Narayan Das

A Factor Analysis Model to Study Variation of Family Planning Performance in India

IN India, the acceptance rate of family planning not only varies between the states, but also between districts within each state. A study of the factors affecting family planning performance by treating a state or a district as a unit of observation seems, therefore, useful for making the family planning programme more effective. The causes of variation in the acceptance of family planning in the states and in the country as a whole are related in the main to socio-economic, demographic and family planning programme execution variables. Such studies of demographic and socio-economic characteristics of the acceptors might inform the programme administrator of needed efforts in terms of educational, motivational and family planning programme execution.

One approach to the study of causes of variation in family planning performance involves the use of regression analysis for measuring the combined effect of a number of variables on a dependent variable. Regression analysis, however, raises intractable problems where multicollinearity characterises the selected variables. Conclusions drawn from regression analysis rest on the assumption that it is possible to quantitatively separate the influence of each of the independent variables. This assumption may prove unrealistic when the selected variables are inter-correlated with each other. Following the lead of a study by Balasubramaniam, Das and others (1970), we propose to approach the problem through factor analysis; this enables us to derive a more fundamental set of orthogonal factors by grouping variables measuring incidence of one or more underlying common factors. The present paper is, thus, an attempt to identify meaningful orthogonal factors for use in studying the causes of variation in family planning performance among various states of India ; it presents an exercise in factor analysis of a correlation matrix generated by the inter-related socio-economic and family planning programme execution variables.

Data

The following socio-economic and programme execution variables have been selected for preparation of correlation matrix which constitutes the basic data for factor analysis:

Percentage of urban population;
Percentage of general literacy for ages 15 +;
Percentage of female literacy for ages 15 +;
Female mean age at marriage;
Per capita income;
Money allocation on family planning programme;
Medical personnel per ten thousand eligible couples;
Paramedical personnel per ten thousand eligible couples;
Urban centres per ten thousand eligible couples;
Primary health centres and sub-centres per ten thousand eligible couples.

The data used for the first four variables and per capita income are collected from the Census of India (1961) and the National Council of Applied Economic Research (1963) respectively. The data relating to money allocation on family planning and the rest of the programme execution variables (last five variables mentioned above) are collected from the Report of the Sixth Family Planning Conference, Bhopal (1969) and the Ministry of Health and Family Planning publications (1970) respectively. Since the figures on medical and para-medical personnel are incomplete with regard to such staff working with the urban centres, the primary health centres and sub-centres, the information about these has been obtained on the basis of certain assumptions (Das and Rama Rao, 1971) and the rates for medical and paramedical personnel are adjusted accordingly.

The acceptance rate of family planning per thousand eligible couples which has been considered as the dependent variable (Yp) is given by

$$Y_p = \frac{T_p}{W_{15-44(p)}} \times 1000, (p = 1, 2, \ldots, N)$$

where, T_p is the total performance in a state *(p)*, comprising the number of sterilizations performed, IUCD's inserted, and the estimated number of users of condom during the year. It has been assumed that a couple has actually used the contraceptive for the whole year. Accordingly, the number of users of condom is estimated by dividing the number of condoms distributed during the year by 72, where 72 is average coital frequency per couple during a year. $W_{15-44(p)}$ is the total number of currently married women in the age group 15-44 in a state *(p)*. The acceptance rates for different states of India have been computed for the years 1966-67 to 1969-70. They have been fluctuating from year to year (Das and Rama Rao, 1971), and so average acceptance rate for the period has been taken for the present analysis.

Model

The model consists of four steps. In the first step the correlation matrix is derived using all the selected variables. The second stage consists of extracting the underlying orthogonal factors from the correlation matrix derived in the first step, using factor analysis. The third step is to obtain linear regression estimate of each orthogonal factor from the factor matrix obtained in the second stage, and the fourth step is the use of these estimates, obtained in the third stage, to explain, through regression analysis, inter-state variation in family planning acceptance rate, in terms of underlying causal factors.

(i) *Preparation of correlation matrix*

The correlation matrix (r_{jk}), $j=l, 2, \ldots, n$; $k=l, 2, \ldots, n$ for the selected socio-economic and programme execution variables X_j, $j=1, \ldots, n$ is obtained by computing the product moment correlation $r_{i!c}$ in each case. Since there are 15 states in India according to 1961 census and the number of selected variables, *n* is eleven, 11x11 correlation matrix is computed using 15 observations (see Table 1).

(//) *Extraction of orthogonal factors*[1]

Factor analysis is used for finding out a weighted combination of the variables which predict a factor best. The model with which we accordingly start may be expressed as follows:

$$X_j = a_{i1} F_1 + a_n F_2 + \ldots + a_{mj} F_m + d_i U_j \quad (j = 1, 2, \ldots, n), \quad m < n$$

where X_j is standardized variable j (i.e. $x_j = \frac{X_j - \bar{X}_j}{\Sigma \sigma_j}$), F_i's are common factors (I=1, 2, . .. , *m*) and *U,* is a unique factor including error for variable *j*.

The following are assumed to hold among the factors:

$$\text{COR}(F_i, V_j) = 0 \quad (i = 1, 2, \ldots, m;\ j = 1, 2, \ldots, n)$$
$$\text{COR}(U_t, U_k) = 0 \quad (j = / k).$$

The basic problem in factor analysis is to determine the factor loading ai,'s of the common factors which account for correlations among variables. This determination could be made by several methods, but, for convenience, only the centroid method of Thurstone (1947) has been used for extracting the loadings of the underlying common factors from the correlation matrix (with unity in diagonal). Later, the factor matrix has been rotated graphically by orthogonal rotation method (Thomson, 1951) for identifying simple and interpretable orthogonal factors (see Table 2).

1 Principal component analysis which might have given a better result to orthogonalize the factors, has not been used since it raises difficulty in the identification of the factors.

TABLE 1

CORRELATION MATRIX OF THE SELECTED SOCIO-ECONOMIC AND FAMILY PLANNING PROGRAMME EXECUTION VARIABLES IN INDIA

Variables (X_j's)	1	2	3	4	5	6	7	8	9	10	11
1. The percentage of urban population1.000	1.000										
2. The percentage of general literacy for ages 15+	0.316	1.000									
3. The percentage of female literacy for ages 15+	0.270	0.971	1.000								
4. Female mean age at marri-age0.078	0.078	0.706	0.725	1.000							
5. Per capita income	0.579	0.522	0.436	0.271	1.000						
6 Money allocation on family planning0.218	0.218	0.379	0.500	0.397	0.155	1.000					
7. Medical personnel*	-0.038	-0.353	+0.236	-0.156	-0.390	0.671	1.000				
8. Para-medical personnel*	0.250	0.500	0.571	0.383	-0.099	0.654	0.357	1.000			
9. Urban centres*	0.491	-0.049	0.033	-0.310	0.246	0.029	0.066	0.097	1.000		
10. Primary health centres*	-0.042	-0.265	-0.217	-0.277	-0.460	0.507	0.875	0.368	0.120	1.000	
11. Sub-centres*	0.214	0.567	0.625	0.432	– 0.108	0.539	0.198	0.980	0.044	0.228	1.000
Mean	17.37	28.36	14.02	16.52	227.20	303.83	0.75	3.98	0.14	0.50	3.04
Standard deviation	6.72	10.37	9.45	1.68	46.41	43.65	0.21	1.17	0.06	0.16	1.09

*Per ten thousand eligible couples.

TABLE 2

THE FACTOR MATRIX OBTAINED FROM THE INTER-CORRELATION MATRIX OF SOCIO-ECONOMIC AND PROGRAMME EXECUTION VARIABLES IN INDIA

Variable (x'_js)	Factor loading $(a_{ij}'s]$ of the respective variable $(x_j's)$ on			Variance contributed by each variable	
	Factor F_1	Factor Ft	Factor F_3	Common	Specific
1. The percentage of urban population	.671	-.030	-.475	.677	.323
2. The percentage of general literacy for ages 15+	.873	-.026	.410	.931	.069
3. The percentage of female literacy for ages 15 +	.854	.093	.432	.925	.075
4. Female mean age at marriage	.604	.034	.539	.657	.343
5. Per capita income	.710	-.423	-.314	.784	.216
6. Money allocation on family planning	.484	.749	-.101	.806	.194
7. Medical personnel*	-.193	.859	-.379	.919	.081
8. Paramedical personnel*	.484	.749	.334	.907	.093
9. Urban Centres*	.374	.000	-.592	.490	.510
10. Primary Health Centres*	-.204	.839	-.329	.854	.146
11. Sub-Centres*	.490	.632	.475	.865	.135
Variance contributed by each factor	3.738	3.156	1.921	8.815	2.185

*Per ten thousand eligible couples.

(iii) *Measurements of orthogonal factors*

To obtain estimate of each orthogonal factor F_f in terms of observed variables x_j's, conventional regression method is employed. If the rank of the factor matrix (aii), $i = 1, 2, \ldots, m; j = 1, 2, \ldots, n$ derived in the second step, is m then any factor F_t may be expressed in the form:

$$F_i = \hat{a}_{i1} x_1 + \hat{a}_{i2} x_2 + \ldots + \hat{a}_{in} X_n \; (i=1,2,\ldots,m),$$

where p_j's are unknown. Since an the variables x_j's are standardized and $a_{ti} = r_{F_i x_j}$, where $r_{F_i}{}^{x}{}_{j}$.is the correlation between the i-th factor and j-th variable, the normal equation for the new model is simply

$$Rb_i = a_i \; (i=!, 2, \ldots, m)$$

where R is the observed correlation matrix (r_{ik}), j - 1, 2,. . . , n; k - 1, 2,. .. «; a_t is the column vector $(a_{i1}\} \; a_{i3}, \ldots, a_{in})$ and b_i's are the least square estimates of B's.

The solution of these equations is easily seen to be

$$b_i = R^{-1} a_i \; (i=1, 2, \ldots m)$$

Thus any factor can be estimated when the correlation of the variables with the factor concerned *(i.e.* the factor loading of each variable on the respective factor) and the correlations among the variables themselves are known (Harman, 1967).

Hence, the estimate of any one of the factors F,-'s may be put in the form:

$$Ft = bn\,Xi + b_{iz}x_a + \dots + b_{in}x_n \,(i — 1, 2, \dots, in).$$

(iv) *Application oforthogonalfactors in regression analysis*

Now, it is possible to 'orthogonalize' a regression situation for any arbitrary set of variables (Kcndall, 1957). According to this model, an observed family planning acceptance rate in standardized form is given by

$$y = \underset{j=1}{\overset{m}{S}} a,/^{1},- + 6$$

where F_t are the new orthogonal factors obtained from the variables *xfs;* a's are the linear functions of the b's; and å is a random variable such that each å is an independent $N(0, \sqrt{\sigma^2})$. Since the factors F_i's are orthogonal, the «'á can be obtained easily by the relation :

$$\alpha_i = \frac{\sum_{p=1}^{N} y_p F_{ip}}{\sum_{p=1}^{N} (F_{ip})^2} \quad \frac{\sum_{p=1}^{N} y_p F_{ip}}{N \lambda_i}$$

where *y,* and F_{ip} are the value of *y* and F_i respectively from the *p-th* state and A, is the variance of i-th factor. Putting the estimate of F_{ip},, and $A_i = 1$ (Since Fi's are standardized):

$$\alpha_i = \frac{1}{N} \sum_{p=1}^{N} \sum_{j=1}^{n} b_{ij}\, y_p\, x_{jp}.$$

Since the variables x_{j1}'s are standardized, their covariances are their correlations *(r)* so that

$$á_i = \sum^{n} b_i\, r_{yxi} \;(i=l, 2, \dots, m) \text{ where } r_{vxi} = \frac{1}{N} \sum_{p=1}^{N} y, x_{ip}.$$

Hence the proportion of variance (*R]*) accounted for by i-th orthogonal factor is given by

$$R_i^2 = \frac{1}{N}\, \alpha_i \sum_{p=1}^{N} y_p F_{ip} \text{ (since variance of } y \text{ is unity)}$$

$$= \alpha_i^2 \;(i = 1, 2, \dots, m).$$

It is therefore, easy to calculate the proportion of total variance accounted for by each orthogonal factor in explaining the causes of variation, with respect to family planning performance. The proportion of total variance explained, R^2 *is* given by

$$R^2 = \sum_{i=1}^{m} R^2{}_1 = \sum_{i=1}^{m} aj \ .$$

The results obtained by the above procedure are discussed in the next section.

Results and Discussion

Table 1 shows mean, standard deviation and correlation matrix for eleven selected variables Examination of the correlation matrix indicates that strong inter-relationship exists between the socio-economic variables. Similarly, among the Family Planning Programme execution variables, money allocation on family planning programme is S correlated to medical and paramedical personnel, primary health centres and sub-centres working for family planning programme. Facilities for family planning programme depend thus on the total amount spent on it in a particular state. At the same time , the performance of the family planning centres and sub-centres depends on the size of medical and paramedical personnel. There is also an overlap between the socio-economic and programme execution variables. Therefore, it would be in order to express all the selected variables in terms of some underlying common factors.

Application of factor analysis through centroid method, using the correlation matrix obtained in Table 1 as basic data and with unity in a row or column as commonality, suggests that there are three underlying common factors among the selected variables. since on each of the centroid factors, there is, more or less, substantial loading for each variable and identification of this factor is difficult, the factor matrix was rotated graphically taking the two factors at a time so that the underlying common factors could be identified The initial factor matrix,* leaving out the third factor (F_3), was rotated clock-wise through 56° to the new positions (see Table 2) so that the factors F_1, and F_2 could be interpreted.

The third factor (F_3) is kept unrotated as the rotation of the third factor (F_3) with either of the rotated factors (F_1 or F_2) could not give any satisfactory result.

The new position of factor matrix is not unique. It is, however, apparent from the rotated factor matrix that new position would be more adequate to identify the meaningful orthogonal factors.

It is evident from Table 2 that three orthogonal factors jointly explain 80 per cent

*The initial factor matrix: is not being presented because of the lack of space, the same can be obtained from author on demand.

of the total variance. The first factor (Fi) explains about 34 per cent of the total variance, the second (F_a) 29 per cent and the third *(F_3)* about 17 per cent. The first five variables and the variable 'urban centres' have high loading on the first (F_1), and almost zero on second factor (F_2). The variables-'money allocation on family planning', 'paramedical personnel' and 'sub-centres', have considerable loading on factor Ft but it is not as substantial as the one shown by the first five variables. Urbanization is obviously associated with economic development, more educational and family planning facilities and enlightened social atmosphere. Therefore, the first factor (F_1) is named 'urban environment'. The second factor (F_2) is called 'Family Planning Facility', since all the manipulative variables (last six variables in Table 2) excepting the variable 'urban centre', have high loadings on factor F_2 and, in common, measure the availabilitay of family planning facilities. The third factor (F_3) remains unidentified; it explains a small part of the total variance. Even so, it may be fruitful to explore these variables in order to understand and identify the third factor.

TABLE 3

THE VALUES OF THE REGRESSION COEFFICIENTS (b_{ij}'s) IN THE EQUATION PREDICTING THE- FACTOR (F_i) AND CORRELATION OF ACCEPTANCE RATE OF FAMILY PLANNING WITH THE VARIABLES (x_j's)

Variables (x_t's)	*Coefficient (b_{ij}'s) of the respective variable (Xj's) on*			*Correlation of acceptance rate with (r_{yij}) x'_j s*
	Factor F_1	*Factor F_2*	*Factor F_3*	
1. The percentage of urban population	.2526	-.0897	-.3237	.516
2. The percentage of general literacy for ages 15+	-.0035	.0903	.0062	.549
3. The percentage of female literacy for ages 15+	.4401	-.3099	.2392	.449
4. Female mean age at marriage	.2596	—.0549	—.1595	.324
5. Per capita income	.1988	—.1220	—.2591	.330
6. Money allocation on family planning	.0774	.4201	-.2278	.523
7. Medical personnel*	-.2033	.0886	-.2034	.237
8. Paramedical personnel*	.5014	.3422	.0069	.644
9. Urban Centres*	.2366	-.0190	-.3120	.422
10. Primary Health Centres*	.0913	.2673	-.1490	.368
11. Sub-centres*	-.4507	.1705	.4979	.607
Proportion of variance explained	.9251	.9749	.9996	

*Per ten thousand eligible couples.

Table 3 shows the standardized coefficients of the respective variables x_js on each of the three orthogonal factors F_is and the proportion of variance accounted for fay each orthogonal factor. The last column of the Table 3 shows the correlation of the family planning acceptance rate with the respective *Xj's*. All the variables are found to be significantly related to the family planning acceptance rate. The only surprising feature is the small magnitude of the correlation coefficient for the variable, medical personnel.

Table 3 gives, in effect, the input for regression analysis for measuring contribution of each orthogonal factor and identifying a factor which can be discarded as unimportant. This analysis shows (see Table 4) that of the three orthogonal factors, Urban Environment (F_1 accounts for 42.23 per cent and Family Planning facility (F_2), for

TABLE 4
THE INDEPENDENT CONTRIBUTION OF EACH OF THE ORTHOGONAL FACTORS (F_i's) ON ACCEPTANCE RATE *(Y)* OF FAMILY PLANNING IN INDIA

Factor *(F_i's)*	*Regression coefficients** *(α_i's)*			*The proportion of variance explained* *(F_i^2)*
F_i = Urban Environments	.64984+			.4223
F_2 = Family Planning facilities	.46102+			.2125
		$R^2_{v}.123$	=	.6348

*To test the significance of a_2's, student *t*-test is applied.
+Significant at .01 level

21.25 per cent of the total variance. Since factor F_3 was not found to be significantly related to family planning acceptance rate, it was dropped from regression analysis. The multiple R for the two orthogonal factors on acceptance rate was found to be 0.79675: therefore, the selected variables explain 63.48 per cent, of the total variance. However, with an R^2 of 63.48 per cent, the prediction equation needs to be improved. New variables should be found to bring R^2 up. Investigation of the residuals may yield some insight into this problem.

The study indicates that urban environment plays a more significant role than that of family planning facility in raising the level of acceptance of family planning programme. Amongst urban environment factors, the influence of education seems comparatively more important, whereas 'medical personnel' plays a major role amongst the family planning facility factors. But the latter factor, family planning facility, cannot be viewed in isolation. An effective programme with all the family planning facilities can, to some extent, contribute in raising the level of acceptance of Family Planning programme. On the other hand, it is doubtful whether the factors of Family Planning facility alone can raise the acceptance rate to a high level, unless accompanied by other changes such as experience of life in urban complexes and massive improvement in educational level.

In India, the states which have done extremely well in family planning are Punjab, Maharashtra and Haryana. The ever-users of family planning methods per 1000 population since inception till March, 1971, are 65 for Punjab, 55 for Maharashtra and 46 for Haryana; the states which have done poorly are Bihar (16 per 1000 population), Rajasthan (20 per 1000 population), Assam (20 per 1000 population) and Uttar Pradesh (22 per 1000 population), while the remaining states occupy middle position (Agarwala, 1972). If we consider the states in the, light of their socio-economic development, the states which are doing well in family planning, are also comparatively better developed in economic and social spheres. Those which are lagging behind in their economic and social development are lagging behind in the sphere of family planning, in spite of the family planning facilities being available to all the states, though not necessarily on the same scale.

Thus it would seem that urban environment helps to create basic awareness about the population problems and congenial climate for nurturing values and attitudes favourable to acceptance of family planning. In order to make the family planning programme of Government of India more effective, it should be properly supplemented by measures of social and economic development. Considering that social and economic development must rest on a long term approach, and that education emerges as having a powerful influence in favour of family planning, massive adult education programme must form a necessary complement of the family planning programme.

ACKNOWLEDGEMENTS

The author is indebted to Prof. (Mrs.) Indira Bhanot, Director, Demographic Research Centre, Baroda and Dr. M. Sreehari, Reader in Statistics, M. S. University, Baroda, for their help in formulating this research and their valuable comments on the paper. Thanks are also due to Miss Jaya Patel who has assisted me in the computational work. The author, however, is solely responsible for any possible errors.

References

1. Agarwala, S. N., 1972, *A Study of Factors Explaining Variability in Family Planning Performance in Different States of India,* paper presented at All India Seminar on Family Planning Problems in India, International Institute for Population Studies (IIPS), Bombay.
2. , 1972, *Family Planning Performance in India,* paper presented at All India Seminar on Family Planning Problems in India, IIPS, Bombay.
J. Balasubramaniam, K.; Das, Narayan; Kumar, A. and Rama Rao, G., 1970, *Analysis of Family Planning Acceptance in India and some Selected States,* Mimeographed, IIPS, Bombay.
4. Census of India, 1961, Cultural and Migration tables. *Registrar General's Office,* I(II-C), India.
5 Das, N. and Rama Rao, G., 1971, *Factors Related to Acceptance of Family Planning Programme in India,* A preliminary project report prepared under the guidance of Dr. S. N. Agarwala, Director, IIPS, Bombay.
6. Harman, H.H., 1967, *Modern Factor Analysis,* Second edition, revised. The University of Chicago, Chicago, 351.

7. Kendall, M. G., 1957, *A Course in Multivariate Analysis, Griffin's Statistical Mimeographs and Courses* London, 70.
8. Ministry of Health and Family Planning, 1970, *Monthly Statements en the Progress of Family Planning Programme in India,* E and I Section, Department of Family Planning, New Delhi.
9. NCAER, 1963, *Inter District and Inter State Income Differentials 1955-56* (Occasional paper No. 6), National Council of Applied Economic Research, New Delhi.
10. Thomson, Godfrey, 1951, *The Factorial Analysis of Human Ability,* The University of London, London, 247-249.
11. Thurstone, L. L., 1947, *Multiple-Factor Analysis. A Development and Expansion of the Victors of Mind,* The University of Chicago, Chicago, 149-170.