 0030-0950/84/0003-0088 $2.00/0

A MARKOVIAN ANALYSIS OF ETHNIC TRANSITION: ST. CLAIR NEIGHBORHOOD IN CLEVELAND, OHIO[1]

VERA K. PAVLAKOVIC, Geography Department, Kent State University, Kent, OH 44242
RICHARD W. JANSON, The Janson Industries, Canton, OH 44706

ABSTRACT. This paper suggests the use of Markovian chain analysis to utilize the data from the United States Census to predict the change of ethnic composition of an areal unit. A method for estimation of the transition matrix is proposed. Computations for the St. Clair neighborhood in Cleveland, Ohio, predict an actual increase in Yugoslavian occupance for the year 2000 when 52% of the neighborhood will probably consist of foreign European stock. This remarkable neighborhood, that has provided a gateway for South Slavs for more than a century, will probably continue as an entry point well into the 21st century.

OHIO J. SCI. 84 (3): 88-94, 1984

INTRODUCTION

In this paper the Markovian chain model is used for the quantitative analysis of demographic changes in an ethnic neighborhood. Specifically, 2 of the properties of the Markovian analysis have been utilized. First, the Markovian model pinpoints those periods wherein the actual development deviates considerably from the stochastic expectation based on observed trends from a previous time interval (Collins 1974); second, since Markovian analysis assumes a closed system (Collins 1976), the predicted relationship among various population segments at the end of each period, when compared with the

[1]Manuscript received 30 November 1981 and in revised form 29 March 1984 (#81-47).

actual composition, emphasizes the impact of the external factors. The paper focuses on the methodological aspects. The interpretation of results in the context of the ethnic neighborhood studied is the focus of a subsequent paper to be published shortly (Pavlakovic and Janson 1984).

MARKOVIAN ANALYSIS AND APPLICATION TO ETHNIC CHANGE

In a Markovian chain model stochastic variables are measured at a point in time, after which they continue to behave randomly in accordance with rules of probability that are specified in a transition matrix (Sprecher 1976). The method of Markovian analysis requires the estimation or observation of this matrix of probabilities that has empirical validity. The use of a constant quantity in transition is a necessary condition for a closed system changing endogenously (Collins 1974).

The Markov model in matrix equation form is simply stated as follows:

$$\underline{E} = P\ \underline{E}' \qquad (1)$$

where $\underline{E}$ is a row vector of the state variable values observed at the initial time; $\underline{E}'$ is a row vector of the same state variable values predicted for the final time; P is the transition matrix of probabilities with several properties that are characteristic of a stochastic Markov matrix. The matrix is square with an order equal to the number of elements in each state vector. Each row of the transition matrix is a probability vector and therefore all element values must be greater than or equal to zero and the sum of the elements of each row must equal 1.00. Final state variables under Markovian assumptions are dependent solely on their values at a previous time (Kemeny and Snell 1976). The assumptions and procedures to estimate the transition matrix are therefore fundamental (Berman and Plemmons 1979).

The state vectors are the system variables being measured to monitor the momentary status or condition of the system. The elements of the state vector are the data. Census data measure state conditions at moments in time at 10-year intervals, not the transitions in state that occur within the system during the interval. A change in form is required, however, from actual census counts to percentage measurements. This is necessary because the variables in transition must be contained within the closed system as they undergo changes of state. The state vectors at the initial time t_0, and at time t_1 after one interval (in this case, one decade between censuses) are observed values, and can be used to constrain the transition matrix that connects the 2 vectors. The first predicted vector is 2 intervals after the initial observations. By comparing the predicted state vector of the closed system at time t_2, with the observed state vector at time t_2, the difference among the elements can be interpreted as responses to exogeneous factors from the environment. This is tantamount to a quantification of history. The approach has wide application and the immense data of the census can be used to predict future relative relationships among variables.

Of course the state vector at time t_2 can be used as an initial vector for another iteration and when multiplied again by the transition matrix will predict the state vector at time t_3. The procedure can be continued into future decades to predict the state vector at times t_4, t_5, t_6, ... If the differences in successive vectors becomes smaller and smaller, a steady state condition is approached (Sprecher 1976).

The time interval between iterations is defined by the interval between 2 observed state vectors that are consistent with the transition matrix. Micro data based on actual observations of the transitions could conceivably be made. For example, all housing units in a census tract could be observed over a 10-year period and an actual count kept of the change of ethnic state that occurs when one family moves out of a specified housing unit and another family moves into the unit. Transition data of these kinds are rarely available and extremely expensive to gather.

Successive iterations can also be used to estimate the survival time of declining ethnic groups under the assumptions of steady state conditions and a closed system. The survival time is measured in Markov intervals, in this case census decades, and may be defined as the number of intervals required for a focal ethnic group to become an arbitrarily small component of a specified neighborhood.

The matrix equation (1) is restated in expanded notation of order 3 to furnish a concrete example of a procedure to estimate a transition matrix appropriate to model ethnic change in an areal unit:

$$[e_1, e_2, e_3]\begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix}$$
$$= [(e_1p_{11} + e_2p_{21} + e_3p_{31}),$$
$$(e_1p_{12} + e_2p_{22} + e_3p_{32}),$$
$$(e_1p_{13} + e_2p_{23} + e_3p_{33})]$$

Each row in the transition matrix controls the disposition of one corresponding element of the initial state vector. That is, every row vector of the transition matrix is a series of elements that sum to 1.00 because the row vectors in the matrix are probability vectors and by definition must sum to one. Each element of the initial state vector is distributed to each element of the final state vector by multiplying the initial element value by each probability in the corresponding row of the transition matrix. The summation of all values to a specified element of the final state vector is the predicted value of the element.

The estimated transition matrix and the hypothesis or theory of the transition must be consistent with the observed values of the census 10 years after the initial observations. An infinity of transition matrices can predict the same final state vector after one iteration, but this fact is not unusual in science. In fact, it is typical of scientific investigation that many theories can explain observed facts, and that the accepted paradigm is inferred.

The theory utilized in this paper is based on ethnic considerations. If an ethnic group is expanding relatively (meaning in percentage terms), this tie to the ethnic neighborhood is a most important geographical consideration. The transition matrix assures that all percentage points in the elements of the initial state vector are maintained and are represented in the final state vector in the full amount.

An ordered procedure for estimating the elements of the transition matrix follows:

P_{ij} denotes an element in the transition matrix; the superscript E refers to an expanding group and the superscript D refers to a declining group. N refers to the order of the matrix, e_i refers to the ith element in the initial state vector; e_i' refers to the ith element in the final state vector.

The elements of the matrix will be estimated in 5 categories: (1) The main diagonal elements of expanding groups; (2) the remaining row elements of expanding groups; (3) the main diagonal elements of declining groups; (4) the remaining column elements of declining groups; (5) the remaining row elements of declining groups, which is the same as the remaining column entries for expanding groups.

The main diagonal elements of all expanding groups are equal to 1.00. This presumes that all percentage points of an element of the initial state vector of an expanding group are transferred to the corresponding element in the final state vector:

$$p_{ii}^{E} = 1.00$$

The off-diagonal elements in the rows corresponding to expanding groups will be zero, because each row is a probability vector and the diagonal element is equal to 1.00.

$$p_{ij}^{E} = 0$$

where $i \neq j$

and where $j = 1, N$

The main diagonal elements for declining groups is such a number less than

1.00 that, when multiplied by the corresponding element of the initial vector will correctly result in the observed value for the corresponding element in the final vector. The implication is that all percentage points of an element of the final state vector come from the corresponding element of the initial state vector; that is:

$$p_{ii}^{D} = \frac{e_i'}{e_i}$$

The remaining column entries for declining groups equal zero. This is because all percentage points come from the focal ethnic group. No other groups contribute to the declining group. Thus,

$$p_{ij}^{D} = 0$$

where $i \neq j$

and where $i = 1, N$

The last group of entries consist of the remaining row entries for declining groups, which are simultaneously the remaining column entries for expanding groups. The assumption made is that the declining groups contribute their respective loss of percentage points to the expanding groups; and that the contribution from each declining group is to each expanding group in proportion to their expansions:

$$p_{ij}^{D} = \frac{\Delta e_j^{E}}{\sum \Delta e_j^{E}}(1 - p_{ii}^{D})$$

where $i \neq j$

and where $j = 1, N$

The Census of Population data on ethnicity were abstracted for selected census tracts and the numbers of each group were converted to decimal proportions of the total population in the tract. In this manner an abstract concept was selected as the medium of transition — the percentage (or decimal fraction) of the areal total population that belonged to a specified ethnic group. The total population of the tract remained a constant 100% even though absolute population was varying. This condition defined a constant quantity in transition over time; namely, 100% of the population. The component percentages of the constituent ethnic groups and their changing proportions were the object of study.

STUDY AREA AND DATA SOURCE

The Markovian approach in this study has been utilized in modeling the ethnic transition in a Slovenian-Croatian neighborhood in Cleveland, Ohio. The study area, known as the St. Clair neighborhood, is one of the 4 main ethnic colonies founded by South Slavic immigrants at the end of the 19th and the beginning of the 20th century in the Cleveland area (Bonutti and Prpic 1977, Cesen 1970, Pap 1973, Ledbetter 1918). The analysis has been limited to the 5 census tracts that represent the core of the neighborhood in the 1970s (Bonutti and Prpic 1977; Levy 1972; Pap 1973). The data from the 1920, 1930, 1940, 1950, 1960, 1970 and 1980 Census of Population are utilized. The official census statistics do not differentiate the immigrants from Yugoslavia by national origin but give the aggregate numbers based on the country of origin. In the Census of 1960 the count of immigrants from Yugoslavia was not published at the census tract level. This required estimations of the Yugoslavian component from the category that included all other Europeans. Detailed data for 1970 were obtained from computer tape and did not require estimation.

One of the difficulties with Census of Population data in interpretation of ethnical information is the changing definitions used in the collection of the data. From 1920 through 1950 the data collected includes a category "foreign born" at the level of the census tract but does not include a category "foreign stock." Foreign stock is defined to include all foreign born plus all native born with foreign or mixed parents, meaning one foreign parent and one native parent (U.S. Bureau of

Census 1942, 1952, 1962, 1972, 1981). However, the ratio of these 2 numbers, defined as the Pav index by the authors, is of interest, and can vary from zero to one.

First, in those census tracts where native born children of foreign or mixed parents leave a neighborhood, the ratio will increase from one time interval to another. Second, in a census tract that functions as a gateway to foreign immigrants, the ratio will increase from one time interval to another. Third, in a census tract that functions as a gateway to foreign immigrants, the ratio will be high because many immigrants come without families, and population figures are skewed toward adults. Finally, a census tract that attracts ethnic groups with high fertility ratios (blacks, Puerto Ricans, and Mexicans, for instance) will have a decreasing ratio.

CONCLUSION

Table 1 is an example of the initial transition matrix 1920-1930 computed from the actual census data under assumptions of transition within a closed system, as discussed earlier. Table 2 exhibits the expected ethnic composition at the end of selected 10-year periods if the determinants of the demographic trend established from 1920-1930 had persisted unchanged. The assumption is unrealistic, of course, but the magnitude of deviation from the actual demographic development pinpoints the relative importance of change in external factors. In fact, restrictions in the U.S. immigration policy, and the effects of the ever changing urban structure, require the computation of successive initial transition matrices. Several examples are presented in table 2. Expected ethnic composition for the year 2000 is based on the trends in the 1970s, calculated from the last census data on the Yugoslavian component available to the authors.

As the analysis demonstrated, pronounced trends were established in the 1920s that include a rapid increase in the group of native white with foreign or mixed parents, and rapid increase in the group of immigrants from Yugoslavia.

TABLE 1

Initial transition matrix 1920-1930.*

FROM \ TO	Native white w/ native parents	Native white w/ foreign/ mixed parents	Foreign born (other than Yu)	Foreign born from Yugo-slavia	Black
Native white w/ native parents	.9102	.0642	.0000	.0236	.0020
Native white w/ foreign/ mixed parents	.0000	1.0000	.0000	.0000	.0000
Foreign born (other than Yu)	.0000	.2128	.7023	.0781	.0068
Foreign born from Yugo-slavia	.0000	.0000	.0000	1.0000	.0000
Black	.0000	.0000	.0000	.0000	1.000

*Row totals may not sum to 1.000 because of rounding error. Defined in terms of 5 census tracts: 1112, 1113, 1115, 1116 and 1117.

The character of the St. Clair neighborhood was being clearly defined as a Slovenian and Croatian ethnic center (Bonutti and Prpic 1977). The other important categories, native white with native parents and other white immigrants, were declining; and the black group was minuscule. By 1970, two groups, Yugoslavs and blacks, were significantly expanding in the neighborhood. That St. Clair is still an entry point for immigrants, mainly Coratian, is remarkable; and if current trends persist immigrants from Yugoslavia will actually increase to 33% by the year 2000 while blacks will increase to 26% by the turn of the century. The black gains as a percentage are even more rapid, but the year 2000 ethnic ties to Europe also include 18% native white with foreign parent(s). Markovian analysis indicates that for the Census of 2000, over 50% of the St. Clair neighborhood

TABLE 2

Ethnic composition of St. Clair neighborhood 1920-1970, and predicted development until year 2000 (in percentages).*

Year	Source	Native white with native parents	Native white with foreign or mixed parents (2nd generation)	Immigrants from Yugoslavia (1st generation)	Other immigrants (foreign born) (1st generation)	Black
1920	Actual	15.4	43.0	18.1	23.4	.1
1930	Actual	14.0	48.9	20.2	16.5	.3
1940	Actual	27.8	42.0	17.3	12.8	.0
1950	Actual	39.9	36.0	12.6	11.0	.3
1960	Actual	44.7	29.5	12.9	11.1	1.7
1970	Actual	35.4	26.0	19.8	9.0	9.8
1980	Actual	—	—	—	—	16.1
1940	T20/30	12.7	53.4	21.8	11.6	.4
1950	T20/30	11.5	56.8	23.0	8.2	.5
1960	T20/30	10.5	59.4	23.9	5.8	.6
1970	T20/30	9.5	61.4	24.6	4.1	.7
1980	T20/30	8.6	62.9	25.1	2.9	.7
1960	T40/50	49.8	30.8	9.2	9.4	.5
1970	T40/50	57.8	26.4	6.7	8.1	.7
1980	T40/50	64.4	22.6	4.9	7.0	.9
1990	T40/50	69.8	19.4	3.6	6.0	1.0
2000	T40/50	74.3	16.6	2.6	5.2	1.1
1980	T60/70	28.0	22.9	25.4	7.3	16.3
1990	T60/70	22.2	20.2	30.0	5.9	21.6
2000	T60/70	17.6	17.8	33.8	4.8	26.0

*Row totals may not sum to 100.0 because of rounding error. Defined in terms of 5 census tracts: 1112, 1113, 1115, 1116 and 1117. T refers to the transition matrix defined by the census years indicated.

will still consist of foreign European stock (4 out of 5, Yugoslavian), over 25% will be black and over 5% will be other races.

The results indicate survival time for declining components of the state vector are on the order of 35-100 years; but the currents of change in American cities are frequently imposed as external forces on neighborhoods, and the results can be devastating, even to stable, long established geographical patterns.

The method of utilizing census data provides insight into the process of change. The Markovian approach is a general method that can be applied to a multitude of demographic and spatial topics where an hypothesis of transition is reasonable.

LITERATURE CITED

Berman, A. and R. J. Plemmons 1979 Non-negative matrices in mathematical sciences. Academic Press, NYC, NY. 316 p.

Bonutti, K. and G. J. Prpic 1977 Selected communities of Cleveland: A socio-economic study. Cleveland Urban Observatory, Cleveland, OH. 278 p.

Cesen, F. 1970 Odlomek iz zgodovine St. Clair avenije. Slovenski izseljenski koledar. Ljubljana: Slovenska izseljenska matica. p. 1-22.

Collins, L. 1974 Estimating Markov transition probabilities from micro-unit data. J. Royal Statist. Soc., Ser. C. 23: 355-371.

——— 1976 An introduction to Markov chain analysis. Concepts and techniques in modern geography, no. 1, Geo. Abstracts Ltd., Univ. East Anglia, Norwich, 35 p.

Kemeny, J. G. and J. L. Snell 1976 Finite Markov chains. Springer and Verleg, New York, NY. 224 p.

Ledbetter, E. E. 1918 The Yugoslavs of Cleveland, with a brief sketch of their historical and political backgrounds. The Mayor's Advisory War Committee, Cleveland, OH. 30 p.

Levy, D. 1972 A report on the location of ethnic groups in greater Cleveland. Cleveland State Univ., Inst. Urban Stud., Cleveland, OH. 83 p.

Pap, M. S. 1973 Ethnic communities of Cleveland. Inst. Soviet and East European Stud. John Carroll Univ., Cleveland, OH. 353 p.

Pavlakovic, V. K. and R. W. Janson 1984 Ethnic transition in Slovenian-Croatian neighborhood in Cleveland, Ohio, 1920-2000: A quantitative approach. Ohio J. Sci. 84(4): In press.

Sprecher, D. A. 1976 Finite mathematics. Harper and Row, Publ., NYC, NY. 366 p.

U.S. Bureau of Census 1942 1940 population and housing statistics for census tracts. Cleveland, Ohio and adjacent area. U.S. Government Printing Office, Washington, DC. p. 5, 66, 67.

——— 1952 1950 population census report. Vol. III, Chap. 12. Census tract statistics. Cleveland, Ohio and adjacent area. U.S. Government Printing Office, Washington DC. p. 22, 23.

——— 1962 U.S. census of population and housing: 1960. Final report PHC (1)-28. Census tracts. Cleveland, Ohio. Standard metropolitan statistical area. U.S. Government Printing Office, Washington, DC. p. 20, 22, 23.

——— 1972 1970 Census of population and housing. Census tracts. Cleveland, Ohio. Standard metropolitan statistical area. U.S. Government Printing Office, Washington, DC. SMSA P-43, Table P-2.

——— 1981 1980 Census of population and housing. Advance reports. PHC 80-U-37. Ohio. U.S. Government Printing Office, Washington, DC. p. 194.