



Mobility Indices in Continuous Time Markov Chains

John Geweke; Robert C. Marshall; Gary A. Zarkin

Econometrica, Vol. 54, No. 6 (Nov., 1986), 1407-1423.

Stable URL:

<http://links.jstor.org/sici?sici=0012-9682%28198611%2954%3A6%3C1407%3AMIICTM%3E2.0.CO%3B2-V>

Econometrica is currently published by The Econometric Society.

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/econosoc.html>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is an independent not-for-profit organization dedicated to creating and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact jstor-info@umich.edu.

MOBILITY INDICES IN CONTINUOUS TIME MARKOV CHAINS

BY JOHN GEWEKE, ROBERT C. MARSHALL, AND GARY A. ZARKIN¹

The axiomatic derivation of mobility indices for first-order Markov chain models in discrete time is extended to continuous-time models. Many of the logical inconsistencies among axioms noted in the literature for the discrete time models do not arise for continuous time models. It is shown how mobility indices in continuous time Markov chains may be estimated from observations at two points in time. Specific attention is given to the case in which the states are fractiles, and an empirical example is presented.

KEYWORDS: Mobility indices, Markov chains, embeddability.

1. INTRODUCTION

MARKOV CHAIN MODELS are widely used in the social sciences to describe the movement of agents across states. If there are n states a Markov chain model has n^2 parameters p_{ij} , arranged in an $n \times n$ transition matrix P , indicating the probability that an agent in state i one period will be in state j the succeeding period. Estimates of the p_{ij} from survey data can be used to address a variety of questions. Given two years of panel data on the income of young females, let state 1 in each year be an income below the poverty line. Then by separating the sample into blacks and whites \hat{p}_{11} for the two groups can be compared. However, to address such questions as "are white young females more mobile than black young females?" requires the comparison of the matrices \hat{P} for the two groups. Questions of this nature have led to the development of mobility indices, which map the matrix P into a scalar $M(P)$. Consider two transition matrices, P_1 and P_2 . Ideally a mobility index should yield the ranking $M(P_1) > M(P_2)$ for any P_1 and P_2 for which reasonable individuals agree that P_1 implies greater mobility than P_2 . Since the concept of mobility comparisons is vaguely defined the problem of comparing mobility indices is ill posed, unless the criteria inherently used by reasonable individuals in the comparison of transition matrices can be quantified. An attractive, systematic approach is to postulate properties which a reasonable index of mobility ought to satisfy, and then evaluate proposed indices against these properties (as Lovell (1962) did for seasonal adjustment procedures). Shorrocks (1978) took this approach, but found logical inconsistencies between apparently plausible criteria and left unresolved the question of whether the criteria are consistent for an interesting subset of all possible transition matrices P . Alternative indices of mobility were also discussed in Shorrocks (1978) and evaluated against these criteria.

In this paper we extend Shorrocks' treatment in several ways. We begin, in Section 2, by suggesting that his criteria fall into two logically distinct categories; criteria within each category are logically consistent but there are several conflicts

¹ The first author acknowledges support from the NSF and the Sloan Foundation. The second author acknowledges support from the NSF through Grant No. SES-8509693. Peter Muoio provided research assistance and Pat Johnson typed the manuscript.

across categories. There are two corresponding categories into which many indices of mobility (including all of those considered by Shorrocks (1978)) can be grouped, and indices are consistent with the corresponding class of criteria and generally inconsistent with the other class of criteria. For transition matrices with nonnegative, real eigenvalues there are no logical conflicts among any of the criteria and many (but not all) indices satisfy the criteria in both categories. In Section 3 we turn to issues of aggregation over time. Many models, and most discussions of mobility, are conditioned on a stipulated time interval. In many applications (e.g. intergenerational occupational mobility) this is natural but in others (e.g., labor force participation and employment status) it is not, and Shorrocks (1978) considered freeing indices of mobility from an arbitrarily chosen time unit. We define indices of mobility for continuous time models, and show that these can be estimated very easily from observations at two discrete points in time. Throughout, we indicate the simplifications that result if the states are defined as fractiles: this can be done if the underlying data are continuous, like income or wealth, but not if they are discrete categories like labor force participation or employment. A continuous time fractile model is developed in Section 4. It is conceptually attractive, because it leads to a one-dimensional mobility function of order n , intermediate between a mobility index (dimension zero) and the transition matrix itself (dimension two). It imposes embeddability as a testable restriction and is much easier to estimate than less restricted models. This model is applied in an empirical example presented in Section 5. The final section discusses some issues for future research.

2. MOBILITY INDICES IN DISCRETE TIME

Let P be the $n \times n$ matrix consisting of the p_{ij} , and let π_j denote the proportion of the population in state j at time t , given the initial distribution $(\pi_{j0}, j = 1, \dots, n)$ at time $t = 0$. If the vector of population proportions $\pi_t = (\pi_{1t}, \dots, \pi_{nt})'$, then $\pi'_{t-1}P = \pi'_0P^t$. Let $\pi^0 = \lim_{t \rightarrow \infty} \pi_t$ if the limit exists; then $\pi^{0*} = \pi'_0 \lim_{t \rightarrow \infty} P^t$. If this limit is the same for all π_0 , denote $\pi = \pi^0$. In the latter case, by setting π_0 to the respective columns of an $n \times n$ identity matrix we see that $\lim_{t \rightarrow \infty} P^t = \underline{1}\pi' = P^+$, where $\underline{1}$ is an $n \times 1$ vector of units. We denote the eigenvalues of P by $\lambda_1, \dots, \lambda_n$, ordered so that $|\lambda_1| \geq \dots \geq |\lambda_n|$, and define the $n \times n$ matrix $A = \text{diag}(\lambda_1, \dots, \lambda_n)$.

A *mobility index* is a function $M(P)$ mapping P into a scalar, and without loss of generality we shall take these to be normalized by $M(I) = 0$. Shorrocks (1978) proposed that mobility indices be evaluated according to several criteria, which we shall cast into three categories. *Persistence criteria* stipulate that an index should be consistent with some simple, intuitively appealing interpretations of the transition matrix P . *Monotonicity* (M) requires that $M(P) > M(P^*)$ if $p_{ij} \geq p_{ij}^*$ for all $i \neq j$, and $p_{ij} > p_{ij}^*$ for some $i \neq j$; The criterion of *immobility*, (I) stipulates $M(P) \geq 0$, and under *strong immobility* (SI) $M(P) > 0$ unless $P = I$.

Convergence criteria stipulate that $M(P)$ should establish an ordering among transition matrices P that is consistent with the rate at which the multiperiod

transition matrices P^t converge to the limiting transition matrix P^+ . These criteria apply only to those P for which P^+ exists, i.e., those for which $|\lambda_2| < 1$.² In the limit, an agent's conditional probability of occupying a state in the next period is the same as the unconditional probability. *Perfect mobility* (PM) requires $M(P^+) \geq M(P)$ for all P , and *strong perfect mobility* (SPM) requires that the inequality be strict unless $P = P^+$. These criteria are motivated by the axiom that perfect mobility is achieved as the number of transitions becomes infinite, and by the related observation that all P^+ have identical rows and so the probability of moving to any state is independent of that originally occupied. The terminology is due to Prais (1955). If it is axiomatic that mobility increases with the number of periods, then we may add to Shorrocks's list the criterion of *monotonicity in period length* (MPL), $M(P^k) \geq M(P^j)$ if $k > j$; *strong monotonicity in period length* (SMPL) requires that this inequality be strict if $P \neq P^+$.

Temporal aggregation criteria remove the influence of the length of the basic time period on comparisons of mobility. The idea that comparisons of rates of convergence should not be reversed by changes in the basic time unit suggests the criterion of *period consistency* (PC): if P and P^* are two transition matrices and $M(P) \geq M(P^*)$, then $M(P^k) \geq M(P^{*k})$, for all integers $k > 1$. Shorrocks suggests that if the index explicitly accounts for the length T of the basic time unit, so that it has the form $M(P; T)$, then we might wish it to satisfy the criterion of *period invariance* (PI), $M(P; T) = M(P^k; kT)$. (Shorrocks adds *normalization* (N), $M(P^+) = 1$; this seems to us less substantive, and in any event all measures consistent with convergence criteria considered here also satisfy N.)

Prais (1955) showed that the mean length of stay in state i is $1/(1 - p_{ii})$, and Shorrocks (1978) suggested the inverse of the harmonic mean of these lengths, scaled by $n/(n-1)$, as an index of mobility:

$$M_P(P) = [n - \text{tr}(P)] / (n - 1).$$

Shorrocks noted that $M_P(P)$ satisfies I, SI, and M. Hence the persistence criteria are internally consistent. Bartholomew's index,

$$M_B(P) = \sum_i \pi_i \sum_j p_{ij} |i - j|,$$

and the unconditional probability of leaving the current state (scaled by $n/(n-1)$)

$$M_U(P) = n \sum_i \pi_i (1 - p_{ii}) / (n - 1)$$

are similar. However these measures violate M, as indicated by the counterexample

$$(2.1) \quad P = \begin{bmatrix} .90 & .05 & .05 \\ .09 & .46 & .45 \\ .09 & .45 & .46 \end{bmatrix}, \quad P^* = \begin{bmatrix} .90 & .05 & .05 \\ .05 & .50 & .45 \\ .02 & .45 & .53 \end{bmatrix},$$

² We appreciate an anonymous referee bringing the sense of this restriction to our attention. It was not imposed on the PM and SPM criteria in Shorrocks (1978).

$$M_U(P) = .3316, \quad M_U(P^*) = .3855,$$

$$M_B(P) = .3789, \quad M_B(P^*) = .4060,$$

$$\pi_1 = .4871, \quad \pi_1^* = .2579,$$

$$\pi_2 = .2577, \quad \pi_2^* = .3651,$$

$$\pi_3 = .2552, \quad \pi_3^* = .3770.$$

(Although the off-diagonal elements p_{21} and p_{31} decrease in going to P^* , this diminishes the probability that an agent in states 2 or 3 will enter state 1, from which exit is difficult. Consequently $M_U(P)$ and $M_B(P)$ increase, and it may be seen that this comes about through changes in the unconditional probabilities π_i .) Since $M_B(P) = M_U(P) = 0$ for any P with a single absorbing state, e.g.

$$(2.2) \quad P = \begin{bmatrix} .50 & .30 & .20 \\ .30 & .50 & .20 \\ .00 & .00 & 1.00 \end{bmatrix},$$

M_B and M_U are inconsistent with SI as well.

It is evident that any function $M(P)$ which can be expressed as a strictly monotonically decreasing function of the moduli of the eigenvalues of P will satisfy all of the convergence criteria, and consequently these are internally consistent. The simplest such index is the *eigenvalue index*

$$M_E(P) = \left(n - \sum_j |\lambda_j| \right) / (n-1).$$

This index is consistent with I and violates SI only if all eigenvalues have modulus 1. Consequently, I, SI, and the convergence criteria are logically consistent. Other indices involving eigenvalues include the *determinant index*

$$M_D(P) = 1 - |\det(P)|,$$

suggested in this form in Shorrocks (1978), and the *second eigenvalue index*

$$M_2(P) = 1 - |\lambda_2|$$

advocated in Sommers and Conlisk (1979) and closely related to the measure of half life discussed in Shorrocks (1978). Observe that $M_D(P)$ violates SPM if $\lambda_n = 0$ and $M_2(P)$ violates SI if $|\lambda_2| = 1$.

There are logical inconsistencies across the persistence and convergence categories: for example M and PM are logically inconsistent as noted in Shorrocks (1978). However if all the eigenvalues of P are real and nonnegative, $M_P(P) = M_E(P)$ since the trace of a matrix is the sum of its eigenvalues. Were there logical conflicts among the persistence and convergence criteria for transition matrices with real and nonnegative eigenvalues, it would not have been possible to find a mobility index satisfying all the criteria. Hence the following.

THEOREM 1: *Within the class of transition matrices with real nonnegative eigenvalues all persistence and convergence criteria are logically consistent.*

It is not known whether the restriction on P in Theorem 1 renders PC consistent with the other criteria; we shall pursue a different approach to this problem in the next section. Shorrocks (1978) shows that M_P is consistent with all persistence and convergence criteria (except MPL, which he did not consider) for quasimaximal diagonal (q.m.d.) transition matrices. (A matrix is maximal diagonal if in each row the diagonal element is largest. It is quasimaximal diagonal if this is true after appropriate scaling of the columns: i.e., there exist μ_1, \dots, μ_n such that $\mu_i p_{ii} \geq \mu_j p_{ij}$ for all i, j .) We find Theorem 1 an attractive complementary result because it extends the class of transition matrices for which a known index is consistent with the persistence and convergence criteria, and evaluation of eigenvalues is straightforward whereas (to our knowledge) the q.m.d. character of a matrix cannot be established algorithmically. None of the indices discussed in Shorrocks (1978) satisfies *all* of the criteria set forth there, even for q.m.d. transition matrices. A counterexample is provided by

$$P = \begin{bmatrix} .19 & .26 & .55 \\ .38 & .59 & .03 \\ .19 & .21 & .60 \end{bmatrix}, \quad P^* = \begin{bmatrix} .20 & .25 & .55 \\ .38 & .59 & .03 \\ .19 & .21 & .60 \end{bmatrix},$$

$$\begin{aligned} M_P(P) &= .8100, M_P(P^2) = .9183, & M_P(P^*) &= .8050, M_P(P^{*2}) = .9201, \\ M_B(P) &= .8466, M_B(P^2) = .9322, & M_B(P^*) &= .8433, M_B(P^{*2}) = .9351, \\ M_U(P) &= .7643, M_U(P^2) = .8930, & M_U(P^*) &= .7614, M_U(P^{*2}) = .8955, \\ M_E(P) &= .7865, M_E(P^2) = .9183, & M_E(P^*) &= .7955, M_E(P^{*2}) = .9201, \\ M_D(P) &= .9905, M_D(P^2) = .9999, & M_D(P^*) &= .9962, M_D(P^{*2}) = 1.0, \\ M_2(P) &= .5965, M_2(P^2) = .8372, & M_2(P^*) &= .6005, M_2(P^{*2}) = .8404. \end{aligned}$$

(The scaling constants that show P and P^* are q.m.d. matrices are 3, 2, and 1 for the first, second, and third columns, respectively; each matrix has a single negative eigenvalue.) We see that M_P , M_B , and M_U violate PC, and M_E , M_D , and M_2 violate M . The restriction of Theorem 1 brings $M_P = M_E$ into agreement with the persistence and convergence criteria; e.g., note the counterexamples (2.1) and (2.2) in which the matrices are all q.m.d. and satisfy Theorem 1. Furthermore even $M_P = M_E$ does not meet PC, as the following counterexample shows:

$$P = \begin{bmatrix} .21 & .24 & .55 \\ .38 & .59 & .03 \\ .19 & .21 & .60 \end{bmatrix}, \quad P^* = \begin{bmatrix} .22 & .23 & .55 \\ .38 & .59 & .03 \\ .19 & .21 & .60 \end{bmatrix},$$

$$\begin{aligned} M_P(P) &= M_E(P) = .8000, & M_P(P^*) &= M_E(P^*) = .7950, \\ M_P(P^2) &= M_E(P^2) = .9219, & M_P(P^{*2}) &= M_E(P^{*2}) = .9236, \end{aligned}$$

in which the transition matrices are q.m.d. and positive definite with real eigenvalues.

3. MOBILITY INDICES IN CONTINUOUS TIME

If the Markov chain model is formulated in continuous time and indices of mobility are functions of the parameters of the continuous time model then the temporal aggregation criteria become irrelevant. We shall argue subsequently that the list of criteria should be shortened even further. It turns out that the continuous time counterparts of three of the six measures discussed for the discrete time model are equivalent, thus further simplifying the analysis. The price paid for this clarification is that continuous time models are not always appropriate when discrete time models are (the case of intergenerational mobility being perhaps the leading instance) although the converse is always true. A further practical difficulty is that for certain configurations of the p_{ij} a corresponding continuous time Markov chain does not exist. This difficulty is symptomatic of the inapplicability of the Markov chain model in a given circumstance rather than a logical problem with the model itself, however.

We begin with a heuristic derivation of the continuous time model and some of its properties from the more familiar discrete time model. (For a rigorous treatment, see Doob (1953, pp. 235-273).) Suppose that there are T time units between observations, and the underlying model is discrete with Δ time units in the transition interval and transition matrix $P(\Delta)$. Without significant loss of generality (Dhrymes, 1978, Propositions 39 and 40) assume that $P(\Delta)$ is diagonalizable. Let the right eigenvectors of P corresponding to its eigenvalues $\lambda_1, \dots, \lambda_n$ be arranged in the columns of the matrix Q , so that $P = Q\Lambda Q^{-1}$. Then $P = P(\Delta)^{T/\Delta}$ and $P(\Delta) = Q\Lambda^{\Delta/T}Q^{-1}$. The transition equation for the true model is $\pi'_t = \pi'_{t-\Delta}P(\Delta)$, so

$$(3.1) \quad (\pi'_t - \pi'_{t-\Delta})/\Delta = \pi'_{t-\Delta}(P(\Delta) - I)/\Delta = \pi'_{t-\Delta}Q(\Lambda^{\Delta/T} - I)Q^{-1}/\Delta.$$

Since $\lim_{\Delta \rightarrow 0} (\lambda_j^{\Delta/T} - 1)/\Delta = \log(\lambda_j)/T$ we have, taking limits of both sides of (3.1), $\dot{\pi}'_t = \pi'_t R$. The intensity matrix $R = QNQ^{-1}$, with $N = \text{diag}(\nu_1, \dots, \nu_n)$, $\nu_j = \log(\lambda_j)/T$. The off-diagonal elements r_{ij} of R indicate probability rates of transition: for a very small time interval δ , the true probability $p_{ij}(\delta)$ that an agent in state i will move to state j is approximately $r_{ij}\delta$, $\lim_{\delta \rightarrow 0} (p_{ij}(\delta)/r_{ij}\delta) = 1$. The diagonal elements of R are the negatives of the rates of transition out of the respective states: since $\lambda_1 = 1$ has corresponding eigenvector $(1, \dots, 1)'$ in the discrete time model, the row sums of R are $\log(\lambda_1) = 0$.

The relationship between the discrete and continuous time models can be complicated by the problems of embeddability and aliasing. The continuous time transition matrix R constructed from P is plausible only if $r_{ij} \geq 0$ for all $i \neq j$. If this condition is satisfied the observed process in discrete time with transition matrix P is said to be *embeddable* in a continuous time process. There is no convenient set of necessary and sufficient conditions for embeddability, although the former have been expanded and the latter reduced over the years (Kingman, 1962; Singer and Spilerman, 1976; Frydman, 1980); however, by performing the calculations outlined in the previous paragraph it is a simple matter to ascertain whether or not a given transition matrix P is embeddable. Embeddability is an essential property of any transition matrix corresponding to a process whose

evolution in continuous time is to be described by a Markov chain model. As a practical matter it is also essential if the formulation is in discrete time and T/Δ is large, since the failure of P to be embeddable will generally mean that $P(\Delta)$ has negative elements.

The aliasing problem is a consequence of the fact that $\exp(\nu_j T) = \exp(\nu_j + i2\pi k/T)T$ for any integer k , and $\log(\lambda_j)$ therefore has many branches. This problem sometimes can be solved arbitrarily always choosing $\log(\lambda_j) = \text{Log}(\lambda_j)$ (the solution with imaginary coefficient smallest in absolute value) or by ruling out some solutions through the requirement of embeddability. In general there exists a high order number of solutions corresponding to $(\text{Log}(\lambda_j) + i2\pi k/T, k=0, \pm 1, \pm 2, \dots), j=1, \dots, n$ for R . The problem arises from the fact that agents can "cycle" between states within an observation interval, while regularly spaced observations always take place at the same point in the cycle. The arbitrary solution is most attractive when $\lambda_j \geq 0$ and least attractive when $\lambda_j < 0$. Since the matrix R is real, in any resolution of the aliasing problem the eigenvalues ν_j must occur in complex conjugate pairs.

Suppose that the transition matrix $P_T = \exp(RT)$ of the discrete time model for observations separated by T time units arises from a continuous time Markov chain. From an index of mobility for the discrete time models $M(P_T)$ we shall construct the index M^* for the corresponding continuous time model by taking $M^*(R) = \lim_{T \rightarrow 0} T^{-1} M(P_T)$. It is straightforward to derive the indices M^* corresponding to the M of the previous section using the fact that π is the same for R and all P_T , and the results $\lim_{T \rightarrow 0} T^{-1}(1 - \exp(zT)) = -z$ and $\lim_{T \rightarrow 0} (1 - |\exp(zT)|) = -\text{Re}(z)$ for any complex z . To render the indices more comparable subsequently, we shall also change the normalization factor of $(n-1)^{-1}$ to n^{-1} in the case of M_P^* and M_E^* , and introduce a normalization factor of n^{-1} for M_D^* . The results are:

$$M_P^*(R) = -\sum_j \nu_j / n = -\sum_j r_{jj} / n = -\log [\det(P)] / n;$$

$$M_B^*(R) = \sum_i \pi_i \sum_j r_{ij} |i - j|;$$

$$M_U^*(R) = -\sum_j \pi_j r_{jj};$$

$$M_E^*(R) = -\sum_j \text{Re}(\nu_j) / n = -\sum_j \text{Re}[\log(\lambda_j)] / n;$$

$$M_D^*(R) = -\sum_j \nu_j / n = -\sum_j r_{jj} / n = -\log [\det(P)] / n;$$

$$M_2^*(R) = -\text{Re}(\nu_2) = -\text{Re}[\log(\lambda_2)].$$

The indices M_P^* and M_D^* are the same, real, and nonnegative. Because R has nonpositive diagonal and nonnegative off-diagonal elements, and $|r_{ii}| = \sum_{j \neq i} |r_{ij}|$, its eigenvalues ν_j all have negative real part (McKenzie, 1960, Theorem 2). Since complex roots occur in conjugate pairs, M_E^* is in fact the same as M_P^* and M_D^* . The eigenvalues of P never have modulus exceeding unity and occur in conjugate pairs if complex; and two necessary conditions for embeddability are that P have

no eigenvalues equal to zero and all negative eigenvalues occur in pairs (Kingman, 1962). Consequently M_P^* , M_D^* , and M_E^* constructed from embeddable P will always be the same. We shall denote the common index M_C^* .

The persistence criteria for mobility indices in discrete time Markov chains apply virtually unchanged in continuous time. Monotonicity requires $M^*(R) > M^*(R^*)$ if $r_{ij} \geq r_{ij}^*$ for all $i \neq j$ and $r_{ij} > r_{ij}^*$ for some $i \neq j$; immobility, $M^*(R) \geq 0$; and strong immobility, $M^*(R) > 0$ unless $R = 0$. Convergence criteria are largely irrelevant. The concept of perfect mobility does not apply to continuous time Markov chains, since P^+ has all eigenvalues but one equal to zero, and $r_{ij} = \infty$ are conceptually inadmissible. (This is one reason for abandoning the normalization factor $(n-1)^{-1}$, introduced by Shorrocks to render $M(P^+) = 1$.) The criterion of monotonicity in period length does not apply directly, although it could be replaced by the criterion of *velocity*, $M^*(kR) > M^*(R)$ for all $k > 1$; velocity is a strictly weaker criterion than monotonicity. Time aggregation criteria are inapplicable by design, but an additional criterion is desirable in the context of continuous time models: *freedom from aliasing* requires that the index be the same for any resolution of the aliasing problem.

All of the indices M^* are nonnegative and satisfy immobility, and M_C^* clearly satisfies strong immobility. The indices M_B^* and M_U^* fail strong immobility for the same reasons as M_B and M_U , and M_I^* fails because $M_I^*(R) = 0$ for any R with $\nu_1 = \nu_2 = 0$. A scaling factor $k > 1$ applied to R multiplies all eigenvalues of R by k but leaves π unaffected, so all indices M^* satisfy the criterion of velocity. Since $M_P^*(R) = -\sum_j r_{jj}$, M_C^* satisfies monotonicity. The indices M_B^* and M_U^* violate monotonicity in much the same circumstances as M_B and M_U . A counterexample is provided by

$$R = \begin{bmatrix} -.05 & .05 & .00 \\ .20 & -.50 & .30 \\ .50 & .20 & -.70 \end{bmatrix}, \quad R^* = \begin{bmatrix} -.05 & .05 & .00 \\ .60 & -.90 & .30 \\ .50 & .20 & -.70 \end{bmatrix},$$

$$M_U^*(R) = .1250, \quad M_U^*(R^*) = .1137,$$

$$M_B^*(R) = .1471, \quad M_B^*(R^*) = .1258,$$

$$\pi_1 = .8529, \quad \pi_1^* = .9194,$$

$$\pi_2 = .1029, \quad \pi_2^* = .0565,$$

$$\pi_3 = .0441, \quad \pi_3^* = .0242.$$

The common index M_C^* exhibits freedom from aliasing, but M_I^* does not. Whether or not M_I^* satisfies the criterion of monotonicity, and M_B^* and M_U^* satisfy freedom from aliasing, are open questions.

We summarize these results, as follows.

THEOREM 2: *The continuous time indices of mobility M_P^* , M_D^* , and M_E^* are the same, and may be expressed equivalently as $-\text{tr}(R)/n$ or $-\log[\det(P)]/n$. The common index satisfies the criteria of monotonicity, strong immobility, velocity, and freedom from aliasing.*

The common mobility index can be calculated so long as $\det(P) > 0$, but this condition is no guarantee that P is embeddable. Since this, or any other index M^* , is defined in terms of R , a check for embeddability should always be made before the index is reported.

4. MOBILITY IN FRACTILE MARKOV CHAINS

While it is a simple matter to determine whether a given transition matrix P is embeddable, it is quite awkward to estimate P subject to the constraint that it be embeddable. The most ambitious attempt appears to be that of Cohen and Singer (1979) who express the likelihood function in terms of the r_{ij} and then maximize it using a grid search procedure in the neighborhood of well-chosen initial estimates. Difficulties arise because gradient methods appear intractable and the intensity matrix R contains so many parameters. One method of resolution is to assume a parsimonious parameterization of R , obtain maximum likelihood estimates by grid search, and compute asymptotic standard errors through numerical differentiation of the likelihood function at the maximum. In this section we describe such a parameterization of R , for the case in which the states are fractiles. The parameterization is chosen not only to expedite maximum likelihood estimation, but also to simplify the interpretation of the intensity matrix and elaborate on indices of mobility through the construction of what we shall call a *mobility profile*.

The fractile model may be constructed to describe the evolution of a cross section of time series realizations of a continuously distributed random variable, like income, earnings, or wealth; it is clearly inappropriate for well-defined discrete states like labor force status. There are both discrete time fractile models (e.g., intergenerational earnings or wealth) and continuous time fractile models (e.g. income or earnings as typically available in panel data). Our treatment applies to both kinds of fractile models, although much of the motivation is provided by the continuous time variant. One of the attractions of the fractile model is that it abstracts completely from distribution, and focuses on mobility. Formally, we shall say that a Markov chain is *fractile* if for all t and all $j = 1, \dots, n$, $\pi_{jt} = n^{-1}$. Equivalently in the discrete time model $n^{-1}\mathbf{1}' = n^{-1}\mathbf{1}'P$, or $n^{-1}\mathbf{1}$ is a left (as well as right) eigenvector of R corresponding to the eigenvalue 1, or column (as well as row) sums are unity; in the continuous time model, $n^{-1}\mathbf{1}$ is a left eigenvector corresponding to the eigenvalue 0. In the discrete time fractile model the mobility indices $M_U(P)$ and $M_P(P)$ are the same, and the indices $M_P(P)$, $M_B(P)$, and $M_U(P)$ all satisfy the persistence criteria. In the continuous time fractile model $M_U^*(R)$ is the same as the common index $M_C^*(R)$; $M_B^*(R)$ is different but satisfies all of the criteria we developed for a continuous time mobility index.

We shall propose a parameterization of the intensity matrix consistent with the probabilities of transition being approximately the same for all pairs of states with the same number of intervening states, and identical in opposite directions: e.g., in the decile Markov chain the rate of transition from the 30th to the 50th

percentile will be the same as from the 50th to the 30th, and about the same as from the 60th to the 80th. We shall consider a more specific, motivating example before proceeding to the general case. Suppose that $a_0 \leq 0$, $a_1 \geq 0$, $a_2 \geq 0$, and $a_0 + 2\sum_{s=1}^2 a_s = 0$. Let the intensity matrix R be of the form

$$(4.1) \quad \begin{bmatrix} a_0 + a_1 & a_1 + a_2 & a_2 & 0 & 0 & 0 & \dots & 0 \\ a_1 + a_2 & a_0 & a_1 & a_2 & 0 & 0 & \dots & 0 \\ a_2 & a_1 & a_0 & a_1 & a_2 & 0 & \dots & 0 \\ 0 & a_2 & a_1 & a_0 & a_1 & a_2 & \dots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\ 0 & \dots & a_2 & a_1 & a_0 & a_1 & \dots & a_2 \\ 0 & \dots & 0 & a_2 & a_1 & a_0 & \dots & a_1 + a_2 \\ 0 & \dots & 0 & 0 & a_2 & a_1 + a_2 & \dots & a_0 + a_1 \end{bmatrix}$$

Notice that the matrix is banded, except in the corners: when a row would extend beyond the matrix, it is "reflected" back and added to the other elements. This form of reflection has been treated in the time series literature (Durbin and Watson, 1951; Anderson, 1970, pp. 284-290). We shall refer to a matrix of the form (4.1) as a *reflectant*.

This parameterization is attractive for several reasons. As will be seen, the eigenvectors of R are the same regardless of the values of the a_s . It also turns out to be the case that the parameterization of the intensity matrix as a reflectant is independent of the choice of n . This is clearly the case in (4.1) if the number of fractiles is reduced from n to $n/2$.

In the general development we shall employ the following definitions.

DEFINITION: An $n \times n$ matrix B is said to be *reflectant* if there exists a sequence $\{a_s\}_{s=-\infty}^{\infty}$ such that

$$b_{kj} = \sum_{s=-\infty}^{\infty} (a_{j-k+2ns} + a_{2ns-(j+k)+1}).$$

DEFINITION: The *Fourier transform* of an absolutely summable sequence $\{a_s\}$ is

$$\hat{a}(\lambda) = \sum_{j=-\infty}^{\infty} a_j \exp(-i\lambda j).$$

THEOREM 3: If B is a reflectant then its eigenvectors consist of the columns of Q , $q_{mj} = \cos[(2m-1)(j-1)\pi/2n]d_{jn}$, and its corresponding eigenvalues are $\hat{a}[(j-1)\pi/n]$, $j = 1, \dots, n$.

PROOF:

$$\begin{aligned} \sum_{m=1}^n b_{km} q_{mj} &= \sum_{m=1}^n \sum_{s=-\infty}^{\infty} (a_{m-k+2ns} + a_{2ns-(m+k)+1}) \\ &\quad \cdot \cos[(2m-1)(j-1)\pi/2n]. \end{aligned}$$

Each element of the sequence $\{a_s\}$ enters the sum exactly once. Either the change of variable $h = m - k + 2ns$ or the change $h = 2ns - (m + k) + 1$ produces the change $\cos [(2m - 1)(j - 1)\pi/2n] = \cos \{[2(h + k) - 1](j - 1)\pi/2n\}$. Hence

$$\begin{aligned} \sum_{m=1}^n b_{km} q_{mj} &= \sum_{h=-\infty}^{\infty} a_h \cos \{[2(h + k) - 1](j - 1)\pi/2n\} \\ &= (1/2) \exp \{i[(2k - 1)(j - 1)\pi/2n]\} \\ &\quad \cdot \sum_{h=-\infty}^{\infty} a_h \exp [i2h(j - 1)\pi/2n] \\ &\quad + (1/2) \exp \{-i[(2k - 1)(j - 1)\pi/2n]\} \\ &\quad \cdot \sum_{h=-\infty}^{\infty} a_h \exp [-i2h(j - 1)\pi/2n] \\ &= \cos [(2k - 1)(j - 1)\pi/2n] \\ &\quad \cdot \sum_{h=-\infty}^{\infty} a_h \cos [h(j - 1)\pi/n] = q_{kj} \tilde{a}(\lambda_j). \end{aligned}$$

The parameterization of an intensity matrix as a reflectant is robust with respect to the arbitrary choice of the number of fractiles in the following sense. Suppose that n_1 and n_2 are alternative numbers of fractiles. Let n^\dagger be the least common multiple of n_1 and n_2 . If the transition or intensity matrix for n^\dagger fractiles is a reflectant, then the corresponding matrix for both n_1 and n_2 fractiles is also a reflectant. This is a consequence of the following result.

THEOREM 4: Suppose that the intensity matrix R is collapsed from n fractile groups to $n^* = n/g$ (n^* and g integer), so that the new $n^* \times n^*$ matrix R^* has entries $r_{km}^* = g^{-1} \sum_{u=1}^g \sum_{v=1}^g r_{g(k-1)+u, g(m-1)+v}$. If R is a reflectant with parameterizing sequence $\{a_s\}$, then R^* is a reflectant with parameterizing sequence $\{a_s^*\}$,

$$a_s^* = \sum_{h=-g}^g [1 - (|h|/g)] a_{gs-h}.$$

PROOF: (Use the fact that for any sequence $\{c_s\}$, $\sum_{u=1}^g \sum_{v=1}^g c_{j+u-v} = \sum_{h=-g}^g (g - |h|) c_{j-h}$.)

$$\begin{aligned} r_{km}^* &= g^{-1} \sum_{u=1}^g \sum_{v=1}^g r_{g(k-1)+u, g(m-1)+v} \\ &= g^{-1} \sum_{u=1}^g \sum_{v=1}^g \sum_{s=-\infty}^{\infty} [a_{(m-k)g+v-u+2ns} + a_{2ns-g(k+m-2)-u-v+1}] \\ &= g^{-1} \sum_{s=-\infty}^{\infty} \sum_{h=-g}^g (g - |h|) [a_{(m-k)g+2ns-h} + a_{2ns-g(k+m-1)-h}] \\ &= g^{-1} \sum_{s=-\infty}^{\infty} \sum_{h=-g}^g (g - |h|) [a_{(m-k+2n^*s)g-h} + a_{[2n^*s-(m+k)-1]g-h}] \\ &= g^{-1} \sum_{s=-\infty}^{\infty} [a_{m-k+2n^*s}^* + a_{2n^*s-(m+k)-1}^*]. \end{aligned}$$

From the definition of the reflectant, it is clear that if a_j and a_{j+2n} are unrestricted then they are not separately identified. The parameterization of a transition or intensity matrix as a reflectant will be more useful to the extent that the number of parameters is kept small. This can be done by truncating $\{a_s\}$, or by a further parameterization of the sequence itself, e.g., $a_s = b|s|^c$ for $|s| \geq m$, $c < 0$ and m small. If $\{a_s\}$ is truncated at some point m , $m < n$, then the sequence provides a characterization of the intensity matrix that is better than an order of magnitude more parsimonious than the matrix itself. We shall refer to the sequence $\{a_s\}_{s=1}^m$ in this case as a *mobility profile*. When m is small relative to n , a_s indicates, effectively, the probabilities or rates of transition across s fractiles; as m becomes a substantial fraction of n , this interpretation is complicated by the reflections in the corners of the intensity matrix. The trace of a reflectant is $n \sum_{s=-\infty}^{\infty} a_{2ns} + \sum_{j \text{ odd}} a_j$. Hence, for $m < n$ we have $M^*(R) = M_U^*(R) = -a_0 - \sum_{j=-m \text{ odd}}^m a_j/n$.

The key features of the continuous time parameterization proposed in this section are the small number of parameters, which facilitates both maximum likelihood estimation and a test against alternatives including non-embeddability; its incorporation of similar transition probabilities to adjacent states; the invariance of the parameterization with respect to the number of states, n ; and computational efficiency owing to the fact that the eigenvectors of the intensity matrix are known. Note that the transition matrix P of a discrete-time fractile model could also be parameterized as a reflectant, using a sequence $\{c_s\}$ in lieu of $\{a_s\}$. The elements of $\{c_s\}$ must lie in the unit interval and sum to unity; Theorems 3 and 4 of course apply to P so parameterized; and for $m < n$ $M_P(P) = M_U(P) = n(1 - c_0)/(n - 1) - \sum_{j=-m \text{ odd}}^m c_j/(n - 1)$.

5. AN EMPIRICAL EXAMPLE

To illustrate these methods we constructed fractile Markov chains using income data from the National Longitudinal Survey of Young Men for the years 1970 and 1971. The data set consists of 654 young men who are employed, married, white, and not enrolled in school in both 1970 and 1971. Noting that 654 is evenly divisible by 6, we selected $n = 6$ fractiles. The unconstrained maximum likelihood estimate of the transition matrix is

$$\hat{P}_6 = \begin{bmatrix} .624 & .229 & .083 & .046 & .009 & .009 \\ .174 & .468 & .211 & .073 & .055 & .018 \\ .101 & .174 & .404 & .211 & .073 & .037 \\ .018 & .101 & .147 & .376 & .284 & .073 \\ .046 & .018 & .101 & .202 & .413 & .220 \\ .037 & .009 & .055 & .092 & .165 & .642 \end{bmatrix},$$

where a typical entry in the i -th row and j -th column is m_{ij}/m_i , with $m_i = 109$ denoting the number of individuals in state i in 1970 and m_{ij} denoting the number in state i in 1970 and j in 1971. For this matrix $M_P(\hat{P}_6) = M_U(\hat{P}_6) = .6148$, $M_B(\hat{P}_6) = .7781$, $M_E(\hat{P}_6) = .6145$, $M_D(\hat{P}_6) = .9960$, and $M_2(\hat{P}_6) = .2698$. The

maximum likelihood estimate of the corresponding three-state transition matrix is

$$\hat{P}_3 = \begin{bmatrix} .748 & .206 & .046 \\ .197 & .569 & .234 \\ .055 & .225 & .720 \end{bmatrix},$$

for which $M_P(\hat{P}_3) = M_U(\hat{P}_3) = .4817$, $M_B(\hat{P}_3) = .3547$, $M_D(\hat{P}_3) = .7591$, and $M_2(\hat{P}_3) = .3150$.

The corresponding intensity matrices, computed as described in Section 3, are

$$\hat{R}_6 = \begin{bmatrix} -.5385 & .4329 & .0603 & .0669 & -.0348 & .0131 \\ .3029 & -.9273 & .5095 & .0052 & .1132 & -.0034 \\ .1655 & .3407 & -1.1121 & .6132 & -.0665 & .0591 \\ -.0577 & .2419 & .2895 & -1.3110 & .8578 & -.0206 \\ .0819 & -.0721 & .1868 & .5202 & -1.1749 & .4582 \\ .0459 & -.0161 & .0659 & .1055 & .3052 & -.5063 \end{bmatrix}$$

for which $M_P^*(\hat{R}_6) = M_U^*(\hat{R}_6) = M_E^*(\hat{R}_6) = M_D^*(\hat{R}_6) = .9283$, $M_B^*(\hat{R}_6) = 2.0256$, $M_2^*(\hat{R}_6) = .3144$, and

$$\hat{R}_3 = \begin{bmatrix} -.3368 & .3290 & .0078 \\ .3106 & -.6946 & .3840 \\ .0262 & .3657 & -.3919 \end{bmatrix}$$

for which $M_P^*(\hat{R}_3) = M_U^*(\hat{R}_3) = M_E^*(\hat{R}_3) = M_D^*(\hat{R}_3) = .4744$, $M_B^*(\hat{R}_3) = .4711$, and $M_2^*(\hat{R}_3) = .3783$.

Since \hat{R}_6 has negative off-diagonal elements \hat{P}_6 is not embeddable, thus calling into question the interpretation of the corresponding estimated mobility measures $M^*(\hat{R}_6)$. The fact that \hat{P}_6 is not embeddable could be interpreted as the failure of a six-state model to exist in continuous time. It could also be interpreted as a small-sample phenomenon. The latter interpretation is motivated by the observation that as the number of observations per fractile decreases the probability of obtaining $\hat{p}_{ij} = 0$ for some i and j increases for any P , and transition matrices with null entries are never embeddable (Singer and Spilerman, 1976); the greater is n for a given sample size, the more prevalent this problem will be.

Consider now estimation of the intensity matrix R subject to the constraint that it be a reflectant. Following the development in Section 4, a given collection of parameters a_0, \dots, a_m define R . Given R , compute its eigenvalues ν_j , $j = 1, \dots, n$, and those of the corresponding transition matrix P , $\lambda_j = \exp(\nu_j)$, $j = 1, \dots, n$. The transition matrix itself is $P(a_0, \dots, a_m) = QAQ^{-1}$, with $A = \text{diag}(\lambda_1, \dots, \lambda_n)$. Grid search over values of a_0, \dots, a_m determines those specific values $\hat{a}_0, \dots, \hat{a}_m$ which provide the maximum value \hat{L}_0 of the log-likelihood function

$$L(a_0, \dots, a_m) = \sum_{i=1}^n \sum_{j=1}^n m_{ij} \log [p_{ij}(a_0, \dots, a_m)].$$

Arithmetic second-differencing of this function around $\hat{a}_0, \dots, \hat{a}_m$ provides the estimated information matrix, and thereby asymptotic standard errors for the \hat{a}_j . The corresponding unconstrained maximum of the log-likelihood function is

$$\hat{L}_1 = \sum_{i=1}^n \sum_{j=1}^n m_{ij} \log(m_{ij}/m_i);$$

under the null hypothesis that the constraints are correct

$$2(\hat{L}_1 - L_0) \stackrel{a}{\sim} \chi^2[(n-1)^2 - m].$$

The maximum likelihood parameter estimates, the likelihood ratio test statistics, and the mobility indices for several reflectant transition matrices are presented in Table I. Consider the 3×3 and 6×6 reflectant matrices below:

$$R_3 = \begin{bmatrix} a_0 + a_1 & a_1 + a_2 & a_2 \\ a_1 + a_2 & a_0 & a_1 + a_2 \\ a_2 & a_1 + a_2 & a_0 + a_1 \end{bmatrix},$$

$$R_6 = \begin{bmatrix} a_0 + a_1 & a_1 + a_2 & a_2 + a_3 & a_3 + a_4 & a_4 + a_5 & a_5 \\ a_1 + a_2 & a_0 + a_3 & a_1 + a_4 & a_2 + a_5 & a_3 & a_4 + a_5 \\ a_2 + a_3 & a_1 + a_4 & a_0 + a_5 & a_1 & a_2 + a_5 & a_3 + a_4 \\ a_3 + a_4 & a_2 + a_5 & a_1 & a_0 + a_5 & a_1 + a_4 & a_2 + a_3 \\ a_4 + a_5 & a_3 & a_2 + a_5 & a_1 + a_4 & a_0 + a_3 & a_1 + a_2 \\ a_5 & a_4 + a_5 & a_3 + a_4 & a_2 + a_3 & a_1 + a_2 & a_0 + a_1 \end{bmatrix}.$$

In row 1 of Table I the reflectant matrix is of the form R_3 where $a_2 = 0$ while in row 2 the constraint $a_2 = 0$ is relaxed. In rows 3 through 7 of Table I the reflectant matrix is of the form R_6 . In row 3 the constraints $a_2 = a_3 = a_4 = a_5 = 0$ are imposed, in row 4, $a_3 = a_4 = a_5 = 0$, and so forth. The standard errors for both the parameter estimates and the continuous time mobility indices, $M_B^*(R)$ and $M_C^*(R)$, are in parentheses below the appropriate value.

The estimate of a_1 in row 1 of Table I indicates that individuals enter an adjoining state at an annual rate estimated to be .3696; that is, in a very small time interval δ years the probability that an individual will enter an adjoining state is estimated to be .3693 δ . The likelihood ratio statistic is $2.039 \stackrel{a}{\sim} \chi^2(3)$, so the reflectant transition matrix is not significantly different from the unconstrained matrix. Since the asymptotic distribution of $M_B^*(R) = .3696$ is intractable we report only the point estimates. The indices $M_B^*(R) = .4928$ and $M_C^*(R) = .4928$ coincide because $a_s = 0$ for all $|s| > 1$. Since they are linear functions of the a_s , their asymptotic standard errors are easily computed. The index $M_C^*(B)$ indicates that the average annual rate at which individuals exit a state is .4928.

The parameter estimates, test statistic, and mobility indices for the tri-diagonal version of R_6 are shown in row 3 of Table I. There is substantial mobility to nonadjoining states in \hat{P}_6 , yet the tri-diagonal reflectant matrix allows for exits only to adjoining states. This explains the relatively high estimate of a_1 , $M_B^*(R)$, and $M_C^*(R)$ in row 3. The null hypothesis that the tri-diagonal parameterization

TABLE I
MOBILITY INDICES, TEST STATISTICS, AND MLE OF PARAMETERS FOR THE REFLECTANT INTENSITY MATRICES

Row #	$\hat{\sigma}_1$	$\hat{\sigma}_2$	$\hat{\sigma}_3$	$\hat{\sigma}_4$	$\hat{\sigma}_5$	$-2(l_1 - l_0)$	Prob Value	$M_1^2(R)$	$M_2^2(R)$	$M_3^2(R)$
3 × 3 Matrix	1	.3696 (.0305)				2.039	.5643	.3696	.4928 (.1645)	.4928 (.1645)
	2	.3300 (.0485)	.0169 (.0182)			1.147	.5636	.3806	.4849 (.0404)	.4737 (.0428)
6 × 6 Matrix	3	.9295 (.0702)				82.067	≈ 0	.2491	1.5492 (.3122)	1.5492 (.3122)
	4	.4067 (.0540)	.1571 (.0247)			46.689	.0025	.2661	1.2015 (.0826)	.9920 (.0756)
	5	.4696 (.0807)	.0320 (.0469)			33.946	.0497	.2887	1.1728 (.0826)	.9557 (.0849)
	6	.4313 (.0681)	.0727 (.0306)	.0380 (.0113)		28.622	.1234	.3023	1.1637 (.0812)	.9402 (.0809)
	7	.4499 (.0684)	.0470 (.0407)	0	.0221 (.0098)	26.540	.1487	.3070	1.1551 (.0807)	.9281 (.0796)

is equal to the unconstrained transition matrix is definitively rejected. In fact, only in rows 6 and 7 is the null not rejected for the six state case. However, when estimating the parameters of R_6 with a_5 constrained to be zero (row 6) the grid search resulted in a negative value for a_3 . Therefore, a_3 was constrained to be zero and maximum likelihood estimates of a_1 , a_2 , and a_4 were obtained. A similar situation arose in estimating the parameters of R_6 without a_5 constrained to be zero (row 7, a_4 was estimated to be less than zero and then constrained to be zero). These results are not surprising given \hat{R}_6 , the unconstrained estimated intensity matrix: since \hat{P}_6 is non-embeddable it is not surprising that the maximum likelihood estimates of the parameters of the reflectant matrix tend toward non-embeddable values. In fact, nearly all of the negative off-diagonal elements in \hat{R}_6 correspond to elements in R_6 which contain either a_3 or a_4 .

As m increases for both the three-state and six-state models, $M_2^*(R)$ increases while $M_C^*(R)$ and $M_B^*(R)$ decrease. The changes are small except for the change from $m=1$ to $m=2$ in the six-state model; $m=1$ for that model is definitively rejected by the likelihood ratio test. For small m the mobility to distant states shown in \hat{P}_3 and \hat{P}_6 can only be interpreted as high exit rates to closely adjoining states. As m is increased these exit rates generally fall and $M_C^*(R)$ and $M_B^*(R)$ decrease.

6. CONCLUSION

This paper extends the axiomatic approach to the development of mobility indices (Shorrocks, 1978) in three ways. First, we show that the axioms can be grouped into three categories: persistence, convergence, and temporal aggregation criteria. We show that for transition matrices with real nonnegative eigenvalues all persistence and convergence criteria are logically consistent. Shorrocks obtained essentially the same result for transition matrices that are quasimaximal diagonal, but this property cannot be verified algorithmically.

Second, the development of continuous time mobility indices removes the influence of the time interval between observations and renders the temporal aggregation criteria irrelevant. Interpretation (although not computation) of continuous time mobility indices requires that the discrete time transition matrix be embeddable in a continuous time Markov chain. We have shown that it is straightforward to determine whether a given transition matrix is embeddable. Since the processes modelled by Markov chains often take place in continuous time, this determination may often be of independent interest, whether or not continuous time mobility indices are to be formulated.

Third, we propose a parsimonious parameterization for the intensity matrix (the continuous time analogue of the transition matrix) that is applicable when the states are fractiles. The respective parameters, a_s , indicate the instantaneous rate of transition across s fractiles. The number of parameters is less than the number of states, so the parameterization conveys more information about mobility than a single index while substantially reducing the number of parameters in the intensity matrix itself. The parameterization places many restrictions on the

intensity matrix; in the illustration provided here using a subsample of size 654 from NLS data for 1970-71, it fails to be rejected.

There are at least two directions in which this work might be extended. First, it should be straightforward to make transitions rates functions of observable, time invariant personal characteristics. This approach has been widely used in proportional hazard models of duration data (e.g., Flinn and Heckmann, 1982). It becomes more complicated when time varying exogenous variables and unobserved heterogeneity are allowed. Second, alternative approaches to mobility in discrete time might be extended to continuous time using the approach taken in this paper. For example, Theil (1972) measures mean first passage time for any state to the limiting distribution, and divides this into travel and waiting time components. It seems likely that such measures are sensitive to temporal aggregation, and a continuous time formulation of the approach might prove fruitful.

Department of Economics, Duke University, Durham, NC 27706, U.S.A.

Manuscript received January, 1985; final revision received February, 1986.

REFERENCES

- ANDERSON, T. W. (1971): *The Statistical Analysis of Time Series*. New York: Wiley.
- COHEN, J. E., AND B. SINGER (1979): "Malaria in Nigeria: Constrained Continuous-Time Markov Models for Discrete Time Longitudinal Data on Human Mixed-Species Infections," *Lectures on Mathematics in the Life Sciences*, 12, 69-133.
- DHRYMES, P. J. (1978): *Mathematics for Econometrics*. New York: Springer-Verlag.
- DOOB, J. L. (1953): *Stochastic Processes*. New York: Wiley.
- DURBIN, J., AND G. S. WATSON (1951): "Testing for Serial Correlation in Least Squares Regression, II," *Biometrika*, 38, 159-178.
- FLINN, C., AND J. HECKMAN (1982): "Models for the Analysis of Labor Force Dynamics," *Advances in Econometrics*, ed. by R. Basmann and G. Rhodes. Greenwich, CT: JAI Press.
- FRYDMAN, H. (1980): "A Structure of the Bang-Bang Representation for 3×3 Embeddable Matrices," *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 53, 305-316.
- KINGMAN, J. F. C. (1962): "The Imbedding Problem for Finite Markov Chains," *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 1, 14-24.
- LOVELL, M. C. (1963): "Seasonal Adjustment of Economic Time Series and Multiple Regression," *Journal of the American Statistical Association*, 58, 993-1009.
- MCKENZIE, L. (1960): "Matrices with Dominant Diagonals and Economic Theory," *Mathematical Methods in the Social Sciences*, ed. by K. J. Arrow, S. Karlin, and P. Suppes. Stanford: Stanford University Press.
- PRAIS, S. J. (1955): "Measuring Social Mobility," *Journal of the Royal Statistical Society, Series A*, Part I, 118, 56-66.
- SHORROCKS, A. F. (1978): "The Measurement of Mobility," *Econometrica*, 46, 1013-1024.
- SINGER, B., AND S. SPILERMAN (1976): "The Representation of Social Processes by Markov Models," *American Journal of Sociology*, 82, 1-54.
- SOMMERS, P. S., AND J. CONLISK (1979): "Eigenvalue Immobility Measures for Markov Chains," *Journal of Mathematical Sociology*, 6, 253-276.
- THEIL, H. (1972): *Statistical Decomposition Analysis*. Amsterdam: North-Holland.