THE EGYPTAION POPULATION AND
FAMILY PLANNING REVIEW.
ISSR,Cairo,Univ.,Vol,27, No.2,1993.

Socio-Economic and Urbanization Profiles of Internal Migration in Egypt: A Canonical Correlation Analysis

M.G. EL-ROUBY
ASSOCIATE PROFESSOR
KING SAUD UNIVERSITY

I. Introduction:

Migration is a multifarious demographic phenomenon which affects and is affected by other socio-demographic, economic, and ecological phenomena; Peterson (1958), Bogue (1959), Elizaga (1970), and Willis (1974). Studies on migration provided evidences on its varying aspects and on the voluminous factors that interact with it; Thomas (1963), Lowry (1966), Lee (1969), Stinner (1969), Vander-Kamp (1971), and Greenwood (1975). Such complexity of the problem has paved the way for the use of multivariate analysis techniques as suitable media for analysing migration data. Cananical correlation analysis is one of such techniques. It is the most general one, but perhaps the least used of them, Tabachnic and Fidell (1983).

Migration in Egypt is one of the vital processes that affect the population growth and structure. Patterns and streams of migration have been the theme for most studies on Egyptian migration; Abdul Hakim (1966), Ibrahim (1982), and El-Saadany (1984). Interaction of migration with other socio-economic variables has also been recognized and investigated; Greenwood (1969), Khalil (1986), and Seifelnesr (1983 and 1992).

The main objective of this paper is to screen some latent profiles of internal migration in Egypt, as evidenced by available national statistics, using a canonical analysis modelling approach. Canonical analysis provides a tool for investigating relationships in research settings of complicated or more general nature. In such settings, two sets of variables are measured or observed and the researcher aims at recognizing how they relate to each other, Tabachnick and Fidell (1983) and Dunteman (1984). To accomplish this, pairs of canonical variates are generated. They are linear combinations of variables in the two sets, one linear combination from each set. Correlation's among these canonical variates are then computed and analysed. By defining canonical variates relating migration variables with other influencing variables one would be able to describe the overlapping effects of the two sets of variables. Also, the various dimensions along which the migration phenomenon is related to socio-demographic and economic phenomena can be meaningfully structured.

The paper consists of five sections including this introduction. Section two describes the data sets that provide the necessary materials for analysis. Section three gives account of research methodology which centres upon the main features of canonical correlation analysis. Section four contains research main findings. Summary of the work is given in the last section.

II. Data:

Amongst the extensive data on particulars of individuals, households and establishments which the 1986 census in Egypt provided are data on migration at various administrative levels of the country. In and out-migration to and from each of the 26 governorates (the largest administrative level) of Egypt were obtained from cross-classifications of the population by governorate of current usual residence and governorate of previous usual residence.[1] In and out-migration rates were then calculated for each governorate and used as the set of dependent variables required for canonical analysis. Information on variables believed to interact with migration were sought through either the population census or other parallel data sources, mainly the statistical year book of Egypt[2]. As many as 2 1 of such variables were considered for meeting the requirements of the present analysis. They are measured mostly in the form of proportions or rates with the intention of standardizing for governorate population size. In the context of canonical analysis they are used as the set of independent variables. They are arbitrarily classified into four distinctive groups summarizing four different profiles of migration, namely, socio-demographic profile (6 variables), economic profile (7 variables), urbanization profile (6 variables), and health profile (2 variables). It is to be noted that the availability of information was a crucial element in determining the number of variables in each of the four categories. The following is a comprehensive list of independent variables included in each of the mentioned groups.

a) <u>Variables summarizing the socio-demographic profile</u>:

1. Proportion of governorate population to total country population (x_1).
2. Governorate population density (x_2)
3. Sex ratio in governorate (x_3)
4. Rate of natural increase in governorate (x_4)

---

1 These data are contained in table (48) of volume 1 of the 1986 Population Census, ref. 861/89/CAPMAS.

2 Statistical yearbook, A.R.E. 1952-1992, CAPMAS, (June, 1993).

5. Proportion of married population to total population fifteen years and over in governorate (x_5)
6. Proportion of illiterate population to total population six years and over in governorate (x_6)

b) **Variables summarizing economic profile:**

1. Proportion of unemployed population to total labour force population in governorate (x_7)
2. Proportion of population engaged in occupations to total population fifteen years and over in governorate (x_8)
3. Proportion of population engaged in industrial and production occupations to total population fifteen years and over in governorate (x_9)
4. Proportion of population engaged in agricultural occupations to total population fifteen years and over in governorate (x_{10})
5. Proportion population engaged in services occupations to total population fifteen years and over in governorate (x_{11})
6. Proportion population engaged in technical occupations to total population fifteen years and over in governorate (x_{12})
7. Average annual expenditure (LE) per individual in governorate (x_{13})

c) **Variables summarizing urbanization profile:**

1. Proportion of urban population to total population in governorate (x_{14})
2. Proportion of asphalted roads to total roads in governorate (x_{15})

3. Proportion of population with university education and above to total population six years and over in governorate (x_{16})
4. Proportion of population working in the manufacturing industries sector to total population fifteen years and over in governorate (x_{17})
5. Proportion of electrically lighted buildings to total buildings in governorate (x_{18})
6. Proportion of modern buildings (using cement and iron in construction) to total buildings in government (x_{19})

d) Variables summarizing health profile:

1. Number of pharmacies per 10,000 of population in governorate (x_{20})
2. Number of beds per 10,000 of population in governorate (x_{21})

III Research methodology:

Canonical analysis is usually initiated through the generation of a correlation matrix R.

Denoting the dependent variables (DV_S) by y and the independent variables (IV_S) by x, we have four correlation matrices:

$R_y \rightarrow$ correlation between the DV_s,

$R_x \rightarrow$ correlation between the IV_s,

$R_{xy} \rightarrow$ correlation between the DV_s and the IV_s

where: $R_{yx} = R'_{xy}$

The correlation matrix to be generated and analysed is then:

$$R = R_y^{-1} . R_{yx} . R_x^{-1} R_{xy} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (1)$$

The analysis proceeds by solving for the eigenvalues and eigenvectors of R. Calculation of eigenvalues and corresponding eigenvectors aims at redistribution of the variance in the original variables into a very few sets, each set capturing a large share of the total variance. These sets, termed canonical variates, are defined by linear combinations of the DV_S on one side and the IV_S on the other, Tabachnick and Fidell (1983, p. 152). Canonical correlation maximizes correlation between each pair of canonical variates, one for the DV_S and one for the IV_S.[1] Solution for eigenvalues λ_i and eigenvectors V_i involves solution of the equation:

$$(R - \lambda I) = 0 \quad \ldots\ldots\ldots\ldots\ldots \qquad (2)$$

The relationship between eigenvalues and canonical correlation is:

$$\lambda_i = \tilde{r}_i^2 \,, \ldots\ldots\ldots\ldots\ldots\ldots \qquad (3)$$

where $\tilde{r}_i$ is the canonical correlation between the ith pair of canonical variates; i = 1, 2, . . . , m and m is the smaller of the number of the DV_S (Ky) or the number of the IV_S (Kx). From equation (3), the canonical correlation between the ith pair of sets of dependent and independent variables is found by taking the square root of the corresponding eigenvalue λ_i . Also $\tilde{r}_i$ may be interpreted as ordinary Pearson product-moment correlation coefficient; when

---

1 Discovery of pairs of canonical variates continues until either no significant linkages between sets remain in the residual correlation matrices or as many pairs of variates have been defined as there are variables in the smaller set, see Tabachnick and Fidell, p. 146.

squared it represents overlapping variance between two variables and λ_i then, represents overlapping variance between a pair of canonical variates.

The significance of one or a set of canonical correlations can be tested using the conventional x^2 or F significance tests. For the x^2 test it is:

$$\chi^2 = -\left[N-1-\left(\frac{K_x + K_{y+1}}{2}\right)\right] l_n \wedge_m \quad \ldots\ldots\ldots\ldots \quad (4)$$

which has a Chi-square distribution with $(K_X \,.\, K_y)$ degrees of freedom; where N is the number of observations (number of governorates in the present analysis), and $\wedge_m$ is defined by:

$$\wedge_m = \prod_{i=1}^{m}(1 - \lambda_i) \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (5)$$

Equations (4) and (5) test whether or not the whole set of canonical correlations differ from zero. If i in equation (5) runs from 2, the equations test whether or not the remaining set of canonical correlations, with the first correlate removed, differs from zero. The test then has a Chi-squared distribution with $(k_{x-1})\,(k_{y-1})$ d.f.

The next step in canonical analysis is to generate two sets of canonical coefficients (canonical weights) for each canonical correlation, one set $\tilde{C}_y$, is needed to combine the DV_S and the other $\tilde{C}_x$ to combine the IV_S. These coefficients are defined by:

$$\tilde{C}_y = (R_y^{-1/2})' \hat{V} \quad \ldots\ldots\ldots\ldots\ldots \quad (6)$$

and

$$\tilde{C}_x = DR_x^{-1} R_{xy} \tilde{C}_y; \quad \ldots\ldots\ldots\ldots\ldots \quad (7)$$

where $\hat{V}$ is the matrix of eigenvectors for the DV_S and D is the diagonal matrix of reciprocals of eigenvalues.

The two sets of canonical coefficients can be used to estimate scores X and Y on canonical variates as:

$$X = Z_x \tilde{C}_x \quad \ldots\ldots\ldots\ldots \quad (8)$$

and

$$Y = Z_x \tilde{C}_x , \quad \ldots\ldots\ldots\ldots \quad (9)$$

where Z_X and Z_Y are standardized scores on the original variables x and Y. Matrices of correlations A_X and A_Y between original variables and canonical variates can be found. They are used as an aid to interpret canonical variates instead of interpreting them via canonical coefficients. A_X and A_Y are calculated by:

$$A_x = R_x \tilde{C}_x , \quad \ldots\ldots\ldots\ldots \quad (10)$$

and

$$A_y = R_y \tilde{C}_y \quad \ldots\ldots\ldots\ldots \quad (11)$$

The proportions of variance extracted from the IV_S and the DV_S by their own canonical variates, $\tilde{V}_x$ and $\tilde{V}_y$, are given by:

$$\tilde{V}_x = \sum_{i=1}^{k_x} \frac{a_{x_i}^2}{k_x} \quad \ldots\ldots\ldots\ldots \quad (12)$$

and

$$\tilde{V}_y = \sum_{i=1}^{k_y} \frac{a_{y_i}^2}{k_y} ; \quad \ldots\ldots\ldots\ldots \quad (13)$$

where axi and ayi are elements of Ax and Ay defined by equations (10) and (11).

The variance of the DV_S explained by the canonical variate of the IV_S (also the variance of the IV_S explained by the canonical variate of the DV_S) is known as the redundarcy in a canonical variate, rd. It is the proportion of variance it extracts from its own set of variables (equations (12) and (13) times the squared canonical correlation for the pair, i.e.,

$$rd = \tilde{V} \cdot \bar{r}^2 \tag{14}$$

IV. The findings:

The data described in section two are subjected to a set of canonical analysis procedures of the sort outlined in section three. The aim is to identify inherent profiles of in and out-migration flows in Egypt as depicted by the latest 1986 population census. To perform this task the computer program CANCORR of SAS software package is used as a suitable means for providing the information necessary for the required analysis. [1] Output from the CANCORR procedure includes detailed information on the mainfold features of a canonical correlation analysis but the intent is not to present here all such details. Rather, the most important of these features are summarized and presented in table (1) below Examination of the contents of this table reveals the following. As the theory of canonical correlation analysis tells[2], two pairs of canonical variates are produced at each application of the CANCORR, since the number

1 For a full description of the CANCORR procedure, reference should be made to SAS user's Guide: Statistics, 1982 edition, SAS Institute Inc., Cary, N.C., USA, 1982.

2 There are as many pairs of canonical variates as the number of variables in the smaller set, see section three of the paper.

of DV_S is 2. The first variate in a pair is a linear combination of the two migration variables (in and out-migration rates). The second variate is a linear combination of the opposing set of variables (the IV_S). The analysis begins by testing for independence between the variates, i.e., to test the null hypothesis:

$$H_N: \tilde{r}_i = \sqrt{\lambda_i = 0} \ , \ i = 1, 2, \ \ldots \ , m$$

against the alternative hypothesis

$$H_A: \bar{r}_i = \sqrt{\lambda_i} \neq 0 \ , \ i = 2, \ \ldots \ , m.$$

using χ^2 or F test.

Significance tests prove that the first three canonical correlations, those relating migration to socio-demographic, economic, and urbanization variables, are different from zero. In other words, the hypotheses of independence between migration and these variables are rejected at high levels of significance. On the other hand, canonical correlation relating migration to health variables does not significantly differ from zero, i.e., H_N for this correlate is accepted. The probability of rejecting H_N is, however, lowest in the case of urbanization variables (pr. > F = 0.0001) and highest in the case of socio-demographic variables (pr. > F = 0.0008). If correlation between the first pair of variates in each case is removed from analysis H_N is not rejected, implying the insignificance of correlation between the second pair of variates. [1]

---

1 Recall that the first significance test concerns correlation's among variabes for all generated pairs.

An assessment of the variance explained by the significant pairs of canonical variates is made through the calculation of variance overlap between each pair, it is the squared canonical correlation $\tilde{r}_i^2$. For the socio-demographic, economic, and urbanization migration profiles, the amount of variance overlap is more than three quarters of total variance in each case. It is highest for the economic profile (85.7%) relative to the socio-demographic (80.5%) and urbanization (77.2%) profiles. Variance overlap for the health profile is only 17% of total variance, signifying very low interaction rate between migration and health variables.

Another way of assessing the importance of canonical variates is through the calculation of the amount of variance a variate extracts from the variables it represents. The migration variate extracts a fairly high proportion of variance from its own variables within each of the four investigated profiles. The highest proportion is recorded within the socio-demographic profile (almost 61%) and the lowest within the urbanization and economic profiles (almost 52% for each).

Regarding the canonical variates of the IV_S , the amount of variance they extract from their own variables is very much lower than that the DV_S extract from their own variables, except within the health profile. The percent of variance explained by own canonical variate is lowest within the socio-demographic profile (19.2) and highest within the health profile (64.7). The only explanation to this is the limited number of variables included in the health set, with higher chance of less variability among the variables.

One of the best indices of the importance of relationships is the redundancy of a canonical variate. It is defined as the amount of variance a canonical variate of one set of variables extract from the variables in the other set. Equivalently, if the score on a canonical variate of one set is known, redundancy of this variate is the amount of uncertainty to be reduced regarding the other set. Table (1) indicates that the DV_S variate in general explains higher proportions of variance from the IV_S than what the IV_S variate explains from the DV_S in all profiles except the one related to health. To demonstrate this, the following figures are quoted. The migration variate extracts a little less than fifty percent of the variance in the socio-demographic variables, 45 and 40% of the variance in the economic and urbanization variables respectively. Whereas it extracts only 10 per cent of variance in the health variables. On the other hand, the urbanization, the economic and the socio-demographic canonical variates extract about 30, 35, and 15% of the variance in the migration variables. The health variate explains only 11% of the variation in the migration variables, signifying again a weaker relationship between migration and health indicators.

Finally, correlations between the canonical variates and the original variables provide further indications about the relative importance of the variables as contributors to the relationship with migration variables. Table (1) contains information on the variables of highest correlation's with the canonical variates. Starting with the DV_S, it is noticed that the in-migration rate plays a more important role in the relationship between migration and determining factors than does the out-migration rate. it has correlation coefficients with own canonical variate over 0.95 within all migration profiles and correlation coefficients with opposite, canonical variates over 0.85 within the first three profiles. The out-migration rate, on the other hand, has very low correlation coefficients with either own or opposite canonical variates except within the socio-demographic profile where correlation coefficient of either type is around 0.5.

Regarding the independent variables a fairly good number of them, especially within the first three migration profiles, correlate highly with both own and opposite canonical variates. Correlations with opposite variates are, however, slightly lower than those with own variates. Within the set of socio-demographic variables, the proportion of illiterate population enjoys the highest correlations, -0.788 and - 0.707 with own and opposite variates respectively. Next to it, comes the rate of natural increase with corresponding correlation coefficients of - 0.466 and - 0.418. Within the set of economic variables, variables related to occupational affiliation have the highest correlations with both own and opposite canonical variates. Correlation coefficients for these variables with own canonical variates range between 0.618 and - .791 and those with opposite variates range between 0.572 and - 0.732.

Within the urbanization group of variables, the proportion of urban population, the proportion of buildings with modern construction style, and the proportion of university education and above enjoy fairly high correlations with own and opposite canonical variates. Correlation coefficients in the first case range from 0.702 to 0.906 and in the second from 0.616 to 0.796.

As for the two health variables, they enjoy high correlations with their own canonical variate (0.939 and 0.643) but low correlations with opposite variates (0.387 and 0.265). This implies low contributions of these variables to the relationships with the migration variables. This is in line with earlier findings concluded through variance analysis.

set. Equivalently, if the score on a canonical variate of one set is known, redundancy of this variate is the amount of uncertainty to be reduced regarding the other set. Table (1) indicates that the DV_S variate in general explains higher proportions of variance from the IV_S than what the IV_S variate explains from the DV_S in all profiles except the one related to health. To demonstrate this, the following figures are quoted. The migration variate extracts a little less than fifty percent of the variance in the socio-demographic variables, 45 and 40% of the variance in the economic and urbanization variables respectively. Whereas it extracts only 10 per cent of variance in the health variables. On the other hand, the urbanization, the economic and the socio-demographic canonical variates extract about 30, 35, and 15% of the variance in the migration variables. The health variate explains only 11% of the variation in the migration variables, signifying again a weaker relationship between migration and health indicators.

Finally, correlations between the canonical variates and the original variables provide further indications about the relative importance of the variables as contributors to the relationship with migration variables. Table (1) contains information on the variables of highest correlation's with the canonical variates. Starting with the DV_S, it is noticed that the in-migration rate plays a more important role in the relationship between migration and determining factors than does the out-migration rate. it has correlation coefficients with own canonical variate over 0.95 within all migration profiles and correlation coefficients with opposite, canonical variates over 0.85 within the first three profiles. The out-migration rate, on the other hand, has very low correlation coefficients with either own or opposite canonical variates except within the socio-demographic profile where correlation coefficient of either type is around 0.5.

Regarding the independent variables a fairly good number of them, especially within the first three migration profiles, correlate highly with both own and opposite canonical variates. Correlations with opposite variates are, however, slightly lower than those with own variates. Within the set of socio-demographic variables, the proportion of illiterate population enjoys the highest correlations, -0.788 and - 0.707 with own and opposite variates respectively. Next to it, comes the rate of natural increase with corresponding correlation coefficients of - 0.466 and - 0.418. Within the set of economic variables, variables related to occupational affiliation have the highest correlations with both own and opposite canonical variates. Correlation coefficients for these variables with own canonical variates range between 0.618 and - .791 and those with opposite variates range between 0.572 and - 0.732.

Within the urbanization group of variables, the proportion of urban population, the proportion of buildings with modern construction style, and the proportion of university education and above enjoy fairly high correlations with own and opposite canonical variates. Correlation coefficients in the first case range from 0.702 to 0.906 and in the second from 0.616 to 0.796.

As for the two health variables, they enjoy high correlations with their own canonical variate (0.939 and 0.643) but low correlations with opposite variates (0.387 and 0.265). This implies low contributions of these variables to the relationships with the migration variables. This is in line with earlier findings concluded through variance analysis.

V Summary:

Data on in and out migration in Egypt by governorate and on variables believed to interact with migration are examined. The purpose is to identify the dimensions through which the basic demographic phenomenon can be rationally structured. This objective is approached through a general multivariate analysis technique known as canonical correlation analysis. It is an appropriate tool for investigating relationships between two sets of variables, one of them may be thought of as a set of dependent variables and the other as a set of independent variables.

The basic step in a canonical correlation analysis procedure is the generation of orthogonal paris of canonical variates. For any pair, the first variate is a linear combination of the dependent variables and the second is a linear combination of the independent variables. Once these variates are generated, a number of criteria are applied to assess the type and strength of relationship between the dependent and the independent variables.

For the purpose of migration analysis, in and out-migration rates by governorate are used as a set of dependent variables. A number of variables believed to act as determinants of migration are used and arbitrarily classified into four groups to represent four distinct migration profiles, namely, socio-demographic, economic, urbanization, and health profiles. The SAS computer program CANCORR is used to generate the information necessary for analysing these profiles.

The analysis shows that the economic profile of migration is the most pronounced one of the four profiles examined. Canonical correlation between the migration and the economic variates is at its highest level (0.926) and so is the amount of variance overlap between them (85.8%). Next to it are the socio-demographic and the urbanization profiles. On the other hand, canonical correlation between migration and health variables does not significantly differ from zero and the percent of variance overlap is very low (17%).

Migration canonical variates within the four specified profiles extract between fifty two and sixty percent of the variance in their own variables and between forty and forty nine percent of the variance in the set of opposite variables. An exception to this last result is the low amount of variance in the health variables explained by the migration variate. This is another indication to the weak relationship between migration and health indicators.

The amount of variance in the independent variables explained by the canonical variates is much lower, ranging from 19% to 65% for own variates and from 11% to 30% for opposite variates.

Finally, there is an indication for several independent variables across the socio-demographic, the economic, and the urbanization profiles to exert larger interaction with migration variates than do other variables. This is evidenced by the high correlations between these variables and both own and opposite canonical variables. Among these variables are the proportion of illiterate population, the proportion of population engaged in occupations, and the type of occupational affiliation.

The above findings, however, summarize the specific migration experience in Egypt in the early eighties. Their validity is hinged upon considerations of quantity and quality of available baseline data.

The procedure of canonical correlation analysis proves to be a very useful technique for illuminating relationships of complicated or general nature such as the ones dealt with in this paper.

TABLE (I)

Summary Indicators for Canonical Correlation Analysis of Internal Migration in Egypt

Indicator	Migration Profile			
	Socio-Demographic	Economic	Urbanization	Health
Canonical corelation ($\tilde{r}_i$)	0.8971	0.9262	0.8786	0.4128
Level of significance (Pr.>F) (P-value)	0.0008	0.0001	0.0003	0.2647
% of variance overlap (λ_i)	80.483	85.777	77.198	17.043
% of variance in DV_S explained by:				
Own canonical variate ($\tilde{V}_y$)	60.67	52.49	52.23	58.96
Opposite canonical variate (rd_y)	48.83	45.02	40.32	10.05
% of variance in IV_S explained by:				
Own canonical variate ($\tilde{V}_x$)	19.17	29.58	38.36	64.72
Opposite canonical variate (rd_x)	15.43	25.38	29.61	11.03
Variables of highest correlations with canonical variates				
(a) Dependent variables:				
INMIGR (y_1) own C.V.	0.957	0.999	0.998	0.978
opposite C.V.	0.858	0.925	0.877	0.404
OUTMIGR (y_2) own C.V.	0.546	0.229	0.220	0.471
opposite C.V.	0.490	0.212	0.193	0.195

(Contd. Table I)

Indicator			Migration Profile			
			Socio-Demographic	Economic	Urbanization	Health
(b) <u>Independent Variables:</u>						
PILLIT (x_6)	own	C.V.	-0.788			
	opposite	C.V.	-0.707			
RNICRSE(x_4)	own	C.V.	-0.466			
	opposite	C.V.	-0.418			
PPOP (x_1)	own	C.V.	-0.403			
	opposite	C.V.	-0.361			
PAGRIC (x_{10})	own	C.V.		-0.791		
	opposite	C.V.		-0.732		
PTECOCC (x_{12})	own	C.V.		0.659		
	opposite	C.V.		0.610		
POCCUP (x_8)	own	C.V.		0.623		
	opposite	C.V.		0.577		
PINDOCC (x_9)	own	C.V.		0.618		
	opposite	C.V.		0.572		
PSERVOC (x_{11})	own	C.V.		0.497		
	opposite	C.V.		0.469		

Cont. (Table I)

Indicator			Migration Profile			
			Socio-Demographic	Economic	Urbanization	Health
UNEMPL (x_7)	own	C.V.		0.471		
	opposite	C.V.		0.436		
PURBAN (x_{14})	own	C.V.			0.906	
	opposite	C.V.			0.796	
PMODBDG (x_{15})	own	C.V.			0.862	
	opposite	C.V.			0.758	
PUNIVED (x_{16})	own	C.V.			0.702	
	opposite	C.V.			0.616	
PINDSEC(x_{17})	own	C.V.			0.651	
	opposite	C.V.			0.572	
PPHARM (x_{20})	own	C.V.				0.939
	opposite	C.V.				0.389
BEDRATIO (x_{21})	own	C.V.				0.643
	opposite	C.V.				0.265

References

1 - Bogue, Donald J. "Internal Migration", The Study of Population: An Inventory and Appraisal, (Philip Hauser & Otis Dudley Duncan, eds., Chicago, 1959), pp. 486-509.

2 - CAPMAS, Population, Housing, and Establishments Census of Egypt, 1986, Final Results, Total Republic, (Cairo, 1989).

3 - ________, Statistical Yearbook, A.R.E. 1952-1992, Cairo, June 1993.

4 - Dunteman, George, H. Introduction to Multivariate Analysis, SAGE Publications , Inc. (1984).

5 - Elisaga, Juan C., "Internal Migration: An overview", The International Migration Review, (1970), pp. 121-46.

6 - El-Saadany, S. The Socio-Economic Correlates of Internal Migration in Egypt, an unpublished M.Sc. Thesis, Cairo Univ., (Cairo, 1984).

7 - Greenwood, M.J., "The Determinants of Labour Migration in Egypt", Journal of Regional Science, 9, 2 (1969), pp. 283-90.

8 - Ibrahim, S. "Research Monograph No. 5, pop. and Family Planning Board, Egypt, (1982).

9 - Illsley, R.A. Finalyson, and B. Thompson, "The Motivation and Characteristics of Internal Migration: a Socio-Economic Study of Young Migrants in Scotland", Milbank Memorial Fund Quarterly, 41, 2 (April, 1963), pp. 115-44 and 3 (July, 1963), pp. 217-48.

10 - Khalil, S.A. "Comparative Analysis of Some Population Characteristics in Cairo Governorate According to 1960 and 1976 Censuses, an M.Sc. Thesis, Cairo Univ., (Cairo, 1986).

11 - Lee, Everett S. "A Theory of Migration", in Migration, Sociological Studies, 2, J.A. Jackson, ed., (1969), pp. 282-97.

12 - Long, Larry H. "Migration Differentials by Education and Occupation: Trends and Variations", Demography, 10, 2, (May, 1973), pp. 243-58.

13 - Lowry, Ira, Migration and Metropolitan Growth: Two Analytical Models (Chanaler, San Francisco, 1966).

14 - Petersen, William, "A General Typology of Migration", American Sociological Review, 23 (1958), pp. 256-66.

15 - SAS Institute Inc., SAS User's Guide: Statistics, 1982 Edition, Cary, N.C., USA (1982).

16 - Seifelnasr, A. " The outflow of Labour from Agriculture: a Framework for Analysing Migration from Rural Areas", L'Egypte Contemporaire, no. 393-394, (1983), pp. 187-224.

17 - __________ , "Association of Reasons of Migration in Egypt with Sex and Level of Education: a Log-Linear Modelling Approach, " The 27th Annual Conference, ISSR, Cairo Univ., Egypt, Vol. 27 (1992), pp. 24-44.

18 - Stinner, William F. and Gordon F. de Jong, "Southern Negro Migration: Social and Economic Components of an Ecological Model," Demography, 64 (1969), pp. 455-71.

19 - Tabachnick, Barbara and Linda S. Fidell, Using Multivariate Statistics, Harper & Row Publishers, New York (1983).

20 - Thomas, Dorothy Swaine, "Selective Internal Migration: Some Implications for Mental Hygiene", in Demographic Analysis: Selected Readings, J.J. Spengler and O.D. Duncan, eds., (1963), pp. 425-31.

21 - Vanderkamp, John, "Migration Flows, their Determinants and the Effects of Return Migration", Journal of Political Economy, 79, 5 (1971), pp. 1012-31.

22 - Willis, Kenneth G. Problems in Migration Analysis. (Saxon House, D.C. Heath Ltd., 1974).