 4(equations I and J)shows how when controlling for other independent variables the relative probability

Table 4: Factor Score Models for Control Groups

Dependent Variable: Ln((IMR Control)/1000-IMR(Contro

MULTIPLIERS

Variable	A IMR in Literates	B IMR in Illiterates	C IMR Caste	D IMR Hindu	E IMR Tap	F IMR Non Tap	G IMR Workers	H IMR Non-Wo
	Relative Probability of Dying/Living							
Birth attendance	-20	-16	-15	-21	1.6*	-22	-14	-25
Lack of Medical Fac	31	17.5	17	20	7	21	13	21
Vaccination	-42	-19.2	-16	-25	-12	-25	-18	-27
Adult Female Literacy			-14	-25	4.5*	-27	-15	-29
Poverty	31	4.8	6.3	9	12	6	7	6
Tap	-36	-7	-7	-12			-10	-11
Hindu	17	12.5	9		-3	14	8	14
Muslim	-2	-12.3	-3.5	-11	5.2	-13	-5	-12
Caste	8	9.8		7	3.5	10	2	8
Tribe	28	9.7	11	13	4.6	12	10	14
Poor Nutrition	0.7	-10*	7	-1*	6	-14*	-6*	16
Crowding	-1*	0.82	-1.6*	2	-8*	0.5	-0.7*	0.5
$\bar{R}^2$	0.59	0.51	0.067	0.62	0.31	0.53	0.08	0.54
Standard error estimate	0.33	0.23	0.316	0.22	0.22	0.28	0.32	0.29
Constant	-2.437	-1.993	-1.943	-2.058	-2.175	-2.054	-1.929	-2.207
σ	0.0894	0.0622			0.0666	0.0749	0.864	0.077
Significance +		2σ				0σ		1σ

+ Nσ is significantly different at N. Standard deviations

β Coefficients of 4 Factor Score regressors are available on request.

of infants dying is 38% greater when mothers marry before the age of eighteen.

The effect of female participation upon infant mortality was investigated with controlled data. Again the constant term was significantly higher for working mothers, thus indicating that female employment status exerts some independent effect upon infant mortality. The relative probability of infants dying was 27% higher among working mothers (Table 4 equations G and H).

The factor approach failed to explain the infant mortality rate of scheduled caste mothers (equation C). The reason for this may be that our factor score model did not contain variables that were relevant to infant mortality in this group. Similarly, for 'mothers drinking tap water', the model was also weak; in this particular model some of the signs on the relative probabilities were not as we would have expected. However, for mothers drinking non-tap water, the factor score model was stronger. In fact, this model and the Hindu controlled model could be readily explained by the factor score approach and were not dissimilar to model A.

Table 4 (equations K and H) also compares the male and female infant mortality models. The constant terms were not significantly different thus indicating that there was no apparent bias in favour of male infants at the all India level.

Functional Forms

The goodness of fit of logistical and linear models cannot be directly compared since the dimensions of the residuals are quite different from each other. The equation standard error of estimate from the linear four-factor model was 21.95 expressed in unit rates of infant mortality, whereas the

Table 5 Comparison of Linear and Logistical Models

	F	R	estimated standard error	transformed standard error (IMR)
Linear	4.73	0.516	21.95	21.95
Logistical	5.42	0.55	0.219	21.3

standard error in the case of the logistical transform of the infant mortality rate was 0.219. To see whether the logistical model is a better description of the data, we have appropriately transformed the fitted values of model into units of infant mortality and calculated the adjusted standard error of the transformed residuals.

As indicated in table 5, this turned out to be 21.3, which is smaller than its linear counterpart of 21.95. However, an F ratio test indicates that at $p = 0.05$ these two standard errors are not statistically significantly different from one another. On the other hand it is noteworthy that the logistical model, as intuition would suggest, fits the data better, but substantially larger samples would be necessary to establish this at conventional levels of confidence.

Conclusions

1. A re-estimation of Jain's regression model showed that DPT vaccination, adult female literacy, poverty, caste and the usage of medical facilities were important determinants of infant mortality in rural India. A better fitting model was achieved by dropping a number of his independent variables and adding others.

2. To assess the relative probability of dying to living attributed to each of the independent variables, a factor score model was estimated, in which all the signs were correct except for 'poor nutrition'. In addition to those

variables indentified by Jain, it was possible to show that tap water, birth attendance and sociological factors were important influences upon infant mortality.

3. Control groups were investigated by factor score modelling techniques. From these, we conclude firstly, that the illiteracy of mothers is a contributory factor to infant mortality and secondly, that the risk of infants dying is significantly higher when mothers marry before the age of eighteen. Mothers who work were also shown to run an increased risk of their infants dying. However no significant difference could be shown in infant mortality rates when mothers with different sources of drinking water were compared.

4. A comparison between male and female infant mortality rates was made through the factor score model. Using this approach no bias in favour of male infants could be detected at the all India level.

5. The choice of functional form was studied. A logistical model of the infant mortality rate was shown to explain the data more accurately than the linear model. However, these differences were not statistically significant.

The Data Appendix

I Survey on Infant and Child Mortality, 1979, The Registrar General, Ministry of Home Affairs, New Delhi, India.

Rural data were collected from 18 of the larger states of India. The Survey formed a sub- sample of the Sample Registration Survey and covered more than 500 thousand house-holds and included over 3 million people. The data were collected on a state-wise basis by a non-medical enumerator. Vital rates from Bihar and West Bengal are generally regarded as unreliable and were therefore excluded from the Survey (Jain, 1985).

The following data were used:-

1. The Infant Mortality Rate 1978. Jain used this for his dependent variable. We also used IMR 1978 to confirm Jain's findings and to estimate models 1 and 2. There were ovservations on 16 states.

2. The Average Infant Mortality Rate. An average infant mortality rate was derived from the results of the Sample Registration Survey between 1972 and 1976, and included in the 1978 IMR. This was considered to be more accurate for the dependent variable. This infant mortality rate was used for the factor score models.

3. IMR Literate. The infant mortality rate among literate mothers.

4. IMR Illiterate. The infant mortality rate among illiterate mothers.

5. IMR Caste. The infant mortality rate among scheduled caste mothers.

6. IMR Hindu. The infant mortality rate among Hindu mothers.

7. IMR Tap. The infant mortality rate among mothers using tap or hand pump as main source of drinking water.

8. IMR Non-Tap. The infant mortality rate among mothers not using tap water as a source of drinking water.

9. The Presence of Medical Facilities. % villages with medical facilities less than 2km distant. (Table 1 only)

10. The Absence of Medical Facilities. % villages with medical facilities more than 5km distant.

11. Usage of Medical Facilities. % distribution of sick children aged 0-6 years receiving attention in medical institution. (Table 1 only).

12. Tap. % population using tap as main source of drinking water.

13. Poverty. % sample households with per capita monthly household expenditure below 50 Rupees.

14. Adult Female Literacy. % literacy in females over the age of 15 years.

15. Total Female Literacy. % literacy in all females.

16. Vaccination. % female infants receiving DPT vaccination.

17. Over-Crowding. % households with one room only.

18. Muslim. % population Muslim.

19. Hindu. % population Hindu.

20. Caste. % population in scheduled caste.

21. Tribe. % population in scheduled tribe.

22. IMR Workers. The infant mortality rate among working mothers.

23. IMR Non-Workers. The infant mortality rate among mothers who do not work.

24. IMR Age at Marriage before 18 years. The infant mortality rate among mothers whose age at marriage was less than eighteen years.

25. IMR Age at marriage greater than 21 years. The infant mortality rate among mothers whose age at marriage was greater than twenty one years.

26. IMR Male. Average Male infant mortality rate 1972,74,76,78.

27. IMR Female. Average female infant mortality rate 1972,74,76 and 1978.

II Levels, Trends and Differentials in Fertility, 1979

The Registrar General, Ministry of Home Affairs, New Delhi, India (1982)

Attendance at birth, this study gave the percentage of rural births attended by trained medical staff. 18 observations were made on a state-wise basis.

III The Sample Registration Survey (1972 to 1976).

The Registrar General, Ministry of Home Affairs, New Delhi, India.
This is an ongoing survey which covers 2,400 sample units in rural areas. Vital rates were collected from 16 states (excluding Bihar and West Bengal) on a state-wise basis.

IV The National Sample Survey

Round 26, (July 1971 to June 1972). Report number 238. Volume 1. 1978. Calorie and protein values of food items in rural areas. The percentage population receiving less than 2,100 calories per diem per capita, in each state was used as the 'poor nutrition' variable. There were observations on 17 states.

References

Beenstock, M., 1980, Health, Migration and Development, Gower Press, Farnborough.

Da Vanzo, J., Habicht, J-p, Hill, K., Preston, S., 1985, Quantitative Studies of Mortality Decline in the Developing World, World Bank Staff Working Paper, Number 638, Population and Development Series, Number 8.

Jain, A.K., 1985, Determinants of Regional Variations in Infant Mortality in Rural India, Population Studies, 39, 407-424.

Krishnan, 1975, Mortality Decline in India, 1951-61: Development v. Public Health Program Hypothesis', Social Science and Medicine, 475-9.

Lawley, D.N., and Maxwell, A.E., 1971, Factor Analysis as a Statistical Method, 2nd edition, Butterworths, London.

Maddala, G.S., 1977, Econometrics, McGraw-Hill Inc., New York.

Ruzicka, L.T., 1984, Mortality in India: Past Trends and Future Prospects', in India's Demography, Essays on the Contemporary Population, edited by T. Dyson and N. Crook, South asian Publishers Pvt. Ltd., New Delhi.

Wyon, J.B. and Gordon J.E., 1971, The Khanna Study: Population Problems in the Rural Punjab. Cambridge, Mass. Harvard University Press.