INTER-STATE MIGRATION IN INDIA

By G. DIVAKARA RAO

Indian Statistical Institute

SUMMARY. The study carried out in this paper demonstrates the applicability of Stouffer's model on 'intervening opportunities and competing migrants, to (i) inter-state migration streams and (ii) a developing country like India. The explanatory power of the model for India would however, be less than that for USA, in view of the inadequate transport and communication facilities prevailing in India. The accuracy of the model may not be enhanced when applied to a more homogeneous sub-group of migrants if such group division cuts across family unit.

1. INTRODUCTION

Stouffer's model belongs to the group of models on the spatial distribution of migrants which centre round the hypothesis : $m = aX/Y$ where m is the number of migrants in a stream, X is the force of attraction and Y the distance. It differs, however, significantly from others in the measure of X and Y. Whereas Zipf (1940) had proposed that when employment and income are uniformly distributed over the areas, X should be the population size of the area and Y should be the shortest transportation distance. Stouffer (1940 and 1960) in his theory of 'Intervening opportunities and Competing migrants' assumes no such relationship between mobility and distance and introduces the concept of intervening opportunities in place of physical distance. A test of Stouffer's hypothesis on inter-city data of 1940 census by Stouffer (1960) and on inter-metropolitan data of 1960 census by Galle and Taeuber (1966) establish the supremacy of Stouffer's model with a coefficient of multiple correlation over 0.95 which is rare in sociological studies.

Stouffer's hypothesis presupposes that migrants have a complete knowledge of opportunities; this may not be realised always. That this model was found to be quite satisfactory in USA was partly due to greater realisation of this assumption. In the less developed countries the imperfect mass media and communication facilities may result in not too good knowledge of opportunities and this is certainly a great factor in reducing the accuracy of the model. This we seek to verify with data on migrants in India. Secondly, Stouffer's hypothesis has not yet been tested on inter-state data. The quite cumbersome and to an extent arbitrary operational definitions of variables, particularly those of intervening opportunities, require to be modified if the hypothesis is to be tested on inter-state migration data. The suitability or otherwise of a few selected operational definitions of the variables have also been tested in this paper.

2. STOUFFER'S MODEL

Much of the attention that Stouffer's model attracted was largely for the reason that it opened a new line in migration studies and at the same time had a high explanatory power. His main contribution had been the alternative measure of 'social

distance' to physical distance which was supposed to be an indispensible factor in earlier studies. In his first enunciation of the theory Stouffer (1940, page 71) proposed that 'the number of persons going a given distance is directly proportional to the number of opportunities at that distance and inversely proportional to the number of intervening opportunities'. The explanatory power of this model as indicated by the multiple correlation coefficient was undoubtedly high but as Stouffer himself had noticed the model did not take into account the direction factor and had a systematic error. A test of this model by Bright and Thomas (1941), Isebell (1944), and Strodtbaek (1949) revealed encouraging conformity between the expected and observed migrants for migration streams of particular distance bands and of particular direction. To correct the deficiencies mentioned above Stouffer (1960) reformulated his original theory by redefining intervening opportunities and adding a variable which he termed as competing migrants.

The reformulated theory of Stouffer can be summarised as follows :

$$Y = \frac{AX_M}{X_B X_C} \qquad \ldots \quad (1)$$

where Y is the number of migrants in a migration stream (say from city P to city Q); X_M, termed size-effect is the product of 'push' (from P) and 'pull' of (Q) which in turn have been measured by number of out-migrants of the place of origin and the number of in-migrants of the place of destination respectively; X_B, invervening opportunities for a migration stream (P to Q) is the agregate of inmigrants to all cities within a circular region, the circle being drawn with extended line PQ (extended 75 miles on either side) as diameter. This excludes city Q but includes city P. It is noteworthy how Stouffer measures opportunities in a region by the number of in-migrants of that region. X_C, competing migrants for a migration stream (P to Q) is the aggregate of out-migrants from all cities within a circular region, the circle being drawn with Q as center and with radius equal to distance between P and Q plus 75 miles. This includes city P but excludes city Q.

Stouffer's model should not be viewed merely as an attempt at predicting the individual cells in a two-way table knowing only the marginal totals. It is a well-formulated theory which takes into account only a few factors in the measurement of which marginal totals have been used and succeeds in explaining a major part of the variation in geographical mobility. The flexibility in the definition of opportunities may offer a wide scope in diverse applications besides spatial distribution of migrants. The explanatory power of Stouffer's model in a country depends largely on the extent of knowledge of opportunities both within and without the place of origin by migrants, which in turn depends on the efficiency of mass media and other communication channels. Once this condition is satisfied the model may prove a good fit. The large measure of success of the model for inter-city or inter-metropolitan migration in USA is due mainly to the fulfilment of this condition. That Gallo and Taeuber (1966) obtained a lower figure for $R(0.05)$ with 1960 census data as compared to the

one, 0.98, obtained by Stouffer (1960) with 1940 census data is presumably due to (Gallo and Taeuber, 1966, page 11) 'marginal differences in the opportunities and costs of moving to city 3 rather than city 2 may have been reduced, so that the explanatory significance of intervening opportunities and competing migrants is reduced. The success of the model also depends on how well the data on migrants is organised by type of opportunities they seek and the degree to which the group of migrants covered by the study are homogeneous with respect to the opportunities they seek.

The criticism that there is 'circularity' in X_M does not detract from the applicability of the model. The part of circularity which comes from (Theodore R. Anderson, 1955, p. 289) 'the number of migrants to be predicted is part of the total migrants used in prediction' is significant only when exactly the same proportion of migrants from a place go to places of particular distance or in other words the same proportion of migrants to a particular place come from places of a particular distance. This situation is unlikely to be the case in majority of migration streams. The other part of 'circularity' involved in the assumption that the number of opportunities is measured by total in-migrants had already been explained by Stouffer himself (1960, page 94).

3. Inter-state migration

We have taken up this study primarily to investigate the applicability of Stouffer's model to inter-state[1] migration which will incidentally demonstrate the usefulness or otherwise of the model in a developing country like India. The latter aspect of the study gains importance in view of the fact that the effectiveness of Stouffer's model depends on the complete knowledge of opportunities for which a primary requisite is an efficient system of mass media and other communication channels. The rate of migration in India is much lower than the rate prevailing in USA and a large part of the movements are seasonal and of short distance, 'marriage migration and marriage-ancillary migration are more significant in India than in many other countries' (Zachariah, 1964, page 251). Migration of adult males is on the whole economically motivated and there exists a relationship between the state of origin and type of industry they choose, while that of females is mostly obligatory or sequential. Migration streams having their origin in rural and urban areas are almost equally prevalent in urbanward migration. Lack of transport and communication facilities and above all illiteracy are some of the factors inhibiting migration in India. Two places may be near in physical distance but not in time taken to travel. It would thus be interesting to study how in a country like India with a different social, economic and communication set-up as compared to USA, the spatial distribution of migrants can be explained by Stouffer's theory.

The data used in this study have been obtained from the Urban Labour Force Survey conducted by National Sample Survey Department[2] of the Indian Statistical Institute and refer to the period July 1965 to June 1966. The data give account of

[1] Only the urbanward migration streams had been aggregated to form the inter-state migration streams.

[2] National Sample Survey is a multi-subject large-scale sample survey organisation in India collecting data from probability samples spread over the entire country on a continuing basis.

all migrations to urban areas having origin in rural or other urban areas. A person had been defined as a migrant if his/her usual place of residence 365 days ago was different from his/her current place of residence. This data on migration are undoubtedly more suitable for a study of Stouffer's model than the data on migration based on place of birth.

There can be no doubt that Stouffer's theory which was found to be satisfactory for explaining inter-city migrations will also be satisfactory for inter-state migrations. What is required to be tested is the adequacy of the operational definitions of the variables evolved for inter-city migration streams to inter-state migration streams. The operational definition of intervening opportunities and competing migrants requiring drawing of circles were given by Stouffer for inter-city migrations only and adherance to these definitions in the study of inter-state migrations is not feasible. Determination of center of circles which is easy in the case of cities because they are represented by dots on maps is not without difficulties for a migration stream between two states. Even if the center of population of states is taken they will lead to two kinds of errors : firstly, the center of population is not necessarily the center of migrating population particularly when migration streams of opposite directions from a state are considered and secondly, in having to include or exclude a state which partly lies within the circles. For these reasons and above all to avoid making the operational definitions of intervening opportunities and competing migrants even more cumbersome we adopted the following modified form of definitions for our study.

Here too as in Stouffer's study we measure opportunities by number of in-migrants. The size-effect, X_{sf}, has been defined as the product of number of out-migrants from the state of origin and the number of in-migrants to the state of destination which is similar to the definition adopted by Stouffer. The circular areas required in the measurement of intervening opportunities and competing migrants have been modified so as to avoid drawing circles. For a migration stream from state P to state Q the circular area for intervening opportunities has been obtained in the following manner. Select the center-state in the shortest railway route from P to Q. If the number of states intervening between P and Q is odd, $2n+1$, then the center-state is the $(n+1)$-th state (excluding P) from P in the route of Q. If the number is even, $2n$, then the center-state is a dot between the n-th and $(n+1)$-th states from P (excluding P) in the route to Q. Having thus chosen the center-state the next step consists in defining circular bands of states. The first circular band includes all states having common border with center-state and the second circular band consists of all states having common border with states in the first circular band and so on. If the states P and Q are in the n-th circular band then the area chosen for intervening opportunities includes all states in the n-th and lower order circular bands including P but excluding Q. Thus for each migration stream there exists a separate area for intervening opportunities. The measure of intervening opportunities (X_B) is then the sum of numbers of in-migrants of states (to all urban areas of a state from other states) in the area defined above.

The area for measuring competing migrants (X_C) for a migration stream, P to Q, has been defined in a similar manner but with the only difference that the state Q has been chosen as the center-state. If the state P lies in the n-th circular band then the area includes all states lying in the n-th and lower order circular bands including P but excluding Q. The measure of competing migrants is then obtained as the sum of out-migrants of all states included in this area.

It is important to note that since urbanward migrations only were used for this study the in-migrants figure for a state relates only to the urban areas of the state while the out-migrants figure for a state relates to the entire area of the state. Hence the variables on size-effect, intervening opportunities and competing migrants should be interpreted as such. Secondly, it has to be pointed out that the operating pull factor for any migration stream is likely to be associated more with a city or urban area than with a state. Possibility of some distortion of results of the inter-state model exists in this case because of the underlying aggregation of urbanward migration streams present in inter-state migration streams. This bias is likely to be small when the range of values for the pull factor of the various urban areas of a state is small but likely to be serious when this range is large.

Expressing (1) in the form,

$$\log Y = \log A + \beta_1 \log X_M + \beta_2 \log X_B + \beta_3 \log X_C \qquad \ldots (2)$$

or

$$Y' = \alpha + \beta_1 X'_M + \beta_2 X'_B + \beta_3 X'_C$$

where Y', X'_M, X'_B and X'_C are the logarithms of Y, X_M, X_B and X_C respectively and $\alpha = \log A$.

The estimated values of regression coefficients obtained by least squares and the results of the test of the model separately on two sets of data —one relating to 'all migrants' including male and female migrants and the other relating to 'male migrants' only—are given in Table 1. This we believe will demonstrate the predictive power of the model for the entire group of migrants vis-a-vis male migrants.

The results of the test given in Table 1 reveal that Stouffer's model could explain only about 50 percent of the variation in the inter-state migration streams which is in contrast to the high values of R^2 (0.95) obtained for inter-city migration (1935-40) in USA. The results thus confirm the expectation that in a developing country like India in view of the inadequate communication and transport facilities and a good proportion of female migrations being obligatory and sequential (Rele, 1969, page 506) the explanatory power of the model will be low. The analysis also reveals that by far the most effective variable in the model is 'size-effect', X_M, and the gain in efficiency of the model by the introduction of the variables-intervening opportunities and competing migrants—is marginal. Part of the explanation may be attributable to the operational definition of the variables—intervening opportunities and competing migrants—used in the Indian context.

TABLE 1. RESULTS OF THE TEST OF STOUFFER'S MODEL SEPARATELY FOR "ALL MIGRANTS' AND 'MALE MIGRANTS'

	all migrants	male migrants
1. $\hat{\alpha}$	4.5909	4.1638
2. $\hat{\beta}_1$	0.6264	0.5416
3. $\hat{\beta}_2$	−0.2942	−0.2877
4. $\hat{\beta}_3$	−0.5261	−0.3632
5. $\bar{R}^2$ *	0.49	0.46
6. standard error of the estimate, S	0.394	0.343
7. partial correlation coefficients :		
(a) $r^2_{Y,X_M \cdot X_B X_C}$	0.38	0.32
(b) $r^2_{Y,X_B \cdot X_M X_C}$	0.06	0.07
(c) $r^2_{Y,X_C \cdot X_M X_B}$	0.08	0.05
8. (a) standard error of $\hat{\beta}_1$	0.087	0.085
(b) standard error of $\hat{\beta}_2$	0.130	0.115
(c) standard error of $\hat{\beta}_3$	0.106	0.108
9. degrees of freedom	84	81

* $\bar{R}^2 = 1-(1-R^2)\left(\frac{n-1}{n-m-1}\right)$ where R is the multiple correlation coefficient, n the number of observations and m the number of independent variables.

This difference in the values of $\bar{R}^2$ for the two countries can be ascribed to several reasons. Prominent among them is the level of transport and communication facilities prevailing in India being lower than that in USA. The more the divergence between the real and apparent opportunities the less would be the efficiency of the model. The second reason involves the size of the units of analysis—city for USA and state for India. Distance measurements tend to lose their meaning when area of the surrounding units is large relative to the distances involved. Undoubtedly this problem exists for the inter-state study. Aside from a kind of attenuation effect that this largeness produces, it can actually distort the image of the underlying regression that is present.

Significantly consideration of a subset of migrating population for testing Stouffer's model has not increased the explanatory power of the model. The value of $\bar{R}^2$ for the male group of migrants has been 0.46 as against 0.49 for all migrants. The reduction in $\bar{R}^2$ for the male migrant group can be attributed mainly to the variable on 'size-effect' whose explanatory power as indicated by square of partial correlation coefficient has been 0.32 for male migrant group as against the value of 0.38 for all migrants. For an understanding of the reasons for this decline in the explanatory power of X_M, size-effect, it is useful to analyse the categories of migrants. The first

category consists of those who migrate with the family which in India includes a chief earner, his wife and children and sometimes his parents, brothers and sisters. The migration of members other than the chief earner is in most cases entirely dependent on the chief earner and as such do not form an independent group for any study of migrations. This category of migrating families often go short distances and in few cases depending on their economic standard go even longer distances. The second category comprises of single migrants who go both short and long distances. A third category consists of those families whose chief earner migrates first and the other members of the family join him in due course. Hence the forces of push and pull which make the size-effect, apply to the family as a whole in case of first and third categories and to the individual in the second category of migrants. By taking only male migrants for the study we have chosen a part of each migrating family by the demographic criterion which is not governed by the laws of migration and this no doubt has disturbed the correspondence between the factor of pull and in-migrants on the one hand and between the factor of push and out-migrants on the other. This is shown by the raw correlation coefficient between Y and the number of out-migrants of the state of origin on the one hand, Y and the number of in-migrants of the state of destination on the other being less for male group of migrants than for all migrants.

Having thus far seen the nature of fit of Stouffer's model to the entire set of migration streams in the country we shall now examine whether the fit is uniformly good for sub-groups of migration streams, the sub-groups having been formed on some well-defined criteria. The fact that its explanatory power is higher for short distance migration streams than that for long distance migration streams is demonstrated by the data. Additionally, we considered three sub-groups of migration streams formed on the basis of geographical zones which in India are a group of contiguous states. The first sub-group included all intra-zonal migration streams; migration streams oriented towards northern and western zones formed the second sub-group while those oriented towards central and eastern zones formed the third sub-group. The results of the test are given in Table 2.

TABLE 2. RESULTS OF APPLICATION OF STOUFFER'S MODEL TO SUB-GROUPS OF MIGRATION STREAMS

	sub-groups		
	first	second	third
1. number of migration streams	23	27	24
2. $\bar{R}^2$	0.60	0.53	0.29
3. partial correlation coefficients			
(a) $r^2_{12\cdot34}$	0.51	0.24	0.22
(b) $r^2_{13\cdot24}$	0.12	0.12	0.014
(c) $r^2_{14\cdot23}$	0.004	0.10	0.06

Numbers 2, 3, 4 stand for $\log X_M$, $\log X_S$, $\log X_O$ respectively.

Undoubtedly there is a noticeable variation in the explanatory power of the model on different sub-groups. The first sub-group comprising of short distance migration streams yields the highest value for $\overline{R}^2(0.60)$: the single contributory factor to this high explanatory power of this sub-group has been the size-effect. The size-effect apparently is less effective in the second and third sub-groups reducing thus the values of $\overline{R}^2$ for these sub-groups. The low value for the square of partial correlation coefficient of competing migrants (0.004) in the first sub-group emphasises the redundancy of the variable and or the unsuitability of the area adopted for migrations between neighbouring states. The disparity in the values of $\overline{R}^2$ for the second and third sub-groups is not so much due to decrease in the explanatory power of the variable on intervening opportunities per se but more due to the unsuitability of the area chosen for measuring intervening opportunities. In fact the size and shape of the area chosen for either intervening opportunities or competing migrants changed depending on the geographical location of the state of origin and or the state of destination resulting in low or high values for these variables without any relation to the size of migration streams. This in general is the reason for the unsatisfactory performance of the variables—intervening opportunities and competing migrants—in the model.

Three alternative definitions of the area for intervening opportunities have been tested—(A) circular area with the state of origin as center state including all circular bands upto but excluding the one containing the state of destination : (B) circular area with the state of origin as center state including all circular bands upto and including the one passing through the state of destination but without the state of destination; (C) area including only those states that lie in the shortest railway route from the state of origin to the state of destination.

The values of coefficient of multiple determination obtained from Stouffer's model with the definitions (A) and (B) for the intervening opportunities were 0.50 and 0.49 respectively which bear negligible difference with 0.51 obtained for Stouffer's model with the original definition for intervening opportunities. These big sized areas for intervening opportunities have been tested with the expectation that they would be in order to explain the observed fast rate of decline in the size of migration streams as the distance increased. But this advantage appears to have been offset by the disadvantage inherent in larger areas in that they are more influenced both in size and shape by such uncontrolled factors as locations of the state of origin and or the state of destination and also length of migration streams. On the other hand the third alternative definition, (C), which is the smallest area thus far considered did yield a value of 0.49 for $\overline{R}^2$. This again is due to the interaction of the aforesaid effects. On final analysis it may be concluded that the original definition of area considered in this study for intervening opportunities which is a close approximation to the one considered by Stouffer for inter-city migration is the optimal one.

An alternative approach to the measurement of push and pull factors implicit in the size-effect has been tried. The push factor has been measured by the number

unemployed and the pull factor by the number unemployed becoming employed; the latter has been crudely measured as the ratio of the product of unemployed last year with the number employed in the current year to the total number of persons. This has failed to be any substitute to the size-effect as defined by Stouffer largely due to the crude measurement of pull factor.

Conclusions

1. With some modifications in the operational definitions of intervening opportunities and competing migrants Stouffer's model is satisfactory in predicting inter-state migration streams. The study, however, could not throw light on how the model fits to inter-state migration streams vis-a-vis inter-city migration streams.

2. In view of the fact that the variables considered in Stouffer's model are to have any effect a good system of transport and communication facilities should prevail in the country and since the transport and communication facilities are relatively less in India than those prevailing in USA, the explanatory power of the model whether for inter-state or inter-city migration streams would be less for India than that for USA.

3. Even though it is implied in Stouffer's theory that the utility of the model will be enhanced when applied to a homogeneous group of migrants, the same is not likely to be fulfilled when such group divisions cut across family unit. This is more true in India where the family as a unit is more strong. A test of the model on male migrant group yields a lower value of $\bar{R}^2$ than that for all migrant group.

4. The measure of size-effect is the dominant factor in the model and this is more so for the group of migration streams between neighbouring states; the large part of variation in the values of $\bar{R}^2$ of the model on different sub-groups of migration streams could be ascribed to this factor alone. The variation in the explanatory power of the measures—intervening opportunities and competing migrant—among different sub-groups of migration streams is partly due to the variation in the size and shape of the area for measuring them which bore no relation to the number of persons forming the migration stream.

5. No substantial increase in the explanatory power of the model can be obtained by changing the operational definition of 'intervening opportunities'. It is, however, probable that some improvement in the model could be obtained by considering two or three types of areas in the measurement of 'intervening opportunities' and 'competing migrants', the type of area to be considered for a particular migration stream depending on the geographical location of the state of origin and or the state of destination of the migration stream.

I am grateful to the Computer Science Unit, Indian Statistical Institute, for the computer time made available to me for carrying out the analysis.

REFERENCES

ANDERSON, THEODORE R. (1955) : Intermetropolitan migration : A comparison of the hypothesis of Zipf and Stouffer. *American Sociological Review*, 20, 287-291.

BRIGHT, MARGARET L. and THOMAS, DOROTHY S. (1941) : Inter-state migration and intervening opportunities. *American Sociological Review*, 6, No. 6, 773-783.

GALLE, O. R. and TAEUBER, K. E. (1966) : Metropolitan migration and intervening opportunities. *American Sociological Review*, 31, No. 1.

ISBELL, ELEANOR COLLINS (1944) : Internal migration in Sweden and intervening opportunities. *American Sociological Review*, 6, No. 6, 627-639.

RELE, J. R. (1969) : Trends and significance of Internal migration in India. *Sankhyā*, Series B, 31, 501-508.

STOUFFER, SAMUEL A. (1940) : Intervening opportunities : A theory relating mobility and distance. *Social Research to Test Ideas*, Chapter 4, 68-91.

——— (1960) : Intervening opportunities and competing migrants. *Social Research to Test Ideas*, Chapter 4, 91-12.

STRODTBACK, FRED L. (1949) : Equal opportunity intervals : A contribution to the method of intervening opportunity analysis. *American Sociological Review*, 14, No. 4, 490-497.

ZACHARIAH, K. C. (1964) : *A Historical Study of Internal Migration in the Indian Sub-continent* 1901-1931, Asia Publishing House.

ZIPF, GEORGE K. (1946) : The P_1P_2/D hypothesis : On inter-city movement of persons. *American Sociological Review*, 11, 677-686.

Paper received : November, 1971.
Revised : March, 1972.