Environment and Planning A, 1983, volume 15, pages 343-355

Analysis of gross migration profiles in England and Wales: some developments in classification

I Bracken
Department of Town Planning, University of Wales Institute of Science and Technology, Cardiff CF1 3NU, Wales
J Bates
John Bates Services, The Old Coach House, Wyke, Gillingham, Dorset, England
Received 9 July 1981, in revised form 4 March 1982

Abstract. In this paper we describe the development and application of two approaches to the analysis of age and sex specific interarea migration. The data relate to local authority areas in England and Wales. These approaches are developments of established migration profile-fitting and cluster analytic methods. The paper shows how these methods have been used to develop complementary classifications from which potentially valuable criteria for migration analysis can emerge. The work described is part of a wider project concerned with the development of methods to improve the official estimation of internal migration in England in the context of local area population forecasting.

1 Introduction

The analysis of interregional population migration has a long history among social scientists (Ravenstein, 1885; 1889; Lively and Taeuber, 1939; Lowry, 1966; Keyfitz, 1972; Stillwell, 1975), and its importance to the broader analysis and explanation of urban and regional systems is widely recognised (Isard, 1969; Rogers, 1968; Rees and Wilson, 1977). Yet the interactions of migration with geographical, economic, and social criteria are of immense complexity that have so far defied the development of general and widely applicable laws and explanations. Systematic knowledge, then, is partial and this is further reflected in the diversity of the various methodological schools through which migration analysis is pursued. These range, for example, at the aggregate level from broad demographic description (Blanchard et al, 1977); directional analysis (Cordey-Hayes and Gleave, 1974); geographic description (Champion, 1976); spatial-interaction modelling (Stillwell, 1977); demographic accounts modelling (Stone, 1970; Rees, 1973; Rees and Wilson, 1977); regression with economic variables (Greenwood, 1973; 1975), or housing and household variables (Hollis et al, 1976); to more complex multiregional analyses and 'demo-economic' modelling represented in the work of Rogers (1975) and Gordon and Ledent (1980), respectively.

In the light of these various approaches, three areas stand out as major contributors to the uncertainty so far inherent in migration analysis. First, and the most widely reported, is that of the deficiency of the data, a feature which becomes highly problematical as the scale of the analysis decreases. Important themes of research in the field have included attempts to maximize the quality of information from official data sources such as the Census (Rees, 1977) and Health Service Records. Related work has attempted to cross-check their validity (Rees and Rees, 1977; Ogilvy, 1979). Further, there has been work to infer greater degrees of disaggregation by entropy maximizing methods (Chilton and Poet, 1973; Willekens et al, 1981). The second area, given the spatial pattern of migration and its well-established 'life-cycle' characteristics (that is, variations in the propensity to migrate relative to age and sex), is to explain migration in terms of these features and thereby deduce general theoretical statements. There are, however, some difficult definitional and methodological problems inherent in this approach. The third area of uncertainty relates to the interaction of policy with migration behaviour, particularly economic and settlement policies. The analysis of these relationships is an important key to any attempt to use migration analysis in a forecasting mode where the future is believed to be in some ways influenced by policy. Again, the lack of well-developed theory poses real problems for planners and policymakers (Rees, 1979).

In this paper we report a specific piece of work carried out to classify local areas according to the shape of their migration profiles—that is, the variation in migration propensity according to age. This work has been done in the context of a much larger programme concerned with improving the official projections of subnational migration, to be used for local area population projections. Although it is not possible to describe the main project here, some details will be useful in setting the current paper in context.

The project was commissioned by the Department of the Environment, and one of its principal aims was to provide a model which would produce migration flows for 116 Local Authorities in England and Wales (that is, Metropolitan Counties and Districts, Shire Counties, and London Boroughs) disaggregated by sex and single years of age. Although data are available at this level of detail from the 1971 Census, it seemed to us impractical to attempt to model individual cells directly, and we therefore sought acceptable simplifications whereby the essence of the patterns could be extracted and used in the model. In any case, the Census migration data are only available on a 10% sample basis, so that the raw data are subject to sample variation, as well as possible random temporal effects due to the particular choice of period to which the Census migration relates. By extracting the general pattern from the data, we can hope to treat the remaining detail as essentially random.

The variation in migration propensity by age is on the whole remarkably constant from one area to another, and a considerable simplification can be made in the model if we can group areas according to the shape of their profiles. The official projection model has required migration data for each area and each year of age. We have been able to reduce the model requirements to twelve representative shapes for each sex, and within each shape have used a mathematical function to smooth out random fluctuations in the data. This permits considerable reductions in the storage required for the operation of the model. We confine the explanation of this work mainly to out-migration and for males only, though similar analyses were made both in respect of females and for in-migration.

2 Classification of migration

In producing a classification of local areas, we drew upon earlier work in the development of hierarchical classifying methods and attempts to capture the underlying patterns of aggregate migration behaviour. This work (Masser and Brown, 1975; Slater, 1976; Bracken, 1976; Rogers et al, 1978; Masser and Scheurwater, 1980) has related both to analysis of spatial variations in migration (whether as bidirectional or net flows) *and* life-cycle characteristics as represented by the age and sex structure of the migrant population. To calculate appropriate migration rates, estimates have to be made of the 'base' population of originating areas and this introduces a further uncertainty for years other than those in which a census is conducted.

Given, first, the very detailed level at which the model had to operate; second, the purpose of the model to estimate *net* migration flows (that is, the residual of the gross flows and therefore a quantity that is highly sensitive to any errors in the gross flow estimation); and third, the 10% sample nature of the Census data, it was vital to establish the maximum confidence in the classification. Two approaches were therefore employed. The first, was by generating migration age profiles after Rogers et al (1978). This is widely accepted as a method having a high degree of analytical power given the apparent systematic relationship between propensity to migrate and the age and sex of the migrant. Broadly, this approach involved fitting a mathematical function to each of the local areas, by selecting appropriate values of the coefficients according to some statistical criterion. Full details of the estimation procedure have been reported in Bates and Bracken (1982). Given an estimate of the accuracy with which these coefficients were estimated, it was possible to set up a statistical measure of 'similarity' between all areas, on a pairwise basis. The procedure adopted enables us to make an allowance for the random variation in the data.

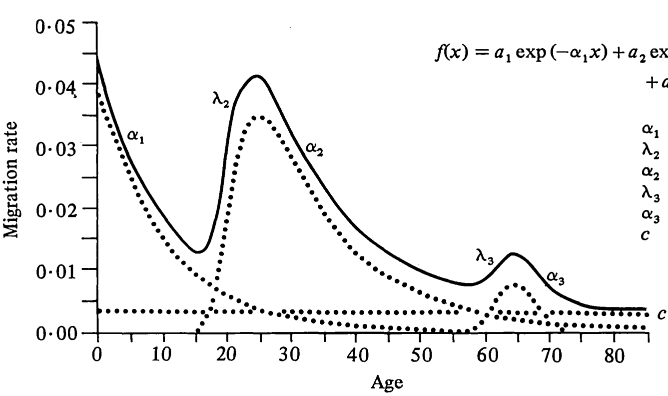
Nonetheless, we felt that there was some danger that by essentially imposing a smoothing process on the data, we might erase some features which were in fact significant for the classification of areas. In addition, the profile fitting approach deals with all ages, and we were concerned that the grouping might be critically affected by the age range considered (though in fact this turned out not to be the case). Finally, given the potential problems in the analysis, we thought it prudent to develop an alternative method, and to see to what extent the two approaches confirmed, complemented, or contradicted each other. Thus we developed the earlier work of Bracken (1976) which involved the use of a cluster analysis algorithm. For this the independent variables used to 'dimension' the clusters were the migration propensities at individual years of age. The approach allowed us to experiment by selection of particular ages to determine the sensitivity of the clustering.

We report on each of these approaches in turn.

3 The curve fitting approach

In an earlier paper in this journal (Bates and Bracken, 1982) we have reported in full our development of the work of Rogers et al (1978) and Castro and Rogers (1979) in applying model schedules to the migration profile. This paper has also made clear the assumptions made about the error structure of the data. In brief, our approach involved the use of the method of maximum likelihood estimation (ML) which, we argue, has some definite advantages. In particular, it produces estimates which are asymptotically unbiased, and which for larger samples are approximately normally distributed with a variance-covariance matrix whose calculation can be shown to be useful in the numerical procedure for estimating the coefficients for the model schedule.

The characteristic shape of the migration age profile, and the mathematical function proposed by Rogers et al (1978) to approximate it, are given in figure 1. The experience of workers in this field has been that the shape of this profile shows



$$f(x) = a_1 \exp(-\alpha_1 x) + a_2 \exp\{-\alpha_2(x-\mu_2) - \exp[-\lambda_2(x-\mu_2)]\} + a_3 \exp\{-\alpha_3(x-\mu_3) - \exp[-\lambda_3(x-\mu_3)]\} + c$$

α_1 rate of descent of pre-labour force curve
λ_2 rate of ascent of labour force curve
α_2 rate of descent of labour force curve
λ_3 rate of ascent of post-labour force curve
α_3 rate of descent of post-labour force curve
c constant

Figure 1. Characteristic migration profile by age and the function used to approximate it (source: Rogers et al, 1978, page 492).

remarkable consistency from one area to another, although the level of migration activity may vary considerably according to the type of area; for instance, urban areas tend to have a higher turnover of migrants, other things being equal, than rural areas. This suggests the need to standardize the profile in some way, to extract the scale factor. One way of doing this, proposed by Rogers et al (1978) is to scale the profile function so that the area under the curve is unity, and to achieve this the raw migration propensities are divided by the so-called gross migraproduction rate (R^{GM}). The process can be summarized as follows. For any area, let the population of age x be represented by $N(x)$, and let the number of those who migrate be $M(x)$. The propensity to migrate can be denoted as $m(x)$, where $m(x) = M(x)/N(x)$. The approach is to express the observed $m(x)$ by the product of a scaling factor γ and the migration age profile function $f(x)$. An acceptable approximation to the value of γ is given by $\gamma = \sum_{x} m(x)$. Thus defined, γ is equivalent to R^{GM} (Rogers et al, 1978). It follows that variations in R^{GM} from one area to another will relate essentially to differing *levels* of mobility.

Although, as above, R^{GM} is obtained by summing the values for $m(x)$ over all possible ages, in practice small sample sizes in the data create problems. In particular the observed values for $m(x)$ can become erratic, and we considered it advisable to restrict the calculation to the age range 1–70 years. Having standardized the observed values of the migration propensity $m(x)$ by R^{GM}, we proceeded to calibrate the function $f(x)$ shown in figure 1 for each local authority area using the methods described in our earlier paper (Bates and Bracken, 1982). In this way, we obtained a set of coefficients particular to each area. In addition, the estimation process produced an estimate of the variance–covariance matrix of the parameters, thus allowing us to make estimates of their accuracy.

As far as this paper is concerned, the principal interest lies in assessing the similarity or dissimilarity of the profiles represented by these parameters. Our approach was to test for similarity between all profiles using a stepwise test, over all pairs of zones. Thus if zone 1 has a parameter vector β_1 and zone 2 has a vector β_2 [these vectors relate to the parameters which define the profile $f(x)$], and the associated variance–covariance matrices are $\mathbf{V}_1$ and $\mathbf{V}_2$, then the scalar quantity K,

$$K = \frac{1}{r}(\beta_1 - \beta_2)^{\mathrm{T}}(\mathbf{V}_1 - \mathbf{V}_2)^{-1}(\beta_1 - \beta_2)\,,$$

will have a F-distribution with n and r degrees of freedom. Here, n is the sample size and r is the number of parameters in the vector. When the sample is large, as in our analysis, the quantity Kr has approximately the chi-squared distribution with r degrees of freedom. Thus if Kr has a value greater than the tabulated value of χ^2, for the appropriate percentage point, it can be concluded that zones 1 and 2 do *not* have the same profile. By a complete pairwise comparison we have been able to group all the profiles in our study.

Figure 2 illustrates how the classification was compiled for just two types. On a literal interpretation of the statistic, 2 represents significant difference between the parameter vectors at the 95% level; 1 represents significant difference at the 90% but not at the 95% level. Otherwise, there is no significant difference (at the 90% level). From the complete set of data from the pairwise test, and by using a three-way criterion, submatrices as in figure 2 were compiled. The aim was to create combinations of areas so that a *maximum* of similarity scores, and a *minimum* of dissimilarity scores appear in each group submatrix. Areas which readily combined were accumulated first, followed by single area addition to the most appropriate group. Some areas were difficult to allocate, and ultimately we relied on subjective judgement, particularly with areas where the sample size was relatively small, leading to a less accurate estimate of the coefficients. When judgement was required, it was based on a consideration of the geographical location, and to a lesser extent, the level of urbanization.

Although we would not claim that the typology is unique, it does account for the most important difference in the shapes of the migration profile. In table 1,

		59	60	61	65	66	67	70	71	73	74	77	78	79	80	83	85	87	62	64
Barnet	59	x																		
Bexley	60	-	x																	
Brent	61	-	-	x																
Ealing	65	-	-	-	x															
Enfield	66	-	-	-	-	x														
Greenwich	67	-	-	-	-	-	x													
Haringay	70	-	2	-	-	-	-	x												
Harrow	71	-	-	-	-	-	-	-	x											
Hillingdon	73	-	-	-	-	-	-	-	-	x										
Hounslow	74	-	-	-	-	-	-	-	-	-	x									
Kingston	77	-	-	-	-	-	-	-	-	-	-	x								
Lambeth	78	-	-	-	-	-	-	-	-	-	-	-	x							
Lewisham	79	-	-	-	-	-	-	-	-	-	-	-	-	x						
Merton	80	-	-	-	-	-	-	-	-	-	-	-	-	-	x					
Richmond	83	-	-	-	-	-	-	-	-	-	-	-	-	-	-	x				
Sutton	85	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	x			
Waltham Forest	87	-	-	-	-	-	-	1	-	-	-	-	2	-	-	-	-	x		
Bromley	62	-	1	-	2	2	2	1	-	2	2	-	2	1	1	2	-	2	x	
Croydon	64	-	-	-	-	2	1	1	-	1	1	-	2	-	-	-	-	1	1	x

		96	98	102	112	113	115	119	20	54	56
Kent	96	x									
West Sussex	98	-	x								
East Sussex	102	-	-	x							
Dorset	112	-	-	-	x						
Glous	113	-	2	1	-	x					
Wilts	115	1	-	-	2	2	x				
Gwynedd	119	-	-	-	-	-	2	x			
North Yorkshire	20	2	-	-	-	2	-	2	x		
Norfolk	55	-	-	-	1	2	-	2	-	x	
Suffolk	56	-	2	2	-	1	2	-	-	2	x

- no significant difference at the 90% level
1 significant difference at the 90% but not at the 95% level
2 significant difference at the 95% level

Figure 2. Selected 'T^2' test matrices for coefficients of males out-migration profiles (for profile groups 2 and 3 of table 3).

Table 1. Estimated coefficients for the twelve out-migration groups given in table 3.

Group	Coefficients										
	c	a_1	α_1	a_2	α_2	μ_2	λ_2	a_3	α_3	μ_3	λ_3
1	0·0066	0·0195	0·1243	0·0479	0·1106	19·4	0·4476	-	-	-	-
2	0·0070	0·0228	0·2099	0·0645	0·1406	21·2	0·3630	0·00081	3·6154	66·9	0·8469
3	0·0063	0·0197	0·1230	0·0539	0·1165	20·1	0·3683	-	-	-	-
4	0·0065	0·0191	0·1256	0·0496	0·1179	20·4	0·4458	0·0013	1·8963	67·1	0·6588
5	0·0059	0·0206	0·1158	0·0405	0·0892	20·4	0·4280	0·0095	0·7604	64·2	11·592
6	0·0064	0·0200	0·1354	0·0562	0·1221	20·5	0·3734	0·000004	8·7470	66·5	1·2410
7	0·0069	0·0215	0·1729	0·0648	0·1460	19·9	0·4712	0·000005	4·4366	67·9	0·6405
8	0·0067	0·0148	0·0902	0·0374	0·0898	18·8	0·5042	0·000001	12·1985	66·0	1·6084
9	0·0061	0·0239	0·1755	0·0675	0·1353	20·4	0·3925	0·00037	1·7496	69·1	0·4468
10	0·0060	0·0250	0·1640	0·0599	0·1268	20·4	0·5893	0·0111	1·6639	65·3	1·1346
11	0·0067	0·0160	0·0704	0·0534	0·1351	21·0	0·5269	0·000002	7·7458	66·0	1·0210
12	0·0064	0·0163	0·0902	0·0468	0·1122	20·1	0·4922	-	-	-	-

we present the parameters for the twelve distinguishable types which emerged, and in figure 3 we show a number of selected profiles to illustrate the range of differences. It should be noted that the retirement effect (represented by the last four parameters) was not significant in all the area types, and even in those which showed evidence of a retirement curve, the parameters were relatively unstable. Figure 4 shows the geographical distribution of the profile types. The membership characteristics are discussed later in this paper. Finally, it should be noted that having grouped areas according to this procedure, data for the areas in any particular group were combined and the model recalibrated on the grouped data. A pairwise test for similarity between the twelve groups showed that the parameters were in all cases significantly different.

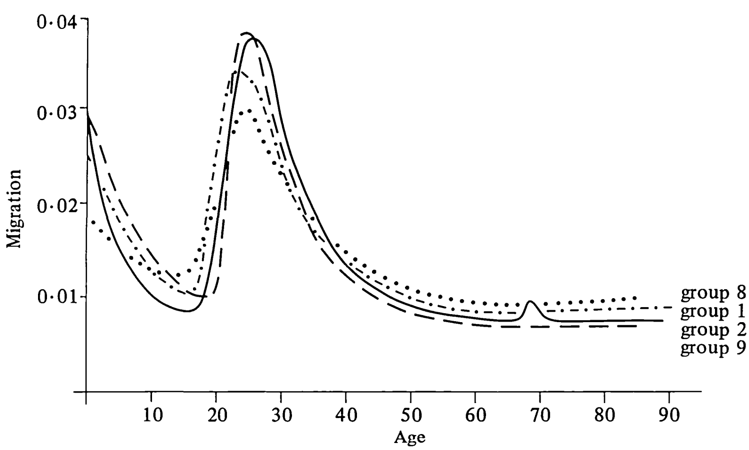


Figure 3. Out-migration profiles for groups 1, 2, 8, and 9 given in table 3.

4 The clustering approach

The general objective of clustering (Ward, 1963) is to rearrange the entities (areas in our case) into groups so that as the number of groups is successively decreased by one, the loss of information concomitant with each cluster join is kept as small as possible. The function selected to be minimized is described below. The technique used, then, is one of optimization and partition, in which mutually exclusive groups of areas are formed by partitioning. It is important, given first, the close similarity of many of the area profiles, and second, the large number of dimensions used in the clustering, to allow for any poor initial allocation of areas to the groups to be corrected by reassignment as the clustering proceeds. By choosing a large number of initial clusters there is a greater probability of the final result being independent of the initial position.

The algorithm used defines distance in terms of ordinary squared euclidean space. Zones are allocated to groups simply by approximating a local minimum for the sum of squared distances between the zone observations and their group cluster centres. An objective assessment of clustering efficiency is obtained by the incorporation into the program of an F-statistic (Beale, 1969) defined by

$$F(c_1, c_2) = \frac{R_{c_1} - R_{c_2}}{R_{c_2}} \Bigg/ \left[\left(\frac{n-c_1}{n-c_2}\right)\left(\frac{c_2}{c_1}\right)^{2/p} - 1\right]$$

where n is the total number of entries (zones); p is the number of observations (age groups) in each entry; c_x is the number of clusters at a given level x; and associated with each level x is a measure of efficiency R_{c_x}, defined as $R_{c_x} = (n-c_x)S_x^2$, where S_x^2 is the mean square deviation within each cluster, summed over all clusters.

Figure 4. Spatial distribution of twelve out-migration profiles for local authority areas in England and Wales (see table 3 for a key to the areas).

The F-statistic is in fact only calculated between *successive* levels, so that $c_1 = c_2 - 1$. A significant result is indicated by an increase in the F statistic between clustering into c_2 groups and the smaller number of groups c_1. A fuller description of the application of this method to migration analysis has been given in Bracken (1976).

In all cases the variables in the cluster analysis were the area values of $m(x)$ at individual ages, standardized by R^{GM}. As noted earlier the main object of the approach was to experiment with variations in the definitions of the age bands. The profiles proved to be fairly insensitive to variations in the age bands chosen, and we discovered that stable typologies resulted even when data for as few as five single years were employed within the range 18–30 years. Given the small value of the migration flows in the age range 40–60 years, these ages were quickly omitted from the analysis.

As an example, we illustrate the results for out-migration (males) from the thirty-two London Boroughs (the City of London was excluded because of its very small population size). Table 2(a) shows that clustering the Boroughs into groups, from the fifteen-group level down, produced a significant indicator of efficient clustering at the eight-group and six-group levels. The greater the relative increase in the F-ratio from one level (n) to the next ($n-1$), the more efficient is the allocation of the Boroughs (zones) to their groups. An assessment was also made in the analysis of the efficiency of the clustering in terms of the 'separation' of the groups in their n dimensional space. Table 2(b) shows that at the eight-group level the groups are well spaced from each other. In terms of proximity, group 7 is the most closely related to the other groups, notably to groups 1, 5, 6, and 8.

The membership of the London Boroughs at the eight-group level is shown in figure 5, and figure 6 illustrates the corresponding mean profile for each group. These profiles, being standardized, represent the comparative propensity to migrate (as earlier defined) according to the mean of the areas allocated to each group. We reproduce only that part of the profile where differences are most significant. Overall, peak movement is achieved at ages 24 or 25, and then followed by steady decline. Not all groups reach a peak at this age; for example, group 2 reaches a peak at age 22 and group 4 at age 23. Given that these scores are standardized by R^{GM} (that is, the area under the curve is unity), the variations in the scores do not directly represent different amounts of migration activity per se, but rather the relative variation in the propensity to move by age. In particular, areas comprising groups 2,

Table 2. Cluster analysis of London Boroughs: out-migration profiles for males 15–34 years of age.

(a) *Clustering efficiency ratio*

Number of groups c	Beale's ratio F
15	0·035
14	0·051
13	0·061
12	0·072
11	0·087
10	0·100
9	0·110
8	0·155*
7	0·184
6	0·227*
5	0·255

* Significant at the 90% level.

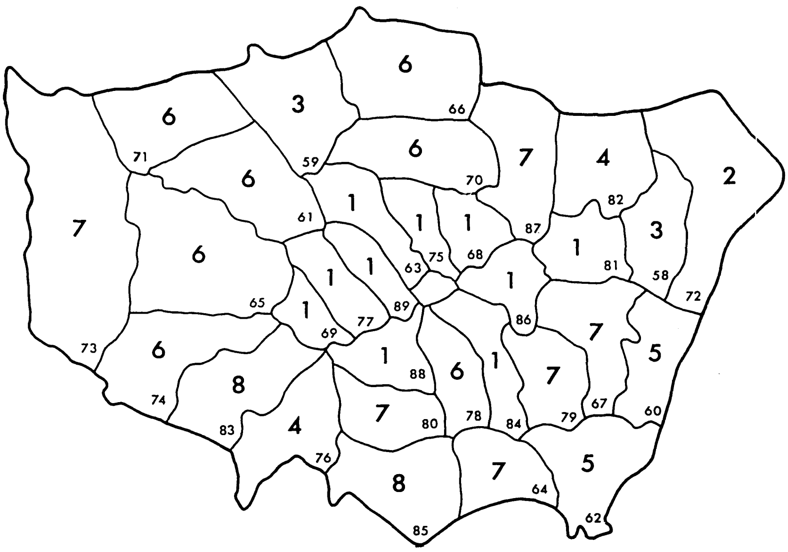
(b) *Relative distances between group centres*

	1	2	3	4	5	6	7	8
1	0·0							
2	0·053	0·0						
3	0·044	0·040	0·0					
4	0·044	0·044	0·040	0·0				
5	0·028	0·041	0·039	0·036	0·0			
6	0·024	0·044	0·034	0·040	0·028	0·0		
7	0·028	0·041	0·027	0·035	0·025	0·021	0·0	
8	0·039	0·048	0·033	0·034	0·031	0·034	0·025	0·0

total intergroup distance = 1·0

3, and 4 have an apparent concentration into the peak, whereas most other areas are notably less peaked.

A similar analysis for in-migration showed that the peak of in-migration propensity occurs earlier by approximately two years, and that for urban areas in particular, the profiles are notably more 'peaked' than for out-migration. Further, there is less *variation* in 'peakedness' over different areas for in-migration than for out-migration.



58 Barking	69 Hammersmith	79 Lewisham
59 Barnet	70 Haringay	80 Merton
60 Bexley	71 Harrow	81 Newham
61 Brent	72 Havering	82 Redbridge
62 Bromley	73 Hillingdon	83 Richmond
63 Camden	74 Hounslow	84 Southwark
64 Croydon	75 Islington	85 Sutton
65 Ealing	76 Kingston	86 Tower Hamlets
66 Enfield	77 Kensington and Chelsea	87 Waltham Forest
67 Greenwich	78 Lambeth	88 Wandsworth
68 Hackney		89 Westminster

Figure 5. Distribution of out-migration groups for males 15-34 years of age in the London Boroughs.

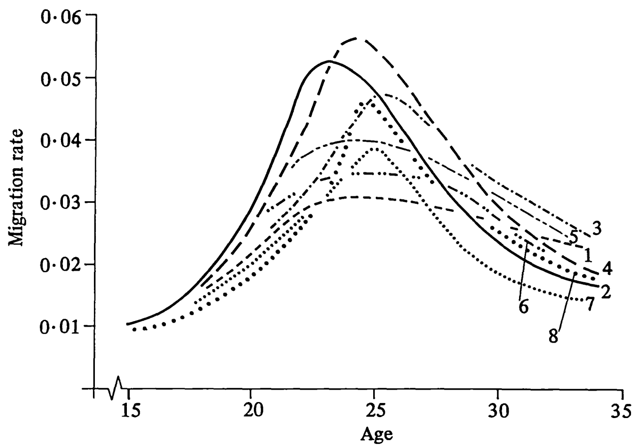


Figure 6. Out-migration profiles for males 15-34 years of age in the London Boroughs (the curves 1-8 refer to the groups 1-8 of figure 5).

In almost all cases, the female migration profiles are very similar to those for males, except that the female profile is generally in advance of the male profile by about two years. Exceptionally, the in-migration profile for females aged 17 and 18 generally rises particularly rapidly.

This method of analysis, for the Greater London area, was then applied to all counties and metropolitan districts in England and Wales. Results were obtained for twelve clusters, and the work at this level enabled a comparison to be made with the results obtained by the curve-fitting approach.

5 A cross-classification of criteria

Although both the methods which we have proposed for classifying areas according to their migration profiles involve some user judgement, they represent essentially independent methods, and it is therefore of interest to see the extent to which they are in agreement. In table 3 we set out the area classification obtained from the curve-fitting method, and against each area is shown the cluster number to which the area is assigned.

If the two methods were in complete agreement, then within any one group obtained by the curve-fitting method, all areas would belong to the same cluster. As might be expected, this is not the case, though the level of agreement is considerable, bearing in mind the sample nature of the data and the variations in the age bands employed. In all groups except one, it is possible to determine which is the predominant cluster. For example, in group 1 which contains twenty-three areas, all but four of the areas fall into cluster number 3. We therefore can clearly identify group 1 with cluster 3, etc. The exception is group 4, which is fairly evenly split between clusters 10 and 11. In both cases, there are some groups or clusters with a

Table 3. Classification of areas for out-migration for males in local authority areas in England and Wales: profile groups and cluster groups. [The numbers in square brackets under 'cluster group' give the predominant cluster group(s).]

Profile group	Zone[a]	Cluster group	Profile group	Zone[a]	Cluster group
1	6 Cleveland	3	2	59 Barnet LB	11
	8 Durham	3		60 Bexley LB	1
	9 Northumberland	3		61 Brent LB	9
	16 Kirklees MD	8		62 Bromley LB	1
	18 Wakefield MD	3		64 Croydon LB	9
	19 Humberside	3		65 Ealing LB	9
	37 Lancashire	3		66 Enfield LB	9
	40 Lincolnshire	3		67 Greenwich LB	6
	41 Northamptonshire	3		70 Haringey LB	9
	42 Nottinghamshire	3		71 Harrow LB	9 [6,9]
	45 Dudley MD	8		73 Hillingdon LB	6
	50 Hereford and Worcester	3 [3]		74 Hounslow LB	9
	51 Salop	3		77 Kingston LB	11
	52 Staffordshire	3		78 Lambeth LB	9
	53 Warwickshire	3		79 Lewisham LB	6
	92 Buckinghamshire	3		80 Merton LB	6
	93 Essex	3		83 Richmond LB	6
	105 Isle of Wight	1		85 Sutton LB	6
	110 Cornwall and Scilly	3		87 Waltham Forest LB	11
	114 Somerset	3			
	116 Clwyd	3			
	118 Gwent	3			
	121 Powys	2			

large number of members, and some residual groups containing those areas which are most difficult to classify. The broad measure of agreement between the two methods, however, suggests that certain salient features common to both approaches have potential value as a basis for classification.

Investigating the classification obtained by the curve-fitting method revealed three criterial characteristics: the presence or absence of retirement migration, the age at which the peak of migration activity occurs, and the extent to which the migration activity is concentrated into the years immediately around the peak. The first of these factors is not explicitly taken into account in the clustering method given the ranges over which the age bands for that analysis were defined, and in any event

Table 3 (continued)

Profile group	Zone[a]	Cluster group	Profile group	Zone[a]	Cluster group
3	20 North Yorkshire	1	5	86 Tower Hamlets LB	5
	55 Norfolk	1		88 Wandsworth LB	5
	56 Suffolk	3		89 Westminster LB	5
	96 Kent	1			
	98 West Sussex	1 [1]	6	39 Leicestershire	1
	102 East Sussex	1		90 Bedfordshire	1
	112 Dorset	9		91 Berkshire	1
	113 Gloucestershire	1		94 Hampshire	1 [1]
	115 Wiltshire	10		95 Hertfordshire	1
	119 Gwynedd	1		97 Surrey	1
				107 Oxfordshire	1
4	1 Gateshead MD	11			
	2 Newcastle upon Tyne MD	11	7	7 Cumbria	10
	3 North Tyneside MD	11		10 Barnsley MD	10 [10]
	4 South Tyneside MD	10		13 Sheffield MD	10
	5 Sunderland MD	10		38 Derbyshire	3
	11 Doncaster MD	10			
	15 Calderdale MD	12	8	14 Bradford MD	8
	17 Leeds MD	1		21 Bolton MD	3
	26 Salford MD	3		24 Oldham MD	8 [8]
	28 Tameside MD	3		25 Rochdale MD	8
	29 Trafford MD	11		27 Stockport MD	3
	30 Wigan MD	10 [10, 11]			
	31 Knowsley MD	10	9	49 Wolverhampton MD	4
	32 Liverpool MD	11		58 Barking LB	7
	33 St Helens MD	10		109 Avon	3
	34 Sefton MD	3		117 Dyfed	10 [4]
	35 Wirral MD	11		120 Mid Glamorgan	4
	36 Cheshire	3		122 South Glamorgan	3
	43 Birmingham MD	11		123 West Glamorgan	4
	44 Coventry MD	3			
	46 Sandwell MD	11	10	72 Havering LB	10
	47 Solihull MD	11		81 Newham LB	11
	48 Walsall MD	3		82 Redbridge LB	6
5	63 Camden LB	5	11	54 Cambridgeshire	1 [1]
	68 Hackney LB	5		111 Devon	1
	69 Hammersmith LB	5			
	75 Islington LB	5	12	12 Rotherham MD	12
	76 Kensington and Chelsea LB	5 [5]		22 Bury MD	12 [12]
	84 Southwark LB	5		23 Manchester MD	11

[a] LB– London Boroughs; MD—Metropolitan Districts.

retirement migration is not only readily identified in the profiles, but is almost wholly revealed only at age 65 at our level of data. The presence of retirement migration as a classifying factor is thus readily determined. The other two factors are dominant, and their interrelationship can be summarized in a two-way cross-classification as suggested in figure 7.

By reference to figure 3 for the purposes of illustration, it can be seen that group 1 was an 'early' peak, with an average amount of concentration into the peak. Group 8 has extremely low concentration, in contrast to groups 2 and 9 which exhibit high degrees of concentration. So far, our investigations have not revealed a pattern which combines a low level of concentration with anything other than a 'normal' peak age. High concentration into the peak is typical of urban areas both for out-migration and in-migration, and areas such as the 'inner' London Boroughs display an early peak. In many cases a high concentration is also found in outer-suburban areas though this is then normally found in conjunction with a normal peak age. Areas of low concentration can occur in both urban and rural locations.

However, it is not the purpose of this paper, to attempt an extensive geographical description. The immediate aim has been to show how criteria for classification have emerged from our work and that this has potential for such further analysis in that considerable confidence can be placed upon the parameter specification for the migration profiles. As with all classification exercises, the results require careful interpretation: yet it is clear that two features stand out as a means of distinguishing aggregate migration behaviour, namely, variations in the *level* of migration activity when standardized by the gross migraproduction rate, R^{GM}, and, secondly, the variations in the peak age for that activity.

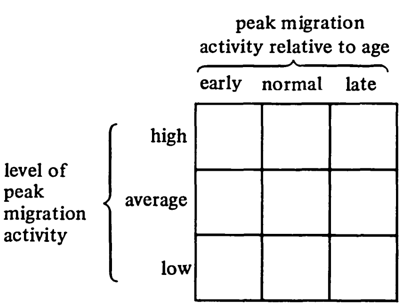


Figure 7. A cross-classification of migration criteria.

Acknowledgement. The publication of this paper has been authorised by the Department of the Environment and has been published by kind permission of the Controller of Her Majesty's Stationery Office. Any opinions and conclusions expressed in the paper should not be taken as reflecting the policy of the Department of the Environment. The authors wish to record their thanks to Richard Stanley and John Spiers of Martin and Voorhees Associates for their help in developing the classifications.

References

Bates J, Bracken I, 1982 "Estimation of migration profiles in England and Wales" *Environment and Planning A* **14**(7) 889-900

Beale E M L, 1969 *Cluster Analysis* Scientific Control Systems, London

Blanchard R D, Bunker J B, Wachs M, 1977 "Distinguishing aging, period, and cohort effects in longditudinal studies of elderly populations" *Socio-Economic Planning Sciences* **11** 137-146

Bracken I, 1976 "Estimation and cluster analysis of net migration age profiles for urban areas in England and Wales" *Regional Studies* **10** 53-69

Castro L J, Rogers A, 1979 "Migration age patterns: measurement and analysis" WP-79-16, International Institute for Applied Systems Analysis, Laxenburg, Austria

Champion A G, 1976 "Evolving patterns of population distribution in England and Wales, 1951–71" *Transactions of the Institute of British Geographers* new series **1** 401–420
Chilton R, Poet R R W, 1973 "An entropy maximizing approach to the recovery of detailed migration patterns from aggregate census data" *Environment and Planning* **5** 135–146
Cordey-Hayes M, Gleave D, 1974 "Migration movements and the differential growth of city regions in England and Wales" *Papers of the Regional Science Association* **33** 99–123
Greenwood M J, 1973 "Urban economic growth and migration: their interaction" *Environment and Planning* **5** 91–112
Greenwood M J, 1975 "A simultaneous-equations model of urban growth and migration" *Journal of the American Statistical Association* **70** 797–810
Gordon P, Ledent J, 1980 "Modelling the dynamics of a system of metropolitan areas: a demo-economic approach" *Environment and Planning A* **12** 125–133
Hollis J, Henderson A, Congdon P, 1976 "Population and household projections for London, 1976" research monograph, RM 506, Greater London Council, London
Isard W, 1969 *Methods of Regional Analysis. An Introduction to Regional Science* (MIT Press, Cambridge, MA)
Keyfitz N, 1972 "On future population" *Journal of the American Statistical Association* **67** 347–363
Lively C E, Taeuber C, 1939 *Rural Migration in the United States* (US Government Printing Office, Washington, DC)
Lowry I S, 1966 *Migration and Metropolitan Growth: Two Analytical Models* (Chandler, San Francisco, CA)
Masser I, Brown P J B, 1975 "Hierarchical aggregation procedures for interaction data" *Environment and Planning A* **7** 509–523
Masser I, Scheurwater J, 1980 "Functional regionalisation of spatial interaction data: an evaluation of some suggested strategies" *Environment and Planning A* **12** 1357–1382
Ogilvy A A, 1979 "Inter-regional migration since 1971: an appraisal of data from two sources—Labour Force Surveys and the National Health Service Central Register" occasional paper, Office of Population Censuses and Surveys, London
Ravenstein E G, 1885 "The laws of migration" *Journal of the Royal Statistical Society* **48** 167–227
Ravenstein E G, 1889 "The laws of migration" *Journal of the Royal Statistical Society* **61** 240–301
Rees G M, Rees T L, 1977 "Alternatives to the Census" *Town Planning Review* **48** 123–140
Rees P H, 1973 "A revised notation for spatial demographic accounts and models" *Environment and Planning* **5** 147–155
Rees P H, 1977 "The measurement of migration from census data and other sources" *Environment and Planning A* **9** 247–272
Rees P H, 1979 "Migration and settlement: 1, United Kingdom" RR-79-3, International Institute for Applied Systems Analysis, Laxenburg, Austria
Rees P R, Wilson A G, 1977 *Spatial Population Analysis* (Edward Arnold, Maidenhead, Berks)
Rogers A, 1968 *Matrix Analysis of Inter-Regional Population Growth and Distribution* (University of California Press, Berkeley, CA)
Rogers A, 1975 *Introduction to Multi-regional Mathematical Demography* (John Wiley, New York)
Rogers A, Racquillet R, Castro L J, 1978 "Model migration schedules and their applications" *Environment and Planning A* **10** 475–502
Slater P B, 1976 "Hierarchical internal migration regions in France" *IEEE Transactions on Systems, Man and Cybernetics* **6** 321–324
Stillwell J C H, 1975 "Models of inter-regional migration—a review" WP-100, School of Geography, University of Leeds, Leeds, England
Stillwell J C H, 1977 "Some historical tests of spatial interaction models for inter-area migration. Part II, Inter-region migration" WP-199, School of Geography, University of Leeds, Leeds, England
Stone R, 1970 *Mathematical Models of the Economy and Other Essays* (Chapman and Hall, London)
Ward J M, 1963 "Hierarchical grouping to optimize an objective function" *Journal of the American Statistical Association* **58** 236–244
Willekens F, Por A, Raquillet R, 1981 "Entropy, multiproportional and quadratic techniques for inferring patterns of migration from aggregate data" in *Advances in Multiregional Demography* RR-81-6, Ed. A Rogers, International Institute of Applied Systems Analysis, Laxenburg, Austria, pp 83–124