The Economic Journal, **86** (*September* 1976), 566–578
Printed in Great Britain

INCOME MOBILITY AND THE MARKOV ASSUMPTION[1]

I. INTRODUCTION

A number of stochastic frameworks have been proposed for the study of income distributions. Although originally conceived as explanations for the characteristic form and stability of observed frequency distributions, the models also offer simple specifications of income mobility. This paper is primarily concerned with the appropriate description of mobility and in particular with the assumption that the process governing income changes is first order Markov over the natural state space.[2] Future income then depends on current income but not on income levels in the past. Objections to the Markov property have previously been made on theoretical grounds[3] but the data necessary for any empirical investigation have been difficult to obtain.[4] Use is made here of tabulations prepared from a sample of 800 male employees, of the same age, whose incomes are known for the financial years ending in 1963, 1966 and 1970. Evidence derived from this sample has already been discussed in Hart (1973).

The examination of the Markov property is undertaken within the context of a Markov chain approach based on Champernowne (1953). The assumptions of the Markov chain model are considered briefly in section II and compared with evidence from the income sample. A simple consistency requirement is violated and this casts doubt on the validity of the Markov assumption. Replacing it with an obvious alternative, and reinterpreting the other assumptions where necessary, produces a generalised version of the Champernowne model. Various aspects of the modified model are discussed in section III, and it is shown that the form of the equilibrium distribution remains substantially unchanged.

Maximum-likelihood estimates of the transition probabilities were computed for both models, when annual income mobility was restricted to a single class movement in either direction. The exact sampling distribution generated by these models was used for estimation and enabled various likelihood ratio tests to

[1] Earlier drafts of the paper were discussed in seminars at Cambridge, Geneva, LSE and Reading. Participants of these seminars, together with Professor Champernowne and a referee, contributed greatly towards improvements in the exposition. Responsibility for the views expressed, however, rests with the author alone.

[2] This assumption is made in almost all stochastic models of size distributions. The only exception appears to be the approach taken by Ijiri and Simon (1964, 1967), who reject "the least acceptable assumption in the simple models: the assumption that the growth rates of individual firms in one period of time are uncorrelated with their growth rates in preceding periods" (1964, p. 79). They propose a non-Markov process and examine equilibrium distributions obtained from simulation studies, commenting that "Stochastic models admitting serial correlation have proved too complex to be solved explicitly in closed form for the equilibrium distributions" (1964, p. 80).

[3] For example, see Lydall (1968), p. 23.

[4] Income levels for the same individuals are required at three or more points of time. An income sample for three consecutive years was obtained by the Oxford Savings Survey and the Markov chain model applied by Vandome (1958). No tests for specification were undertaken apart from a comparison between the actual distribution and the predicted equilibrium.

be performed. The results of these tests, reported in section IV, show the modified model to be an acceptable specification of the income generating process and a significant improvement over the Markov version.

II. THE MARKOV CHAIN APPROACH

Data on mobility can be conveniently summarised in a table cross-classifying income levels in two years. Using income classes constructed on a logarithmic scale, tabulations have been made for the sample of male employees. The following table, in which the rows have been divided by the row sums to give a mobility matrix M_1, corresponds to the period 1963–5.[1] It is such a matrix which

		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

forms the basis for a Markov chain model. The matrix element m_{jk} represents the proportion of individuals originally in class j whose income in the second year falls within class k, and can be regarded as an estimate of the transition rate p_{jk} defined as the probability of a movement from class j to class k. Making the assumptions:

(A 1) (Population homogeneity.) The same transition rates apply to all individuals in the group being studied.

(A 2) (First-order Markov.) Individuals in income class j at time t have the same transition probabilities regardless of their past history.

(A 3) (Time homogeneity.) The transition rates p_{jk} remain constant over time.

the process is described by the equation $\boldsymbol{n}(t_m) = \boldsymbol{n}(t_{m-1})P$, where $\boldsymbol{n}(t_m)$ is the row vector indicating the expected numbers in each of the income classes m periods after the process begins.[2] A weak condition[3] on the *transition matrix P* guarantees that the distribution converges to a unique stationary state or equilibrium distribution $\boldsymbol{n}^*$ which depends only on the transition probabilities and not on the initial distribution $\boldsymbol{n}(t_0)$.

In Champernowne's formulation there are an infinite number of income

[1] In defining the income classes an attempt was made to eliminate the overall growth effects and isolate the relative changes. Incomes were converted to their logarithms (base 10) and allocated to classes arranged about the mean log income for that year, so the end points of the intervals shift upwards over time. Apart from the extreme classes which are open-ended, the intervals here have length 0·1 and within any class the highest incomes remain approximately 25 % greater than the lowest. Minor adjustments to the matrix elements have been made where rounding errors caused the row sum to differ from unity.

[2] This is the closed form of the model which excludes entry or exit from the population.

[3] The matrix has to be irreducible and aperiodic, for which a sufficient condition is that the three main diagonals have non-zero elements.

classes of equal logarithmic length above some minimum level. A further assumption

(A 4) (Law of proportionate effect.) The probability distribution of proportional income changes is independent of current income,

is interpreted as meaning that the probabilities p_{jk} depend only on the magnitude of the transition $k-j$ and are otherwise independent of the current class occupied so that[1] $p_{jk} = q_{k-j}$. With an infinite number of income classes there is no guarantee that a stationary state exists, since the distribution could shift continuously into higher income classes with no upper bound on mean income. His "stability" condition[2] eliminates this possibility and ensures that the process converges to an equilibrium, which is always asymptotically geometric. He goes on to show that a geometric distribution over logarithmic income classes is equivalent to an exact Pareto distribution for income.[3]

Any comparison of the above assumptions with evidence from mobility matrices inevitably runs into a variety of statistical problems. Given the limitations on data there is little option but to accept the assumption of population homogeneity (A 1) at the outset. This requires an appropriate selection of income ranges and in particular the intervals should not be too wide, since the transition probabilities may then be significantly different for the upper and lower incomes within each group. However, reducing the size of the interval spreads the sample over a larger number of classes and makes the true proportions transferring between classes more difficult to estimate accurately.[4] For the empirical work reported in section IV, some attempt to overcome this problem was made by allowing the class length to vary in the hope that the data would suggest an optimal value, but on this point the results were inconclusive.

A second major difficulty arises from the fact that when evidence is found to conflict with the model, it is virtually impossible to isolate those assumptions which are invalid. It is important to realise that there is a hierarchy of assumptions. The choice of state space and population homogeneity (A 1) are essentially prior to the Markov property (A 2), which in turn is required before time homogeneity (A 3) and proportionate effect (A 4) are well defined. Thus evidence apparently inconsistent with either (A 3) or (A 4) may more appropriately reflect a violation of (A 1) or (A 2).

There is one specification test which does not depend on the validity of (A 3) or (A 4) and is therefore of particular significance. If the transition matrices for the

[1] Champernowne (1953, p. 322) assumes only that the law applies to the "rich", i.e. for income classes j greater than some positive integer J. In the special case where the law applies to all classes except the lowest and where incomes can only change by one class in each time period, the process becomes equivalent to a simple random walk with a "reflecting barrier". See Cox and Miller (1965), pp. 25–46.

[2] Champernowne (1953), p. 324.

[3] *Ibid.* p. 326. The assumption of a strictly positive minimum income is crucial to the form of the equilibrium distribution resulting from the "law" of proportional effect. *Without* that assumption the process tends to generate a lognormal (or "Gibrat") distribution, but *with* that assumption the distribution becomes asymptotically Pareto for large incomes.

[4] An estimate for the standard error of m_{jk} is given by $\sqrt{\frac{m_{jk}(1-m_{jk})}{n_j - 1}}$, where $n_j = \sum_k m_{jk}$. This gives a standard error of about 0·03 for the majority of the elements of the matrix M_1.

periods 1963–6, 1966–70 and the combined period 1963–70 are denoted by P_1, P_2 and P_3 respectively, then

$$\boldsymbol{n}(1966) = \boldsymbol{n}(1963).P_1,$$
$$\boldsymbol{n}(1970) = \boldsymbol{n}(1966).P_2 = \boldsymbol{n}(1963).P_1P_2,$$
$$\boldsymbol{n}(1970) = \boldsymbol{n}(1963).P_3,$$

so $P_3 = P_1P_2$. The corresponding mobility matrices M_2 and M_3 for the income sample are

$$\begin{array}{cc} 1966\text{–}70 & 1963\text{–}70 \\ M_2 = \begin{bmatrix} 0{\cdot}78 & 0{\cdot}15 & 0{\cdot}07 & 0{\cdot}00 & 0{\cdot}00 \\ 0{\cdot}22 & 0{\cdot}50 & 0{\cdot}24 & 0{\cdot}03 & 0{\cdot}01 \\ 0{\cdot}05 & 0{\cdot}23 & 0{\cdot}45 & 0{\cdot}25 & 0{\cdot}02 \\ 0{\cdot}00 & 0{\cdot}05 & 0{\cdot}23 & 0{\cdot}45 & 0{\cdot}27 \\ 0{\cdot}01 & 0{\cdot}00 & 0{\cdot}05 & 0{\cdot}19 & 0{\cdot}75 \end{bmatrix}, & M_3 = \begin{bmatrix} 0{\cdot}69 & 0{\cdot}22 & 0{\cdot}09 & 0{\cdot}00 & 0{\cdot}00 \\ 0{\cdot}26 & 0{\cdot}40 & 0{\cdot}23 & 0{\cdot}10 & 0{\cdot}01 \\ 0{\cdot}05 & 0{\cdot}25 & 0{\cdot}38 & 0{\cdot}26 & 0{\cdot}06 \\ 0{\cdot}01 & 0{\cdot}08 & 0{\cdot}31 & 0{\cdot}37 & 0{\cdot}23 \\ 0{\cdot}00 & 0{\cdot}02 & 0{\cdot}08 & 0{\cdot}18 & 0{\cdot}72 \end{bmatrix} \end{array}$$

and

$$M_1M_2 = \begin{bmatrix} 0{\cdot}57 & 0{\cdot}25 & 0{\cdot}14 & 0{\cdot}03 & 0{\cdot}01 \\ 0{\cdot}24 & 0{\cdot}36 & 0{\cdot}27 & 0{\cdot}10 & 0{\cdot}03 \\ 0{\cdot}09 & 0{\cdot}25 & 0{\cdot}34 & 0{\cdot}24 & 0{\cdot}08 \\ 0{\cdot}03 & 0{\cdot}11 & 0{\cdot}26 & 0{\cdot}34 & 0{\cdot}26 \\ 0{\cdot}01 & 0{\cdot}03 & 0{\cdot}12 & 0{\cdot}26 & 0{\cdot}58 \end{bmatrix}.$$

The diagonal elements of M_1M_2 are consistently below those of M_3, implying that the number of individuals whose position in the distribution is substantially unchanged over the seven year period would be underestimated on the basis of the subperiod matrices. This regular deviation from predicted values has invariably occurred when Markov chains are applied to other related problems, notably occupational mobility; and it has been shown that either population inhomogeneity or the non-Markovian nature of the process is sufficient to account for this discrepancy.[1] Unfortunately it is impossible to discriminate further without more detailed data, although using a sample of the same sex and age (30 years old in 1963) eliminates some of the potential causes of inhomogeneity. The Markov assumption is therefore at least suspect and on balance the more obvious candidate for replacement.

The law of proportionate effect suggests that the values along any diagonal of a mobility matrix should be approximately equal. Neglecting the extreme (open-ended) income classes, this is broadly true for the matrices M_1, M_2 and M_3 with deviations within the range attributable to sampling fluctuations. About time homogeneity nothing can be said as the matrices cover different time periods and are consequently not comparable. In any case, had the observed transition rates appeared to conflict with these assumptions, relaxing the Markov condition and making the required changes to (A 3) and (A 4) may be sufficient to re-establish consistency. This will be demonstrated in the next section.

Given that the Markov assumption is to be abandoned, the simplest modifica-

[1] See Bartholomew (1973), pp. 34–42. One suggested solution is the "mover–stayer" model applied to income mobility by McCall (1971). This assumes that separate subgroups of the population have different transition rates – the "movers" who can change their income class and the completely immobile "stayers". On the other hand, the "cumulative inertia" model abandons the Markov assumption. Another possible explanation is the inappropriate selection of income classes. For example, it is known that if the states of a Markov chain are grouped together, the resulting process will not in general be Markov.

tion is to allow transition rates to depend on both current income and immediate past history. Replacing (A 2) with

(A 2′) Individuals having been in income class i at time $t-1$ and now in class j at time t have the same transition rates regardless of their income history prior to $t-1$.

the process is now second-order Markov[1] and probabilities have to be assigned to all possible three state sequences $i \rightarrow j \rightarrow k$. Denoting by p_{ijk} the probability of a transition from j to k in the period $(t, t+1)$ given that state i was occupied at time $t-1$, the time homogeneity assumption (A 3′) will now require that p_{ijk} does not change over time. A natural replacement for the law of proportionate effect is

(A 4′) For all individuals who have experienced the same income growth in the recent past, the probability distribution of proportional size changes in income is independent of current income.

Keeping the equi-logarithmic income classes, this can be interpreted as requiring that p_{ijk} depends only on the size of the transitions $j-i$ and $k-j$ in the two periods $(t-1, t)$ and $(t, t+1)$. Thus

$$p_{ijk} = q_{k-j}^{j-i}$$

and the Markov case considered by Champernowne corresponds to the restriction that the q terms do not depend on the superscripts.

For theoretical reasons it seems likely that the probability of moving to a higher relative income position is inversely related to recent changes in rank and vice versa.[2] In terms of the transition rates this will mean that $q_{+}^{-} > q_{+}^{0} > q_{+}^{+}$ and $q_{-}^{-} < q_{-}^{0} < q_{-}^{+}$, where the signs indicate the direction of the consecutive income movements. One argument is based on the distinction between permanent and transitory income made by Friedman (1957). Even if permanent income follows a Markov process, the process governing actual income will be non-Markov[3] unless the permanent and transitory parts are perfectly correlated. High relative income growth in one period will normally be associated with a large positive transitory component and there will be a deterioration in the relative income position next period if transitory income growth is not repeated. A second possibility involves the changes in income associated with occupational movements through promotion/demotion, acquisition of new skills, completion of periods of

[1] The process is second-order Markov with reference to the original natural choice of a state space, i.e. one-dimensional income intervals. Many simple non-Markov processes can be converted into Markov processes by redefining the state space. The modification proposed here is one such case. If the states of the system are two-dimensional vectors giving the income classes occupied in successive periods, the first-order Markov condition is again applicable. This is essentially the method for determining the equilibrium distribution adopted in the next section.

[2] If only the Markov assumption is violated, income growth rates in consecutive periods will not be independent. However, the converse does not apply. For example, models replacing proportional effect with "regression towards the mean" also give correlated growth rates and in practice it is difficult to discriminate between these two explanations (Creedy (1974), pp. 409–10). Thatcher (1976) attempts a comparison of the "regression" formulation with a simple model derived from Friedman (1957), which is non-Markov. His results suggest that the "regression" model does rather poorly and he concludes: "it is for discussion whether the data described in this paper really provide any evidence that the earnings of manual men follow a Markov process at all".

[3] Lydall (1968, 1974) has correctly pointed this out on a number of occasions, although he appears to regard it as a criticism of stochastic models in general rather than one particular assumption.

training and so on. Since those who have recently received a significant income increment due to promotion are unlikely to be considered for further promotion in the near future, they will tend to experience lower income changes than the average of their contemporaries, some of whom are being promoted.

Other factors may contribute to violations of the Markov assumption; for example, incremental salary scales and the regular revisions of wage and salary rates, when occupational groups may catch up after periods in which their relative incomes have been eroded. The proposed generalisation to a second-order process is clearly only a simple attempt to capture some of the non-Markovian aspects omitted in Champernowne's original exposition.

III. A MODIFIED MODEL ALLOWING TRANSITION RATES TO DEPEND ON BOTH CURRENT INCOME AND INCOME IN THE PREVIOUS PERIOD

The logarithmic income classes will be numbered 0, 1, 2, 3, ... and for convenience it will be assumed that individuals can move up or down only a single income class in each period.[1] The appropriate transition rates can then be selected from the tables

$$\begin{array}{lc} & \text{Future transition } (t, t+1) \\ & \overbrace{\begin{array}{ccc} -1 & 0 & 1 \end{array}} \\ \begin{array}{l}\text{Past}\\ \text{transition}\\ (t-1,t)\end{array} \left\{\begin{array}{r} -1 \\ 0 \\ 1 \end{array}\right. & \begin{array}{ccc} q_{-1}^{-1} & q_0^{-1} & q_1^{-1} \\ q_{-1}^{0} & q_0^{0} & q_1^{0} \\ q_{-1}^{1} & q_0^{1} & q_1^{1} \end{array} \end{array} \quad \text{(when the class occupied is not the lowest),}$$

$$\begin{array}{lcc} & 0 & 1 \\ \begin{array}{l}\text{Past transition}\\ (t-1,t)\end{array} \left\{\begin{array}{r} -1 \\ 0 \end{array}\right. & \begin{array}{c} 1-q_1^{-1} \\ 1-q_1^{0} \end{array} & \begin{array}{c} q_1^{-1} \\ q_1^{0} \end{array} \end{array} \quad \text{(for the lowest class).}$$

The rows in each table sum to unity and will be identical if the (first order) Markov assumption holds.

Let $n_{ij}(t)$ be the expected number of individuals in class i at time $t-1$ and class j at time t and denote the total number of individuals (invariant over time) as N. Then

$$n_{jk}(t+1) = \sum_i n_{ij}(t)\, p_{ijk} \quad \text{for all} \quad j, k \geqslant 0,$$

$$\sum_k n_{jk}(t+1) = \sum_i n_{ij}(t) \sum_k p_{ijk} = \sum_i n_{ij}(t),$$

$$\sum_j \sum_k n_{jk}(t+1) = \sum_i \sum_j n_{ij}(t) = N.$$

The distribution n_{ij}^* will be an equilibrium distribution for this process if

$$n_{jk}^* = \sum_i n_{ij}^* p_{ijk} \quad \text{for all} \quad j, k \geqslant 0,$$

which gives

$$n_{00}^* = n_{00}^* p_{000} + n_{10}^* p_{100} = n_{00}^*(1-q_1^0) + n_{10}^*(1-q_1^{-1}),$$

$$n_{01}^* = n_{00}^* p_{001} + n_{10}^* p_{101} = n_{00}^* q_1^0 + n_{10}^* q_1^{-1} \quad \text{when} \quad j = 0,$$

$$n_{jk}^* = n_{j-1,j}^* q_{k-j}^1 + n_{j,j}^* q_{k-j}^0 + n_{j+1,j}^* q_{k-j}^{-1} \quad \text{when} \quad j > 0.^2$$

[1] This restriction was made partly to simplify the exposition and also to keep the computations of section IV within reasonable limits. It may not be unreasonable if the time periods are relatively short and the income classes sufficiently wide. It seems unlikely that relaxing the assumption would substantially change any of the results in this section.

[2] When $|k-j| > 1$, $n_{jk}^* = 0$, since $q_{k-j}^u = 0$.

The solution to these equilibrium equations, obtained by successive substitutions and the identity $\sum_k n^*_{jk} = \sum_i n^*_{ij}$, satisfies

$$n^*_{j+1,j} = n^*_{j,j+1} = \left[\frac{\Delta_1}{\Delta}\right]^j n^*_{10} \quad (j \geqslant 0),$$

$$n^*_{jj} = \frac{\Delta_2}{\Delta}\left[\frac{\Delta_1}{\Delta}\right]^{j-1} n^*_{10} \quad (j > 0),$$

$$n^*_{00} = \frac{1 - q_1^{-1}}{q_1^0} n^*_{10},$$

where

$$\Delta = \begin{vmatrix} 1 - q_1^{-1} & q_1^0 \\ q_0^{-1} & 1 - q_0^0 \end{vmatrix}, \quad \Delta_1 = \begin{vmatrix} q_1^1 & -q_1^0 \\ q_0^1 & 1 - q_0^0 \end{vmatrix}, \quad \Delta_2 = \begin{vmatrix} 1 - q_1^{-1} & q_1^1 \\ -q_0^{-1} & q_0^1 \end{vmatrix}.$$

The equilibrium expected number in the jth income class N^*_j is derived from the sum $\sum_i n^*_{ij}$ or equivalently $\sum_k n^*_{jk}$ giving

$$N^*_0 = n^*_{00} + n^*_{10} = \left\{\frac{1 + q_1^0 - q_1^{-1}}{q_1^0}\right\} n^*_{10},$$

$$N^*_1 = n^*_{01} + n^*_{11} + n^*_{21} = \left\{\frac{\Delta_1 + \Delta_2 + \Delta}{\Delta}\right\} n^*_{10},$$

$$N^*_j = n^*_{j-1,j} + n^*_{jj} + n^*_{j+1,j} = \left\{\frac{\Delta_1}{\Delta}\right\}^{j-1} N^*_1 \quad (j > 0).$$

Making the assumption that $\Delta > \Delta_1$ corresponding to Champernowne's "stability condition", the N^*_j sum to a finite multiple of n^*_{10} ensuring that a unique equilibrium distribution exists. The value of n^*_{10} is obtained by equating $\sum_{j=0}^{\infty} N^*_j$ with N, but, whatever its value, N^*_j is clearly distributed geometrically above the lowest class and in this range, therefore, income will have an exact Pareto distribution.

The results can be compared with the first-order Markov case by setting $q^{j-i}_{k-j} = q_{k-j}$, which makes

$$\Delta = q_{-1}, \quad \Delta_1 = q_1, \quad \Delta_2 = q_0, \quad \Delta + \Delta_1 + \Delta_2 = 1$$

and

$$N^*_0 = \frac{n^*_{10}}{q_1},$$

$$N^*_1 = \frac{n^*_{10}}{q_{-1}} = \frac{q_1}{q_{-1}} N^*_0,$$

$$N^*_j = \left(\frac{q_1}{q_{-1}}\right)^{j-1} N^*_1 = \left(\frac{q_1}{q_{-1}}\right)^j N^*_0 \quad (j \geqq 0).$$

Since this is geometric over all income classes, the generalisation to a second-order process has only altered the form of the distribution at the lowest income class.

If the mobility matrix M were constructed for the period $(t, t+1)$ the elements m_{jk} will be given approximately by

$$m_{jk} \simeq \frac{n_{jk}(t+1)}{\sum_i n_{ij}(t)} = \frac{n_{jk}(t+1)}{N_j(t)} = \frac{\sum_i n_{ij}(t)\, p_{ijk}}{N_j(t)}$$
$$= \sum_i w_{ij}(t)\, q_{k-j}^{j-i},$$

where
$$w_{ij}(t) = \frac{n_{ij}(t)}{N_j(t)} \quad \text{and} \quad \sum_i w_{ij}(t) = 1.$$

The expected value of m_{jk} is therefore a weighted sum of q_{k-j}^{-1}, q_{k-j}^{0} and q_{k-j}^{1}, with the weights determined by the proportions of individuals in state j at time t who occupied each of the possible classes in the previous time period. These proportions are likely to vary in different time periods and the expected value of m_{jk} will change over time. Thus, time homogeneity for a second-order process does not ensure constancy of the mobility matrix, even when the sample is arbitrarily large. Not only is evidence from mobility matrices inappropriate in the examination of time homogeneity but the corresponding notion of transition rates p_{jk} is no longer well defined. However, if equilibrium is ever established the weights $w_{ij}(t)$ will have their constant equilibrium values $w_{ij}^* = n_{ij}^*/N_j^*$, the probability p_{jk} becomes unambiguous and variations in the elements m_{jk} will all be attributable to sampling.

Apparent violations of the law of proportional effect may be similarly misleading. If the weights $w_{ij}(t)$ at any time t depend only on the class change $j-i$, the expected value of m_{jk} will be identical for all elements for which $k-j$ is the same. However, equality between, for example, $w_{ii}(t) = n_{ii}(t)/N_i(t)$ and $w_{jj}(t) = n_{jj}(t)/N_j(t)$ cannot be guaranteed, unless the process is in equilibrium when

$$w_{j-1,j}^* = \frac{\Delta}{\Delta+\Delta_1+\Delta_2},$$
$$w_{jj}^* = \frac{\Delta_2}{\Delta+\Delta_1+\Delta_2},$$
$$w_{j+1,j}^* = \frac{\Delta_1}{\Delta+\Delta_1+\Delta_2} \quad (j > 0).$$

Apart from this special case elements on the same diagonal of M may differ for reasons other than sampling variation, even though the law of proportional effect (A 4′) is in operation.

IV. SOME EMPIRICAL RESULTS

In this section maximum-likelihood estimates of the transition probabilities are obtained and a comparison made between the first- and second-order Markov versions of the Champernowne model. It was assumed that the time interval appropriate for the transition rates was one year, during which individuals could move up or down no more than a single income class. In the three-year period 1963–6, therefore, the cumulated change is restricted to no more than three classes in either direction, and in 1966–70 to no more than four.

To estimate the transition rates use has been made of the exact sampling distributions which these stochastic models generate. To avoid unnecessary complications it was assumed that the minimum income class was never occupied. For the first-order Markov version, characterised by two parameters q_{-1} and q_1, the probability of a change of j classes in a period of n years $r_j^{(n)}$ is then given by the coefficient of x^j in the expansion of $(q_{-1}/x+1-q_{-1}-q_1+q_1x)^n$. Since the first-order Markov assumption here ensures that class changes within disjoint time periods are independent, the probability ρ_{ij} of changing i classes in 1963–6 and j classes in 1966–70 is

$$\rho_{ij} = r_i^{(3)} r_j^{(4)}$$

with

$$\sum_{i=-3}^{3} \sum_{j=-4}^{4} \rho_{ij} = 1.$$

The observed class changes within the two periods are then arranged in a similar way so that t_{ij} is the number of individuals who experienced a change of i classes in 1963–6 and j classes in 1966–70. With a class length of 0·1 the sample of 800 was allocated as follows[1]

		Class change 1966–70								
		−4	−3	−2	−1	0	1	2	3	4
Class change 1963–6	−3	.	.	.	.	3	1	.	1	.
	−2	.	.	.	3	6	4	1	1	.
	−1	.	.	1	22	86	47	7	1	.
	0	.	1	12	92	202	103	15	1	1
	1	1	.	11	54	67	31	5	.	.
	2	.	.	7	2	7	.	.	.	.
	3	.	.	1	3	.	.	.	.	.

The likelihood is then calculated from

$$\log_e L = \sum_i \sum_j t_{ij} \log_e \rho_{ij}$$

and using a standard computer routine the values of q_{-1} and q_1 which maximise $\log_e L$ can be found.

The first test performed was to compare the maximum-likelihood L_1 with the completely unconstrained maximum-likelihood L_0, obtained by choosing any values of ρ_{ij} over the 63 cells, subject only to the ρ_{ij} summing to 1. This maximum occurs when

$$\rho_{ij} = \frac{t_{ij}}{\sum_{i,j} t_{ij}} = \frac{t_{ij}}{800}$$

so

$$\log_e L_0 = \sum_i \sum_j t_{ij} \log_e \left(\frac{t_{ij}}{800}\right).$$

Since $-2\log_e (L_1/L_0)$ is asymptotically a χ^2 statistic with 60 degrees of freedom, the value of $\log_e L_0/L_1$ can be compared with the critical values 39 and 44 corresponding to the 5 % and 1 % significance levels. For the full sample of 800,

[1] With this class length one individual experienced class changes of (−5, 3) and another (4, −1). Observed changes, such as these, outside the permissible range were assigned the probability of the nearest attainable cell.

Table 1 shows that the first-order model performs badly and would be rejected for all seven class intervals considered. A second run was performed including only those whose log incomes exceeded the mean log income in each of the three years.[1] With this restricted sample of 270 in the higher income bracket, the first-order model is rejected only for the smallest class interval 0·03.

Table 1. *Likelihood Ratios*

	Full sample (800)			Restricted sample (270)		
Class length	$\log_e L_0/L_1$	$\log_e L_0/L_2$	$\log_e L_2/L_1$	$\log_e L_0/L_1$	$\log_e L_0/L_2$	$\log_e L_2/L_1$
0·03	173	62	111	61	47	14
0·05	66	55	11	31	25	6
0·07	66	50	16	38	30	8
0·10	61	40	21	31	20	11
0·13	55	27	28	20	4	16
0·16	54	20	34	21	8	13
0·20	57	13	44	39	16	23
Critical values						
5 %	39	37	4·7	39	37	4·7
1 %	44	42	6·7	44	42	6·7

In the second-order version there are six parameters to be estimated and the computation is a little more complicated. The method used involved identifying the 3^7 possible sequences of annual class changes between 1963 and 1970, and assigning to each the appropriate probability.[2] The ρ_{ij} are obtained by summing the probabilities of those sequences with an aggregate class change i in the period 1963–6 and j in 1966–70. The values of the parameters, q_{-1}^{-1}, q_1^{-1}, q_{-1}^0, q_1^0, q_{-1}^1 and q_1^1 were then chosen to maximise the likelihood L_2. The likelihood ratios given in Table 1 show that for the full sample $\log_e L_0/L_2$ exceeds both the 5 % and 1 % critical values when the class interval is less than 0·1, but above 0·1 the second-order model is not rejected at either significance level. With the restricted sample, the second-order version is rejected only for the lowest class interval of 0·03.

Perhaps the most interesting test is a direct comparison of the first- and second-order models. Since the former is a special case of the latter, with four further restrictions on the parameters, $2 \log_e L_1/L_2$ is approximately χ^2 with four degrees of freedom. This gives critical values for $\log_e L_1/L_2$ of 4·7 (5 %) and 6·7 (1 %). Likelihood ratios computed with the full sample and restricted sample exceed both of these critical values at all class sizes, except for the one case in which it exceeds only the 5 % level. The second-order process therefore appears to be a significant improvement over the first-order model regardless of the class interval chosen.

The poor performance of the first-order version is attributable to the inability of the model to generate large enough values for the probabilities ρ_{ij} when i and j

[1] It was intended that this subsample would correspond more closely to Champernowne's postulate that the law of proportional effect applies only to the incomes of the rich: see p. 568, fn. 1.

[2] Since the initial probability depends on the unknown income class for 1962, the arbitrary assumption of no class change in the year 1962–3 was made.

have opposite signs. To see this more clearly the maximum likelihood values of ρ_{ij} corresponding to the 0·1 class length have been multiplied by 800 to give the predicted values comparable to the matrix $[t_{ij}]$ given earlier in this section.

First order

i \ j	−4	−3	−2	−1	0	1	2	3	4
−3	.	.	.	.	.	.	.	.	.
−2	.	.	1	5	10	5	1	.	.
−1	.	1	7	35	75	36	7	1	.
0	.	2	18	93	200	96	19	2	.
1	.	1	7	36	76	36	7	1	.
2	.	.	1	5	10	5	1	.	.
3	.	.	.	.	.	.	.	.	.

Second order

i \ j	−4	−3	−2	−1	0	1	2	3	4
−3	.	.	.	.	1	.	.	.	.
−2	.	.	.	3	8	4	1	.	.
−1	.	.	5	27	75	37	10	1	.
0	.	1	17	89	210	92	19	2	.
1	.	1	14	50	78	27	5	.	.
2	.	.	2	6	9	3	1	.	.
3	.	.	.	1	1	.	.	.	.

A comparison of the aggregate numbers in the corner groups of cells confirms that the second-order process provides more accurate predictions.

			Predicted	
i	j	Observed	First order	Second order
−	−	26	49	35
−	+	63	50	53
+	−	79	50	74
+	+	36	50	36

With regard to the most appropriate choice of class length, the evidence does not conclusively point to any particular value. Both versions appear to prefer a class length of 0·13 or 0·16, for which the ratio of highest to lowest incomes in any class would be about 1·4. On the other hand, the first-order model performs *relatively* better for a class length of 0·05, for which the highest–lowest income ratio is 1·12. It should be remembered, however, that the single class jump restriction may tend to bias these results against smaller class intervals, and could account for the consistently poor performance of the smallest class length.

Finally, some mention should be made of the estimated transition probabilities. Table 2 gives the values computed to three decimal places when the class lengths are 0·05 and 0·1. For the smaller interval the second-order transition rates do not appear to vary systematically with the past change in class. This is not true, however, for the larger interval when there is a significantly lower probability of successive movements in the same direction. In fact, for both the full and restricted samples, the probability of a positive class change in one period is inversely related to the past transition and vice versa. The results therefore correspond to the expected variation predicted in section II. It is also interesting to note that the Champernowne stability assumption $q_{-1} > q_1$ was violated for all seven class intervals and both samples, apart from the single exception (full sample – 0·05 class length) given in the table. This occurs even though the main inflation and income growth effects were eliminated in the construction of the income classes. The transition rates for the second-order process satisfy the modified stability condition $\Delta > \Delta_1$ in six of the full sample cases (0·07 being the exception) but again fail with the restricted sample for all seven class lengths.

Table 2. *Estimated Transition Rates*

			Full sample			Restricted sample		
			Future transition u			Future transition u		
			−1	0	1	−1	0	1
Class length 0·05	First order	q_u	0·317	0·367	0·316	0·206	0·509	0·285
	Second order	q_u^{-1}	0·323	0·384	0·293	0·232	0·329	0·439
		q_u^0	0·240	0·458	0·302	0·190	0·557	0·253
		q_u^1	0·436	0·209	0·355	0·343	0·305	0·352
Class length 0·1	First order	q_u	0·106	0·785	0·109	0·080	0·796	0·124
	Second order	q_u^{-1}	0·096	0·399	0·505	0·064	0·288	0·648
		q_u^0	0·129	0·698	0·173	0·117	0·689	0·194
		q_u^1	0·637	0·263	0·100	0·535	0·427	0·038

Certain qualifications to these results should be made. The restriction to a single class movement in either direction within a year produces a rather simplified Markov model and could be relaxed with little difficulty. However, the parameters of the corresponding second-order process increase in number very rapidly, and estimation will become rather time consuming. If suitable data are available, it would also be useful to examine whether a time period other than a year was more appropriate for the second-order process and the degree to which the transition rates varied with age and time.

V. CONCLUSIONS

Evidence derived from observed transition matrices for a sample of male employees suggests that the process governing income mobility is not first-order Markov. Theoretical considerations reinforce this view. If the process is assumed instead to be second-order Markov, the framework proposed by Champernowne (1953) can be suitably generalised, and the form of the equilibrium distribution remains virtually identical. The modified version was compared with the original by estimating the transition probabilities for each and calculating the corresponding likelihood ratios. Tests reveal that the second-order variant is a significant improvement; and when the transition rates deviate systematically from the requirements of a first-order process, they do so in a way consistent with theoretical expectations.

A. F. SHORROCKS

London School of Economics

Date of receipt of final typescript: November 1975

References

Bartholomew, D. J. (1973). *Stochastic Models for Social Processes*, 2nd ed. London: Wiley.
Champernowne, D. G. (1953). "A Model of Income Distribution." Economic Journal (June). Reprinted in Champernowne (1973).
—— (1973). *The Distribution of Income between Persons*. Cambridge University Press.
Cox, D. R. and Miller, H. D. (1965). *The Theory of Stochastic Processes*. London: Methuen.
Creedy, J. (1974). "Income Changes over the Life-Cycle." *Oxford Economic Papers*, vol. 26 (November).
Friedman, M. (1957). *A Theory of the Consumption Function*. Princeton University Press.
Hart, P. E. (1973). "The Comparative Statics and Dynamics of Income Distributions." *University of Reading Discussion Papers*, series A, no. 49.
Ijiri, Y. and Simon, H. A. (1964). "Business Firm Growth and Size." *American Economic Review* (March).
—— —— (1967). "A Model of Business Firm Growth." *Econometrica*, vol. 35 (April).
Lydall, H. F. (1968). *The Structure of Earnings*. Oxford University Press.
—— (1974). Review of Champernowne (1973). *Journal of Economic Literature*, vol. 12 (September).
McCall, J. J. (1971). "A Markovian Model of Income Dynamics." *Journal of the American Statistical Association*, vol. 66 (September).
Thatcher, A. R. (1976). "The New Earnings Survey and the Distribution of Earnings." In *The Personal Distribution of Income* (ed. A. B. Atkinson). London: Allen and Unwin.
Vandome, P. (1958). "Aspects of the Dynamics of Consumer Behaviour." *Bulletin of the Oxford University Institute of Statistics*, vol. 20 (February).