Ginsburg, Norton.
1961 Atlan of Economic Development. Chicago: University of Chicago Press.
Huntington, Samuel P.
1965 "Political development and political decay." World Politics 17 (April):386–430.
Lerner, Daniel.
1958 The Passing of Traditional Society. Glencoe: The Free Press.
Lenski, Gerhard E.
1966 Power and Privilege: A Theory of Social Stratification. New York: McGraw-Hill.
Lipset, Seymour Martin.
1960 Political Man. Garden City, N.Y.: Doubleday and Co.
Marsh, Robert M. and William L. Parish.
1965 "Modernization and communism: a re-test of Lipset's hypotheses." American Sociological Review 30 (December):934–942.
von der Mehden, Fred R.
1964 Politics of the Developing Nations. Englewood Cliffs, N.J.: Prentice-Hall.
Nettl, J. P.
1967 Political Mobilization. New York: Basic Books.
Pye, Lucian W. (ed.).
1963 Communications and Political Development. Princeton: Princeton University Press.
Riggs, Fred W.
n.d. "The comparison of whole political systems." University of Minnesota Center for Comparative Political Analysis, mimeographed.
Rostow, W. W.
1960 The Stages of Economic Growth. London: Cambridge University Press.
Russett, Bruce M., et al.
1964 World Handbook of Political and Social Indicators. New Haven, Conn.: Yale University Press.
Sawyer, Jack.
1967 "Dimensions of nations: size, wealth, and politics." The American Journal of Sociology 73 (September):145–172.
de Schweinitz, Karl, Jr.
1964 Industrialization and Democracy. New York: The Free Press.
Shannon, Lyle W.
1958 "Is level of development related to capacity for self-government?" American Journal of Economics and Sociology 17 (July): 367–382.
1959 "Socio-economic development and political status." Social Problems 7 (Fall):157–169.
Simpson, Dick.
1964 "The congruence of political, social, and economic aspects of development." International Development Review 6 (June): 21–25.
United Nations.
1960 to present Statistical Yearbook.
1963 Compendium of Social Statistics.
1963 Report on the World Social Situation.
Walinsky, Louis J.
1963 The Planning and Execution of Economic Development. New York: McGraw-Hill.

# A STOCHASTIC MODEL OF SOCIAL MOBILITY *

ROBERT MCGINNIS
*Cornell University*

*Stochastic probability processes are considered as models of social mobility. Such processes are extremely similar to, and hence useful in the study of human mobility. However, the best known of these models, the stationery Markov chain, provides a poor representation of mobility, largely because of its one-step dependency axiom. An elaboration of the Markov chain is suggested, using an additional axiom to the effect that one's probability of remaining in a social or geographic location increases monotonically with increases in duration of prior residence in it. This axiom generates a dynamic stochastic model of mobility. Formal properties of the model are displayed. Relevant computer simulation experiments and empirical research are described.*

SPACE, time, and motion are inextricably related concepts.[1] In the description of one of them, the other two must be invoked, at least implicitly. Since mobility, in its most skeletal form, is just motion through a property space, its analysis clearly must include a temporal component. To study mobility is to study paths of points. Each *point* represents an element's location

* A number of people have been involved in the development of the model that is discussed here, especially John Pilger, Syracuse University, George Masnick, Brown University, Alan Hershberg, Jet Propulsion Laboratory, Neil Sloane, Cornell University, and George Myers, World Health Organization. This line of investigation was suggested originally by Leo F. Schnore. The research reported here was supported in part by the National Science Foundation Grant Number GS 1429.

[1] For a detailed discussion, see Whitrow, 1961, expecially Chapters 5 and 6; and Margenau, 1950, especially Chapter 7.

in a property space at an instant in time. Each *path* represents its motion through time in this space.

In this sense of the term, mobility analysis is simply the study, both empirically and in the abstract, of families of *temporal* functions. Such functions connect measures of location in property space.[2] If a temporal function is to represent social mobility, the conditioning information is critically important and must itself be time bound. This is because *social* mobility is an historical process in a deeper sense than that of its measurement in time. In social mobility, it can safely be assumed that where one is today depends both on where he was yesterday, and on certain characteristics of the path that brought him there. Thus, the conditioning information of a mobility function should describe some relevant part of earlier history of the element, of the social system itself, or of both.

At a suitably abstract level, then, theories about social mobility are similar to theories about temporal functions, virtually to the point of their being interchangeable. Despite this fact, the mobility literature has been rather less than saturated with explicit functions of time. Only in recent years have sociologists given serious consideration to such functions in their studies of mobility and change.

Some portents of change in the sociological analysis of mobility occurred roughly fifteen years ago. In 1954, T. W. Anderson proposed a class of temporal functions for the analysis of changing political attitudes. Anderson used the stochastic process functions known as Markov chains to represent changes in political preferences among a panel of respondents.[3] A year later, Blumen, Kogan and McCarthy (1955) published a major study of industrial mobility in the American labor force using the same form of temporal function.

More recently, this family of temporal functions has been used to characterize such diverse phenomena as changes in conformity behavior in an Asch experiment, (Cohen, 1963) patterns of human migration, (Ter Heide, 1963; De Cani, 1961) reproductive behavior, (Perrin and Sheps, 1964) and intergenerational occupational mobility (Hodge, 1966).

All of these studies may be taken as evidence of a healthy development in sociological theory. One discouraging note must be sounded, however. Few applications of these temporal functions worked especially well when tested against data, and some of them were howling failures. In each case, Markov chain theory was applied, but in no case did it prove to be a particularly good representation of social phenomena. The reason for this failure evidently rests in the fact that the Markov chain lacks the necessary detail with which to represent accurately the social phenomena under study. The "Cornell Mobility Model," described below, employs a more elaborate form of Markov chain that may reduce this defect somewhat. In order to characterize this model and to contrast it with the Markov chain, it will be useful to review briefly the latter's structure of axioms and to evaluate them as they pertain to social mobility.

*Markov Chains and Social Mobility. A Markov process* is a stochastic, or time-dependent probability function that is characterized by a set of states (which can be considered finite for purposes of this paper), $S=\{s_1, s_2 \ldots . , s_m\}$; a probability distribution over the states at each time t, $B(t)=[b_1(t), \ldots . , b_m(t)]$; and a square m x m matrix $P(t)=[p_{ij}(t)]$. The number $b_i(t)$ is real, non-negative and subject to $\sum_i b_i(t)=1$, all t. Hence, it is interpretable as the probability that an element, x, is in a state $s_i$ at time t, noted $b_i(t)=\Pr\{x(t)\epsilon\, s_i\}$. The typical element of P(t) also is real and non-negative, but subject to $\sum_j P_{ij}(t)=1$. Thus, $P_{ij}(t)$ is interpretable as the *conditional* probability that an element is in state j at time t given that it was in state i at t—1, noted $P_{ij}(t)=\Pr\ \{x(t)\epsilon s_j|$

[2] A temporal function can be represented by the form $f(x,t,\gamma)=\alpha$, where x is an element, t is a real-valued measure of time, $\gamma$ is conditioning information, and $\alpha$ is a vector of location in property space.

[3] The later sections of this paper require some familiarity with Markov Theory. An adequate review is contained in Kemeny and Snell, 1960. A more thorough treatment is provided in Feller, 1957, and Parzen, 1962. From this point, we shall consider only discrete measures of time, and denumerable state sets.

$x(t-1)\epsilon s_i\}$. Such a structure is called a transition or a stochastic matrix.

The Markov process is characterized, and distinguished from other stochastic processes, by the following *axiom:*

> A stochastic process is a Markov process
> $$\leftrightarrow$$
> $\Pr\{X(t)=s_j|X(t-1)=s_i\}=$
> $\Pr\{X(t)=s_j|X(t-1)=s_i \text{ and } \gamma(t-k)\}$
> where $k=2, 3, \ldots, t$ and $\gamma(t-k)$ (1)
> is any additional conditioning information about the prior history of X.

That is, a Markov process is any stochastic process such that the outcome at time t depends on the outcome at time $t-1$ and on nothing that occurred at any earlier point in time. It is for this reason that the Markov process sometimes is called a one-step dependency process.

As any useful simplifying assumption must, condition (1) has a payoff. If we *define*

$$\overset{(n)}{P}_{(t)}=[\overset{(n)}{p}_{ij}(t)], \text{ where}$$
$$\overset{(n)}{p}_{ij}(t)=\Pr\{x(t)\epsilon s_j|x(t-n)\epsilon s_i\}, \qquad (2)$$
$$n=1, 2, \ldots, t,$$

then $\overset{(n)}{P}_{(t)}$ can be interpreted as the matrix of n-step transition probabilities whose typical element $\overset{(n)}{p}_{ij}(t)$, is the probability of going from $s_i$ at time $t-n$ to $s_j$ at time t. Then it follows from (1) and the theorem on joint occurrences of statistically independent events (see Feller, 1957:115) that, for any Markov process,

$$\overset{(n)}{P}_{(t)}=\prod_{k=t-n+1}^{t} P_{(k)}, \text{ where } n\leqslant t. \qquad (3)$$

Since the distribution vector of probabilities in any stochastic process can be expressed as

$$B(t)=B(t-1)\,P(t), \qquad (4)$$

the product of the vector at $t-1$ and the transition matrix at t, it follows that

$$B(t)=B(t-n)\,\overset{(n)}{P}(t). \qquad (5)$$

In particular, the distribution vector at t is just the product of the initial vector and the sequence of transition matrices:

$$B(t)=B(0)\prod_{k=1}^{t} P_{(k)}.$$

A *Markov chain* is a further simplification of a Markov process given by the following *axiom:*

> A Markov chain is any Markov process with transition matrix, P(t), such that, for some stochastic matrix P,
> $$P(t)=P, \text{ all } t, \qquad (6)$$

so that P is independent of and constant in time. It follows immediately that, for any Markov chain

$$\overset{(n)}{P}_{(t)}=P^n, \text{ the nth power of P.} \qquad (7)$$

A rich variety of formal consequences flows from (6). In fact, the properties of Markov chains are so many and complex that the bulk of published research in Markov processes is restricted to the subclass of chains. This is unfortunate for sociologists, so much of whose data are intrinsically stochastic, since the Markov chain conditions appear not to correspond well to many temporal problems in sociology. To understand why this is the case, we must consider what the Markov axioms require and to what extent these requirements are satisfied by the process of social mobility.

First, it must be clear that a Markov process is an independence model, and is similar in this respect to the binomial or any other so-called independent trials model. The one difference is that in independent trials models, the *events* themselves must be statistically independent, where the Markov requirement is that *transitions* among events must be indepedent.[4] The Markov chain axiom (6) makes it yet more similar to the binomial or multinomial model in requiring a constant probability, not of an event, but of transitions among events. Condition (1) requires that the probability of a move between two states be independent of every historical fact other than that of the location at time $t-1$. This is not so extreme an assumption as it is sometimes taken to

[4] An independent trials model is the degenerate case of a Markov chain in which $p_{ij}=p_{kj}$, $k=1, 2, \ldots, m$, that is, in which each row of p is equal to every other row.

be. It does not require that the effects of history "die" after a single interval. The probabilities of *sequences* of moves, even with common origins and termini such as $s_i \rightarrow s_k \rightarrow s_j$ as against $s_i \rightarrow s_r \rightarrow s_j$, might be quite different. What is required is just that any two elements occupying a common state at time $t-1$ must have identical probabilities of moving to a specified state at time t, regardless of their possibly divergent prior histories.

Within the range of sociological phenomena, this is a patently unrealistic condition. Consider the mobility among a set of states, $S=\{s_i\}$, of a cohort of people. The effects of history frequently are cumulative in such a way as to give quite different mobility probabilities to individuals, despite the fact that they happen to share a common state at a particular time. It is probably not reasonable, for example, to assign equal probabilities of recidivism to two persons classed as being on probation at a particular time. This is clear in the case in which one is a first offender while the other is a ten-time loser. Nor is it necessarily reasonable to assign equal probabilities of shifting an ideological stance to two people, one of whom has held the position steadfastly for many years and the other of whom has just adopted it.

Condition (6) is equally unrealistic for sociological purposes, but in a different way. In contrast to (1), which is an assumption about the entities that make up a system, condition (6) is an assumption about the system itself, requiring essentially that it be closed and thus undisturbed by external forces in its transitional structure. Although this condition is both major and unrealistic, it is somehow easier to cope with in sociological analysis than is that of (1). Anderson, for example, showed how the stationarity condition of (6) was violated by a panel of respondents in their patterns of switching preferences among possible candidates for the United States' Presidency in the months prior to an election. (Anderson, 1954.) The transition matrices that characterized the panel changed markedly as the result of an external force, the occurrence of the two parties' nominating conventions. Stationarity, in fact, can take the status of an hypothesis rather than that of an untested assumption.[5]

The problem reduces to this: time-dependent probability theory seems to be an altogether natural framework for the analysis of social mobility. But the most thoroughly studied of these, Markov theory, makes assumptions that are probably violated in most, if not all, applications to human mobility. In particular, the assumption that history acts itself out in single independent steps is not the way in which processes of social change seem to operate.

The failure of the Markov model shows up in a peculiar and characteristic way that might be called "lumping on the diagonals." That is, observed transition matrices often display markedly higher diagonal values, representing non-movers, than are predicted from the model. This reflects the obvious fact that a propensity to be a stayer hinges on more than the simple condition of current location, contrary to the first Markov axiom.

The beginnings of one possible solution to this problem are sketched out here. These consist of building a more complicated role for history in the mobility process, incorporating this reconstruction in Markov theory, and of studying some consequences of these alterations. The crux of this approach is that the fate of the system (or of an element in it) at time t hinges *jointly* on its location at $t-1$ and on the duration of its prior residence there. Clearly this is but one of many possible ways of building history more closely into stochastic models of social mobility. However, results thus far suggest that this approach may merit further study.

### THE CORNELL MOBILITY MODEL [6]

Earlier Markov chain applications did not give especially good representations of social mobility. But it does not necessarily

[5] For test procedures, see Anderson and Goodman, 1957; Suppes and Atkinson, 1960.

[6] This label is due to Dr. Leroy Stone, Dominion Bureau of Statistics, who suggested it in a thus far unpublished paper, "Some Methods for Investigating Limit Properties of the Cornell Stochastic Process Model." In this paper, Dr. Stone suggested several interesting approaches to the investigation of limiting distributions of the model described here. This paper contains the groundwork for proof of Theorems (14.5) and (14.6).

follow that Markov theory ought to be abandoned in this context. In fact, its basic structure, involving the probabilistic motion of elements through time from state to state, strongly resembles the basic structure of social mobility. For this reason, and because it is so well-studied, there may be greater merit in developing a more elaborate Markov theory so as to make its behavior similar to the mobility behavior of humans than in the alternative of discarding it.

The procedure that led to the Cornell Mobility Model was to impose an additional axiom giving a more complicated role to time than in any earlier model. The axiom was rooted in substantive sociological considerations rather than in those of mathematics. It was begun with the observation that people are not necessarily homogeneous in their tendencies to be mobile even though they may be in a common location at a particular time. A number of sources made it seem equally plausible that movement out of a status position (or any other social location) is constrained chiefly by one's ties to that position.[7] Moreover, the strength of these ties normally should be expected to grow with the passage of time.

These observations suggested the following simple axiom about motion through time in social space:

*Axiom of Cumulative Inertia.* The probability of remaining in any state of nature increases as a strict monotone function of duration of prior residence in that state. (8)

Thus, the axiom implies that not all elements in state $s_i$ at time t are governed by a single law of mobility. In particular, those who have been there longer have a greater probability of remaining than do relative newcomers.

This axiom yields a broad range of testable propositions. For example, the longer one's criminal career, the more likely one is to remain a criminal; the longer a person resides in a community, the more likely he is to remain there; or the more extensive a history of emotional disturbance, the more likely is an observation of disturbance. These hypotheses about elements of a system seem plausible enough, but only if the system is closed.[8]

The Cornell Mobility Model is a closed system that incorporates the Axiom of Cumulative Inertia and which generates a "two-dimensional" Markov chain. The basic definitions and formal axioms of the model are given below. These are followed by a description of certain properties of the model, preliminary results of simulation experiments, and suggestions for further development. First, however, it would be useful to consider intuitively how Markov theory must be elaborated so as to include the Axiom of Cumulative Inertia.

Stochastic processes incorporate time as history of the system. The cumulative inertia axiom involves time as the history of *elements* in a system. Clearly, these two aspects of time are closely related, but they are not the same thing. System time corresponds to the usual aging process, but individual state-occupation time differs in that a form of rebirth occurs whenever an element moves from one state to another. Thus, two time scales, one for the system and one for elements in the system must be involved. These scales will be indexed by the letters $t$ and $d$ respectively.

In traditional stochastic models, a population is distributed at a point in time in a vector of states, $S=[s_i]$. Mobility is accomplished through open channels among some or all of the states according to a probability schedule given by the transition matrix, $P(t)=[p_{ij}(t)]$. In the Cornell Mobility Model, a population is distributed at a point in time in a two-dimensional matrix of states. Mobility then is governed by the axioms of the model, described below.

Intuitively, the Cornell Model partitions

[7] We need not be concerned about the nature of the ties, whether of sentiment, of coercion, or of other sources. However, such considerations certainly could be incorporated formally into the Cornell Mobility Model.

[8] As with the stationary Markov chain, an important use of such a closed system model is as a benchmark for testing hypotheses about the consequences of intervention in an open social system. The axiom is probably a bad one insofar as mobility is concerned in a hierarchy with an ordinary promotional system. In such a case, the relation between tendency to remain in a state and prior residence in it probably is a parabolic function, with a maximum point whose first derivative is zero and whose second derivative is negative.

a population into four cells, as in Table 1. Note that any occupant of state ${}_1s_j(t+1)$ cannot have been in state $s_j$ at time t. Similarly, an occupant in state ${}_ds_i(t+1)$, for $d>1$, must have been in that same state at time t and must, in addition, have been in that state for $d-1$ prior consecutive time intervals. The Model is more fully represented in Figure 1.

TABLE 1. MOBILITY PATTERNS BY RESIDENTIAL STATUS IN THE CORNELL MOBILITY MODEL

| Residential status at time t | Mobility Status in the Interval (t, t+1): Stayer | Mover |
|---|---|---|
| Continuing resident | ${}_ds_i(t)\rightarrow{}_{d+1}s_i(t+1)$ | ${}_ds_i(t)\rightarrow{}_1s_j(t+1)$ |
| Newcomer | ${}_1s_i(t)\rightarrow{}_2s_i(t+1)$ | ${}_1s_i(t)\rightarrow{}_1s_j(t+1)$ |

where $d>1$
$i\neq j$

*Basis and Axioms of the Cornell Model.* Let a population of elements be partitioned at each indexed point in time into a (mutually exclusive and exhaustive) finite set of states, $S=\{s_i, i=1,2,\ldots,m\}$. Let each state of S be subpartitioned by the index set $D=\{1,2,\ldots,d,\ldots\}$, with the resultant doubly-partitioned set

$$_DS=\{{}_1s_1,\ {}_1s_2,\ \ldots,\ {}_1s_m,\ {}_2s_1,\ {}_2s_2,\ \ldots,\ {}_2s_m,\ \ldots\}. \quad (9)$$

The sentence "$x(t)\epsilon_ds_i$" is interpreted to mean "the element x is in the state $s_i$ at time t and has been in that state for d consecutive prior time periods." Since $_DS$ is a partition, it follows that, for each x and t, there exists one and only one ordered pair $\langle d, i\rangle$, such that the sentence "$x(t)\epsilon_ds_i$" is true.

For each t, let a countably infinite sequence of transition matrices, ${}_dP(t)$, be given by

$$\begin{gathered}{}_dP(t)=[{}_dp_{ij}(t)],\\ \text{where } {}_dp_{ij}=\Pr\{x(t)\epsilon_ks_j|x(t-1);\epsilon_ds_i\}; \quad (10)\\ d,k=1,2,3,\ldots\\ i,j=1,\ 2,\ \ldots,\ m.\end{gathered}$$

The structure ${}_dP(t)$ is the *duration-specific*

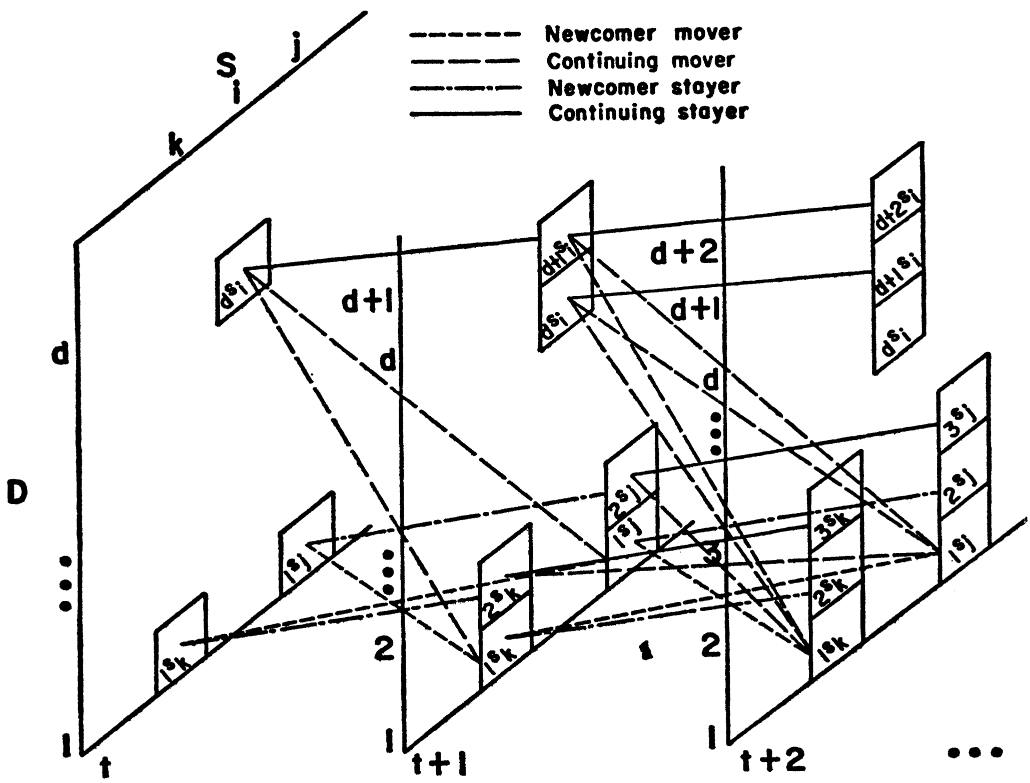


FIGURE 1. REPRESENTATION OF LOCATION AND MOBILITY PATTERNS IN THE CORNELL MOBILITY MODEL

transition matrix governing the behavior at time t of elements which at time t–1 had been in their respective states for d consecutive prior time units. Clearly, this is not the same thing as P(t), the gross Markovian transition matrix. It should be the case, however, that the gross matrix is recoverable from the duration-specific matrices.

Finally, let ${}_dP(t)$ be partitioned by

$${}_dS(t)=[{}_ds_{ij}(t)],$$

where

$${}_ds_{ij}(t)=\begin{cases}{}_dp_{ij}(t) & \text{if } i=j\\ 0 & \text{otherwise}\end{cases}, \qquad (11)$$

and

$${}_dM(t)={}_dP(t)-{}_dS(t).$$

${}_dS(t)$ is just the diagonalization of ${}_dP(t)$, commonly called the stayer matrix. ${}_dM(t)$ is the matrix that governs the behavior of movers.

With this notational basis, the Model is governed by the following:

*Axioms for the Cornell Mobility Model.*

1. $\Pr\{x(t)\epsilon_k s_j|x(t-1)\epsilon_d s_i\}=\Pr\{x(t)\epsilon_k s_j|x(t-1)\epsilon_d s_i \text{ and } \xi\}$; where $\xi$ is any additional information about x prior to time t.
2. ${}_dP(t)={}_dP$, all t.
3. $\Pr\{x(t)\epsilon_k s_j|x(t-1)\epsilon_d s_i\}=0$ if $\begin{cases}i=j \text{ and } k\neq d+1\\ i\neq j \text{ and } k\neq 1\end{cases}$. (12)
4. ${}_dS<{}_{d+1}S$; $\lim_d {}_dS=I$, the identity matrix.
5. There exists a stochastic matrix, $R=[r_{ij}]$, subject to $r_{ii}=0$, all i, such that ${}_dM=(I-{}_dS)R$.

Axiom (12.1) is the first Markov axiom applied to duration-specific rather than to gross transition matrices, while (12.2) is the second Markov axiom similarly applied. As a result of (12.2) time subscripts may be dropped from ${}_dP$, ${}_dS$, and ${}_dM$. Note that this Axiom does not imply that the gross matrix, P(t), necessarily is stationary. The third Axiom regulates individual time relative to that of the system and, basically, sets their clocks at an identical scale. By this Axiom, a stayer gains one unit of duration time for each increment in t and a mover regains duration status d=1.

Axiom (12.4) simply formalizes (8), the Cumulative Inertia Axiom. In addition, it specifies that, in the "long run," any sufficiently dogged stayer finally will be stopped in his tracks. Evocations of the graveyard aside, this may not be an empirically realistic treatment of the later behavior of the non-mobile. It is a problem that is being investigated separately. (Henry et al., 1968)

The final axiom of this system is, in a sense, a dummy postulate. It asserts that a *mover* from $s_i$ will go to any given state, $s_j$, with probability that is indepedent, both of his prior residential history, and of time, the probability depending only on his location and on the particular destination. This is a dummy axiom in the sense that it ignores the "pull" part of various "push-pull" hypotheses, especially the notions of attractive mass and intervening opportunities. Again, differential attraction phenomena are being investigated separately. (McGinnis and White, (1967)

*Some Properties of the Model.* Mathematically, this turns out to be a rather complicated Model, due in large part to the fact that a newcomer may have been a resident of any other state with any prior duration of residence there. Nonetheless, certain properties of the Model have been located analytically. To describe them, some additional definitions are necessary. Let

1. $A(t)=[a_{id}(t)]$, where $a_{id}(t)=\Pr\{x(t)\epsilon_d s_i\}$;
2. $A_d(t)$ be the dth column vector of A(t);
3. $B(t)=(A(t)\zeta)^T$, $b_i(t)=\Pr\{x(t)\epsilon s_i\}$, where $\zeta$ is a conformable column vector of ones, and where $^T$ indicates transposition. (13)
4. $C(t)=[c_{id}(t)]$, $c_{id}(t)=\dfrac{a_{id}(t)}{b_i(t)}$;
5. $\overline{C}_d(t)$ be a diagonalization of the dth column of C(t).

Thus A(t) is a rectangular matrix that contains the joint probability of location in a state at time t and of prior residence in it; B(t) is a 1 x m vector that contains the marginal probability of being in each state at time t; C(t) is the same size as A(t), but contains the conditional probability of having d residence units given that an element is in state $s_i$. The square m x m matrix $\overline{C}_d(t)$ has entries consisting of the dth column of $C(t)$ on the diagonal and zeroes in all off-diagonal cells.

With this additional notation, it can be shown that[9]

*Theorem.* With the preceding definitions and noting the transpose of ${}_dM$ by ${}_dM^T$,

1. $A_d(t)=\begin{cases}\sum_d {}_dM^T A_d(t-1) & \text{if } d=1\\ {}_{d-1}S\ A_{d-1}(t-1) & \text{if } d>1.\end{cases}$
2. $(B(t)=\sum_d {}_dp^T A_d(t-1)$
3. $P(t)=\sum_d \overline{C}_d(t-1)\,{}_dP$ (14)
4. $P(t)$ is non-stationary in t.
5. There exists a stochastic matrix, P, such that $\lim_t P(t)=P$.
6. There exists a distribution vector, B, such that $\lim_t B(t)=B$.

Parts (14.1) through (14.3) display the basic algebraic structure of the Cornell Model. It can be seen from (14.2) and (14.3) that the stationary Markov chain is the degenerate case of the Cornell Model in which the ${}_dP$ matrices are constant in d (recognizing that $\sum_d A_d(t-1)=B(t-1)$ and that $\sum_d \overline{C}_d=I$, the identity matrix). The non-stationarity of the gross transition matrix (14.4) is an immediate consequence of (14.3) and the fact that $\overline{C}_d(t)$ varies in time. While (14.5) and (14.6) show that the Cornell Model converges to an equilibrium state, they do not imply that its behavior in the limit is like that of a Markov chain. In particular P, the limit of $P(t)$, does not have the property of uniform column vectors as does the limit of $P^n$ in a stationary chain.

*Computer Simulation Experiments.* In order to obtain further insights into the behavior of the Cornell Model, a program of 40 simulation experiments was conducted.[10] Monte Carlo processes were not used. Instead, each experiment consisted of iterating the system from an initial condition to convergence.[11] Necessary inputs for the experiments were m, the number of states; $A(0)$, the initial duration-specific distribution matrix; ${}_1S$, the newcomer stayer matrix; a function to generate ${}_dS$ from ${}_1S$ subject to (12.4); R, the stochastic matrix that produces ${}_dM$ from ${}_dS$ by Axiom (12.5).

Experiments were limited to a five-state system. For simplicity, $B(0)$ was set equal to the vector $(A_1(0))^T$, which is equivalent to investigating a cohort of newcomers at $t=0$. Five experimental values of $B(0)$ were used, $[1,0,0,0,0]\ldots[0,0,0,0,1]$, so that the entire cohort initially was located in a single state.[12]

The ${}_dS$ generator function chosen for these experiments was suggested by results of empirical research (discussed below). It was the geometric sequence

$${}_dS=I-(1-\frac{1}{a})^{d-1}(I-{}_1S),\ a>1, \quad (15)$$

which clearly satisfies both conditions of (12.4). Note that (12.5) and (15) yield

$${}_dM=(1-\frac{1}{a})^{d-1}{}_1SR. \quad (16)$$

The effect of this function is seen more clearly by reexpressing (15) as the equivalent recursive equation

$${}_dS={}_{d-1}S+\frac{1}{a}(I-{}_{d-1}S). \quad (17)$$

That is, with each increment of time, the probability of staying is increased by a fixed fraction, $1/a$, of the remaining range, $I-{}_{d-1}S$. Clearly, the larger the value of $a$, the slower the system ought to converge. Four experimental values were chosen, $a=2$, 4, 10, 20. Since these experiments were not concerned with attraction theories of mobility, R was loaded uniformly with the value .25 in the off-diagonal cells. In effect, this says that a mover is equally likely to go to any of the four potential destinations.

Four values of $a$, five $B(0)$ vectors and two ${}_1P$ matrices generated 40 simulation experiments. The major value of these experiments was to provide insights into the convergence behavior of the system with respect to $B(t)$ and $P(t)$; to learn when

[9] Proofs are omitted throughout this paper, but are available on request from the author.

[10] An I.B.M. 360/65 was used. Maximum time per experiment was less than 60 seconds; average time was approximately 5 seconds.

[11] Convergence time was defined as that value of t such that $|p_{ij}(t)-p_{ij}(t-1)|<.0005$.

[12] These values were used in anticipation of a need for independent simultaneous equations and in the suspicion that the limit of $B(t)$ is independent of $B(O)$.

and to what forms they converge and how these patterns are affected by initial conditions. Although it is impossible to reproduce the full set of experimental summaries in this space, partial results for eight experiments are displayed in Table 2. The only parameters that vary in these eight experiments are B(0) and $\alpha$. The two B(0) vectors, labeled 1 and 5, load the cohort into states $s_1$ and $s_5$ respectively at $t=0$. Results for each of the four experimental values of $\alpha$ are reproduced in Table 2. Each column contains the major results of a single experiment. For example, the first column gives results for the inputs $\alpha=2$ and $B(0)=[1,0,0,0,0]$. Convergence was reached at $t=17$. Although the entire cohort was located in state $s_1$ at $t=0$, 32.6 percent was in this state at $t=17$. The initial retention probability in this state was ${}_1p_{11}(0)={}_1p_{11}=.5$. At convergence time, the gross retention probability was almost unity, $p_{11}(17)=.999$.

A comparison of the first two columns of Table 2 shows vividly that the limit distribution can be affected by the initial distribution. To shift the initial location of the cohort from state $s_1$ to $s_5$ reduces the final proportion in $s_1$ by a factor of nearly 10 and increases it in $s_5$ by a factor of about 3.5. A continued comparison of this sort across columns of Table 2 makes it equally clear that the effects of initial conditions disappear as $\alpha$ becomes large. That is, the slower ${}_dP$ converges in d to its limit, I, the more independent the limit distribution becomes of initial conditions. The more rapid the convergence, the more elements become entrapped in an initial state.

The lower block of Table 2 contains the final diagonal values of the gross transition matrix, ${}_dP$. These results show that for sufficiently low values of $\alpha$ and high values of ${}_1p_{ii}$, P(t) converges to the identity matrix, with the result that equilibrium becomes stability; that mobility ceases altogether. Again, as $\alpha$ increases sufficiently, P(t) converges to a matrix such that equilibrium does not imply an absence of motion.

Although these results prove nothing in an analytical sense, they strongly suggest the following

*Conjecture.*

1. There exist $\alpha$ sufficiently large that $\lim_t B(t)$ is independent of the initial distribution, B(0). (18)
2. There exist $\alpha$ sufficiently small and ${}_1S$ sufficiently large that $\lim_t {}_dP=I$.

*Preliminary empirical results*. Three studies of migration in highly diverse populations have yielded similar results, all of which lend partial support to the Cornell Model. Myers, McGinnis and Masnick have shown that migration patterns of selected

TABLE 2. PARTIAL SUMMARY OF COMPUTER SIMULATION RESULTS WITH ${}_1P$ DIAGONAL VECTOR=[.5, .6, .7, .8, .85]

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Input * | $\alpha$ | 2 | 2 | 4 | 4 | 10 | 10 | 20 | 20 |
| | B(O) | 1 | 5 | 1 | 5 | 1 | 5 | 1 | 5 |
| Output ** | lim(t) | 17 | 18 | 26 | 26 | 114 | 114 | 305 | 298 |
| | 1 | .326 | .034 | .129 | .031 | .022 | .016 | .046 | .041 |
| | 2 | .110 | .046 | .096 | .055 | .044 | .037 | .062 | .054 |
| lim $b_i(t)$ | i=3 | .148 | .062 | .164 | .092 | .109 | .091 | .098 | .088 |
| | 4 | .194 | .082 | .268 | .151 | .306 | .253 | .258 | .246 |
| | 5 | .222 | .776 | .343 | .671 | .519 | .603 | .536 | .571 |
| | 1 | .999 | .996 | .493 | .976 | .735 | .697 | .531 | .531 |
| | 2 | .997 | .997 | .991 | .987 | .858 | .848 | .647 | .645 |
| lim $p_{ii}(t)$ | i=3 | .998 | .998 | .995 | .992 | .945 | .942 | .783 | .782 |
| | 4 | .999 | .999 | .997 | .996 | .982 | .979 | .925 | .923 |
| | 5 | .999 | .999 | .998 | .999 | .990 | .992 | .967 | .967 |

* $\alpha$ is the parameter of equation (15) that determines the rapidity of increase in ${}_dS$. The larger the value of $\alpha$, the slower the rate of increase. B(O) is the initial distribution vector. 1=[1, 0, 0, 0, 0]; 5=[0, 0, 0, 0, 1].

** lim B(t) is the converged value of the distribution vector. lim $p_{ii}(t)$ is the limit value of the diagonal vector of lim P(t). Off diagonal elements are constant across columns within each row.

families in the state of Washington satisfy the Axiom of Cumulative Inertia during the period of the study. (Myers et al., 1967.) Migration histories obtained from 1,700 Seattle high school students permitted incidence of migration to be examined as a function of prior residence. The authors conclude that "although the data are not ideal, they indicate a definite trend that tends to support the axiom of cumulative inertia." (Myers et al., 1967:125.)

In a much more comprehensive study, P. Morrison, using a random sample of 5,000 residential histories drawn from the population registration system of the Netherlands, concluded that these data were consistent with the axiom of cumulative inertia. (Morrison, 1967.) However, Morrison's analysis suggested that yet another temporal variable, biological age, also played a strong role in determining probabilities of migration. "First, within specific categories, the probability of migrating declines as duration status increases. Second, the exact form of the relationship differs from one age to another suggesting that age is an interacting variable." (Morrison, 1967.) Morrison found that the quadratic family provided the best regression equations for migration probabilities on the log of prior duration. This is roughly consistent with the ${}_dP$ generator function used in the simulation experiments described above.

K. Land replicated and extended Morrison's study, using a stratified random sample of 1,640 cases drawn from the Monterrey, Mexico Metropolitan Area. (Land, 1967). The relevant population consisted of resident males between the ages of 21 and 60 years. Despite the cultural, economic and ecological differences between the Mexican and Dutch populations, Land's results were strikingly similar to Morrison's. The effects of duration of residence clearly operates in both samples in a manner that is consistent with the Axiom of Cumulative Inertia. In fact, they suggest that the Axiom might be specified with the statement that the strict monotone function is non-linear.

*Next steps*. Clearly this is an interim report on work in progress. Analytical work remains to be done, primarily the characterization of limiting behavior as a function of initial conditions. Empirical studies to date suggest that the Cornell Model may deserve further consideration, but at the same time, that it needs elaboration. In particular, the ${}_dP$ generator must be transformed to a multivariate function of biological age as well as of the ${}_1P$ matrix. Further research may show that position in the family life cycle, in addition to or instead of age, may need to be taken into account.

It has been suggested that a competing model might yield results similar to, but more efficient than, those of the Cornell Model.[13] It may be that despite the results reported above, the Axiom of Cumulative Inertia is false; that, instead, individual psychological propensities to mobility that are invariant in time generate the observed differential migration patterns. This possibility needs to be checked out both by further simulation studies and by more empirical research. Finally, the Model needs further refinement in yet other ways. To convert it from a cohort to a general population model, birth and death processes must be built into it. Less trivially, Axiom (12.5) should be replaced by one that reflects hypotheses about differential flows of movers.[14] Only when these steps have been taken can the value, if any, of this mobility model be assessed.

## REFERENCES


Anderson, Theodore W.
1954 "Probability models for analyzing time changes in attitudes." In Paul F. Lazarsfeld (ed.), Mathematical Thinking in the Social Sciences, Glencoe: The Free Press.
1955 "Intermetropolitan migration: A comparison of the hypotheses of Zipf and Stouffer." American Sociological Review 20:287–291.

Anderson, Theodore W. and Leo A. Goodman.
1957 "Statistical inference about Markov chains." Annals of Mathematical Statistics 28:89–110.

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, N.Y.: Cornell University.


[13] This suggestion was advanced initially by J. S. Coleman in personal conversation.

[14] For example, in applications to migration the theories of Zipf and Stouffer can be built into the Cornell Model as further axioms on the ${}_1p$ matrix. See Anderson, 1955; Stouffer, 1940; Ter Heide, 1963; and Zipf, 1946.

Cohen, Bernard.
1963 Conflict and Conformity: A Probability Model and Its Application. The M.I.T. Press.
De Cani, John S.
1961 "On the construction of stochastic models of population growth and migration." Journal of Regional Science 3:1-13.
Feller, William.
1957 Probability Theory and Its Applications. Second edition. New York: Wiley.
Henry, Neil W., Robert McGinnis and Heinrich W. Tegtmeyer.
1968 "A finite model of mobility." Paper presented to the Population Association of America, annual meetings.
Hodge, Robert W.
1966 "Occupational mobility as a probability process." Demography 3(1):19-34.
Kemeny, John and J. Laurie Snell.
1960 Finite Markov Chains. Princeton: Van Nostrand.
Land, Kenneth C.
1967 "The evaluation of a stochastic model of migration: a replication and extension of findings." Paper presented to the Population Association of America, annual meetings.
Margenau, Henry.
1950 The Nature of Physical Reality. New York: McGraw-Hill.
McGinnis, Robert and John White.
1967 "Simulation experiments on a stochastic attraction model." Unpublished ms.
Morrison, P. A.
1967 "Duration of residence and prospective migration: the evaluation of a stochastic model." Demography 4(2):559-560.
Myers, George C., Robert McGinnis and George Masnick.
1967 "The duration of residence approach to a dynamic stochastic model of internal migration: a test of the axiom of cumulative inertia." Eugenics Quarterly 14:121-126.
Parzen, Emanuel.
1962 Stochastic Processes. San Francisco: Holden-Day.
Perrin, E. B. and M. C. Sheps.
1964 "Human reproduction: a stochastic process." Biometrics 20 (March):28-45.
Stouffer, Samuel A.
1940 "Intervening opportunities: a theory relating to mobility and distance." American Sociological Review 5:945-957.
Suppes, Patrick and Richard C. Atkinson.
1960 Markov Learning Models for Multiperson Interactions. Stanford: Stanford University Press.
Ter Heide, H.
1963 "Migration models and their significance for population forecasts." Milbank Memorial Fund Quarterly 41 (January):57.
Whitrow, Gerald J.
1961 The Natural Philosophy of Time. London: Thomas Nelson and Sons.
Zipf, George K.
1946 "The $p_1p_2/d$ hypothesis: on the intercity movement of persons." American Sociological Review 11:377-386.

# SOCIAL PARTICIPATION AND SOCIAL STATUS *

ROBERT W. HODGE
*University of Chicago*

DONALD J. TREIMAN
*University of Wisconsin*

***The relationships between various aspects of social participation—voluntary organization memberships, church attendance, and informal association with friends—and a number of social status and social background factors are examined using data from a representative sample of residents of a suburban county adjacent to Washington, D.C. In particular, the role of direct intergenerational transmission of participation patterns in determining levels of social participation is investigated by using the technique of path analysis to derive estimates of the effects of parents' participation patterns (for which no direct measurements are available) upon those of their offspring. For both males and females, membership in voluntary organizations appears to be at least as strongly influenced by parent's level of participation in such organizations as by respondent's socioeconomic status. In the case of church attendance, however, a strong direct intergenerational effect is found only for females, and not for males. Church attendance of males appears to be strongly influenced by their spouses' attendance patterns, a result which is consistent with the role of women as expressive leaders of families.***

THE positive association between membership in voluntary organizations and socioeconomic status is one of the best documented relationships in the sociological literature (a standard reference is Wright and Hyman, 1958). Numerous independent

* This research was supported by a grant from the National Science Foundation (NSF #G85, "Occupations and Social Stratification"), whose support is gratefully acknowledged. We are greatly indebted to Norman M. Bradburn who made the data utilized in this report available to us. Although