TWO STOCHASTIC MODELS OF INCOME MOBILITY

by

Peter MITTER

Research Memorandum No.125

January 1978

ABSTRACT

Guided by two Markov chains with tridiagonal transition matrices of distinct structure, dynamic inequality characteristics of an income propagation process are compared with the static characteristics of the corresponding stationary distribution, i.e. the result of the process. The two models reflect the concepts of proportionate effect and regression towards the mean, respectively. Moreover the influence of small process parameter changes (persevering perturbations) on the stationary distribution is studie.

List of symbols

Generally

Subscript/superscript letters are always lower case.Letters in matrices are always lower case.Games of matrices are always upper case

cf p.7

$$P = \begin{pmatrix} 1-p & p & & & & \\ q & 1-(p+q) & p & & & \\ & q & 1-(p+q) & p & & \\ & & \ddots & \ddots & \ddots & \\ & & & q & 1-(p+q) & p \\ & & & & q & 1-q \end{pmatrix}$$

↑ upper case   all p,q: lower case

1 in formulas/symbols: always one (no el)

0,o in formulas/symbols:always zero (no oh)
with one exception on p.17

Sometimes symbols are primed: P', $(P^h)'$, x_i'
Sometimes they are barred: $\bar{\lambda}$, $\bar{c}$

p.2

K_1, K_k, K_i subscripts: 1(one), lower case k,i

Δ, Δ^{i-1} superscript: i-1

$>$, $\rightarrow$, ∞

p.3

$g, (g_1, \ldots, g_k), g_i$ subscripts: 1(one), lower case k,i

$\forall i$

$p_{ij}^{(h)}$ subscripts: lower case i,j
superscript: lower case h in parentheses

p(h), p(0) in last parentheses: zero

P'p(h-1) first P: upper case with prime

$\leqslant$, $\sum$ (upper case Greek sigma - sum sign)

p.4

$\sqrt{}$ square root sign

$\neq$

λ lower case Greek lambda

p.5

μ lower case Greek m

ε, [ ]

p.6

$\|$, $|$ (pairwise only)

$\approx$

α lower case Greek alpha

/ fraction-stroke (slanting)

p.7

$c = p/q$ lower case letters

$\frac{1}{k}$ fraction-stroke (horizontal) p.9

σ lower case Greek sigma

p.13

d^{i^2} superscript i^2 with superscript 2

$d^{(i-1)^2}$ analogously

p.16

$\hat{d}$

p.17

ε lower case Greek epsilon

$O(\varepsilon^2)$ first letter: upper case "oh"

1. INTRODUCTION

Markovian Models (esp. finite Markov chains) are frequently used instruments for the investigation of mobility phenomena. Applications to income propagation go back to Champernowne. He showed that in the case of transition probabilities which only depend on the number of jumps upwards (or downwards) the resulting stationary distribution is Pareto, provided that a stability condition is fulfilled. Empirical and theoretical defects of the simple Markov model in other application fields, notably occupational mobility, led to the development of refined models (for a summary see e.g. Stewman [6]), which subsequently have been used to describe income distributions, too. Shorrocks [5] studied a second order Markov chain (with transition rates depending on the present income and on that of the preceeding period), which also yields the Pareto as the stationary distribution. McCall [4] applies the mover-stayer model where the population is considered to fall into two subpopulations with different mobility characteristics.

The core of those models is the simple Markov chain. In this paper two processes of that type are studied. Both models have the special feature that in a single time step there is only the possibility to jump into a neighbour income class. In section 2 this type of process which can be interpreted as a random walk with state-dependent transition rates is studied in detail. Model 1 which is a special case of Champernowne's model is discussed in section 3. Characteristics of the stationary distribution (the result of the process) are stated in terms of the parameters which govern the process. Section 4 deals with model 2 where the chance to decrease income is positively, the chance to increase income negatively correlated with the size of income. This model can be considered as a generalization of model 1. It yields stationary distributions which are qualitatively different from the Pareto. In section 5 the influence of small parameter changes on the stationary distribution is studied.

2. INCOME DISTRIBUTION PROCESSES AS RANDOM WALKS

Markovian Models (esp. finite Markov Chains) are frequently used instruments for the investigation of mobility phenomena. However there are several assumptioms implied by the models which are satisfied only approximately in reality. The discussion of those problems is beyond the scope of this paper (see e.g. Bartholomew [1]), thus the underlying assumptions are only stated here explicitely.

(i) In each period h=0,1,2,... an income takes on one of k dinstinct states $K_1, \dots K_k$. the states are distributed equiproportionally, i.e. $K_i = K_1 \cdot \Delta^{i-1}$ with some $\Delta > 1$. The case $k \to \infty$ is considered at the end of section 3. In this case the stationary distribution as well as some inequality measures only exist if further conditions (e.g. Champernowne's stability condition) are satisfied.

(ii) The process describing the dynamic change of an income is a finite homogeneous Markov chain. Thus it is necessary to consider well-defined states (e.g. midpoints of income classes) instead of aggregates of states (e.g. income classes).

(iii) In a single time step there are only two possibilities: to jump to a neighbour state (with nonzero probability!) or to remain in the state. Therefore the matrix of transition probabilities is tridiagonal with nonzero adjacent diagonals (i.e. the chain is irreducible).

THEOREM 2.1. If the matrix of transition probabilities P of a finite homogeneous Markov chain is tridiagonal with nonzero diagonals and nonzero adjacent diagonals, then

(i) all states are ergodic (i.e. aperiodic persistent with finite mean recurrence time)

(ii) there is one and only one stationary distribution $g = (g_1,\ldots,g_k)$, moreover, $g_i > 0 \quad \forall i$

(iii) the higher transition probabilities $p_{ij}^{(h)}$ (entries of P^h) converge to g_j irrespective of i

(iv) absolute distributions $p(h) = P'p(h-1)$ converge to g irrespective of the initial distribution p(0)

<u>Proof</u>. It is easy to see that the matrices P^h are strictly positive for $h > k-2$. Therefore no state can be periodic. As the chain is irreducible, all states must be persistent non-null states and thus ergodic. The remaining statements follow from the ergodic theorem.

The unique stationary distribution g is the eigenvector of P' associated with the eigenvalue 1 and can thus be determined by $P'g=g$, i.e. by the set of linear equations

$$g_1 = p_{11}g_1 + p_{21}g_2$$

$$g_2 = p_{12}g_1 + p_{22}g_2 + p_{32}g_3$$

$$\vdots$$

$$g_k = p_{k-1,k}\, g_{k-1} + p_{kk}\, g_k$$

The g_i therefore can be computed recursively by

$$g_i = \frac{p_{12}\, p_{23} \cdots\cdots p_{i-1,i}}{p_{21}\, p_{32} \cdots\cdots p_{i,i-1}} \cdot g_1 \quad 1 < i \leq k$$

where $g_1 > 0$ must be chosen so that $\sum_{i=1}^{k} g_i = 1$.

The special shape of the matrix faciliates not only the evaluation of the dominating eigenvector, but also of other characteristics. This is summarized in the sequel.

As the matrix P has positive adjacent diagonals, it is similar to a symmetric tridiagonal matrix. This transformed matrix has the same diagonal elements like P, the adjacent diagonal elements are replaced by $\sqrt{p_{i+1,i}\, p_{i,i+1}}$ (cf. Wilkinson [7] ,p.335). Thus all eigenvalues of P are real, moreover, they are all distinct (Wilkonson [7] , p.300). If, additionally, P has a dominating diagonal. i.e. $p_{ii} > \sum_{j\neq i} p_{ij}$ for all i (as it usually happens to be in the case of income mobility matrices), all eigenvalues are positive due to Gershgorin's theorem. The leading principal minors of $P-\lambda I$

$$D_r(\lambda) = \det \begin{pmatrix} p_{11}-\lambda & p_{12} & & & \\ p_{21} & p_{22}-\lambda & p_{23} & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & p_{r-1,r} \\ & & & p_{r,r-1} & p_{rr}-\lambda \end{pmatrix}$$

are of great importance for the investigation of the structure of P. As one can proof by induction, the following recursion formula holds

$$D_o(\lambda) = 1$$

$$D_1(\lambda) = p_{11} - \lambda$$

$$D_r(\lambda) = (p_{rr} - \lambda)D_{r-1}(\lambda) - p_{r,r-1}p_{r-1,r}D_{r-2}(\lambda)$$

Note that $D_k(\lambda)$ is the characteristic polynomial of P.

The functions $D_r(\lambda)$ can be used to estimate the eigenvalues of P. Let the quantities $D_o(\mu), D_1(\mu), \ldots D_k(\mu)$ be evaluated for some $\mu \in [-1,1]$. Then the number of agreements in sign of consecutive members of this sequence is the number of eigenvalues of P which are strictly greater than μ (if $D_r(\mu)=0$ for some r, then $D_r(\mu)$ is taken to have the opposite sign to that of $D_{r-1}(\mu)$). The eigensystems of P and its transpose can also be determined using the $D_r(\lambda)$. The components x_{ij} of the (unnormalized) eigenvector x_j of P associated with the eigenvalue λ_j are given by the formula

$$x_{1j} = 1$$

$$x_{ij} = (-1)^{i-1} \frac{D_{i-1}(\lambda_j)}{p_{12}\, p_{23} \cdots p_{i-1,i}}$$

See e.g. Gantmacher and Krein [3], p.8o. There are j-1 changes in sign in the sequence of the components of x_j. To obtain the eigenvectors y_j of the transpose P', the subscript of the factors in the denominator have to be interchanged.

At the end of this section, assume an arbitrary distribution p(o) over the k income states. In h periods this distribution is transformed by the process to $p(h) = (P^h)'p(o)$ which converges to the stationary distribution g. In fact the transition probabilities alter with time, too. Now the equilibrium concept of a unique stationary distribution only makes sense if the change of the parameters which govern the process is slow in comparison with the dynamics of the process itself. The stationary state will be approximated adequately only if the convergence $p(h) \rightarrow g$ is sufficiently fast to make the assumption of a constant transition matrix P (and with that a constant equilibrium g) plausible.[1] What determines the speed of this convergence $p(h) \rightarrow g$?

THEOREM 2.2. Let λ_2 be an eigenvalue of second largest modulus of P. Then for almost all p(0) the speed of convergence is asymptotically the same like that of a geometric series with quotient $|\lambda_2|$ i.e. for large h it holds that

$$\|p(h+1) - g\| \approx |\lambda_2| \cdot \|p(h) - g\|$$

where $\|.\|$ denotes an arbitrary norm.

Proof. For simplicity assume λ_2 to be the only eigenvalue of second largest modulus. Denote by $y_1 = g, y_2, \ldots y_k$ the eigenvectors of P' (associated with the eigenvalues $\lambda_1 = 1 > |\lambda_2| > \ldots \geq |\lambda_k|$), then p(0) can be represented by a linear combination $\sum \alpha_i y_i$, and

$$p(h) = (P^h)' \sum_{i=1}^{k} \alpha_i y_i = \sum_{i=1}^{k} \lambda_i^h \alpha_i y_i = \alpha_1 g_1 + \sum_{i=2}^{k} \lambda_i^h \alpha_i y_i$$

As $p(h) \rightarrow g$, α_1 must be equal 1. Then

$$\frac{\| p(h+1)-g \|}{\| p(h) - g \|} = |(\lambda_2)| \frac{\|\alpha_2 y_2 + \sum_{i=3}^{k} (\lambda_i/\lambda_2)^{h+1} \alpha_i y_i\|}{\|\alpha_2 y_2 + \sum_{i=3}^{k} (\lambda_i/\lambda_2)^{h} \alpha_i y_i\|}$$

If $\alpha_2 \neq 0$, the right hand side tends to $|\lambda_2|$. If $\alpha_2 = 0$ the speed is even higher, but the initial distributions distinguished by $\alpha_2 = 0$ form a subspace of measure zero.

An important measure in this context is the half-life of the process, the time necessary to half the distance between actual and stationary distribution. As can be seen from theorem 2.2., the half-life is asymptotically equal $\log 0.5 / \log|\lambda_2|$.

For the matrix given at the end of section 4 $\lambda_2 = .783$. In this case the half-life of the process is 2.82 periods of 3 years each, i.e. 8.5 years.

3. MODEL 1: PROPORTIONATE EFFECT

The "law of proportionate effect" states that the probability distribution of proportional income changes is independent of current income. Together with the equiproportionally increasing income states under consideration one would expect that the chances of upward or downward income changes do not depend on the present state. Therefore the matrix of transition probabilities is of the form

$$P = \begin{pmatrix} 1-p & p & & & & \\ q & 1-(p+q) & p & & & \\ & q & 1-(p+q) & p & & \\ & & \ddots & \ddots & \ddots & \\ & & & q & 1-(p+q) & p \\ & & & & q & 1-q \end{pmatrix}$$

with $p > 0$, $q > 0$, $p+q \leq 1$. Let $c=p/q$ be the ratio

$$\frac{\text{chance of moving up}}{\text{chance of movind down}}$$

If $c=1$, the stationary distribution is rectangular: $g_i = \frac{1}{k} \; \forall i$.
If $c \neq 1$, it is a geometric distribution

$$g_i = \frac{1-c}{1-c^k} c^{i-1} \qquad 1 \leq i \leq k \; .$$

Because of $K_i = K_1 \, \Delta^{i-1}$ this means

$$\log g_i = (\log c / \log \Delta) \log K_i + \text{const}$$

If $c < 1$ this is the same form as the Pareto, but there is a maximal income K_k. Note that g does not depend on the absolute values of the probabilities of upward or downward mobility but on their ratio c only. Median and mean of the stationary distribution are monotone increasing with c. If c is altered, the corresponding Lorenz curves do not alter monotonely. This can be concluded by the fact that the slope of the Lorenz curve at x=0 is

$$\frac{y_1}{x_1} = \frac{(1-\Delta c)(1-c^k)}{(1-(\Delta c)^k)(1-c)}$$

the slope at x=1 is

$$\frac{1-y_{k-1}}{1-x_{k-1}} = \frac{(1-c^k)(1-\Delta c)}{(1-(\Delta c)^k)(1-c)} \Delta^{k-1} = \Delta^{k-1} \frac{y_1}{x_1}$$

If now c is altered by a small amount Δc, the slopes at the two endpoints of the curve alter in the same direction. Thus the curves corresponding to c and c + Δc intersect at least once in some interior point. The shape of g nevertheless suggest that concentration measures like variance or Gini coefficient are monotone increasing for $c < 1$ and monotone decreasing for $c > 1$, but generally this is only true for sufficiently small or large c.

Table 1

Measures of the Model 1 stationary distribution

Median	m	$\dfrac{\ln \frac{1+c^k}{2}}{\ln c}$ [a]
Mean	μ	$K_1 \dfrac{\sum(\Delta c)^i}{\sum c^i}$
Variance	σ^2	$K_1{}^2 \left[\dfrac{\sum(\Delta^2 c)^i}{\sum c^i} - \left(\dfrac{\sum(\Delta c)^i}{\sum c^i} \right)^2 \right]$
Coefficient of Variation	$\frac{\sigma}{\mu}$	$\sqrt{\dfrac{\sum(\Delta^2 c)^i \sum c^i}{(\sum(\Delta c)^i)^2} - 1}$
Lorenz curve (vertices)	x_i	$\dfrac{1-c^i}{1-c^k}$
	y_i	$\dfrac{1-(\Delta c)^i}{1-(\Delta c)^k}$
Gini coefficient		$\dfrac{(1-c)(1-\Delta c)}{(1-c^k)(1-(\Delta c)^k)} \sum_{i=0}^{k-1} \sum_{j=i}^{k-1} c^{i+j}(\Delta^j - \Delta^i)$

[a] or the following integer

If not stated explicitely, sum indices range from 0 to k-1.

Let now P_k be the Model1-matrix with k rows (columns) and $\lambda_2(k)$ the eigenvalue of second largest modulus of P_k. Due to theorem 2.2 the speed of convergence of an absolute distribution p(h) essentially depends on $\lambda_2(k)$. Assume $\lambda_2(k) > 0$ which is evidently true if P has a dominating diagonal (cf. section 2).

THEOREM 3.1. If $k_1 < k_2$, then $\lambda_2(k_1) < \lambda_2(k_2)$

Proof. The proof is shown for $k_1=k$, $k_2=k+1$, $\bar{\lambda} := \lambda_2(k)$. $P_k - \bar{\lambda}I$ has leading principal minors $D_o(\bar{\lambda})$, $D_1(\bar{\lambda}), \dots D_k(\bar{\lambda})=0$. As there is only one eigenvalue strictly greater than $\bar{\lambda}$, there is exactly one agreement in sign of consecutive members of $D_o(\bar{\lambda}), \dots D_{k-1}(\bar{\lambda})$. $P_{k+1} - \bar{\lambda}I$ has leading principal minors $D_o(\bar{\lambda}), \dots D_{k-1}(\bar{\lambda})$, $E_k(\bar{\lambda})$, $E_{k+1}(\bar{\lambda})$.

$$E_k(\bar{\lambda}) = (1-(p+q)-\bar{\lambda})D_{k-1}(\bar{\lambda})-pq\ D_{k-2}(\bar{\lambda})$$

$$= -p\ D_{k-1}(\bar{\lambda}) + (1-q-\bar{\lambda})D_{k-1}(\bar{\lambda})-pq\ D_{k-2}(\bar{\lambda})$$

$$= -p\ D_{k-1}(\bar{\lambda}) + D_k(\bar{\lambda}) = -p\ D_{k-1}(\bar{\lambda})$$

Because of $D_k(\bar{\lambda}) = 0$ it holds that $D_{k-1}(\bar{\lambda}) \neq 0$, thus $E_k(\bar{\lambda})$ has the opposite sign of $D_{k-1}(\bar{\lambda})$

$$E_{k+1}(\bar{\lambda}) = (1-q-\bar{\lambda})(-p.D_{k-1}(\bar{\lambda}))-pq\ D_{k-1}(\bar{\lambda})$$

$$= -p(1-\bar{\lambda})\ D_{k-1}(\bar{\lambda}) = (1-\bar{\lambda})E_k(\bar{\lambda})$$

$E_{k+1}(\bar{\lambda})$ has the same sign like $E_k(\bar{\lambda})$, therefore there are exactly two agreements in sign in the sequence $D_o(\bar{\lambda}), \dots D_{k-1}(\bar{\lambda})$, $E_k(\bar{\lambda})$, $E_{k+1}(\bar{\lambda})$, i.e. there are exactly two eigenvalues of P_{k+1} which are strictly greater than $\bar{\lambda}$.

Calculating $\lambda_2(k)$ for k=2,3,4,5 yields the following corollary.

COROLLARY.

$$k = 2 : \lambda_2 = 1 - (p+q)$$

$$k = 3 : \lambda_2 = 1 - (p+q) + \sqrt{pq}$$

$$k = 4 : \lambda_2 = 1 - (p+q) + \sqrt{2}\sqrt{pq}$$

$$k = 5 : \lambda_2 = 1 - (p+q) + \frac{\sqrt{3+\sqrt{5}}}{2}\sqrt{pq}$$

$$k > 5 : |\lambda_2| > 1 - (p+q) + \frac{\sqrt{3+\sqrt{5}}}{2}\sqrt{pq} \; ,$$

If $p \approx q$ it holds that $|\lambda_2| > 1-0.39\,p$ for $k > 5$. If e.g. there are equal chances of upwards and downward change and staying, resp., then the speed of convergence is not faster than proportional to 0.87^h, the half-life of the process is at least 5 periods.

Model 1 can be modified by considering infinetly many states of type $K_i = K_1 \Delta^{i-1}$, Of course, the stationary distribution exists only if $c < 1$ (Champernowne's stability condition). To guarantee the existence of the measure of this distributions, further conditions have to be satisfied. They are stated in table 2. As opposed to the finite case, the Lorenz curve alters monotonely with c. Fix the index i of a vertex and consider this vertex to be a function of c: $(x_i(c), y_i(c))$. If c goes from 0 to $1/\Delta$, $(x_i(c), y_i(c))$ moves along a straight line from (1,1) to $(1-1/\Delta^i,0)$. Take now some fixed $\bar{c}$ and a corresponding Lorenz curve $L(\bar{c})$. Because of the convexity of $L(\bar{c})$, the line $(x_i(c), y_i(c))$ lies below $L(\bar{c})$ if $c > \bar{c}$. As this holds for all indices i, the Lorenz curve $L(\bar{c})$ lies wholly inside the Lorenz curve corresponding to some $c > \bar{c}$.

Table 2

Measures of the Model 1 stationary distribution
(infinitely many income states)

Median	m	$\frac{\ln 2}{-\ln c}$ [a]	
Mean	μ	$K_1 \frac{1-c}{1-\Delta c}$	if $\Delta c < 1$
Variance	σ^2	$K_1^2 \left[\frac{1-c}{1-\Delta^2 c} - \left(\frac{1-c}{1-\Delta c}\right)^2\right]$	if $\Delta^2 c < 1$
Coefficient of Variation	$\frac{\sigma}{\mu}$	$\sqrt{\frac{(1-\Delta c)^2}{(1-c)(1-\Delta^2 c)} - 1}$	if $\Delta^2 c < 1$
Lorenz curve (vertices)	x_i	$1 - c^i$	
	y_i	$1 - (\Delta c)^i$	if $\Delta c < 1$
Gini coefficient	G	$\frac{c(\Delta - 1)}{1 - \Delta c^2}$	if $\Delta c < 1$

[a] or the following integer

4. MODEL 2: REGRESSION TOWARDS THE MEAN

Main assumption: The higher the state, the lower the chance of further increase and the higher the chance of fall. In Model 2 the chances of moving upward are assumed to decrease geometrically with factor $d < 1$, the chances of moving downward increase with factor $1/d$.

$$P = \begin{pmatrix} 1-p & p & & & \\ d^{k-2}q & 1-dp-d^{k-2}q & dp & & \\ & d^{k-3}q & 1-d^2p-d^{k-3}q & d^2p & \\ & & \ddots & \ddots & d^{k-2}p \\ & & & q & 1-q \end{pmatrix}$$

with $p > 0$, $q > 0$, $p+q \leq 1$. Here $c=p/q$ is the ratio chance of moving up from lowest state/chance of moving down from highest state. Model 1 in the limiting case $d > 1$. For the stationary distribution we have

$$\frac{g_{i+1}}{g_i} = c.d^{2i-k}$$

thus with $S = \sum_{i=0}^{k-1} \left(\frac{c}{d^{k-1}}\right)^i d^{i^2}$ we obtain

$$g_i = \frac{1}{S} \left(\frac{c}{d^{k-1}}\right)^{i-1} d^{(i-1)^2}$$

Because of the equiproportionally chosen income states this means

$$\log g_i = \frac{\log c/d^{k-1}}{\log \Delta} \log K_i + \frac{\log d}{(\log \Delta)^2} (\log K_i)^2 + \text{const}$$

The continuous analogue of this has

$$-\frac{d(\log g)}{d(\log K)} = \frac{\log d^{k-1}/c}{\log \Delta} + \frac{2 \log 1/d}{(\log \Delta)^2} \quad \log K$$

As d was assumed to be smaller than one the frequency tends faster to zero than the Pareto which would have the right hand side constant.

For small values of c ($c \leq d^{k-2}$) the frequencies g_i are monotonically decreasing, for large values ($g \geq d^{2-k}$) they are increasing. For moderate values $d^{k-2} < c < d^{2-k}$ (which obviously holds if $p \approx q$, i.e. if $c \approx 1$) the frequencies firstly are monotonically increasing, and, after passing a single maximal value (in limit cases two adjacent maximal values), the g_i are monotonically decreasing. The mode is the state No. $\left[\frac{1}{2}(k+2-\ln c/\ln d)\right]$, this index is increasing in c (d fixed), with decreasing d (c fixed) it tends to the central state $K_{\left[\frac{n+1}{2}\right]}$. The mode corresponds exactly to that income state which is distinguished by the smallest difference between the chances of moving upward and moving downward.

Because of $g_{i+1}/g_i = c.d^{2i-k}$ the stationary distribution concentrates at the central state with decreasing d (c fixed). Thus in the case of c fixed, $c \approx 1$, the author conjectures that concentration measures like Gini's coefficient are monotonely increasing functions of d (though the corresponding Lorenz curves may intersect, as can be seen from examples).

The results of Model 2 differ qualitatively from those of Model 1. One may wonder if Model 2 can be generalized e.g. by introducing different change rates d_1 (change of upward mobility chances) and d_2 (change of downward mobility chances). Qualitatively this concept yields the same stationary distribution, because now

$$\frac{g_{i+1}}{g_i} = c \; \frac{d_1^{\,i-1}}{d_2^{\,k-i-1}} = c \; (\sqrt{d_1/d_2})^{k-2} \sqrt{d_1 d_2}^{\,2i-k}$$

But this stationary distribution may be derived by a Model 2 process with $\tilde{c} = c\,(\sqrt{d_1/d_2})^{k-2}$, $\tilde{d} = \sqrt{d_1 d_2}$.

Table 3

Measures of the Model 2 stationary distribution

Mean	μ	$\frac{K_1}{S} \sum \left(\frac{c\,\Delta}{d^{k-1}}\right)^i d^{i^2}$
Variance	σ^2	$K_1^2\left[\frac{1}{S} \sum \left(\frac{c\,\Delta^2}{d^{k-1}}\right)^i d^{i^2} - \left(\frac{1}{S} \sum \left(\frac{c\,\Delta}{d^{k-1}}\right)^i d^{i^2}\right)^2\right]$
Coefficient of Variation	$\frac{\sigma}{\mu}$	$\sqrt{\frac{\left(\sum \left(\frac{c}{d^{k-1}}\right)^i d^{i^2}\right)\left(\sum \left(\frac{c\,\Delta^2}{d^{k-1}}\right)^i d^{i^2}\right)}{\left(\sum \left(\frac{c\,\Delta}{d^{k-1}}\right)^i d^{i^2}\right)^2} - 1}$
Lorenz curve (vertices)	x_i	$\frac{1}{S} \sum_{j=0}^{i=1} \left(\frac{c}{d^{k-1}}\right)^j d^{j^2}$
	y_i	$\frac{\sum_{j=0}^{i=1} \left(\frac{c\,\Delta}{d^{k-1}}\right)^j d^{j^2}}{\sum_{j=0}^{k-1} \left(\frac{c\,\Delta}{d^{k-1}}\right)^j d^{j^2}}$
Gini coefficient	G	$\frac{1}{S} \frac{\sum_{i=0}^{k-1} \sum_{j=i}^{k-1} \left(\frac{c}{d^{k-1}}\right)^{i+j} d^{i^2+j^2}\left(\Delta^j - \Delta^i\right)}{\sum_{i=0}^{k-1} \left(\frac{c\,\Delta}{d^{k-1}}\right)^i d^{i^2}}$

If not stated explicitely, sum indices range from 0 to k-1.

Shorrocks [5] presents 3 empirical transition matrices, (from a sample of male employees), one of them is shown below.

		1966					
		1	2	3	4	5	Sample size
1963	1	0.64	0.29	0.04	0.03	0.00	76
	2	0.14	0.56	0.26	0.03	0.01	212
	3	0.02	0.22	0.54	0.21	0.01	256
	4	0.01	0.04	0.27	0.54	0.14	164
	5	0.00	0.01	0.05	0.27	0.67	92

The stationary distribution is given by

$$g_1 = 0.12 \quad g_2 = 0.25 \quad g_3 = 0.30 \quad g_4 = 0.22 \quad g_5 = 0.11$$

Form $g_{i+1}/g_i = c.d^{2i-k}$ we obtain the identy $\frac{g_{i+2}\, g_i}{g_{i+1}^2} = d^2$.

This formula allows an estimate $\hat{d} = 0.8$. In connection with the assumption c=1 this estimated value corresponds to a hypothetical stationary distribution

$$g_1 = 0.12 \quad d_2 = 0.23 \quad g_3 = 0.30 \quad g_4 = 0.23 \quad g_5 = 0.12$$

which very well fits the stationary distribution given above.

5. PERTURBATION THEORY

One may wonder how the stationary distribution alters if some transition probabilities are changed, e.g. by reinforcement or by reduction of mobility barriers or by creation of favoured transitions. If those variations are relatively small (perturbations), the effects can be estimated satisfactory by first order approximations. As in our case the unperturbed matrix P' (tridiagonal, positive adjacent diagonals) has k distinct real eigenvalue $\lambda_1=1,\lambda_2,\ldots\lambda_k$ corresponding to eigenvectors $g,y_2,\ldots y_k$, the resulting formulas are quite simple.

Assume now that the matrix P is replaced by a perturbed matrix $P + \varepsilon B$, $\|B\| = \|P\|$. Of course, B must have row sums 0 and ε must be so small that $P + \varepsilon B$ is a stochastic matrix.

THEOREM 5.1. Denote the eigenvalues of P with $1,\lambda_2,\ldots\lambda_k$, the corresponding eigenvectors with $x_1,x_2,\ldots x_k$, the corresponding eigenvectors of P' with $g,y_2,\ldots y_k$. Then the matrix $P + \varepsilon B$ defines a Markov chain with a stationary distribution $g(\varepsilon)$ which satisfies

$$g(\varepsilon) = g + \varepsilon \sum_{i=2}^{k} \frac{x_i' Bg}{(1-\lambda_i)x_i' y_i} \cdot y_i + O(\varepsilon^2)$$

Remark. The first order approximation

$$\tilde{g} = g + \varepsilon \sum_{i=2}^{k} \frac{x_i' Bg}{(1-\lambda_i)x_i' y_i} y_i$$

is a probability distribution, i.e. $\tilde{g}_i \geqslant 0$, $\sum_{i=1}^{k} \tilde{g}_i = 1$.

Proof. See Wilkinson [7] , p.69.

Proof of the remark. If ε is sufficiently small, $\tilde{g}_i \geq 0$ is satisfied (remember $g_i > 0$!). Denote the components of y_i with $y_{1i},\dots,y_{ki}$. As $x_i'y_j=0$ if $i\neq j$ (eigenvectors of P and P' which are associated to different eigenvalues are orthogonal), we have $x_1'y_i=0 \quad \forall i>1$. But $x_1=(1,1,\dots,1)'$, thus $\sum_{j=1}^{k} y_{ji}=0 \quad \forall i > 1$ and we obtain the result

$$\sum_{j=1}^{k} \tilde{g}_j = \underbrace{\sum_{j=1}^{k} g_j}_{=1} + \underbrace{\varepsilon \sum_{j=1}^{k} \sum_{i=2}^{k} \frac{x_i'Bg}{(1-\lambda_i)x_i'y_i} y_{ji}}_{\sum_{i=2}^{k} \frac{x_i'Bg}{(1-\lambda_i)x_i'y_i} \underbrace{\sum_{j=1}^{k} y_{ji}}_{=0}} = 1$$

Unfortunately this formula requires the calculation of the complete eigensystem of P and P', but it is valid for arbitrary perturbations B (provided row sums 0). If the perturbation only concerns one single origin state, a simpler formula can be derived.

THEOREM 5.2. If for some i $(1 < i < k)$ $p_{i,i-1}$, p_{ii}, $p_{i,i+1}$ are replaced by $p_{i,i-1}+\varepsilon_1$, $p_{ii}-\varepsilon_1-\varepsilon_2$, $p_{i,i+1}+\varepsilon_2$, resp., the stationary distribution g is altered to

$$\begin{aligned}
\tilde{g}_1 &= a.g_1 \\
&\vdots \\
\tilde{g}_{i-1} &= a.g_{i-1} \\
\tilde{g}_i &= \frac{p_{i,i-1}}{p_{i,i-1}+\varepsilon_1}\, a\, g_i = \frac{p_{i,i+1}}{p_{i,i+1}+\varepsilon_2}\, b\, g_i \\
\tilde{g}_{i+1} &= b.g_{i+1} \\
&\vdots \\
\tilde{g}_k &= b.g_k
\end{aligned}$$

In first approximation the factors a,b are given by

$$a = 1 + \sum_{j=i}^{k} g_j \frac{\varepsilon_1}{p_{i,i-1}} - \sum_{j=i+1}^{k} g_j \frac{\varepsilon_2}{p_{i,i+1}}$$

$$b = 1 - \sum_{j=1}^{i-1} g_j \frac{\varepsilon_1}{p_{i,i-1}} + \sum_{j=1}^{i} g_j \frac{\varepsilon_2}{p_{i,i+1}}$$

Proof: by straightforward computation.

FOOTNOTE

[1] Generally in equilibrium models a small change of the parameters not necessarily results in a small change of the corresponding equilibrium (cf. the literature on Catastrophe theory). In our situation, however, this is true, as can be seen from theorem 5.1 in section 5.

REFERENCES

1. D.J.Bartholomew, "Stochastic Models for Social Processes," 2nd ed., Wiley, London 1973.

2. D.G.Champernowne, A Model of Income Distribution, Econ.J. 63 (1953), 318-351.

3. F.R.Gantmacher, M.G.Krein, "Oszillationsmatrizen, Oszillationskerne und kleine Schwingungen mechanischer Systeme," Akademie-Verlag, Berlin (East) 196o.

4. J.J.McCall, A Markovian Model of Income Dynamics, J.Amer.Statist.Assoc., 66 (1971),

5. A.F.Shorrocks, Income Mobility and the Markov Assumption, Econ.J. 86 (1976), 566-578.

6. S.Stewman, Markov Models of Occupational Mobility, Theoretical Development and Empirical Support, J.Mathematical Sociol. 4 (1976), 2o1-278.

7. J.H.Wilkinson, "The Algebraic Eigenvalue Problem", Clarendon Press, Oxford 1965.