80-70

INSTITUTE FOR RESEARCH ON POVERTY

THE ANALYSIS OF MOBILITY PROCESSES BY THE INTRODUCTION OF INDEPENDENT VARIABLES INTO A MARKOV CHAIN

by

Seymour Spilerman

DISCUSSION PAPERS

THE UNIVERSITY OF WISCONSIN, MADISON, WISCONSIN

THE ANALYSIS OF MOBILITY PROCESSES BY THE INTRODUCTION OF INDEPENDENT VARIABLES INTO A MARKOV CHAIN*

Seymour Spilerman
University of Wisconsin

*The research reported here was supported by funds granted to the Institute for Research on Poverty at the University of Wisconsin by the Office of Economic Opportunity, pursuant to provisions of the Economic Opportunity Act of 1964. I wish to thank David Dickens for his assistance with the statistical computations, and Arthur Goldberger, David McFarland, and Hal Winsborough for their comments. A draft of this paper was presented at the annual meetings of the American Sociological Association, Washington, D.C., September 1-3, 1970.

ABSTRACT

A major drawback with the traditional Markov formulation of social mobility is that it assumes homogeneity among persons in an origin state with regard to their transition behavior. This requirement has led to a deemphasis in consideration of the ways in which the transition probabilities vary among individuals. In this paper a regression procedure is introduced which allows a heterogeneous population to be examined within a Markov framework. The advantages of this formulation are threefold: (a) it allows the sources of variation in the population transition probabilities to be determined; (b) it facilitates distinguishing between over-time change in the transition probabilities which results from shifts in the population on particular attributes and genuine structural change in which the rules governing transitions have altered; (3) it enables individual-level transition matrices to be constructed which are necessary for projecting to the k-step population matrix in the presence of heterogeneity. The techniques developed in this paper are applied to an analysis of geographic migration.

THE ANALYSIS OF MOBILITY PROCESSES BY THE INTRODUCTION OF INDEPENDENT VARIABLES INTO A MARKOV CHAIN

Seymour Spilerman
University of Wisconsin

1. INTRODUCTION

The use of Markov chains as a structure for analyzing change in the distribution of a population has achieved considerable popularity in recent years. Although rather stringent assumptions must be met by the data for the Markov model to be formally applicable [see McFarland (1970) for a discussion of these requirements], it will reproduce most social processes reasonably well for a few time periods even when the requirements are not satisfied. This fact has encouraged its application in projecting growth and migration (Fuguitt 1965; Tarver and Gurley 1965), analyzing social mobility (Prais 1955; Matras 1960; Lieberson and Fuguitt 1967), and measuring social distance (Beshers and Laumann 1967), in addition to its more conservative use as a base line model for assessing the extent to which a process diverges from the Markov assumptions (Hodge 1966).

Aside from computational simplicity, the attractiveness of this model derives from its focus upon interrelationships as a <u>system</u>. Over-time change in the state distribution of a population is viewed as a consequence of the interdependencies acting through time and, perhaps, changing in the process as well. This formulation leads to a consideration of properties of dynamic systems such as "equilibrium" and "rate of convergence," concepts which rarely arise in causal analyses of change yet are fundamental to an understanding of adaptive social systems. Unfortunately, several difficulties emerge when applying Markov models to the study of social mobility, even when each <u>individual</u>

in the population conforms to the central assumption of a first order Markov process, namely that his probabilities of making particular transitions are determined solely by his present state and are independent of past history. These difficulties arise because traditional Markov theory is concerned with state changes by a single individual, while in applications to social mobility we consider the movements of an entire population.

(a) Heterogeneity. As several researchers (Blumen, Kogan, and McCarthy 1955; Hodge 1966; McFarland 1970) have noted, the Markov model fails to provide correct projections in the context of population heterogeneity, a condition which is common in many instances of social mobility. If individuals differ in either their rates of mobility or in the transition matrices which govern their movements, the Markov model will underpredict the observed proportion of the population who fail to move in subsequent time periods. There is consequently a need for devising projection methods that are suitable for a heterogeneous population.

(b) Change in the transition probabilities. Change over time in the elements of a transition matrix is commonly attributed to structural alterations in the underlying relationships. For example, successive intergenerational occupational mobility matrices may differ because technological developments have altered the demand for particular vocational skills. This is usually what is meant by a "non-stationary" process, namely that the rules which condition a son's occupational alternatives in terms of his father's position have changed. However, change in the elements of the population transition matrix can also result from heterogeneity. As McFarland (1970:469) has indicated, the p_{ijt} elements in the one-step population transition matrix at time t,

$$P_t(1) = \begin{pmatrix} p_{11t} & \cdots & p_{1mt} \\ \vdots & & \vdots \\ p_{m1t} & \cdots & p_{mmt} \end{pmatrix} \tag{1}$$

where $\sum_j p_{ijt} = 1$ for all i, represent averages of the origin state rows from the individual-level transition arrays. Thus $P_{t_o}(1)$ and $P_{t_1}(1)$, the one-step population transition matrices at times t_o and t_1, may be different because the distribution of individuals (and hence the weighted sum of the individual-level transition matrices) is different at the two time points <u>even though the rule governing transitions has remained unchanged for every person</u>. Consequently, in the presence of heterogeneity there is a need to distinguish between genuine structural change (where the rules of the game have altered) and changes in $P_t(1)$ which result solely from demographic shifts in the population.

(c) <u>The analysis of heterogeneity and change</u>. Related to the above considerations is the necessity for making sociologically meaningful statements about the components of heterogeneity and change. Since the Markov formulation focuses upon the <u>results</u> of a process (the population transition probabilities and the state distribution of the population) rather than upon the determinants of change, it has not been very useful as an analytic tool, as distinct from a projection device. With the explicit consideration of heterogeneity, however, two types of analyses become feasible: (1) the p_{ij} elements in the population transition matrix can be analyzed in terms of characteristics of individuals to determine which attributes are responsible for the heterogeneity, and (2) the extent to which demographic shifts over time in the one-step population transition matrix are the result of particular individual characteristics can be ascertained. Thus, by relaxing the requirement that all persons need transfer according to a single transition array, we are led to a consideration of the impact of individual differences on the population-level process, and to the importance of distinguishing between two sources of change in the aggregate matrix, that which results from structural alterations and from demographic shifts.

Having argued that these analyses become conceptually meaningful when heterogeneity is formally considered within a Markov framework directs our concerns to the manner by which they could be carried out. These questions are examined in this paper. A regression procedure is first introduced which allows the population heterogeneity to be attributed to particular characteristics of persons. This analysis also provides a framework for examining the change over time in the population transition matrix, and apportioning this change between demographic shifts and structural alterations. These topics are discussed in section 2. In section 3 consideration is given to the construction of individual-level transition matrices, and to projecting from these to the k-step population matrix.

2. HETEROGENEITY IN THE STATE TRANSITION PROBABILITIES

In order to contend with population heterogeneity within the Markov framework, Blumen, Kogan, and McCarthy (1955) proposed dividing the population into two types of persons--movers and stayers. The former are considered to be homogeneous in their transition behavior and to follow a Markov chain, while the latter are viewed as permanent residents in their origin states. More recently, Spilerman (1970) has extended this model to a continuous distribution of persons in terms of mobility rates. Each time an individual undergoes a state change, however, his destination is determined by a transition matrix M which is the same for all persons. Thus, in these models, the burden of explaining population heterogeneity is cast entirely upon variations in the rate of movement since, by assumption, the transition matrix is invariant across individuals.

An alternative approach to heterogeneity would be to assume that each individual c makes a single transition in a time interval but follows an M_c matrix which is unique to him, specifying his probabilities for transferring to the various destination states. Here, heterogeneity would mean that the taste or opportunity for making a particular transition is influenced by individual characteristics. Thus, in place of locating the impact of individual differences in the rate at which persons move, the present approach sees social characteristics as differentially affecting the transition probabilities.

Like the prior extension of the mover-stayer model (Spilerman 1970), this formulation is also a generalization of the Blumen, Kogan, and McCarthy (1955) approach to heterogeneity. One may conceive of the mover-stayer process as consisting of two types of individuals, all of whom move in each time interval. "Stayers" simply transfer according to the identity matrix. The present model is not, however, a generalization of the mover-stayer extension (Spilerman 1970). Were it not for the stochastic nature of that extension, it could be viewed as a process in which every person undergoes a transition in each time interval, with the heterogeneity being expressed by permitting persons to follow different matrices. Thus, while some follow the matrix I, others would transfer according to M, M^2, M^3, or some higher power of M. Now, if a person always moved according to the same M^p matrix, the present formulation would be a generalization of the mover-stayer extension since the individual transition arrays could be either powers of M or other matrices. However, the mover-stayer extension only requires the expected number of transitions to be constant for an individual; the actual number (and hence the power of M) varies over the time units in a random manner.

Thus, the present generalization of the mover-stayer model, which attributes heterogeneity to differences in the individual-level transition matrices, is not a generalization of the prior extension, but represents a conceptually different approach to heterogeneity. There are instances where either model could apply--industrial mobility and geographic migration are examples. There are also social processes for which only the present model would be applicable such as inter-generational mobility, where the notion of individual differences in the number of transitions made would seem conceptually inappropriate.

The problem of analyzing heterogeneity. Casting the problem of heterogeneity into this framework directs attention toward explaining the individual differences in the transition probabilities. If persons follow unique transition matrices, this fact should be attributable to differences among them in personal characteristics. Consistent with this reasoning, the common procedure for ascertaining the way by which transition probabilities vary with social attributes is to construct separate arrays for subpopulations. For example, Rogers (1966) and Tarver and Gurley (1965), using Markov formulations to analyze geographic migration, disaggregate the population to produce separate transition matrices by age and race. To subdivide a population in this manner is highly inefficient, however, since even with a large number of observations it is usually not possible to control on several attributes simultaneously and retain sufficient cases for estimating the p_{ij}'s of the transition matrix. Moreover, the generation of subpopulation matrices complicates the analytic problem of attributing the heterogeneity to particular factors since we must compare arrays which represent complex patterns of change, rather than single elements. Thus, even if it were possible to control on several variables simultaneously, the problem of disentangling the effects of these attributes would be immense.

The usual manner for handling complexities of this nature is to resort to regression analysis. In this section we present a method for incorporating independent variables into a transition matrix via a regression formulation and, using the resulting equations, constructing separate transition matrices for subpopulations.[1] More generally, we wish to consider the reverse of the problem addressed recently by McFarland (1970). Given individual-level transition matrices, McFarland shows how they may be aggregated to obtain the population transition matrix. Our problem, instead, is to efficiently <u>disaggregate</u> the transition matrix for a heterogeneous population. This is also a question of greater analytic importance since we seek to expose the effects of heterogeneity on the structure of the transition matrix.

<u>The regression model</u>. Consider the following approach. Let $P_{t_o}(1)$ be the observed t_0 to t_1 transition matrix for a population (see equation 1; the subscript t will henceforth be omitted whenever the time referent is clear). For the purpose of illustration we will assume that the substantive problem concerns geographic migration although this analysis would be equally applicable to industrial or occupational mobility. Each cell entry p_{ij} in P(1) therefore represents the proportion of individuals in system state (geographic region) S_i at time t_0 who have migrated to S_j by t_1. Since P(1) is a population transition matrix it has been estimated from the movements of individuals who presumably differ from one another in their probabilities of making a particular transition. For example, the transition from state S_i to S_j may be a function of age, race, occupation, region of birth, or of other variables.

Instead of the usual procedure of disaggregating the population into separate transition arrays to reduce the heterogeneity on these variables, we construct m^2 regression equations, where m equals the number of system states. For each origin state S_i at t_0 define an individual level variable y_{ij} which equals one if a person moved to S_j by t_1, and is zero otherwise. Corresponding to each origin state S_i, m such variables can be defined $(y_{i1}, y_{i2}, \ldots, y_{im})$. Exactly one element of this vector will equal one for an individual, specified by his destination state. Now, using as observations all individuals in S_i at t_0, regress

$$y_{ij} = a_{ij} + \sum_{k}^{K} b_{ijk}X_k + e_{ij} \qquad \text{for } j = 1, \ldots, m \tag{2}$$

where the X_k's are individual characteristics which are expected to explain the heterogeneity in the transition probabilities, and e_{ij} is the error term for the ij^{th} equation. This procedure will yield m equations for each origin state S_i, and m^2 equations for the full array.

Consider the ij^{th} entry of the array P(1). For an individual c in S_i at t_0, having attribute profile $(\acute{X}_{1c}, \ldots, \acute{X}_{Kc})$,

$$\hat{y}_{ijc} = \hat{a}_{ij} + \sum_{k} \hat{b}_{ijk}\acute{X}_{kc} \tag{3}$$

is an estimate of his probability of making the i-j transition. This probability interpretation results from the dependent variable having been coded 0-1, and will be of importance in subsequent sections. With the dependent variable coded in this manner, $\hat{y}_{ijc}$ can also be interpreted as the expected number[2] of transitions by individual c from S_i to S_j.

The expected number interpretation is important because, unlike probabilities, these values can be summed over observations. Thus, for the subpopulation in S_i at t_0, the expected number of S_i to S_j transitions is given by $\sum_{c=1}^{N_{i.}} \hat{y}_{ijc}$, where $N_{i.}$ is the number of persons in location S_i at t_0. Now, letting

N_{ij} equal the number of these persons who have moved to state j by t_1, we have

$$\sum_{c}^{N_{i.}} \hat{y}_{ijc} = \sum_{c}^{N_{i.}} y_{ijc} = N_{ij}$$

where the first equality derives from the least squares procedure of fitting a regression line, and the second from the definition of y_{ij}. Since the standard estimation formula for the transition probabilities[3] of a Markov chain is $p_{ij} = \frac{N_{ij}}{N_{i.}}$ we obtain

$$\frac{1}{N_{i.}} \sum_{c}^{N_{i.}} \hat{y}_{ijc} = \frac{N_{ij}}{N_{i.}} = p_{ij}$$

Therefore, substituting from equation 3,

$$p_{ij} = \frac{1}{N_{i.}} \sum_{c}^{N_{i.}} \hat{y}_{ijc} = \hat{a}_{ij} + \sum_{k}^{K} \hat{b}_{ijk} \left(\frac{\sum_c X_{kc}}{N_{i.}} \right) \qquad (4)$$

where the terms $\sum_c X_{kc}/N_{i.}$ for $k = 1, \ldots, K$ can be interpreted as the "typical" profile of an individual in S_i at t_0.

Equation (4) gives the desired decomposition of p_{ij} in terms of the conditioning variables $(X_1, \ldots, X_K)$. We therefore have a matrix of m^2 equations,

$$Y = \begin{pmatrix} y_{11} = \hat{a}_{11} + \Sigma\hat{b}_{11k}X_k, & \ldots, & y_{1m} = \hat{a}_{1m} + \Sigma\hat{b}_{1mk}X_k \\ \vdots & & \vdots \\ y_{m1} = \hat{a}_{m1} + \Sigma\hat{b}_{m1k}X_k, & \ldots, & y_{mm} = \hat{a}_{mm} + \Sigma\hat{b}_{mmk}X_k \end{pmatrix} \tag{5}$$

in which the b's which are significant in an equation indicate the social characteristics responsible for heterogeneity with respect to that transition. In an equation where all b's are insignificant, $y_{ij} = a_{ij}$ and no heterogeneity is present.

The matrix Y of regression equations therefore reveals the components of heterogeneity in the population transition matrix. Moreover, P(1) can be calculated directly from Y. When the regressions are evaluated on the "typical" individual profile at each origin state S_i for $i = 1, \ldots, m$ we obtain[4] by equation (4) $P(1) = \hat{Y}$.

Disaggregating the population transition matrix. In addition to reproducing the P(1) matrix, the array Y of regression equations enables transition matrices to be constructed for subpopulations [$P(1)_s$ matrices]. For example, if one of the regression variables were a dummy for race (e.g., $X_R = 1$ for non-whites, zero for whites) then the non-white transition matrix $P(1)_{nw}$ could be estimated from Y by evaluating the regression equations on individuals having $X_R = 1$ --

$$\hat{P}(1)_{nw} = \left(\hat{p}_{ij} = \frac{1}{N_{i.}} \sum_{c_R} \hat{y}_{ijc} \right) \tag{6}$$

where $N_{i.}$ is the number of non-whites in S_i at t_0, and $\hat{y}_{ijc}$ is specified by equation (3). This estimate of the non-white transition matrix will not be identical to the observed non-white matrix $P(1)_{nw}$ since the error term $\hat{e}_{ij}$ satisfies $\sum_c \hat{e}_{ijc} = 0$ for the entire population that is used to estimate the regression surface, but not necessarily for a subset of the observations.

The estimates of p_{ij} in $\hat{P}(1)_{nw}$ have been made under the assumption that non-whites have b-weights which are identical to those of whites on all variables except for a term for race (and factors which are interactions with race if these are present). The $\hat{p}_{ij}$ value for non-whites will therefore differ from its white counterpart as a result of the additive effect of the race term b_R and because of racial differences in the individual profiles of attributes (X_k's). In this manner, the matrix Y of regression equations can be used to construct a transition array for persons with any given attribute or combination of characteristics.

When dummy variables are used with a single dimension (such as age) or where dummy variable regressors are provided for the main effects and all possible interactions among several dimensions (such as when categorical variables are used for age deciles and for race), then the evaluation of the Y matrix on individuals having a particular attribute combination (e.g., 20 to 30 years old and Negro) will yield identical results to what would be obtained from constructing a separate transition matrix from the observed movements of this subpopulation. The advantages of the regression format are threefold: First, it allows a variable such as age to be treated as continuous. This means a loss of fewer degrees of freedom, and an opportunity to simultaneously consider several variables. Second, as a result of using a continuous variable formulation it becomes possible to extrapolate to values for which few or no observations are present. Third, subject to the assumptions of the regression equation, it is possible to construct a separate transition matrix for each <u>person</u> in the population. This point will be developed in section 3.

Geographic migration. Using data made available by Karl Taeuber from his analysis of residential mobility in the United States (Taeuber et. al., 1968), the above procedures were applied to inter-regional transitions by males. The Taeuber data were collected in 1958 from retrospective reports about prior residences and are described in detail elsewhere (Taeuber et. al., 1968).[5] For the purpose of this study four geographic regions were defined as states of the process: (1) Northeast, (2) North-central, (3) South, and (4) West. The time points that were used are $t_0 = 1937$, $t_1 = 1944$, $t_2 = 1951$, and $t_3 = 1958$. These were selected to provide residence histories for the adult years of this cohort. The t_0-t_1 transition matrix for the population, calculated from the observed movements, is presented in Table 1, together with the number of persons represented by each transition probability.

Table 1 about here

Individual-level data are available from this study for a number of population characteristics: birthplace, race, age, occupation in 1958, class of worker in 1958, city size, and duration of stay at each residence. For this illustration the effect of each variable was assumed to be additive. Also, with the exceptions of age and duration of residence, for which there is evidence for non-linear effects (Land 1969:139; Morrison 1970:11, 14), all relationships were assumed to be linear or binary. The following regression model was therefore used,[6] the observations being all persons in origin state S_i in 1937:

$$y_{ij} = a_{ij} + \sum_{k=1}^{17} b_{ijk}X_k + e_{ijk} \qquad \text{for } i,j = 1, \ldots, 4 \quad (7)$$

where $y_{ij} = 1$ if a person was in S_i in 1937 and S_j in 1944, and zero otherwise,

TABLE 1. Population Transition Probabilities for Inter-regional Migration, 1937-44

A. Transition Matrix

$$P_{1937}(1) = \begin{pmatrix} .9705 & .0096 & .0121 & .0078 \\ .0069 & .9473 & .0154 & .0304 \\ .0111 & .0280 & .9383 & .0226 \\ .0027 & .0150 & .0168 & .9655 \end{pmatrix} \quad \begin{array}{c} \underline{\sum_j P_{ij}} \\ 1.000 \\ 1.000 \\ 1.000 \\ 1.000 \end{array}$$

B. Number of Individuals

$$N_{1937}(1) = \begin{pmatrix} 3358 & 33 & 46 & 27 \\ 29 & 3991 & 65 & 128 \\ 47 & 119 & 3984 & 96 \\ 3 & 17 & 19 & 1090 \end{pmatrix} \quad \begin{array}{c} \underline{N_{i.}} \\ 3460 \\ 4213 \\ 4246 \\ 1129 \end{array}$$

and X_1 = 1 if upper white collar, zero otherwise

X_2 = 1 if lower white collar, zero otherwise

X_3 = 1 if upper blue collar, zero otherwise

X_4 = 1 if lower blue collar, zero otherwise

(the farming trades were taken as the base category for the occupation variables[7])

X_5 = 1 if privately employed, zero otherwise

X_6 = 1 if government employee, zero otherwise

("self-employed" was taken as the base category for the class-of-worker variables)

X_7 = 1 if non-white, zero otherwise

X_8 = 1 if resident of a large city in 1937, zero otherwise

X_9 = 1 if resident of a medium-sized city in 1937, zero otherwise

("small city" was taken as the base category for the city size variables)

X_{10} = age in 1937

X_{11} = years at current residence as of 1937

X_{12} = number of residences as of 1937

X_{13} = age in 1937 (squared)

X_{14} = years at current residence (squared)

X_{15} = 1 if born in region 2, zero otherwise

X_{16} = 1 if born in region 3, zero otherwise

X_{17} = 1 if born in region 4, zero otherwise

(birthplace in region 1 was taken as the base category for the birthplace variables[8]).

The most important from among these variables, as judged by a partial F criterion, was selected from each equation by means of a stepwise regression procedure.[9] The use of this method is justified when one is concerned with obtaining an efficient set of predictors by deleting superfluous variables, as is our interest here, rather than with testing theories about particular variables. The resulting equations are presented in Table 2.

Table 2 about here

It is evident from the R^2 values in the bottom row that these 17 variables do not explain very much of the variation among individuals in migration behavior. None of the entries exceeds .05 in magnitude. Despite these small values it is possible that the most important demographic variables which influence migration have been included in the analysis. When dealing with individual-level data relating to the occurrence or non-occurrence of an event, a very large idiosyncratic component may be involved so that even small increases in the explained variance require the consideration of many additional variables [see Stinchcombe (1968:67-68) on the relationship between explanations at the individual and at the aggregate level]. In fact, in order for the homogeneity requirement of the Markov model to be satisfied, all explanatory variables would have to be insignificant so that the migration differences among individuals become chance factors, unrelated to their social attributes.

It must also be remembered that the R^2 values depend upon how well the independent variables distinguish among the movements of all persons in an origin state. For example, if Negroes were to make a particular transition, but whites were as likely to do so as not, the R^2 value would be small even though the regression equation would predict perfectly the migration behavior of every Negro. It is the b-weights for particular variables and not the R^2 values which must concern us. Using these we could distinguish the migration behavior of Negroes from that of whites and, in general, the differences among any subgroups.

TABLE 2. Unstandardized Regression Coefficients Estimated from Inter-regional Transitions, 1937 - 44

Independent Variable	Equation for[a]-- y_{11}	y_{12}	y_{13}	y_{14}	y_{21}	y_{22}	y_{23}	y_{24}
Constant	.942	.0299	.0585	.0240	.107	.862	.0394	.0668
X1	-.0236	.00988	.0115		.00908	-.0231	.0148	
X2								
X3								.0172
X4								
X5						-.0250		
X6			.0112			-.0453		.0245
X7	.0417	-.0209	-.0220			.0409	-.0256	-.0296
X8		-.00970						
X9								
X10			-.00312		-.00300	.00138	-.000674	
X11		-.000714				.00271	-.000479	
X12	.00905	-.00613		-.00450		.0113		-.0107
X13			.000436		.0000350			
X14	.0000573			-.0000239				-.0000696
X15	-.157	.133		.0302	-.0475			
X16	-.0440	.0265	.0348		-.0472		.0275	
X17						-.171	.123	
R^2	.030	.049	.007	.005	.019	.024	.015	.011

[a] Each reported entry has a t-value greater than 2.00.

TABLE 2 (con't)

Independent Variable	y_{31}	y_{32}	y_{33}	y_{34}	y_{41}	y_{42}	y_{43}	y_{44}
Constant	.106	.0232	.912	.0225	.0620	.0673	.0534	.824
X1		.0280	−.0340					
X2		.0337	−.0397					
X3		.0308	−.0360					
X4		.0248	−.0354	.0165				
X5	.00682		−.0202	.00937				
X6								
X7	.0217	.0316	−.0636					
X8					−.0104			
X9			−.0174	.0120	−.00864			
X10								
X11		−.000761				−.00583	−.00144	.00884
X12		−.00817			−.00342	−.0109	−.00873	.0222
X13	−.00000562		.0000239	−.0000111				
X14					−.0000165	.000152		−.000202
X15	−.0941	.0601			−.0437	.0406		
X16	−.0999		.0647		−.0444		.0382	
X17					−.0435			.0411
R^2	.026	.016	.032	.008	.038	.042	.017	.043

Turning to a consideration of the determinants of population heterogeneity in geographic migration, the equations in Table 1 exhibit the following main patterns:

(1) Upper white collar persons are more likely to make inter-regional transitions, irrespective of their region of origin (except for the West).

(2) Negroes in the North and Midwest are less likely to move; Negroes in the South are more likely to change region. On the average, a southern Negro has a probability of migrating that is .06 larger than the value for a white man in this region with an identical attribute profile (note the coefficient for X_7 in equation y_{33}).

(3) The probability of not making a regional change increases with duration at a residence.

(4) There is a strong tendency for persons to return to their region of birth, irrespective of which region this is or where they currently reside.

These appear to be the main sources of population heterogeneity with respect to regional change. Not only does the matrix of regression equations provide this information, it also permits separate transition matrices to be constructed for each subpopulation.[10] This is accomplished by evaluating Y over individuals having the appropriate attributes, in the manner described in connection with equation 6.[11] Matrices for Negroes and for upper white collar persons were estimated and are presented in Table 3, alongside the observed transition arrays for these groups. It is evident that this simple regression model, which assumes the effects of the attributes to be entirely additive, reproduces the observed subpopulation matrices reasonably well. These arrays should also be compared with the transition matrix for the entire population (Table 1). The important comparisons involve the main diagonal elements since the number of observations in the off-diagonal cells is very small. These comparisons suggest much the same findings as were reported above in points (1) and (2). The results are not identical, however, since the equations present the net effects of particular attributes, while other characteristics of individuals are not controlled for in the subpopulation transition matrices. Finally, as an example of a highly mobile subgroup, the matrix for white, white collar persons, residing outside their regions of birth in 1937, is presented in row C of Table 3.

Table 3 about here

3. ANALYZING CHANGE OVER TIME

Two problems arise in the analysis of mobility processes in a Markov formulation. There is the problem of determining the extent to which demographic shifts in the population are responsible for changes in the one-step population transition matrix, and there is the question of how to project forward in time to the k-step transition matrix in the presence of heterogeneity.

Change in the one-step transition matrix over time. The first problem is an analytic one, and can be posed in the following way. We have profiles of persons which contain, among other attributes, their locations at equally spaced time points, t_0, t_1, t_2, t_3,.... Using this information, we can construct one-step transition matrices $P_{t_0}(1)$, $P_{t_1}(1)$, $P_{t_2}(1)$, etc. These arrays will normally differ from one another as a result of two processes. First, the influence of particular attributes on an individual's probability of making a given transition may change; for example, the demand for skilled labor in a geographic region may decrease, so the contribution from this occupational affiliation to a decision to migrate there would be lowered. Change of this nature

TABLE 3. Predicted and Observed Transition Matrices for Subpopulations, 1937-44

Predicted[a]				$\sum_j \hat{p}_{ij}$	Observed				$N_{i.}$
				A. Negroes					
.9862	.0001	.0068	.0098	1.003	.9687	.0000	.0067	.0067	150
.0045	.9827	.0058	.0058	.999	.0058	.9825	.0058	.0058	171
.0275	.0497	.8885	.0289	.995	.0275	.0496	.8881	.0348	947
.0020	.0112	.0478	.9372	.996	.0000	.0000	.0000	1.0000	12
				B. White collar					
.9496	.0185	.0207	.0082	.997	.9501	.0185	.0206	.0109	921
.0137	.9321	.0266	.0257	.998	.0138	.9318	.0267	.0277	1012
.0072	.0278	.9461	.0184	1.000	.0122	.0277	.9456	.0144	901
.0018	.0168	.0161	.9645	.999	.0000	.0239	.0269	.9493	335
				C. White, white collar, residing outside region of birth in 1937					
.8422	.1011	.0302	.0264	1.000	.7963	.1481	.0370	.0185	54
.0379	.9208	.0450	.0293	1.033	.0313	.9064	.0469	.0156	64
.0426	.0593	.8934	.0178	1.013	.0385	.0641	.8974	.0000	78
.0037	.0321	.0185	.9447	.999	.0000	.0403	.0268	.9329	149

[a] Predictions obtained by evaluating Y on the subpopulation with the requisite attributes.

means that the b-coefficients in the regressions (equation 5) will differ in successive time periods. This question could be pursued by reestimating the matrix Y of regression equations from the transitions in subsequent time intervals, and examining the sequence of b-coefficients corresponding to each of the variables.

However, the $P_t(1)$ matrix could also change over time even though the factors which determine the movements of individuals, the coefficients in the regression equations, remain constant. This would result from a shift in the distribution of individuals among the system states or, more generally, from a change in the demographic structure of the population, that is to say, in the way that particular attributes are distributed in the population. For example, as our society becomes more educated the average earnings of individuals will increase, even though the income return from education may remain unchanged. It is a shift in the distribution of the population on the education variable which is responsible for the alteration.

An example of where these considerations would be relevant can be drawn from the work of Robert W. Hodge (1966). Using Gladys Palmer's study of intragenerational occupational mobility in six cities (Palmer 1954), Hodge computes transition matrices for males during 1940-44 and 1945-49. These matrices are not identical, so one may wish to assess the extent to which the differences are due to changes in factors such as age of the cohort, level of educational attainment, and industry of employment.

These questions could be examined by estimating the matrix Y of regression equations from data in the initial time interval, then using Y together with the profile of individual characteristics at t_1 (1945) to generate $\hat{P}_{t_1}(1)$, and the individual profiles in subsequent time periods, if these are available, to estimate later transition matrices. A measure of the effect of demographic shifts on the p_{ij} elements, under the assumption of a constant relationship between the p_{ij}'s and the individual characteristics, can be obtained by comparing the observed t_1 to t_2 transition matrix, $P_{t_1}(1)$, with $\hat{P}_{t_1}(1)$, the matrix estimated from Y. The contribution from population shifts is indicated by the extent to which $\hat{P}_{t_1}(1)$ provides a better estimate of $P_{t_1}(1)$ than is given by $P_{t_0}(1)$, the estimator that would be used with a stationary Markov chain. This procedure therefore measures the change which should occur in a transition matrix from demographic shifts alone.

The above discussion can be illustrated with the geographic migration data. For each time interval, 1944-51 and 1951-58, the matrix Y of regression equations was evaluated on the profiles of individual characteristics from the appropriate origin year. This operation permits transition matrices to be constructed which are based on the relationships between individual attributes and migration propensities <u>which existed during the time interval 1937-44</u>. By evaluating the matrix Y on the individual profiles at subsequent time points we are able to attribute all of the resulting change in the matrix elements to shifts in the distribution of the population on particular attributes.

The results of this analysis are presented in Tables 4 and 5. In Table 4, the predictions from the above procedure are reported alongside the observed transition matrices for the intervals 1944-51 and 1951-58. Comparable arrays are not presented for the period 1937-44 since, by the requirements of ordinary least squares regression, the predicted population transition matrix will be identical to the observed one for this period. The 1937-44 matrix is presented in Table 1.

<u>Tables 4 and 5 about here</u>

TABLE 4. Predicted and Observed One-Step Transition Matrices for 1944-51 and 1951-58

Predicted[a]				$\Sigma\hat{p}_{ij}$	Observed				N_i
A. 1944-51									
.9733	.0084	.0110	.0061	.999	.9645	.0087	.0122	.0145	3437
.0029	.9621	.0107	.0251	1.001	.0048	.9575	.0120	.0257	4160
.0090	.0256	.9502	.0172	1.002	.0114	.0255	.9494	.0136	4110
.0014	.0144	.0163	.9668	.999	.0082	.0291	.0157	.9475	1341
B. 1951-58									
.9748	.0077	.0141	.0054	1.002	.9803	.0047	.0091	.0059	3393
.0025	.9758	.0059	.0218	1.006	.0022	.9750	.0082	.0147	4157
.0058	.0229	.9651	.0104	1.004	.0057	.0134	.9701	.0107	4015
.0009	.0095	.0131	.9756	.999	.0013	.0088	.0067	.9831	1483

[a] Predictions obtained by evaluating Y on the profiles of individual attributes at the origin year.

TABLE 5. Indices of Dissimilarity Between Observed and Predicted Transition Matrices for 1944-51 and 1951-58

Region of origin	Index of Dissimilarity[a]: Between observed 1944-51 transition matrix and-- Observed 1937-44 Matrix	Between observed 1944-51 transition matrix and-- Predicted 1944-51 Matrix	Between observed 1951-58 transition matrix and-- Observed 1937-44 Matrix	Between observed 1951-58 transition matrix and-- Predicted 1951-58 Matrix
S_1^b	.7	.9	1.0	.8
S_2	1.0	.4	2.8	.5
S_3	1.2	.3	3.2	.8
S_4	2.0	2.1	1.8	.8
All origin states combined[c]	1.3	.6	2.3	.8

[a] $ID = 1/2\ \Sigma |P_i - \hat{P}_i|$. See Taeuber and Taeuber (1965:236) for a discussion of this index.

[b] Each entry gives the percentage of persons who would have to change their destination states to bring the observed distribution into accord with the predicted.

[c] The average of the indices, weighted by the number of observations in an origin state, indicates the proportion of all persons whose destinations would have to be changed to bring the observed distribution into accord with the predicted.

Comparing the observed transition arrays for the three time periods, it is evident that the amount of change in $P_t(1)$ is not very great. Nevertheless, some trends are apparent, the most pronounced being the tendency for the main diagonal elements to increase in size over time. As was suggested, this could be a consequence of either structural change--an alteration in the attractiveness of regional migration for individuals who have a particular array of attributes, or of demographic shifts--change in the distribution of the population over the profile of attributes, including region of residence. However, the predicted arrays in Table 4 make it evident that the pattern of change noted with the observed data can be accounted for by the relationships between regional migration and individual characteristics which existed during 1937-44. Thus, the increase in the size of the main diagonal elements appears to be largely a consequence of demographic shifts, and not due to structural change.

Normally, to estimate the one-step transition matrix $P_{t_i}(1)$ at some time point t_i subsequent to t_0 we would use $P_{t_0}(1)$. That is, having no knowledge about how the transition matrix is changing, we assume a stationary Markov chain. To investigate the change in $P_t(1)$ over time, and the proportion of this change which can be attributed to demographic shifts on the 17 variables considered in this analysis, Indices of Dissimilarity were computed between several pairs of arrays. These compare the observed transition matrices at 1944 and 1951, in Table 4, with $P_{1937}(1)$ from Table 1 to obtain measures of the total change in this array. Also, the observed and predicted matrices for 1944 and 1951 (from Table 4) are compared in order to measure the amount of residual change, which is not accounted for by population shifts. The results are presented in Table 5.

In the first four rows index values are reported for each region of residence in 1944 and 1951. Each entry has an interpretation as the percentage of persons in the origin state who would have to change their destinations to bring the observed distribution of moves into accord with the predicted distribution (Taeuber and Taeuber 1965:236). For example, the first two entries in row 2 indicate that a 1.0 percent change has taken place in the pattern of movement of individuals residing in region 2 during the period 1937 to 1944, but only 0.4 percentage points cannot be accounted for by demographic shifts. The comparable values for changes in the row 2 probabilities between 1937 and 1951 are 2.8 for the total amount of change, and 0.5 for the residual change which is not attributable to demographic shifts on the 17 variables.

Averages of the entries in each column, weighed by the number of persons in an origin state, are presented in row 5. These values report the proportions of the total population who would have to change destinations in order to bring the pair of matrices being compared into agreement. Although the total amount of change in the observed matrix is small, 1.3 percent between $P_{1937}(1)$ and $P_{1944}(1)$, 2.3 percent between $P_{1937}(1)$ and $P_{1951}(1)$, less than half of these values is attributable to structural alterations (0.6 percentage points for the first pair of arrays, 0.8 percentage points for the second pair). Thus, a major proportion of the change in the migration behavior of this cohort is due to shifts in the distribution of the population, and not to alterations in underlying relationships.

The projection problem. In the absence of heterogeneity, P(k), the observed k-step transition matrix, may be estimated from the Markov property

$$\hat{P}(k) = \prod_{i=1}^{k} P_{t_i}(1) \tag{8}$$

which reduces to the familiar relation, $\hat{P}(k) = P^k$ when P is constant over time.

The projection problem arises because equation (8) does not hold in the presence of population heterogeneity. In this situation, the main diagonal elements of $P(k)$ will exceed the corresponding entries on the diagonal of $\hat{P}(k)$.

However, if we assume that each person independently is following a Markov process, then, using the matrix Y of regression equations, we can obtain an estimate of $P(k)$ in the presence of heterogeneity.[12] This is accomplished in the following manner: For each individual c a separate transition matrix $M_c(1)$ is constructed by evaluating the 16 equations[13] in Y on his attribute profile in 1937.[14] Now, let N_c be a matrix with 1 on the main diagonal of the i-th row and zero in all other cells, where i denotes individual c's region of residence in 1937. We then have (McFarland 1970:469)

$$P(1) = N^{-1} \sum_c N_c M_c(1) \qquad (9)$$

where N is a diagonal matrix, $N = \Sigma N_c$. Since each person is assumed to follow a Markov chain with his individual transition matrix $M_c(1)$, the k-step population transition matrix is given by

$$P(k) = N^{-1} \sum_c N_c \, M_c^k(1) \qquad (10)$$

Equation (10) holds under the assumption that the individual transition matrices $M_c(1)$ are constant over time. However, using the methods of the previous section we can relax this requirement and incorporate over time changes in the elements of the $M_c(1)$ arrays <u>which affect all persons in an identical manner</u>. The most common example of this is the effect of age. Change on this variable can be handled by generating time specific transition matrices $M_{ct}(1)$ in which the age contribution for an individual depends upon the time period t to which the matrix pertains. The k-step transition matrix then becomes

$$P(k) = N^{-1} \sum_c N_c \prod_t M_{ct}(1) \qquad (11)$$

Although equation (11) represents a time-varying Markov process, all relationships and profile data are from the initial time period. The changing impact of age on an individual's transition probabilities is estimated from the cross-sectional relationships in this time interval. Consequently, projection from equation (11) is comparable to projection in a time-homogeneous Markov process in that both models require data at only two time points for parameter estimation. This method departs from the Markov formulation by explicitly incorporating the effects of population heterogeneity and, for this reason, individual-level data are required.

The projection procedure can be illustrated with the geographic migration data. Using the individual profiles from 1937 together with a matrix Y' of regression equations, three transition matrices[15] were estimated for each individual: $M_{c1937}(1)$, $M_{c1944}(1)$, and $M_{c1951}(1)$. The array Y' differs from Y, the matrix of regression equations presented in Table 2, in that variables whose values change with the occurrence of a transition have been removed. Without making further assumptions about the migration process we have no knowledge concerning characteristics such as city size, duration of residence, or number of residences for an individual subsequent to 1937.[16] For the purpose of projection, the regressions reported in Table 2 were therefore reestimated with variables X_8, X_9, X_{11}, X_{12}, and X_{14} deleted. By this procedure, the three $M_{ct}(1)$ matrices for an individual will differ only as a result of the effect of the seven year-age interval between successive evaluations of Y'.

Estimates of $P_{1937}(2)$ and $P_{1937}(3)$ were constructed by using the $M_{ct}(1)$ matrices with equation (11). These arrays are presented in row 3 of Table 6, below the observed two and three-step transition arrays (row 1), and the estimates of these matrices from a time-dependent Markov chain in which the observed one-step population transition matrices for successive time intervals have been multiplied together following equation (8) (row 2). Comparing the main diagonal entries it is evident that the estimates from the present model are superior to the Markov projections even though the latter are based upon observed transitions covering all time periods. This impression is confirmed by computing Indices of Dissimilarity between the observed two and three-step arrays and the projected matrices. These results are reported in Table 7, separately for each origin state in 1937 and for comparisons between entire matrices. The entries in the last row indicate that while the observed two-step matrix is in disagreement with the Markov projection by 1.6 percent, the discrepancy with the projected array obtained from the present method is 1.0 percent. The corresponding index values for the three-step projections are 2.4 and 1.8 percent. Thus, with the geographic migration data one can do better in projecting from an assumption of heterogeneity and using only data from the initial time interval, than by assuming a Markov process at the population level and using the observed one-step transition matrices from each time interval.

Tables 6 and 7 about here

4. CONCLUSIONS

A major drawback with the traditional Markov formulation of social mobility is that it assumes homogeneity among persons in an origin state with regard to their transition behavior. This requirement leads to an underestimation of the main diagonal entries when projecting forward in time to the k-step matrix in the presence of heterogeneity, but what is more important conceptually, the

TABLE 6. Observed and Projected Two and Three-Step Population Transition Matrices for 1937-44 and 1937-51

A. Observed Matrices

$$P_{1937}(2) = \begin{pmatrix} .9477 & .0145 & .0194 & .0185 \\ .0083 & .9238 & .0176 & .0503 \\ .0177 & .0445 & .9074 & .0304 \\ .0035 & .0230 & .0186 & .9548 \end{pmatrix} \qquad P_{1937}(3) = \begin{pmatrix} .9344 & .0176 & .0251 & .0288 \\ .0088 & .9084 & .0199 & .0629 \\ .0203 & .0520 & .8900 & .0377 \\ .0035 & .0239 & .0177 & .9548 \end{pmatrix}$$

B. Projected Transition Matrices Using Markov Formulation (equation 8)

$$\hat{P}_{1937}(2) = P_{1937}(1)P_{1944}(1) = \begin{pmatrix} .9362 & .0180 & .0235 & .0219 \\ .0116 & .9084 & .0266 & .0534 \\ .0217 & .0515 & .8916 & .0351 \\ .0108 & .0430 & .0313 & .9150 \end{pmatrix}$$

$$\hat{P}_{1937}(3) = P_{1937}(1)P_{1944}(1)P_{1951}(1) = \begin{pmatrix} .9180 & .0226 & .0318 & .0276 \\ .0136 & .8865 & .0337 & .0662 \\ .0266 & .0627 & .8658 & .0449 \\ .0120 & .0504 & .0370 & .9006 \end{pmatrix}$$

C. Projected Transition Matrices Using Method of Equation (11)[a]

$$\hat{P}_{1937}(2) = \begin{pmatrix} .9486 & .0179 & .0218 & .0153 \\ .0086 & .9147 & .0238 & .0535 \\ .0181 & .0479 & .8950 & .0381 \\ .0061 & .0290 & .0320 & .9414 \end{pmatrix} \qquad \hat{P}_{1937}(3) = \begin{pmatrix} .9336 & .0254 & .0333 & .0225 \\ .0098 & .8978 & .0268 & .0706 \\ .0211 & .0616 & .8669 & .0477 \\ .0101 & .0424 & .0461 & .9288 \end{pmatrix}$$

---

[a]Negative estimates of individual probabilities were set equal to zero, values greater than one were set equal to one.

TABLE 7. Indices of Dissimilarity Between Observed and Projected Two and Three-Step Transition Matrices

Region of origin	Index of Dissimilarity[a]: Between $P_{1937}(2)$ and-- $\hat{P}_{1937}(2)$ (Markov)	$\hat{P}_{1937}(2)$ (equation 11)	Between $P_{1937}(3)$ and-- $\hat{P}_{1937}(3)$ (Markov)	$\hat{P}_{1937}(3)$ (equation 11)
S_1[b]	1.2	.5	1.8	1.2
S_2	1.5	.8	2.2	1.3
S_3	1.6	1.2	2.4	2.1
S_4	4.0	1.8	5.4	4.0
All origin states combined[c]	1.6	1.0	2.4	1.8

[a] $ID = 1/2\ \Sigma|P_i - \hat{P}_i|$.

[b] Each entry gives the percentage of persons who would have to change their destinations to bring the observed distribution into accord with the projected distribution.

[c] The average of the indices, weighted by the number of observations in an origin state, indicates the proportion of all persons whose destination states would have to be changed to bring the observed distribution into accord with the projected.

Markov model diverts attention from a consideration of the determinants of the transition probabilities. With the explicit consideration of heterogeneity, however, we are able to examine the transition behavior of the population in terms of the variables which condition the individual propensities. Thus, the question of who moves and where receives a sociological answer.

Previous attempts to incorporate heterogeneity within a Markov framework have taken the direction of permitting individual variations in the rate at which transitions occur (Blumen, Kogan, McCarthy 1955; Spilerman 1970). At each move, however, a transition matrix which is identical for all persons is used. The problem of heterogeneity was examined from a different perspective in this paper. Instead of requiring that it be expressed in individual differences in the rate of mobility, the strategy here was to permit the transition probabilities themselves to vary among persons. Each individual is assumed to follow a Markov process, but in accordance with his own transition matrix, which is estimated by the regression procedure introduced in section 2. With these individual-level matrices, projection to the k-step array may be accomplished in the context of population heterogeneity by employing the method proposed recently by McFarland (1970).

For sociological analysis a more fundamental matter than projection concerns the analysis of change. The regression formulation contributes to this objective by suggesting the necessity for distinguishing between two sources of change in the transition probabilities. Change may occur as a result of demographic shifts in the population, in which the rules governing individual transitions remain intact but the distribution of relevant attributes in the population has altered, or because of structural change, in which the rules themselves have altered. In examining geographic migration we found that the former process was primarily responsible for the over-time variations in the transition probabilities. Although the focus in this study was explicitly on accounting for demographic shifts in the population, with structural change being relegated to the residual category, structural change could also be analyzed directly. The most evident procedure would be to reestimate the regression model in successive time intervals and compare the changes in each regression coefficient over the time sequence.

Having now considered two alternative approaches to the incorporation of population heterogeneity within a Markov framework--by permitting variations in the rate of mobility or in the transition probabilities--what can be said about their comparative merits? Conceptually, it appears that there are some social processes for which only the latter model could apply, although both formulations are plausible for a wide range of mobility processes. The present model is preferable in studies of inter-generational mobility since the notion of different rates at which these transitions occur has little conceptual merit. With respect to intra-generational occupational mobility, change in industry affiliation, or geographic migration, both formulations seem plausible: persons may differ in the rate at which they move; also, individual characteristics may influence the probability of making particular transitions.[17] Thus, the higher geographic mobility of white collar persons may be a rate of mobility effect since this pattern was noted in all regions, while the differential mobility of Negroes, depending upon whether they were residing *in* the South or elsewhere, could be ascertained only by the model of this paper.

In terms of an ability to make sociologically interesting statements, the model developed in this paper seems more promising. It enables the different sources of change to be distinguished; also, the notion that individual characteristics relate to the likelihood of making *particular* transitions is appealing. Moreover, the detail at which information is provided about the process exceeds that in the alternative formulation. In this model, n^2 regressions are carried out,[18] one for each pair of origin and destination states; in the mover-stayer extension only a single regression is appropriate to the model (with the dependent variable being the number of moves made by an individual). Thus, while greater detail is obtained in the mover-stayer extension for distinguishing among persons according to their probabilities of moving, no information is forthcoming on the variation with regard to making particular transitions. Nevertheless, where both formulations are conceptually plausible each will complement the other in contributing to an understanding of the process.

APPENDIX 1

The regression model presented in the text (equation 5) for estimating the population transition matrix involves a comparison between individuals making a transition from i to j and a mixture of two other groups: persons who fail to transfer out of the origin state i, and persons who make a transition but not to j. Although both groups are assigned the same value on the dependent variable, $y_{ij} = 0$, it is probably the case that different conditioning variables are important in distinguishing the i-j movers from each of these groups. For example, in geographic migration, duration at a residence is important for differentiating movers from non-movers, but probably not for distinguishing between i-j movers and migrants to a different region. To avoid this composite character of the reference category it would seem desirable to compare i-j movers with only one of the above groups. A comparison with non-movers seems preferable since the category "movers to a state other than j" will itself change as the destination state of interest is varied.

The most direct procedure would be to perform regressions like those in equation 2 for each off-diagonal element, but to take as observations only individuals who make either an i-j or an i-i transition. Thus, y_{ij} would equal one for persons making the transition of interest and zero for the reference group, which now contains only individuals exhibiting a common behavior. (The equation for y_{ii} would use the same observations as before since the reference category for non-movers would be all persons transferring out of state i.) However, this formulation presents a problem in that the observations for each regression in row i will no longer be the same (and equal to $N_{i.}$). As a result, the simple relationship of equation 4, which allows P_{ij} to be derived from $\hat{Y}_{ij}$, no longer pertains.

How, then, can this procedure be used for estimating the transition probabilities of a Markov chain?

Consider the following approach: Construct regression equations as indicated in the preceding paragraph. For each transition probability on the main diagonal of P(1) we have from equation 4

$$p_{ii} = \frac{1}{N_{i.}} \sum_c \hat{y}_{iic} = \hat{Y}_{ii} = \frac{N_{ii}}{N_{i.}}$$

where the term $\hat{Y}_{ij}$ is introduced for notational convenience and will equal $\sum_c \hat{y}_{ijc}$ divided by the number of observations in the particular regression. For each off-diagonal equation y_{ij} we now obtain, summing over the predicted values from the appropriate individual profiles,

$$\hat{Y}_{ij} = \frac{1}{N_{ij} + N_{ii}} \sum_c \hat{y}_{ijc} = \frac{N_{ij}}{N_{ij} + N_{ii}}$$

Therefore,

$$\frac{\hat{Y}_{ij}\hat{Y}_{ii}}{1 - \hat{Y}_{ij}} = \frac{\left(\frac{N_{ij}}{N_{ij} + N_{ii}}\right)\frac{N_{ii}}{N_{i.}}}{1 - \frac{N_{ij}}{N_{ij} + N_{ii}}}$$

$$= \left[\frac{N_{ij}N_{ii}}{(N_{ij} + N_{ii})N_{i.}}\right]\left[\frac{N_{ij} + N_{ii}}{N_{ii}}\right]$$

$$= \frac{N_{ij}}{N_{i.}}$$

$$= p_{ij}$$

Consequently, an alternative computation of P(1) is given by

$$P(1) = \begin{pmatrix} \hat{Y}_{11} & \cdots\cdots & \hat{Y}_{1m}\hat{Y}_{11}\Big/\left(1-\hat{Y}_{1m}\right) \\ \vdots & & \vdots \\ \hat{Y}_{m1}\hat{Y}_{mm}\Big/\left(1-\hat{Y}_{m1}\right) & \cdots\cdots & \hat{Y}_{mm} \end{pmatrix}$$

While the regression equations corresponding to this model result from more valid comparisons than the approach used in the text, the row sums of probabilities derived from these estimates tend to exhibit a greater deviation from the value one when subpopulation matrices are estimated. It therefore seems preferable to use the former method for constructing transition matrices when the off-diagonal elements are small, since the reference observations would then consist largely of non-movers.

NOTES

[1]Somewhat related approaches to heterogeneity have been presented by Coleman (1964:Chap. 6) and Orcutt _et al_. (1961). Coleman decomposes the transition intensities of a continuous-time Markov process by assuming individual flows between states in a structured manner. The states are constructed to correspond to combinations of the independent variables, consequently the rate of flow between pairs of states provides an indication of the importance of particular variables. By tying the independent variables to states of a discrete-state process Coleman is unable to incorporate variables in a continuous formulation. Also, in contrast with the present analysis, the number of transition probabilities to be estimated increases with the number of variables considered. Orcutt _et al_. use regression techniques to estimate probabilities but, while they discuss Markov processes (1964: 286-94), they fail to combine the two models and treat them instead as alternative formulations.

[2]For a random variable $x, E(x) = \sum_{0}^{\infty} xP_x$. If $x = 1$ with probability $\hat{y}_{ijc}$ and zero with probability $1 - \hat{y}_{ijc}$, $E(x) = 0(1-\hat{y}_{ijc}) + 1(\hat{y}_{ijc}) = \hat{y}_{ijc}$.

[3]Observed frequencies and estimates of these which are constrained to yield precisely the observed values will be written as population values (without hats).

[4]Since $\sum_c \hat{e}_{ijc} = 0$ for each regression equation, this estimation of P(1) will yield p_{ij}'s which are identical to the observed frequencies, $\frac{N_{ij}}{N_{i.}}$

[5]From the perspective of substantive findings, the reader is cautioned that the data used here are biased. Histories were collected only for the

four most recent residences of an individual and for his birthplace. As a result, persons with more than five locations had gaps in their histories and had to be excluded from the analysis. Since individuals with many residences are likely to have made regional changes, this exclusion results in a reduction in the size of off-diagonal elements, and in the extent of population heterogeneity.

[6]It is possible to object to this model because of the heterogeneous nature of the $y_{ij} = 0$ category. Individuals are assigned this value if they failed to move, or if they made a transition to a state other than j. It seems likely that the individual attributes which relate to these two types of behavior will be different, and consequently persons who make an i-j transition ($y_{ij} = 1$) should be compared to one or the other of the reference groups, not to a mixture of both. In terms of the number of individuals in each group in the present example, the regression model is tantamount to comparing $y_{ij} = 1$ persons with the subpopulation that fails to move (see Panel B of Table 1). If the sizes of the off-diagonal elements in the transition matrix were substantial, a formulation which explicitly compares the i-j movers with only individuals who make an i-i transition should be used. (See Appendix 1 for the development of such a model.) In practice, that model would entail a cost of increasing the discrepancy between the row sums in estimated transition matrices and the value one.

[7]The occupation categories were constructed as follows: Upper white collar--professional and managerial; Lower white collar--clerical and sales; Upper blue collar--craftsmen and operatives; Lower blue collar--

private household workers, service, and laborers; Farm occupations--farmers and farm laborers.

[8]The region variables were defined as follows: Region 1--New England and Middle Atlantic states; Region 2--East North Central and West North Central; Region 3--South Atlantic, East South Central and West South Central; Region 4--Mountain states and West.

[9]The backwards option of the stepwise regression program was used. All variables were initially entered into the equation. They were removed and reentered one at a time according to the test criterion. A .05 significance level was used for cutoff.

[10]This procedure assumes there is at least one person from the sub-population in each geographic region at t_o so the regression equations can be evaluated on an attribute profile.

[11]Because of the requirement that $\sum_j p_{ij} = 1$ in a transition matrix, only m-1 equations are actually required for estimating the m row probabilities. However, since the error terms of the regressions do not necessarily sum to zero over a subgroup, the row sums of the matrix will not equal one exactly. All m equations were therefore used to estimate the p_{ij}'s in order to assess the magnitude of the discrepancy. From the resulting row sums, which are presented in Table 3, it is evident that the error term is quite small.

[12]If none of the b's in equation (5) is significant this suggests that the process is Markovian, in which case equation (8) could be used to estimate P(k).

[13]While only $m(m-1) = 12$ regression equations are necessary for estimating the individual p_{ij}'s, $m^2 = 16$ equations were used in order to obtain an indication of the extent to which the row sums will deviate from $\sum_j p_{ij} = 1$.

[14]In contrast to the estimation of the population transition matrix, $\hat{P}_{t_1}(1)$, where only the profiles of individuals <u>at an origin state</u> were used with a row of equations, here each person's profile is inputed to all equations in Y. This method therefore assumes that an individual's characteristics summarize all relevant information about his behavior. If he were to change states his transition propensities would conform to that of others in the new state who have profiles which are identical to his.

[15]Since the p_{ijc} elements in $M_c(1)$ are estimated from regression equations which have 0-1 dummy dependent variables, it is possible for the predicted values to exceed one or be less than zero. There are meaningless estimates for a probability, so values outside the 0-1 range were truncated at the probability limits. These adjustments had a negligible effect on the results. On the average, only one out of each 110 estimates fell outside the 0-1 range by as much as .01. Likewise 96 percent of the time the sum of the row probabilities for an individual was within the range .98 - 1.02. Procedures which constrain the estimated probabilities to the 0-1 range, such as logit and probit techniques (Thiel 1970), are available. However, the adjustments would have been minor and these methods lack the simplicity of interpretation available to the b-coefficients in a linear regression model which uses a dichotomous dependent variable.

[16]Individual attributes which are contingent upon a state change cannot be handled by these computations. Although an $M_c(1)$ matrix indicates probabilities for making various transitions, an individual will make one or another, and his value on such a variable will be altered by the resulting transition. Variables of this nature can therefore be handled only through a simulation procedure in which each person is assigned to a destination state by comparing a random number with his p_{ijc} values, and has characteristics which depend upon the transition made assigned to him in a similar fashion. While information on such variables are available from the Taeuber study and were used in the analysis of over-time change in $P_t(1)$, we are concerned here with projection into the future and assume the unavailability of these data.

[17]It should be noted that an i-to-i transition has different meanings in the two models. In the mover-stayer extension the moves occur randomly in time. An i-i transition can therefore be interpreted as a within-region change of residence, if the problem is one of regional migration, or a within-industry job change if the process concerns industrial mobility. In the present generalization, an i-i transition does not distinguish between the absence of a move and an occurrence in which a change of state fails to occur.

[18]In fact, n^2 systems of equations could be estimated in order to incorporate complex structural assumptions as to the determinants of each p_{ij} in the transition matrix.

REFERENCES

Beshers, J. M., and E. O. Laumann

1967 "Social distance: A network approach." American Sociological Review 32:225-236.

Blumen, I., M. Kogan, and P. J. McCarthy

1955 The Industrial Mobility of Labor as a Probability Process. Cornell Studies of Industrial and Labor Relations, vol. 6. Ithaca, New York: Cornell University.

Coleman, James S.

1964 Introduction to Mathematical Sociology. New York: Free Press.

Hodge, R. W.

1966 "Occupational mobility as a probability process." Demography 3:19-34.

Fuguitt, G.

1965 "The growth and decline of small towns as a probability process." American Sociological Review 30:403-411.

Land, Kenneth C.

1969 "Duration of residence and prospective migration: Further evidence." Demography 6:133-140.

Lieberson, S., and G. V. Fuguitt

1967 "Negro-white occupational differences in the absence of discrimination." American Journal of Sociology 73:188-200.

Matras, J.

1960 "Comparison of intergenerational occupational mobility patterns." Population Studies 14:163-169.

McFarland, David D.

1970 "Intra-generational social mobility as a Markov process: Including a time-stationary Markovian model that explains observed declines in mobility rates over time." American Sociological Review 35:463-476.

Morrison, Peter A.

1970 Movers and Stayers: An Empirical Analysis Based on Two Longitudinal Data Files. Unpublished MS, The RAND Corporation, P-4374.

Orcutt, Guy H., M. Greenberger, J. Korbel, and A. Rivlin

1961 Microanalysis of Socioeconomic Systems: A Simulation Study. New York: Harper and Bros.

Palmer, G.

1954 Labor Mobility in Six Cities. New York: Social Science Research Council.

Prais, S. J.

1955 "Measuring social mobility." Journal of the Royal Statistical Society 118 (Series A):55-66.

Rogers, Andrei

1966 "A Markovian analysis of migration differentials." Proceedings, American Statistical Association, Social Science Section 1966.

Spilerman, Seymour

1970 "Extensions of the mover-stayer model." A paper presented at the Third Conference on the Mathematics of Population, University of Chicago, July 20-24, 1970.

Stinchcombe, A.

1968 Constructing Social Theories. New York: Harcourt, Brace and World.

Taeuber, Karl E., L. Chiazze, Jr., and W. Haenszel

1968 Migration in the United States. Washington: Government Printing Office.

Taeuber, K. E., and A. E. Taeuber

1965 Negroes in Cities. Chicago: Aldine.

Tarver, James D., and William R. Gurley

1965 "A stochastic analysis of geographic mobility and population projections of the census divisions in the United States." Demography 2:134-139.

Thiel, Henri

1970 "On the estimation of relationships involving qualitative variables." American Journal of Sociology 76:103-154.