WORKING PAPER 537

MIGRATION ANALYSIS BASED ON NATIONAL HEALTH SERVICE CENTRAL REGISTER DATA: TRENDS AND MODELS

John C.H. Stillwell

Paper prepared for the British Society for Population Studies Conference on 'Migration: Trends, Models, Interpretations', held at the London School of Economics on Thursday, 28 June, 1990

School of Geography
University of Leeds
Leeds LS2 9JT

June 1990

MIGRATION ANALYSIS BASED ON NATIONAL HEALTH SERVICE CENTRAL REGISTER DATA: TRENDS AND MODELS

1 Introduction

It is generally agreed that migration involves the change of an individual's place of residence. However, population analysts frequently draw a distinction between the event of migration, the move, and the person involved in the migration, the migrant. Nam et al. (1990) indicate that in the majority of countries of the world, statistical information is available for migrants and not for moves, since most countries rely on Censuses as their primary sources of data. In Great Britain, questions in the last four Censuses of Population have provided transition data on internal migrants between regions, counties, districts, wards and enumeration districts. However, researchers and planners in Britain also benefit increasingly from the availability of data from the Central Register of National Health Service (NHSCR) patients, a source from which counts of moves between areas can be derived.

This paper seeks to demonstrate the value of the continuous time series movement data provided by the NHSCR for identifying changes taking place between Censuses, and in particular, for updating the picture of internal migration provided by the last Census at the beginning of the 1980s. In Section 3, changes in the volume, spatial distribution and age structure of migration over time are examined, and in Section 4, two examples of model-based research using the NHSCR data are presented. The empirical and modelling sections of the paper follow a short review of the characteristics of the data and an assessment of its reliability based on a comparison with data from the 1981 Census.

2 NHSCR migration data

The National Health Service Central Register is an important source of information about the redistribution of the population between those areas which Family Practitioner Committees (FPCs) utilize for the administration of national health services (Ricketts 1988). Whenever an NHS patient re-registers with a doctor in a new FPC area, information about the patient's date-of-birth, sex and origin FPC area are recorded and this information is maintained nationally in a central register. In England and Wales, for example, the FPC areas are represented by the metropolitan districts of provincial cities; the non-metropolitan or shire counties; and an aggregated set of London boroughs (Figure 1; Table 1). There are separate registers in Scotland and Northern Ireland.

Studies of internal migration in the UK based on data from this source have been undertaken since the 1950s (Newton and Jeffery 1951, Rowntree 1957). More recently, the value

Figure 1: Family Practitioner Committee areas in England and Wales plus other UK zones

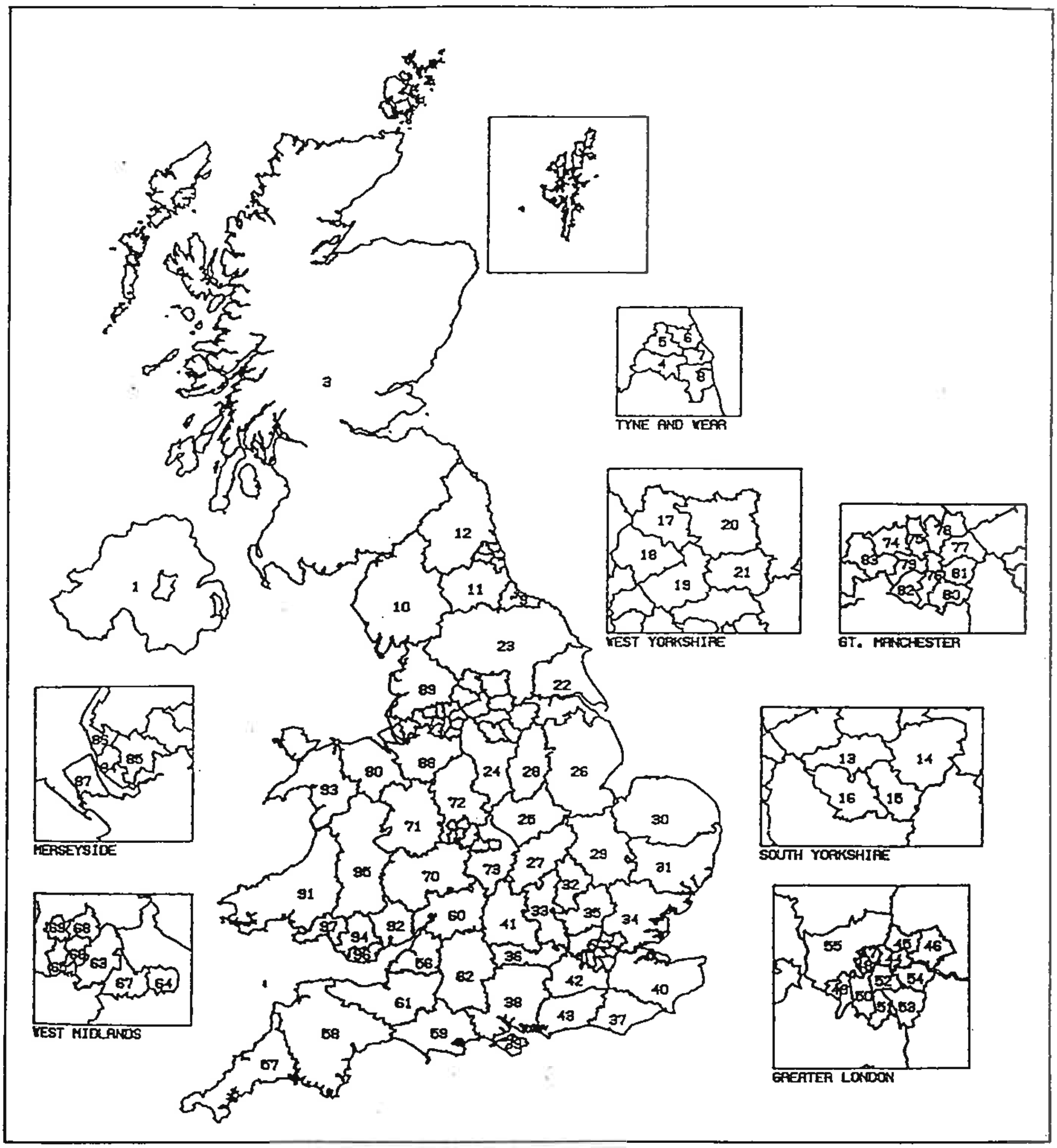


Table 1: Family Practitioner Committee areas in England and Wales plus other UK zones

Code	FPCA name	Code	FPCA name	Code	FPCA name
1	Northern Ireland	37	East Sussex	63	Birmingham
2	Isle of Man	38	Hampshire	64	Coventry
3	Scotland	39	Isle of Wight	65	Dudley
4	Gateshead	40	Kent	66	Sandwell
5	Newcastle	41	Oxfordshire	67	Solihull
6	North-Tyneside	42	Surrey	68	Walsall
7	South-Tyneside	43	West Sussex	69	Wolverhampton
8	Sunderland	44	City, Hackney,	70	Hereford &
9	Cleveland		Newham & Tower		Worcestershire
10	Cumbria		Hamlets	71	Shropshire
11	Durham	45	Redbridge &	72	Staffordshire
12	Northumberland		Waltham Forest	73	Warwickshire
13	Barnsley	46	Barking &	74	Bolton
14	Doncaster		Havering	75	Bury
15	Rotherham	47	Camden &	76	Manchester
16	Sheffield		Islington	77	Oldham
17	Bradford	48	Kensington,	78	Rochdale
18	Calderdale		Chelsea &	79	Salford
19	Kirklees		Westminster	80	Stockport
20	Leeds	49	Richmond &	81	Tameside
21	Wakefield		Kingston	82	Trafford
22	Humberside	50	Merton,Sutton	83	Wigan
23	North-Yorkshire		& Wandsworth	84	Liverpool
24	Derbyshire	51	Croydon	85	St. Helens &
25	Leicestershire	52	Lambeth,Southwark		Knowsley
26	Lincolnshire		& Lewisham	86	Sefton
27	Northamptonshire	53	Bromley	87	Wirral
28	Nottinghamshire	54	Bexley & Greenwich	88	Chesire
29	Cambridgeshire	55	Middlesex	89	Lancashire
30	Norfolk	56	Avon	90	Clwyd
31	Suffolk	57	Cornwall	91	Dyfed
32	Bedfordshire	58	Devon	92	Gwent
33	Buckinghamshire	59	Dorset	93	Gwynedd
34	Essex	60	Gloucestershire	94	Mid-Glamorgan
35	Hertfordshire	61	Somerset	95	Powys
36	Berkshire	62	Wiltshire	96	South Glamorgan
				97	West Glamorgan

Note:
Middlesex consists of the London Boroughs of Barnet, Brent, Harrow, Ealing, Hammersmith, Hounslow, Enfield, Haringey and Hillingdon.

of the NHSCR has been recognized by the Office of Population Cenuses and Surveys (OPCS) and the Department of the Environment (DOE), as a source of data for updating 1981 Census migration figures in the procedures for generating the net migration assumptions that are used in the model for subnational population projection (Martin, Voorhees and Bates 1981).

As the the period of time since the last Census lengthens, so the value of alternative and more up-to-date data sets increases. A key advantage of the NHSCR data is its availability on a continuous basis. Between 1975 and 1984, OPCS extracted a 10% sample of moves from the register and produced quarterly tabulations of total moves into and out from each FPC area, disaggregated by five year age group and sex, together with aggregate moves between FPC areas in England and Wales. Since 1984, however, 100% counts of moves have been available on a quarterly basis in the form of coded records of individual moves known as Primary Unit Data (PUD).

The problems and limitations of the NHSCR data such as the omission of short distance moves where no re-registration takes place, or the undercounting of young adult males, are now well documented (Ogilvy 1980, Devis 1984, Devis and Mills 1986, Bulusu 1988, Boden 1989). OPCS estimates that there is an average time lag of about three months between a move and the subsequent re-registration. It is therefore necessary when using NHSCR data for research to make the assumption that transfers recorded between September 30 of one year and the next are the appropriate data for conducting a mid-year to mid-year analysis. Data for twelve month periods are used to avoid seasonal fluctuations that occur in the quarterly figures.

The problems associated with the NHSCR data inevitably create doubts about its reliability as an indicator of inter-area population movement, and this has stimulated a number of studies comparing data from the NHSCR against data from the Censuses of 1971 (Ogilvy 1980) and 1981 (Thomson 1984, Devis and Mills 1986, Boden el al. 1987; 1988). In this context, it is helpful to classify the differences which exist between the information from these two sources into three categories.

Firstly, there are <u>conceptual</u> differences between the two types migration data (Courgeau 1980, Ledent 1980, Rees 1984). Registration data can be referred to as 'movement' data because it is a record of all moves taking place during a particular time period, rather than an account of the number of persons migrating during that period, which is the measure of migration used in the Census. The 1981 Census records 'transitions', i.e. persons whose usual residence on census date was different from that one or five years previously, and the age-time plans of observation differ between the two types of migration. This conceptual difference is very important in explaining different levels of mobility between the two sources when compared over a similar time period which are attributable to multiple or return movements by the same individual. Furthermore, a person who changes residence during the period yet dies before the end of the period will be registered in the NHSCR but excluded from the Census. Thus, ceteris paribus, we would expect the NHSCR counts to exceed the counts from the Census.

The second category of differences includes operational measurement differences, and in this case, emphasis is focused on the accuracy and comprehensiveness of the two methods of collection. There are problems with both sets of data which involve non-statement or mis-reporting of details, under-enumeration and sampling error. For example, Devis and Mills (1986) estimate that the 1981 Census failed to record roughly 172,000 migrants because of operational measurement difficulties. The problems of this type associated with the NHSCR include the non-recording of migrants who dont bother to register with a new doctor in their new FPC area. Different groups of the population have different registration and re-registration propensities according to their age, sex and family status characteristics.

Differences between populations at risk of being captured make up the third set of differences, since certain sub-groups of the population are handled differently in each data source. The three most important sub-groups in this respect are students, members of the Armed Forces and their dependents, and infants. The movement of students is excluded from the 1981 Census data but included in the NHSCR, provided that the student re-registers at his or her place of education. The Census records the recruitments and internal postings of Armed Forces personnel by origin and destination, but discharges are recorded as civilian migrations. The NHSCR, on th other hand, records only recruitment and discharge by origin and destination FPC area respectively, and no postings data is available. Infant migrants aged less than one at the end of the pre-Census period are excluded from the transition data but are recorded by the NHSCR. In making comparisons between the Census and the NHSCR migration data sets, it is possible to exclude the infant migrants from the latter. Comparative analysis is also enhanced by making adjustments to age-disaggregated data for age-time plan differences, by adopting estimation methods to assign non-stated flows and by assigning NHSCR Armed Forces recruitments and discharges to FPC areas based on the relative sizes of their Armed Forces populations (see Boden et al. 1988 for further details).

Four points can be identified which summarize the systematic comparison of the Census data and the NHSCR PUD for 1980-81. Firstly, while NHSCR/Census ratios indicate that the NHSCR flows between FPC areas overall are 37% higher than the Census counts, there is evidence of a scale effect (Table 2) with the difference reaching 54% at the standard region level and 19% for flows between FPC areas

Table 2: NHSCR and Census migration flows and ratios at different spatial scales

Migration flows between FPCAs	Total moves (NHSCR)	Total transitions (Census)	Difference	Ratio
(1) Between standard regions	967,224	629,915	337,309	1.54
(2) Between MNM regions	1,289,451	881,826	407,625	1.46
(3) Between MNM regions, within standard regions	322,227	251,911	70,316	1.28
(4) Within MNM regions	524,917	442,918	81,999	1.19
(5) All flows	1,814,368	1,324,744	489,624	1.37

Source: unpublished NHSCR and Census data supplied by the Office of Population Census and Surveys.

Notes:
Relationship between row items:-

row(3) = row(2) - row(1)
row(4) = row(5) - row(2)
row(5) = row(1) + row(3) + row(4)

within the old metropolitan and non-metropolitan (MNM) county boundaries; i.e. there is a greater discrepancy for longer distance migrations. Secondly, when ratios are compared across FPC areas, distinct variations occur. For example, metropolitan areas have the highest inflow ratios, illustrating the importance of the inmovement of students to higher education establishments. Thirdly, there is further evidence of the importance of the student factor when age and sex differences in the ratio at the FPC area scale are observed (Figure 2). Both male and female profiles have high ratio values at ages 15-19. Non-surviving migrants will be particularly important in the 75+ age category and this will partly explain the high ratios for this age group. Finally, high correlation coefficients for outflows, inflows and net flows at different spatial scales (Figure 3) suggest that the general spatial patterns of gross and net migration captured in the two data sources are very similar.

The conclusion which therefore emerges from this brief summary is that the NHSCR does have credibility as an alternative data source, although clearly its interpretation must be handled with due regard to the features that distinguish it from Census transition data. It has the primary advantage of providing information which allows us to monitor the changing patterns of migration over time.

3 Trends in internal migration

A mid-year to mid-year time series of aggregate, inter-FPC area migration data has been assembled, together with FPC area age-disaggregated gross inflows, gross outflows and mid-year population estimates for males and females. The series, which commences in 1975-76 and finishes in 1985-86, provides the opportunity to investigate geographical patterns of migration and changes in migration activity at a variety of spatial scales. Data for the 97 zones that constitute the base system of interest (Figure 1) can be consolidated for examination at coarser spatial scales: metropolitan counties/non-metropolitan regions; standard regions; macro regions.

Dominating the pattern of aggregate net migration at the FPC area scale in England and Wales are the losses occurring from most metropolitan areas and the gains experienced by the majority of non-metropolitan counties during the 1975-86 period. The map of net migration rates for the final year of the series (Figure 4) shows positive balances for very few metropolitan districts or London boroughs. Merseyside, in particular, experienced higher rates of loss during the period than the other provincial conurbations, and there is little evidence of any improvement taking place. On the other hand, net losses from Greater London declined significantly during the first half of the period from over 80,000 in 1975-76 to below 40,000 in 1982-83, before rising again to over 50,000 in

Figure 2: NHSCR/Census inflow ratios by age and sex

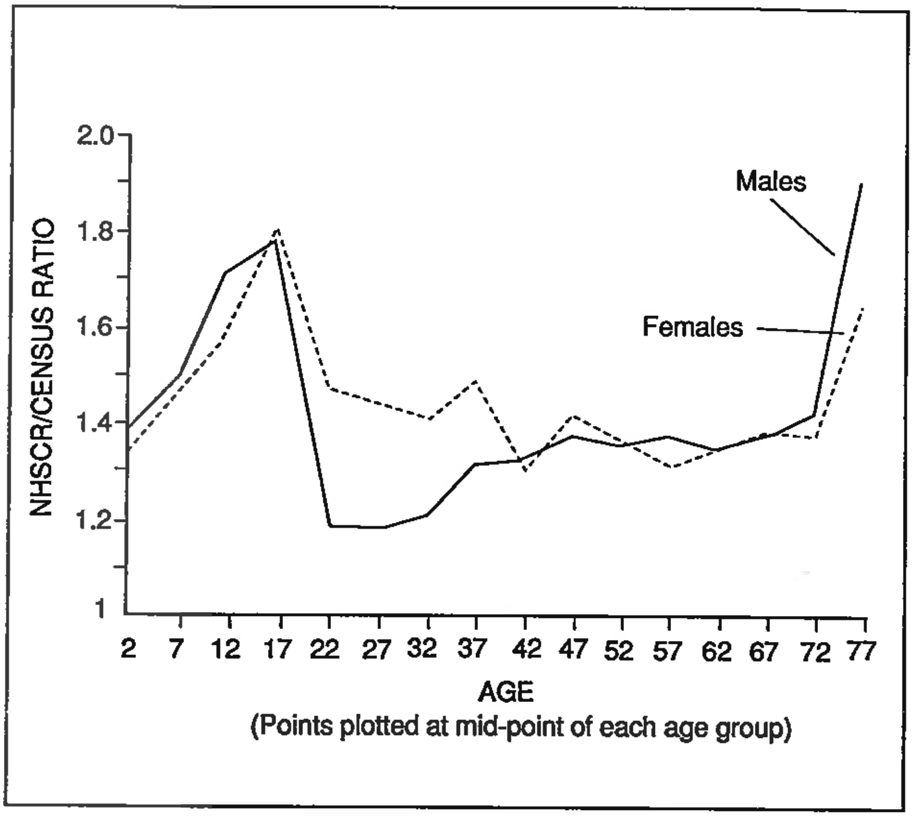


Figure 3: Scatterplots of NHSCR re-registration rates against Census migration rates at different spatial scales

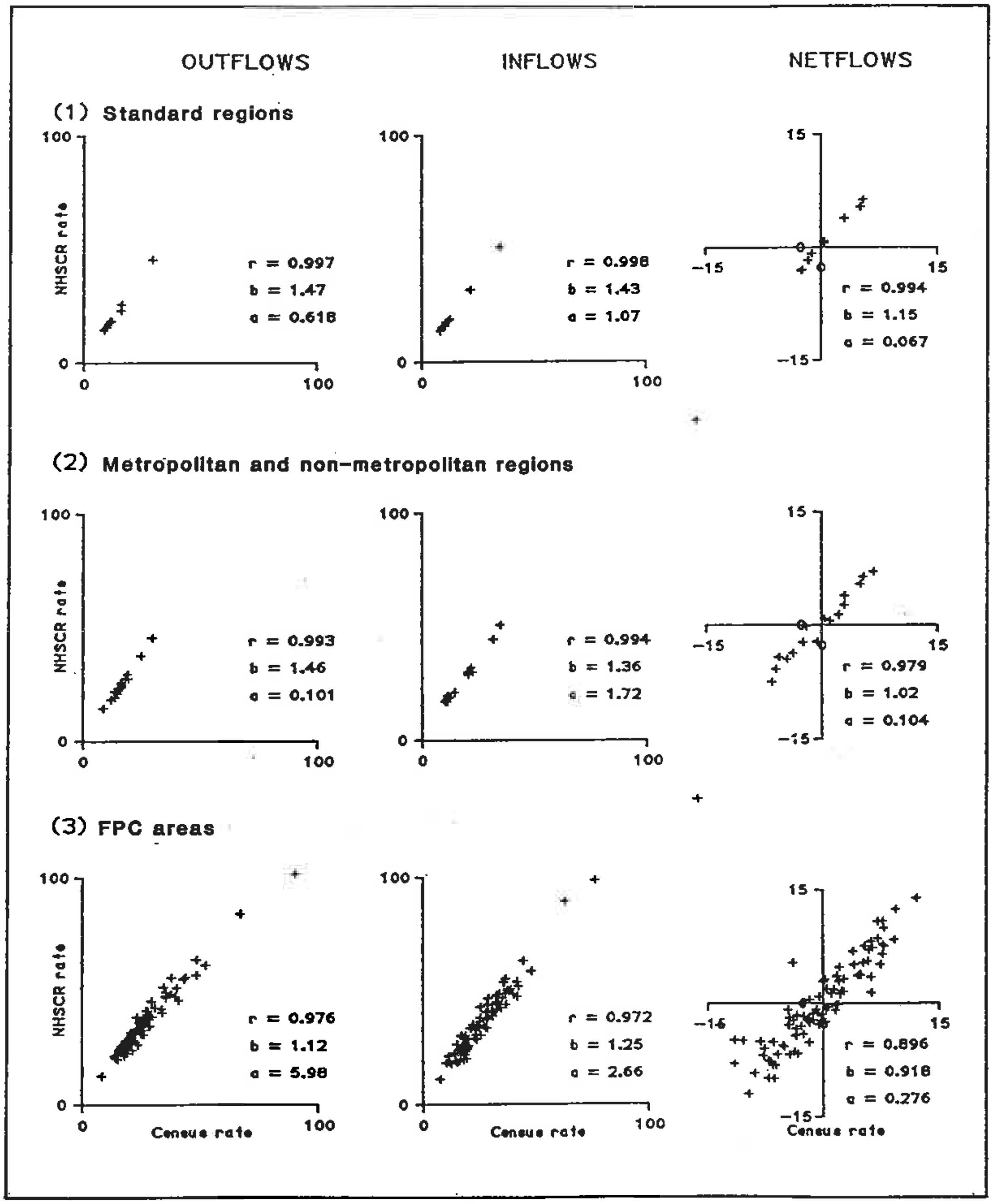
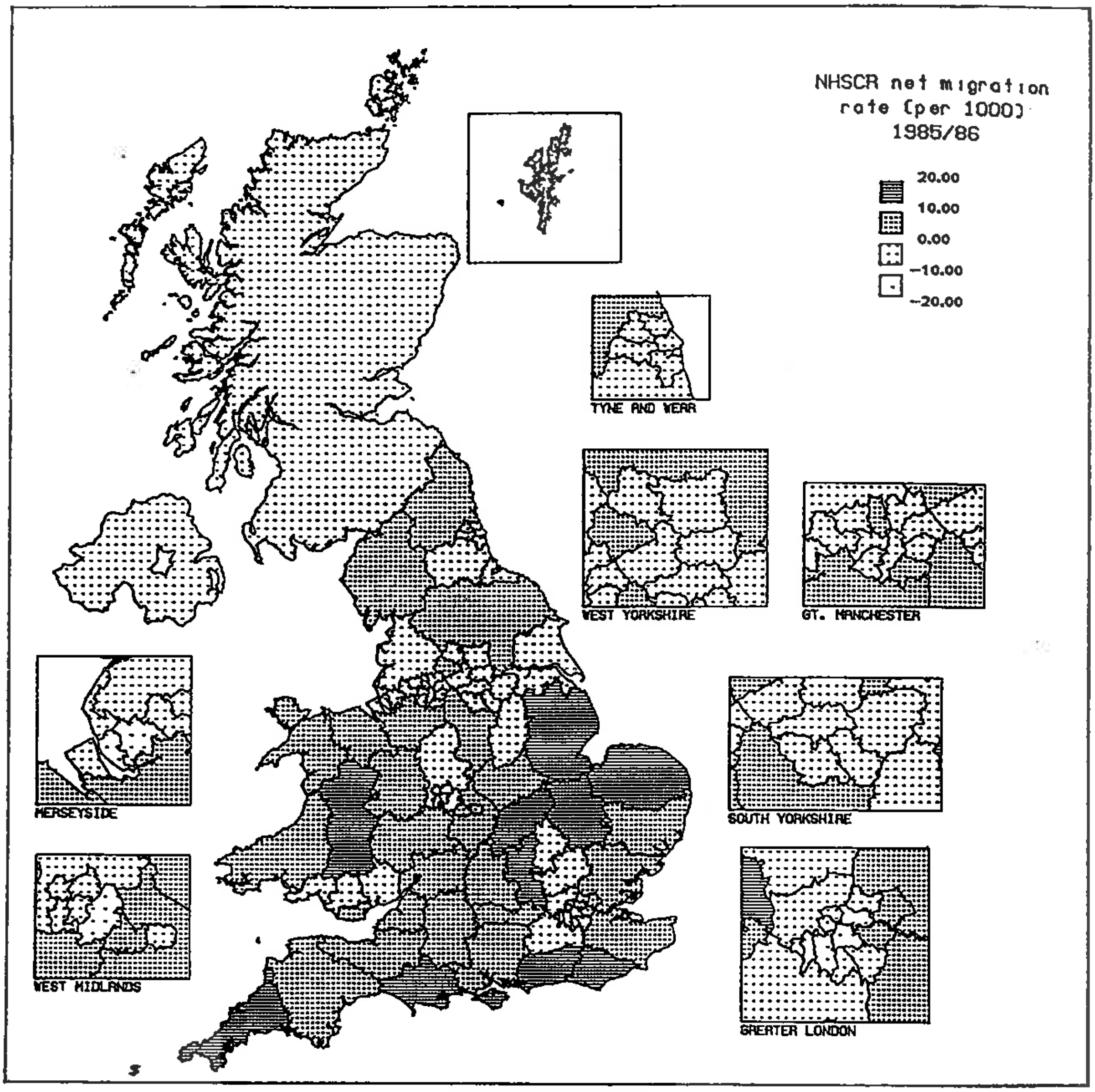


Figure 4: Net migration rates, FPC areas, 1985-86



1985-86. Figure 4 indicates that exceptions to the general pattern do occur, with some shire counties experiencing losses. These include counties that are part of the outer metropolitan area surrounding Greater London as well as counties such as Cleveland and Durham in the North, West and Mid Glamorgan and Gwent in Wales, and Humberside, Staffordshire and Nottinghamshire, all of which suffered major manufacturing declines during the 1980s. The counties with the highest rates of gain are the rural areas including Powys, Lincolnshire, Norfolk, Cambridgeshire and several counties on the south coast, where retirement migration is likely to be particularly important.

Inevitably, the patterns of age-specific net migration differ significantly from the aggregate distribution. The 20-24 years age group comprised 20% of total inter-FPC area migration in both 1980-81 and 1985-86 and the changing distribution of movements for this group between these two years is illustrated in Figure 5 which shows how the pattern of net gains had become more concentrated by 1985-86. The balances of almost all the FPC areas in the North and Midlands have become negative with exception of Greater Manchester. Thus it is apparent that the migration activity of this most mobile age group has become increasingly directed towards the high density areas of the South East at the expense of the remainder of the country.

A statistical comparison of the inter-FPC area migration matrices for broad age groups between 1980-81 and 1985-86 using dissimilarity indices (Boden et al. 1990) indicates that, in relative terms, changes in distribution patterns are most marked for the 70+ age group. The map of percentage change in inmigration rates (Figure 6) shows most FPC areas experiencing increased rates during the 80s. A number of metropolitan FPC areas in Greater Manchester, Merseyside and the West Midlands showed significantly higher inmigration rates in 1985-86 but the more important increases occurred in the non-metropolitan counties of Devon and Cornwall in the South West, Bedfordshire and Kent in the South East and Dyfed and Powys in Wales. Most Greater London FPC areas experienced declines in the rate of inmigration and increases in the rate of outmigration, but the most significant decreases in outmigration rates for the elderly (Figure 7) occurred from the more rural areas including Hampshire, East Sussex, Kent, Lincolnshire, Lancashire, Durham, Scotland, and the two Welsh FPC areas of Mid and South Glamorgan. As with inmigration, most FPC areas experienced outmigration increases.

Since the volume of information available to examine spatial patterns and temporal change in migration at the FPC area scale is very large, it is appropriate to consider aggregating the data for analytical purposes. Two approaches are illustrated here. The first approach has involved grouping FPC areas into four categories on the basis of a ranking of their densities (high, medium-high, medium-low, low) using 1981 Census usually resident populations (Stillwell et al. 1990). Density is regarded as

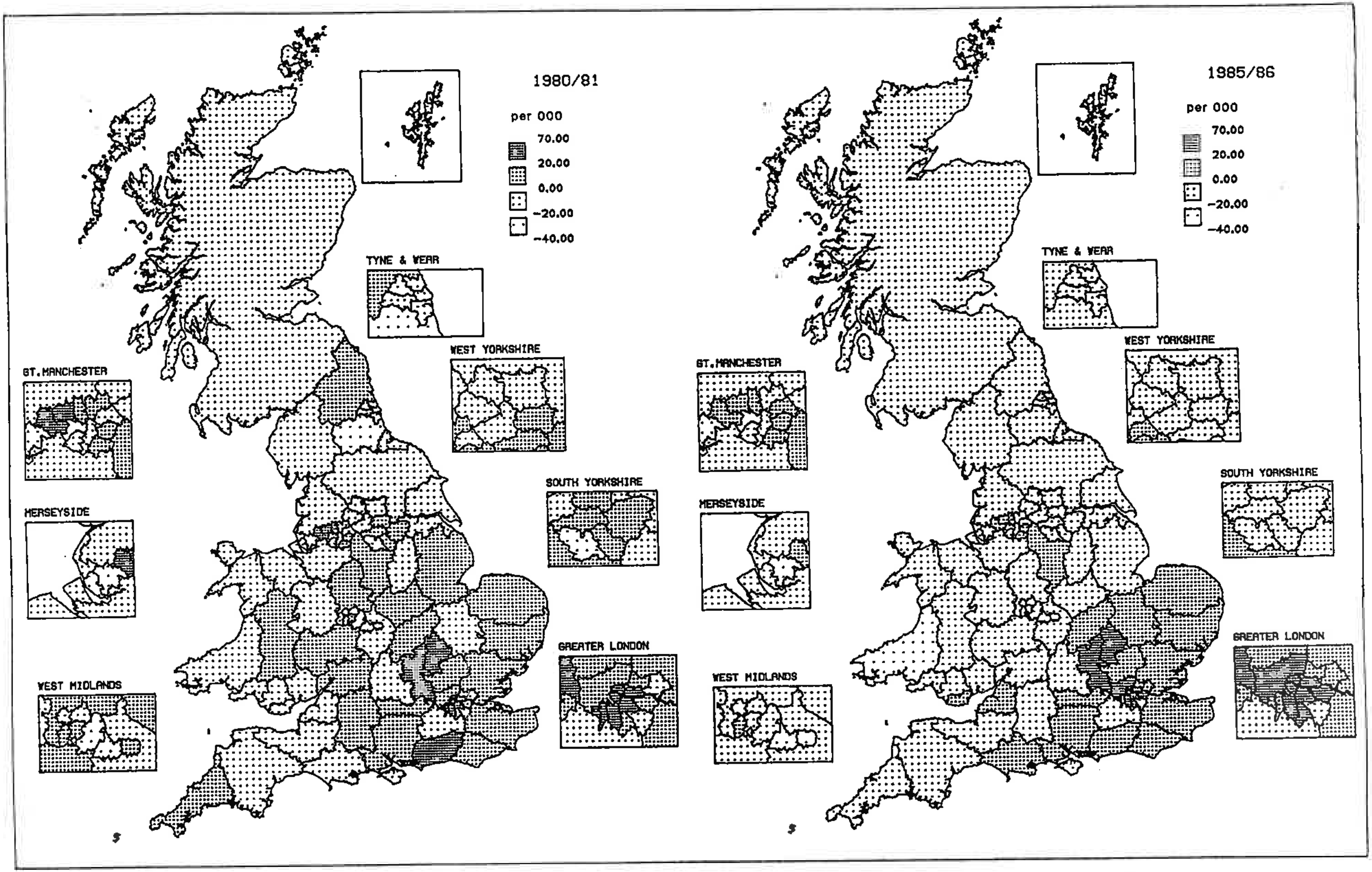


Figure 5: Net migration rates, FPC areas, 20-24 age group, 1980-81 and 1985-86

a proxy for urbanization. FPC areas in each category which are located in the North and South of the country are distinguished to form an eight group classification.

Figure 8 demonstrates how the North-South 'divide' in migration terms has widened between 1975 and 1986. Rees (1990) shows how the tempo of movement from North to South has resulted in a net outflow increasing in 1979-83 by more than double its 1975-79 value. In the North, net losses have occurred from both high and medium-high density areas without the compensatory gains by FPC areas in medium-low and low density categories that have occurred in the South. Northern FPC areas in the medium-low density class have actually suffered losses since 1980-81 and low density areas also recorded a negative balance in 1985-86. Thus, whilst northern metropolitan areas have experienced considerable net losses, the outflows have not resulted in corresponding gains in less urbanized parts of the North. Countermetropolitanization in the North is less important than the movement from North to South. In the South, the time series profile for Greater London's FPC areas shows a deceleration in the volume of loss to 1982-83 without a comparable decline in the gains by the low and medium-low density FPC areas. Since 1980-81, the decline that was taking place in the most rural counties has been reversed so that by 1985-86, these areas gained over 60,000 migrants. Countermetropolitanization is certainly a continuing feature of migration behaviour in southern Britain in the 1980s.

The graphs illustrated in Figure 9 summarize the age structure of net migration rates for FPC areas falling in: (a) the North and South; (b) metropolitan and non-metropolitan areas; and (c) the eight density categories. The graphs depict single year of age net migration rates (extracted from the PUD) for 1980-81 and 1985-86. The North/South profiles reflect the net losses from the North of persons in all age groups, but particularly those in the labour force age range. The peak in the age schedule comes at age 23 in both periods and the difference between the schedules for 1980-81 and 1985-86 demonstrates the growing attraction of the South to this younger section of the population. The metropolitan/non-metropolitan classification shows that it is only in the age range 17-23 that the metropolitan areas as a whole gain migrants. The peak of these schedules is at age 19, possibly reflecting the importance of student movement to higher education establishments as well as the availability of jobs. Furthermore, the rates of net gain are lower for those aged 17-19 in 1985-86 than in 1980-81. Movement away from metropolitan areas at or around retirement age is also evident in both years. These patterns are mirrored by the non-metropolitan area profiles, although the rates are lower in the case of the latter. Interesting differences between age-specific net migration patterns in the North and South of the country are evident from the profiles by density category. Northern high density areas show a single peak whereas a double peak can be observed for Greater

Figure 6: Percentage change in inmigration rates, FPC areas, 70+ age group, 1980-81 to 1985-86

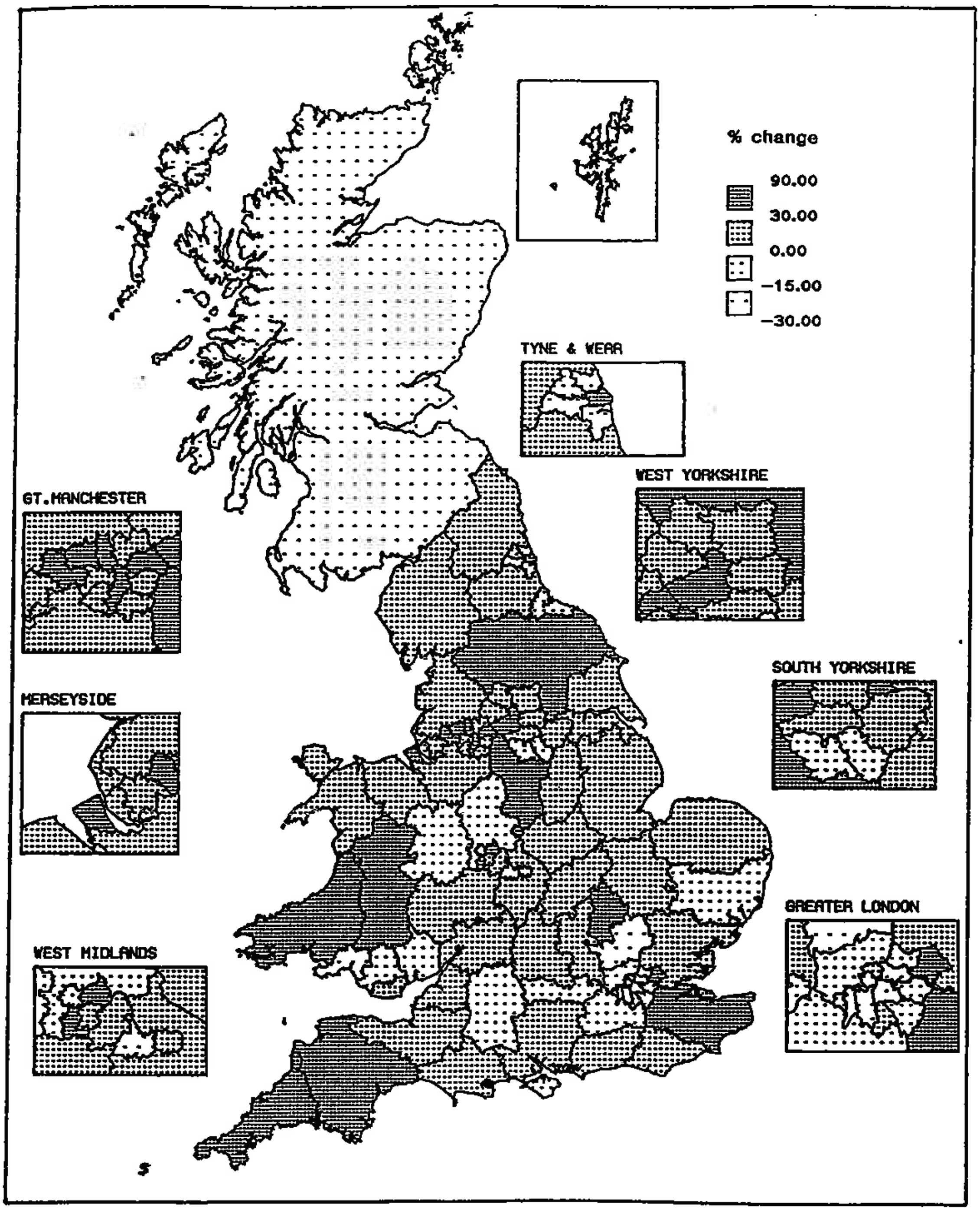


Figure 7: Percentage change in outmigration rates, FPC areas, 70+ age group, 1980-81 to 1985-86

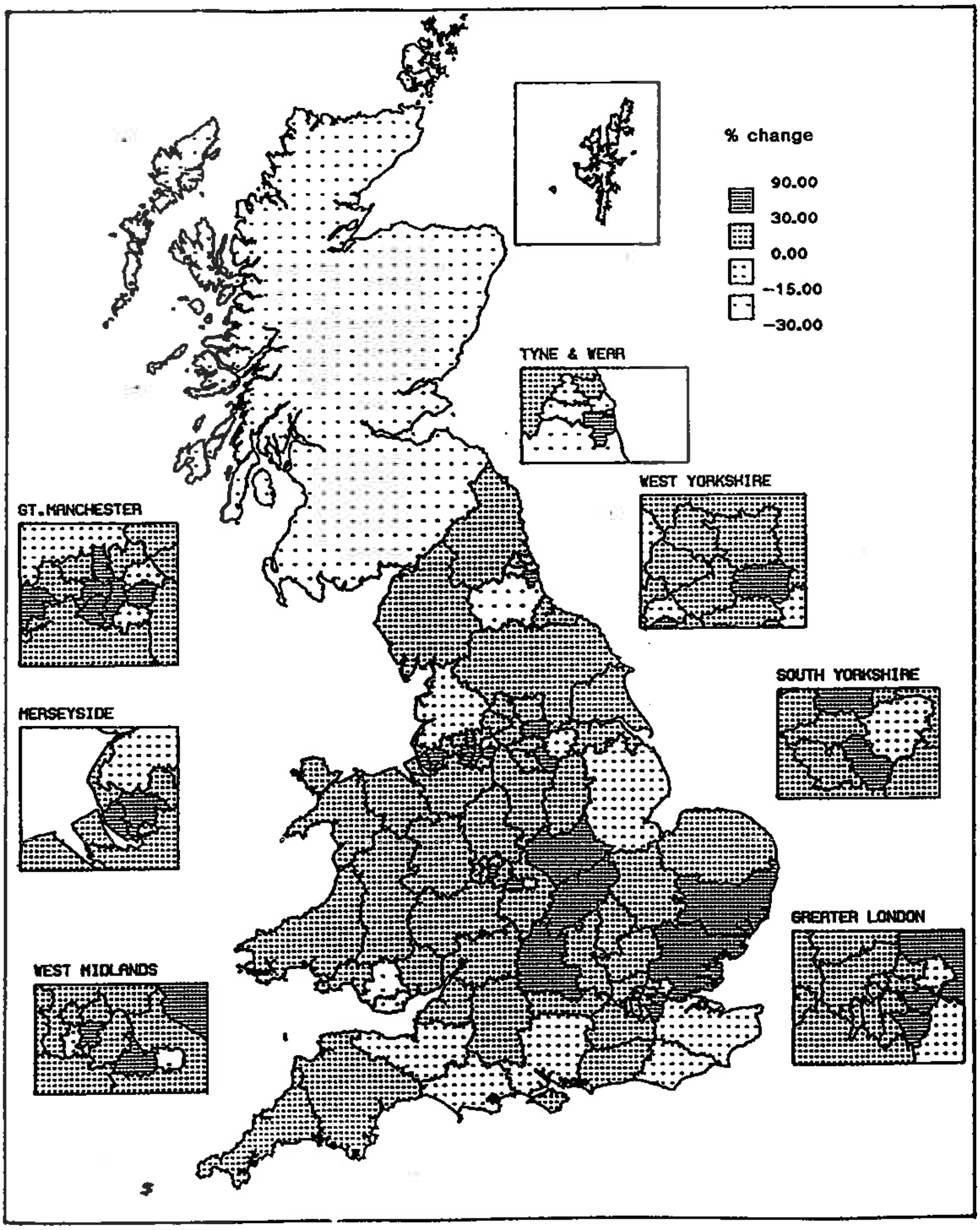
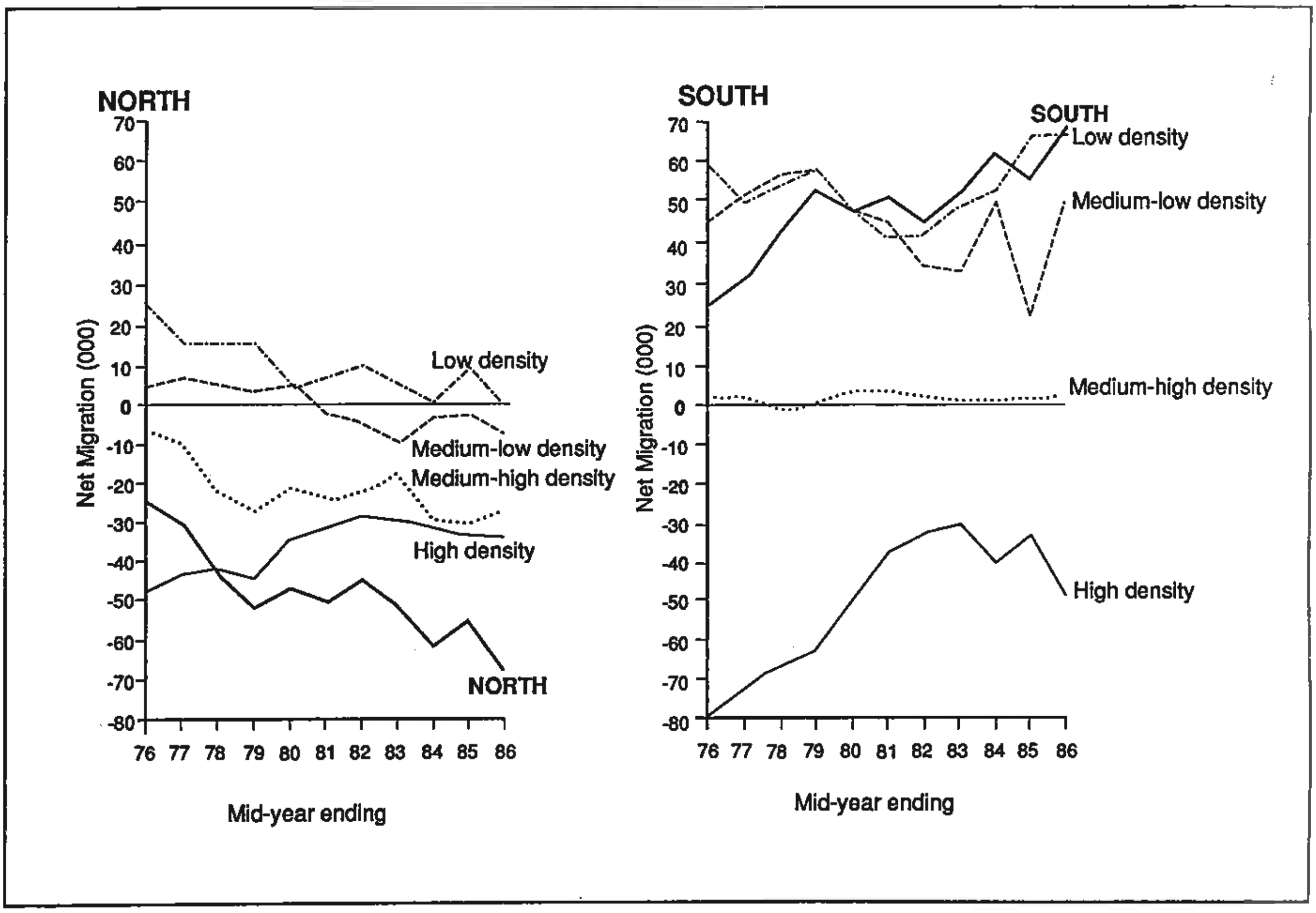


Figure 8: Net migration balances for FPC area density categories in the North and South, 1975-86



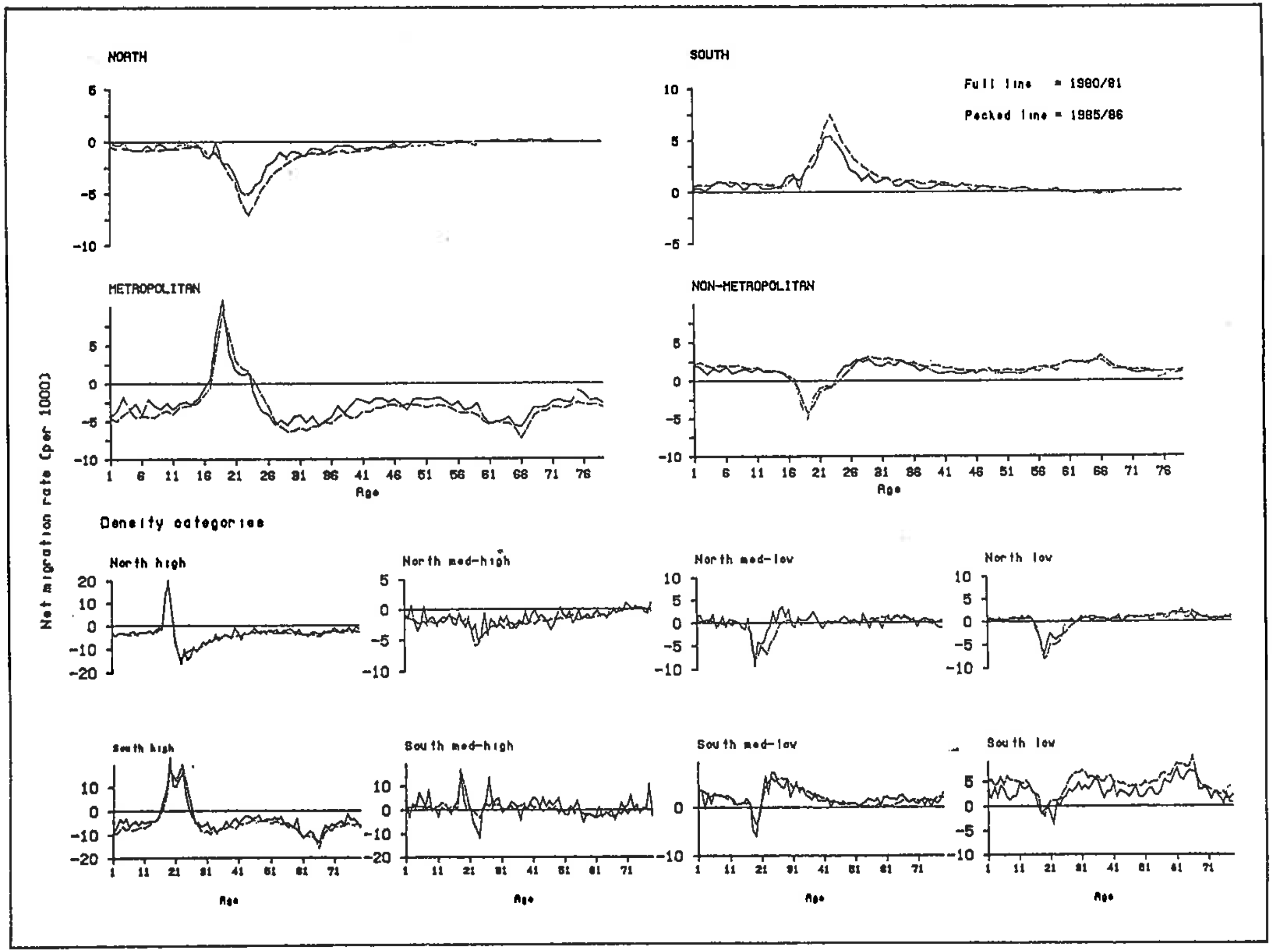


Figure 9: Age-specific net migration rates by North/South, metropolitan/non-metropolitan and population density divisions, 1980-81 and 1985-86

London. Low density FPC areas in the South continue to attract migrants, particularly in the retirement age range, whereas areas of similar density in the North have much lower rates of gain.

The second approach that has been used to analyse trends in migration is the components of migration framework outlined by Willekens and Baydar (1986), in which the migration flow between an origin and a destination is separated into four components:

(i) a level component which measures the overall level of migration in the system;
(ii) a generation component which measures the relative importance of each area as a generator of outmigrants; i.e. the proportion of total migration that originates from that area;
(iii) an attraction component which measures the relative importance of each area as a recipient of inmigrants; i.e. the proportion of total migration that is attracted to that area; and
(iv) a distribution component which measures either the proportion of total outmovement from each origin area that arrives at each destination area, or the proportion of total inmigration to each destination area that comes from each origin area.

The remaining comments in this section of the paper relate to the first three components, whilst the distribution component is considered in a modelling context in section 4.

The decline in the overall level of internal migration in Britain during the 1970s has been well documented (see Ogilvy 1982, Devis 1984, Stillwell and Boden 1989, for example). In order to examine changes in the 1980s, data for two further years, 1986-87 and 1987-88, has been added to the initial 1975-86 time series. Figure 10 illustrates the total number of moves per annum recorded between FPC areas, expressed also as a percentage of the base year volume. The level of movement fell by about 17% between 1975-76 and 1981-82 but increased steadily thereafter so that by 1987-88, it was at a level higher than at any time in the preceding 12 years. Whilst it is clear that the overall propensity to migrate has fluctuated significantly, the extent of temporal instability in the shares of migration into and from particular regions can be guaged from the generation and attraction components. Time series indices of the two components have been computed for a set of UK metropolitan and non-metropolitan regions (including Northern Ireland). Figure 11 indicates that as a general rule, the generation components are more stable than the attraction components, with South Yorkshire being the principal exception. Largest declines in the attraction index are evident for the regions of Northern Ireland, Scotland, the remainder of the North and Merseyside, whilst Greater London has the most consistent increase. On the other hand, Greater London experiences a significant

Figure 10: Changes in the level component of inter-FPC area migration, 1975-88

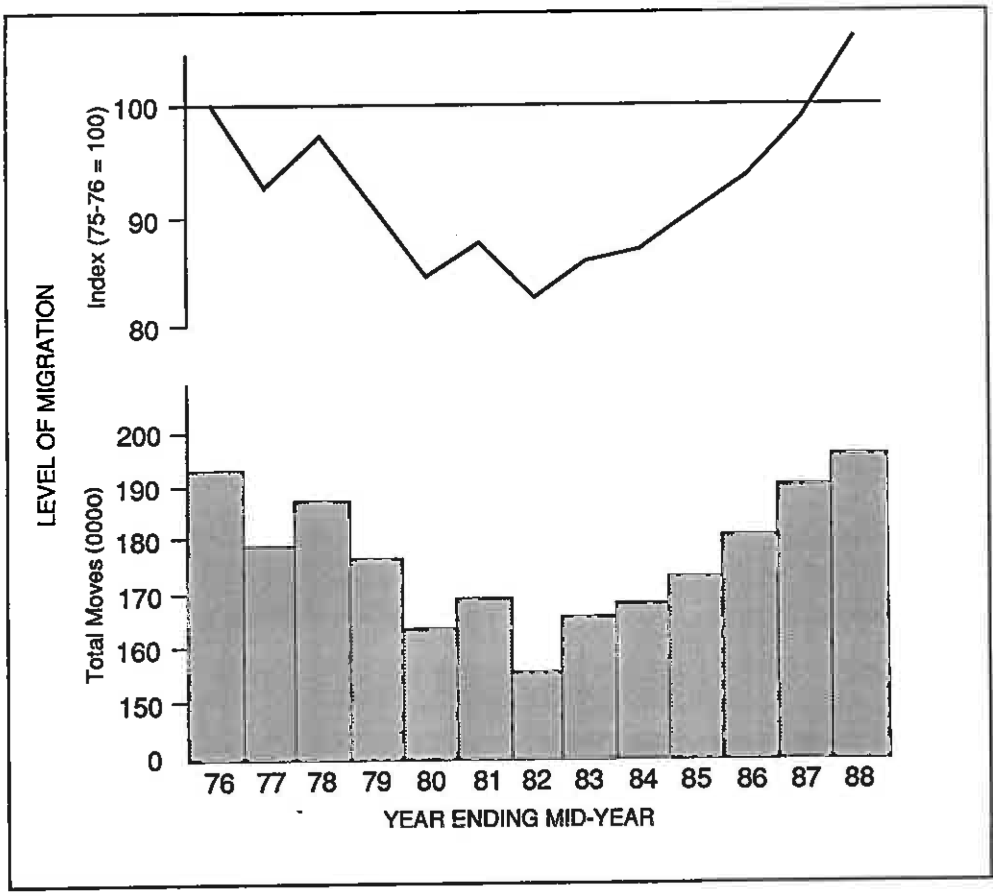
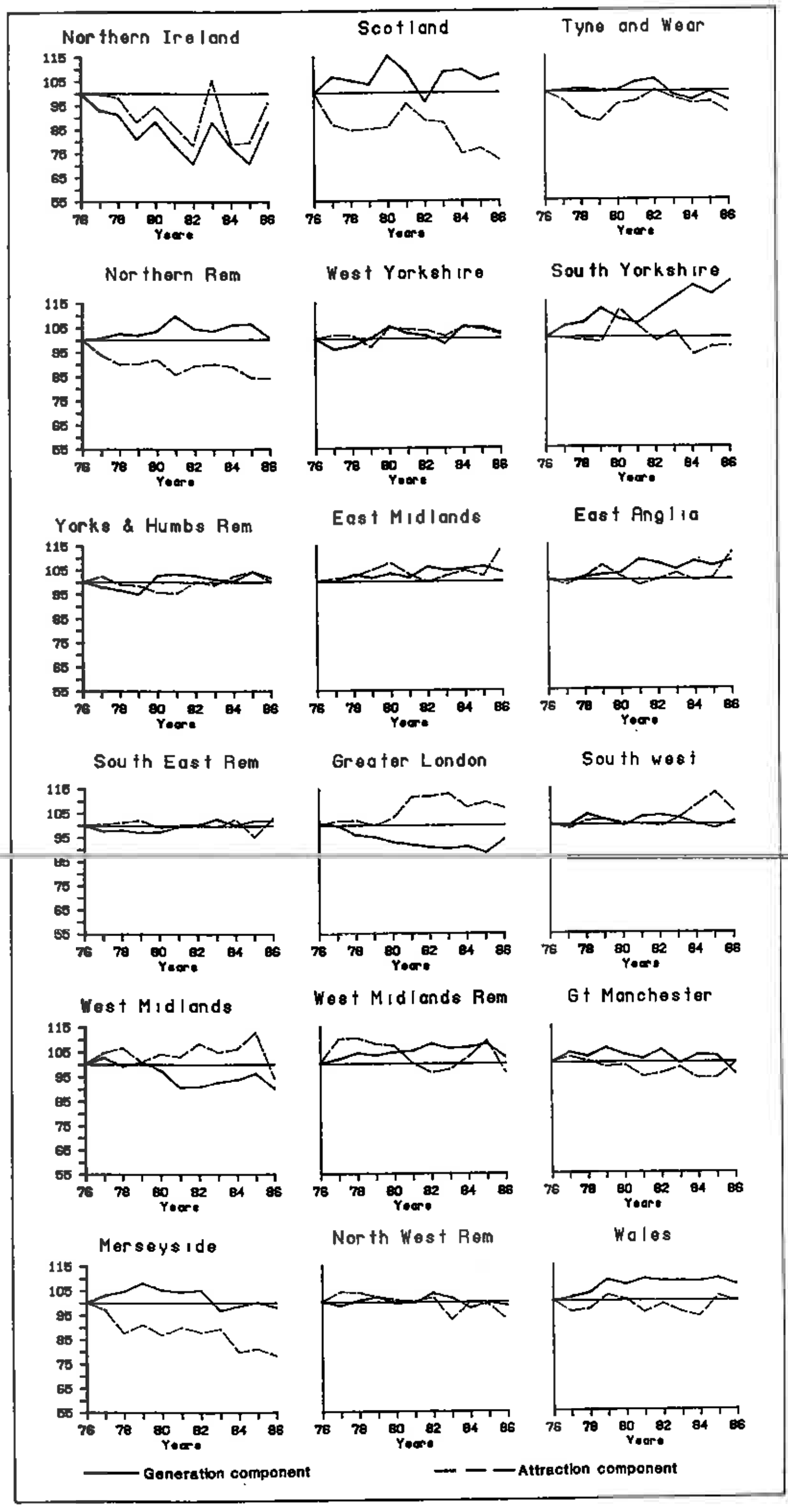


Figure 11: Changes in the generation and attraction components, metropolitan and non-metropolitan regions, 1975-1986



decline in its share of total outmigration, unlike Northern Ireland, where less contact with the rest of the UK is occurring in both directions.

4 Model-based analysis

Modelling provides a framework through which the type of migration analysis reported in the previous section may be extended to investigate cause and effect or to estimate unknown information in historical or future contexts. In this paper, two applications of modelling techniques using the NHSCR data are presented for exemplification. In the first instance, we focus on the role of distance and examine its frictional effect on migration between FPC areas using constrained spatial interaction models. Secondly, migration age schedule modelling is used to identify the characteristics of age profiles for groups of FPC areas.

Spatial interaction models derived from entropy maximizing methods (Wilson 1974) enable knowledge about gross zonal outmigration and inmigration to be incorporated in the form of constraints into traditional unconstrained gravity models. Thus a doubly constrained spatial interaction model of flows between the 94 FPC areas of England and Wales plus Scotland included as a single zone, can be written as:

$$M_{ij} = A_i B_j O_i D_j d_{ij}^{-b_*} \qquad (1)$$

where O_i = total outmigration from FPC area i;

D_j = total inmigration to FPC area j;

$d_{ij}^{-b_*}$ = a linear function representing the inverse relationship between migration and distance from area i to area j, where b_* is the calibrated distance decay parameter;

and where A_i and B_j are endogenous balancing factors which ensure that out- and inmigration constraints are satisfied.

Inter-area distances have been calculated using national grid references for centres of local authority districts, counties and regions (OPCS 1984) as follows:

$$d_{ij} = \sqrt{(e_i - e_j)^2 + (n_i - n_j)^2} \qquad (2)$$

where e_i and e_j are easting coordinates of origin i and destination j; and

n_i and n_j are northing coordinates of origin i and destination j.

The optimum generalized decay parameter, b_*, has been estimated with a Newton-Raphson iterative search routine using a software package known as IMP (Stillwell 1984) and can be interpreted as a measure of the general propensity to migrate over distance. Thus, higher parameter values indicate that distance has a more pronounced effect on migration.

When a doubly constrained model is fitted to each of the 11 inter-FPC area movement matrices, the resulting parameters show that the propensity to migrate over distance has fluctuated. It fell gradually from 1975-76 to 1980-81 (b_* dropped from 1.24 to 1.19), rose to 1.22 in 1983-84 and fell again to 1.20 by 1985-86. The average distance over which migration has taken place increased from 125.7km in 1975-76 to 129.2km in 1981-82, declined to 127.8km by the end of the period.

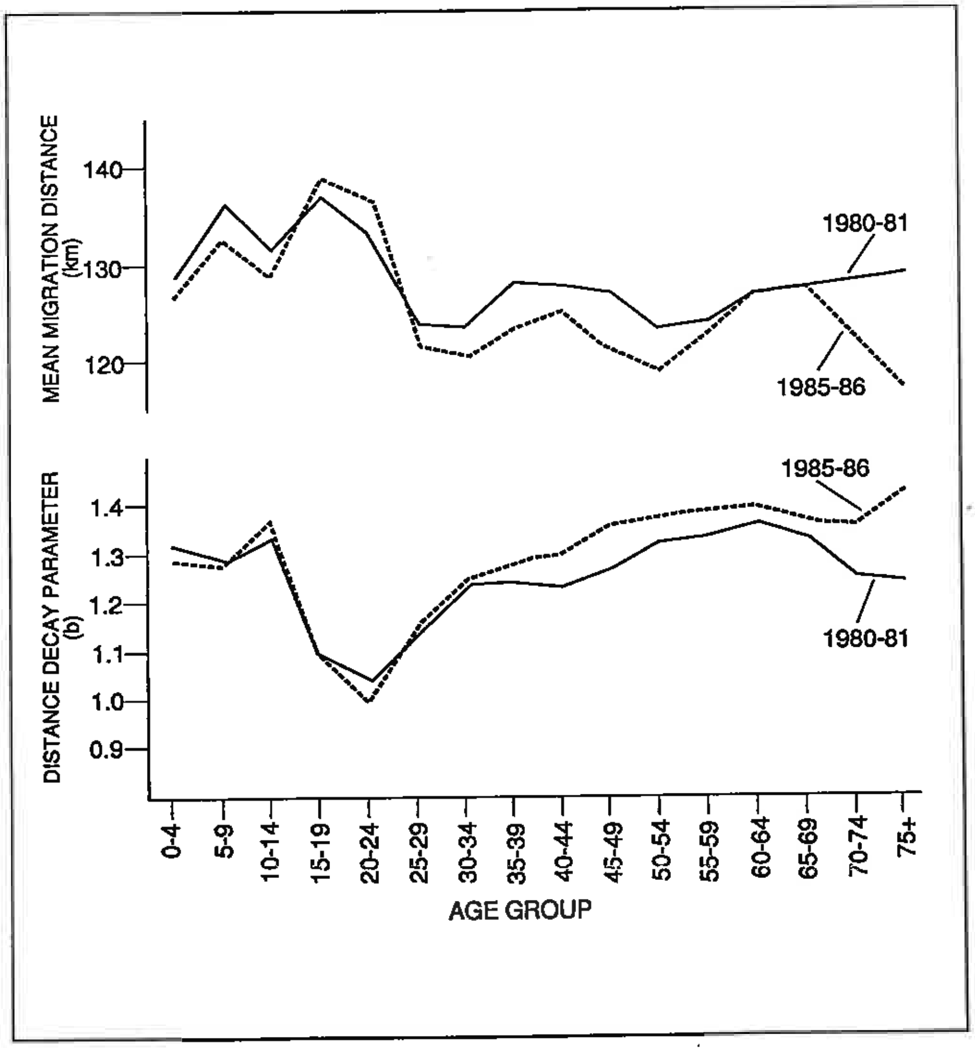
The generalized decay parameter gives an indication of change in the overall propensity to migrate but it is very likely that aggregate movement from specific origins and to specific destinations will show spatial, temporal and age-related variation. Models with origin- and destination-specific parameters have been calibrated which demonstrate the extent of this variation. The influence of distance on outmovement and inmovement does vary between zones as shown in Table 3, where the FPC areas have been ranked on the basis of the value of their decay parameters. The parameter values range from over 2.0 in the case of outmigration from and inmigration to Powys and Gwynedd to less than 1.0 in the case of outmigrants from Kent and inmigrants to several of the London FPC areas. There is no significant relationship between the decay parameter value and the mean migration distance; i.e. the propensity to migrate over distance is not necessarily determined by the accessibility of the zone to the rest of the system, as measured by the mean migration distance. There are areas whose migrants show a similar propensity to migrate over distance but whose migrants actually move over very different distances. Likewise, there are areas with simliar mean migration distances but with very different decay parameters.

It can be argued that the propensity to migrate over distance depends as much on age as it does on the region of origin or destination. Figure 12 illustrates mean movement distances and generalized decay parameters for inter-area migrants in 16 age groups (0-4, ..., 70-74, 75+) for two 12 month periods: 1980-81 and 1985-86. The profiles reflect some of the characteristics of individuals in particular life course groups. In the early age groups, distance exerts most influence on migrants aged 10-14 and least influence on those aged 20-24. Thereafter, the influence of distance increases steadily to around retirement age before decreasing. Mean distances of migration tend also to

Table 3: Mean migration distances and area-specific decay parameters for FPC areas, 1985-86

FPC AREA	MMD	ORIGIN PARAMETER	FPC AREA	MMD	DESTINATION PARAMETER
POWYS	139.57	2.25	POWYS	142.96	2.16
GWYNEDD	182.37	2.08	GWYNEDD	179.36	2.00
LINCS	141.24	1.71	HEREFORD	108.38	1.75
SANDWELL	54.13	1.62	STAFFS	99.36	1.71
HEREFORD	119.26	1.60	DUDLEY	55.26	1.70
DYFED	201.12	1.60	WALSALL	54.43	1.69
MID-GLAM	117.87	1.59	SUFFOLK	148.92	1.66
CLWYD	129.70	1.59	ROTHERHM	69.20	1.65
WALSALL	73.58	1.56	NORFOLK	179.11	1.65
SUFFOLK	146.93	1.54	LINCS	143.30	1.63
DUDLEY	77.93	1.54	SANDWELL	44.17	1.63
HUMBERSD	165.36	1.53	DYFED	205.03	1.63
NORFOLK	176.44	1.53	HUMBERSD	157.00	1.62
ROTHERHM	87.99	1.51	MID-GLAM	121.81	1.62
BARNSLEY	91.18	1.49	CLWYD	117.30	1.61
WAKEFELD	94.81	1.48	W-GLAM	160.78	1.59
W-GLAM	161.11	1.48	BARNSLEY	79.73	1.54
GLOUCS	124.71	1.47	SOLIHULL	59.55	1.52
SALOP	132.18	1.46	TAMESIDE	56.05	1.52
STAFFS	116.60	1.46	DERBYSHR	100.43	1.51
ST-HELEN	80.98	1.46	SALOP	123.79	1.51
DONCASTR	114.13	1.45	LON-BH	52.84	1.49
DERBYSHR	109.86	1.43	SEFTON	85.85	1.49
WOLVERHN	88.46	1.43	ST-HELEN	69.38	1.47
SOLIHULL	76.89	1.42	WOLVERHN	76.34	1.46
WIGAN	86.55	1.41	WIGAN	72.42	1.46
CUMBRIA	213.77	1.40	GWENT	129.23	1.46
BIRMINGH	92.02	1.40	SOMERSET	153.57	1.45
TAMESIDE	74.13	1.40	NTHMBLND	149.87	1.43
SOMERSET	145.53	1.38	W-SUSSEX	114.18	1.43
GWENT	127.96	1.38	WAKEFELD	90.61	1.42
NOTTS	118.36	1.37	N-YORKS	145.00	1.40
LON-BH	64.24	1.37	BURY	66.69	1.37
SALFORD	70.93	1.36	E-SUSSEX	135.33	1.38
DORSET	159.25	1.35	ROCHDALE	73.05	1.38
NTHMBLND	184.44	1.33	CAMBS	137.37	1.36
CAMBS	135.31	1.33	ESSEX	112.46	1.36
MANCHSTR	81.61	1.32	DORSET	171.09	1.36
SHEFFELD	113.37	1.31	GLOUCS	133.30	1.36
ROCHDALE	88.42	1.31	S-GLAM	140.91	1.36
N-YORKS	165.49	1.29	DONCASTR	112.47	1.35
WILTS	128.98	1.29	OLDHAM	69.53	1.35
BURY	86.64	1.29	STOCKPRT	69.60	1.35
LVERPOOL	105.76	1.29	NTHANTS	120.18	1.34
BEDFORDS	102.92	1.27	CORNWALL	300.62	1.34
OLDHAM	86.85	1.27	BOLTON	79.18	1.33
TRAFFORD	81.38	1.27	CALDERDL	90.06	1.31
CALDERDL	105.13	1.26	WILTS	134.61	1.31
CORNWALL	284.10	1.26	BIRMINGH	92.89	1.31
SEFTON	123.43	1.25	TRAFFORD	67.62	1.31
S-GLAM	142.62	1.25	BEDFORDS	103.75	1.29
LON-CHNT	54.95	1.24	CUMBRIA	202.37	1.29
E-SUSSEX	141.90	1.23	NOTTS	121.67	1.28
BRADFORD	118.01	1.22	DEVON	246.34	1.28
STOCKPRT	91.71	1.22	WARWICKS	91.34	1.28
N-TYNESD	137.15	1.21	CHESIRE	110.17	1.28
LEEDS	126.73	1.21	KENT	136.91	1.27
NTHANTS	123.86	1.21	KIRKLEES	99.02	1.25
WARWICKS	103.90	1.21	N-TYNESD	112.11	1.24
CHESIRE	127.00	1.21	LVERPOOL	99.63	1.24
KIRKLEES	114.34	1.20	SALFORD	75.02	1.23
W-SUSSEX	124.86	1.20	GATESHED	95.04	1.21
LEICS	127.98	1.19	LEICS	126.30	1.21
OXFORDSH	119.34	1.19	WIRRAL	106.77	1.20
LON-RWF	66.13	1.19	LANCS	137.82	1.20
GATESHED	117.06	1.18	SHEFFELD	113.63	1.19
DEVON	233.00	1.18	BUCKS	103.26	1.19
BOLTON	103.30	1.18	AVON	155.29	1.19
DURHAM	183.85	1.17	DURHAM	161.99	1.18
AVON	145.96	1.17	BRADFORD	112.15	1.17
SUNDRLND	161.71	1.15	LON-BROM	69.86	1.16
NEWCSTLE	160.32	1.13	SUNDRLND	140.64	1.15
S-TYNESD	148.83	1.13	OXFORDSH	125.34	1.15
BUCKS	106.89	1.12	LEEDS	122.52	1.14
COVENTRY	96.81	1.12	HERTS	94.65	1.14
CLEVELND	194.13	1.11	LON-CRDY	65.63	1.14
LON-LSL	67.40	1.11	MANCHSTR	91.33	1.14
LANCS	159.17	1.11	CLEVELND	176.62	1.13
LON-MSW	70.82	1.10	HANTS	153.92	1.13
LON-BG	78.90	1.10	IOWIGHT	154.72	1.13
LON-CROY	75.66	1.09	S-TYNESD	131.96	1.10
LON-CI	63.42	1.07	LON-RWF	68.52	1.10
LON-KCW	61.09	1.07	SURREY	102.35	1.07
ESSEX	132.44	1.06	LON-BG	77.23	1.07
LON-BROM	82.30	1.06	NEWCSTLE	150.55	1.06
HANTS	149.51	1.05	COVENTRY	103.66	1.04
LON-MIDD	85.02	1.05	BERKS	120.37	1.00
IOWIGHT	150.23	1.02	LON-MSW	81.00	.95
HERTS	105.00	1.01	LON-RK	88.02	.91
BERKS	116.17	1.01	LON-MIDD	103.10	.86
WIRRAL	143.70	1.01	LON-CHNT	93.57	.83
SURREY	107.13	1.00	LON-LSL	94.90	.83
LON-RK	82.03	1.00	LON-CI	90.91	.79
KENT	156.26	.99	LON-KCW	106.85	.68
SCOTLAND	441.26	.49	SCOTLAND	424.50	.43

Figure 12: Mean migration distances and generalized distance decay parameters, inter-FPC area migration, by age group, 1980-81 and 1985-86



fluctuate with age. There are peaks in the profile associated with movers aged 5-9 and 15-19. A comparison of the figures for 1980-81 with those for 1985-86 indicates that the friction of distance became less important for the 20-24 age group but more important for those at other ages.

How well do spatial interaction models of this type fit aggregate NHSCR data sets in a historical context? Goodness-of-fit can be measured using a variety of statistical indices which compare matrices of predicted with observed migration. Two commonly computed measures, the mean absolute (percentage) deviation (MAD) and the coefficient of determination (R2), have been selected to illustrate the accuracy with which observed flows can be predicted. No attempt has been made to improve the spatial interaction model fits by adjustment of any of the distance values. The computed statistics in Table 4 show relatively little fluctuation from year to year in the historical models and demonstrate that origin-specific parameter calibration reduces the average deviation from around 40% to 38%. Further small improvements in fit are obtained when destination-specific parameters are calibrated.

The OPCS/DOE subnational migration projection methodology requires the calibration of model migration schedules for the standardized outmigration and inmigration rate profiles of each FPC area in 1980-81 (Martin, Voorhees and Bates 1981, Bracken and Bates 1983, Bates and Bracken 1987). Seven groups emerge when areas are then grouped together according to model parameters.

An alternative classification method is suggested here which derives groupings of areas on the basis of observed single year-of-age migration rate schedules for 1985-86. Age-specific rates within each cluster are then aggregated and a single cluster model migration schedule is calibrated. The initial step in the process is the computation of the Squared Euclidian Distance (SED) between two areas i and j:

$$SED_{ij} = \sum_{a} (m_i^a - m_j^a)^2 \qquad (3)$$

where m_i^a, m_j^a = the standardized inmigration rates by age group a for areas i and j.

The 'average linkage between groups' method is used where the distance (D) between two clusters, I and J, is the average distance between all pairs of FPC areas in which one member of the pair is from each cluster:

$$D_{IJ} = \sum_{i \in I} \sum_{j \in J} SED_{ij} \,/\, n_I n_J \qquad (4)$$

where n_I and n_J = the number of areas in clusters I and J.

Table 4: Goodness-of-fit statistics for inter-FPC area migration models, 1975-76 to 1985-86

YEAR	Generalized		Origin-specific		Destination-specific	
	MAD	R2	MAD	R2	MAD	R2
1975-76	40.35	0.7873	37.79	0.8052	36.46	0.8218
1976-77	40.37	0.7834	37.99	0.8006	36.70	0.8206
1977-78	40.32	0.7856	37.92	0.8023	36.62	0.8205
1978-79	40.34	0.7831	37.89	0.7999	36.66	0.8148
1979-80	40.13	0.7910	37.69	0.8098	36.54	0.8250
1980-81	40.16	0.7934	38.61	0.8109	36.28	0.8253
1981-82	40.41	0.7928	38.05	0.8068	36.51	0.8248
1982-83	40.61	0.7853	38.18	0.8037	36.84	0.8206
1983-84	39.49	0.7881	37.17	0.8015	35.94	0.8182
1984-85	38.25	0.7868	35.46	0.8045	34.11	0.8220
1985-86	39.06	0.7693	36.29	0.7865	34.92	0.8044

Type of Spatial Interaction Model; Type of fit statistic

Notes: (i) MAD is the mean absolute % deviation for the total number of non-zero cells defined as:

$$100 \left(\left(\sum_i \sum_j \left| \left(M_{ij}^{Obs} - M_{ij}^{Pred} \right) / M_{ij}^{Obs} \right| \right) / \left(N^2 - N \right) \right)$$

where N = the number of zones in the system

(ii) R2 is the coefficient of determination which measures proportion of total variation in the observed matrix which is explained by the predicted matrix

The agglomeration schedules produced in the clustering of the inmigration profiles indicate that the increase in the distance coefficient is fairly constant until around stages 75-80 of the procedure (15-20 cluster). After this point, increases become larger and less uniform suggesting that the association of areas becomes more irregular. A 'break' in the series can be identified as a point at which an optimum number of clusters has been reached and Table 5 outlines the FPC area composition of 15 clusters for inmigration. A different grouping is produced for outmigration. The age-specific migration rates for areas within each cluster are aggregated to form a 'cluster profile'. Finally, characteristic features of profiles can be quantified by calibrating a model migration schedule for each cluster using a version of the MODEL package developed by Rogers and Planck (1984) and operationalized at Leeds by Stillwell et al. (1987).

The MODEL package requires pre-selection of a schedule for each cluster with or without a retirement component. This choice proved difficult as some schedules showed evidence of a peak only at age 65. A full model schedule with a retirement component is defined as:

$$
\begin{aligned}
m_I^a = \; & b_1 \exp(-\alpha_1 a) \\
& + b_2 \exp\{-\alpha_2 (a - \mu_2) - \exp(-\lambda_2 (a - \mu_2))\} \\
& + b_3 \exp\{-\alpha_3 (a - \mu_3) - \exp(-\lambda_3 (a - \mu_3))\} \\
& + c \qquad (5)
\end{aligned}
$$

where m_I^a = the rate of migration into or out of one cluster I in age group a,

where the profile of the schedule is defined by seven of the eleven parameters $\alpha_1, \alpha_2, \mu_2, \lambda_2, \alpha_3, \mu_3, \lambda_3$.

and where the level of the schedule is determined by the remaining parameters b_1, b_2, b_3 and c.

A complete model schedule was only fitted to certain of the cluster profiles after careful inspection of the observed rates. A seven parameter model was fitted in other cases. The goodness-of-fit of each model schedule is measured using the E statistic defined as:

$$
E = 100 \sum_a |m_I^a(\text{mod}) - m_I^a(\text{obs})| \,/ \sum_a m_I^a(\text{obs}) \qquad (6)
$$

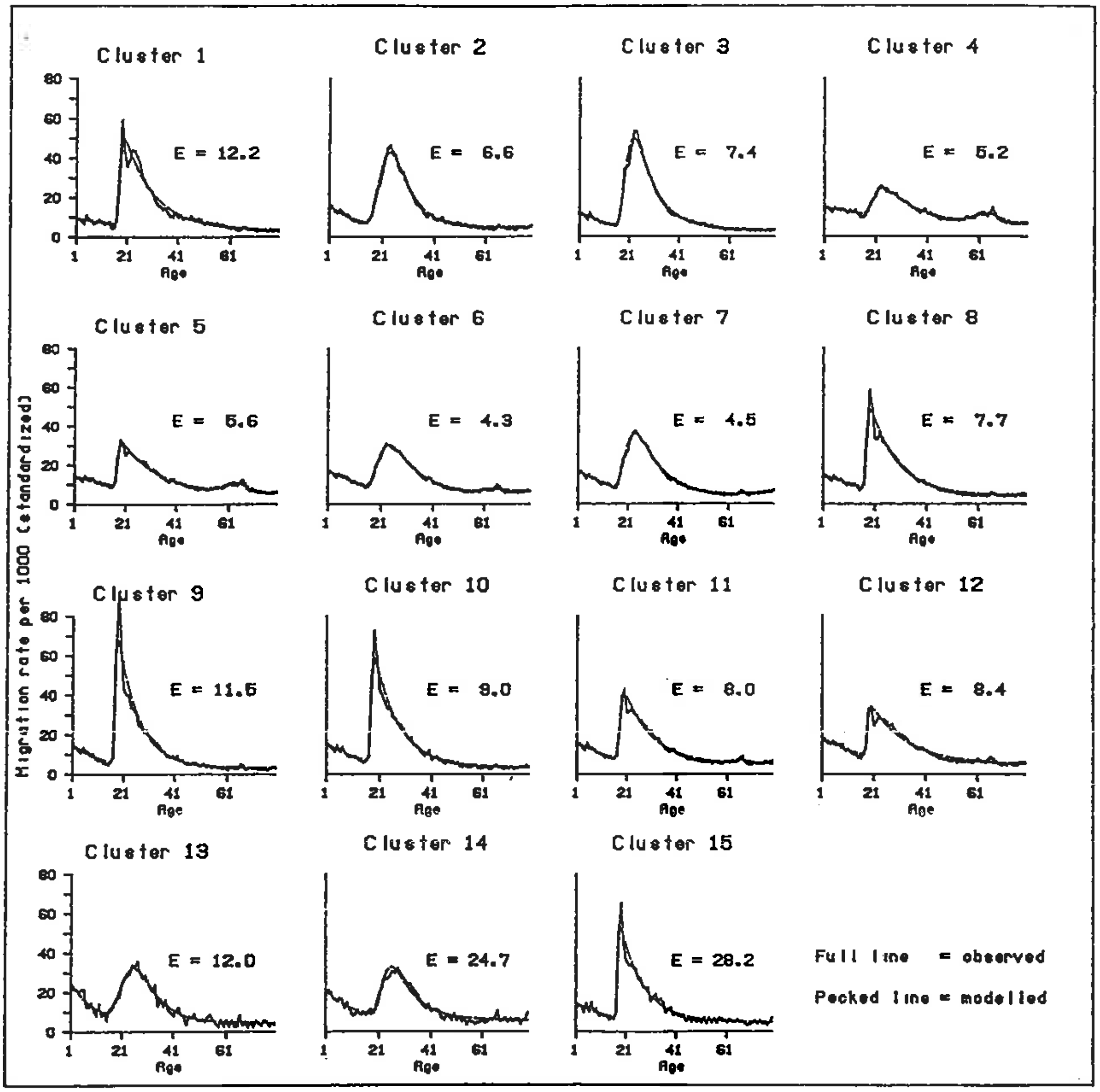
where $m_I^a(\text{mod})$, $m_I^a(\text{obs})$ are predicted and observed migration rates for age group a.

Figure 13 illustrates the model schedules superimposed on

Table 5: FPC area classification based on 1985-86 inmigration rates

1. Camden, Islington; Kensington, Chelsea, Westminster
2. Redbridge, Waltham Forest; Croydon; Bexley, Greenwich
3. City, Hackney, Newham, Tower Hamlets; Richmond, Kingston; Merton, Sutton, Wandsworth; Lambeth, Southwark, Lewisham; Middlesex
4. Lincolnshire; Suffolk; Isle of Wight; W. Sussex; Cornwall; Dorset; Somerset; Hereford; Salop; Clwyd; Powys
5. Humberside; Norfolk; E. Sussex; Devon; Lancashire; Dyfed; Gwynedd
6. Cumbria; Northumberland; Rotherham; Calderdale; Wakefield; Derbyshire; Northamptonshire; Buckinghamshire; Essex; Kent; Gloucestershire; Wiltshire; Warwickhire; Bolton; Oldham; Rochdale; Tameside; Sefton; Wirral; Cheshire; Gwent
7. Gateshead; N. Tyneside; Barnsley; Bedfordshire; Hertfordshire; Berkshire; Barking, Havering; Bromley; Dudley; Sandwell; Walsall; Bury; Stockport; Trafford; Wigan; St. Helens
8. Durham; Leicestershire; Nottinghamshire; Oxfordshire; Avon; Birmingham; Wolverhampton; Manchester; Salford; S. Glamorgan
9. Newcastle; Sheffield; Coventry
10. Leeds; Liverpool
11. Sunderland; Cleveland; Bradford; Kirklees; Cambridgeshire; Staffordshire; Mid-Glamorgan
12. Scotland; Doncaster; N. Yorkshire; Hampshire
13. S. Tyneside
14. Solihull
15. W. Glamorgan

Figure 13: Observed and model inmigration schedules for 15 clusters, 1985-86



the observed rates for inmigration. Apart from where model schedules are fitted to a single area, the goodness-of-fit is very good, with E values below 10.0 in most cases. The parameter values for each of the 15 cluster profiles are presented in Table 6.

The London FPC areas are divided into three distinct groups on the basis of the inmigration data. Camden, Islington, Kensington, Chelsea and Westminster, with a double peaking evident in the observed labour force curve, form the first group. The model schedule smoothes the curve and produces a relatively low μ_2 parameter, illustrating the importance of inmigration of those in their late teens. The high value emphasises the sharp increase in the rate of migration at age 18. The other two groups of London FPC areas have profiles with a high but later labour force peak. The labour force curves are rather more symmetrical with a less emphatic jump in the rate of migration on the upward slope. Clusters 4, 5 and 6 contain many of the rural counties and all show some evidence of a small peak in retirement inmigration. The labour force peaks for these groups are generally at a lower level than for the London groups, and significantly, these clusters exhibit a considerably higher child dependency index ($\gamma = b_1/b_2$) than the London groups. In cluster 7, a number of constituent areas show clear evidence of an upward retirement slope: Bedfordshire, Hertfordshire, Berkshire, Surrey, St. Helens, Trafford and Barking/Havering.

The schedules of clusters 8, 9 and 10 are dominated by very high peaks in their labour force curves at an early age and no retirement component. The sharp increase in the rate of inmovement around age 18 is emphasized by the relatively high λ_2 values giving significant asymmetry. The student factor may be important in determining the shape of the schedule for these three clusters which are all major university cities or counties. Clusters 11 and 12 contain rather unusual combinations of areas. Cambridgeshire stands out as having a unique profile with a sharp peak at age 19 (student factor again) and quite a significant retirement component. Areas of group 12 show evidence of a relatively flat labour force curve but Scotland, North Yorkshire and Hampshire have a localized peak at age 19, perhaps partly reflecting the inmigration of Armed Forces recruits. The only three areas failing to combine were South Tyneside, which had a very late labour force peak; Solihull, having a unique profile shape for a metropolitan area with a rounded labour force component; and West Glamorgan.

5 Conclusion

Despite its shortcomings, data from the NHSCR provides a wealth of valuable, up-to-date information on how migration propensities and patterns are changing from year to year. The sort of insights that have been gained from different types of empirical and model-based analysis reported in this paper are sufficiently interesting to warrant a more

Table 6: Parameters and parameter ratios for immigration clusters

PARAMETER	CLUSTER 1	2	3	4	5	6	7	8
b_1	0.007	0.013	0.010	0.010	0.010	0.014	0.016	0.012
α_1	0.039	0.108	0.091	0.031	0.046	0.066	0.058	0.059
b_2	0.055	0.087	0.088	0.026	0.027	0.043	0.058	0.050
μ_2	18.319	22.028	20.434	20.899	18.186	20.949	21.024	18.129
α_2	0.087	0.148	0.123	0.096	0.071	0.096	0.104	0.098
λ_2	1.514	0.289	0.404	0.403	1.580	0.352	0.335	1.978
b_3	.	.	.	0.000	0.000	0.004	0.002	.
μ_3	.	.	.	81.157	80.985	60.257	0.000	.
α_3	.	.	.	0.545	0.557	0.047	0.051	.
λ_3	.	.	.	0.099	0.101	0.384	0.000	.
c	0.003	0.005	0.003	0.005	0.005	0.004	0.002	0.004
γ	0.132	0.155	0.119	0.383	0.389	0.332	0.280	0.246
β	0.447	0.729	0.735	0.325	0.650	0.686	0.553	0.610
ε	17.492	1.960	3.275	4.195	22.114	3.655	3.214	20.280

PARAMETER	CLUSTER 9	10	11	12	13	14	15
b_1	0.014	0.014	0.013	0.015	0.024	0.018	0.011
α_1	0.091	0.081	0.077	0.064	0.105	0.095	0.078
b_2	0.076	0.064	0.042	0.029	0.067	0.050	0.054
μ_2	17.949	18.081	18.539	18.771	24.210	23.300	17.900
α_2	0.121	0.107	0.089	0.070	0.137	0.121	0.105
λ_2	2.745	2.159	1.304	2.512	0.210	0.400	3.060
b_3	.	.	0.003	0.005	.	.	.
μ_3	.	.	70.000	86.825	.	.	.
α_3	.	.	0.500	0.500	.	.	.
λ_3	.	.	0.227	0.623	.	.	.
c	0.003	0.003	0.005	0.005	0.004	0.006	0.005
γ	0.181	0.224	0.306	0.523	0.358	0.360	0.204
β	0.752	0.756	0.863	0.917	0.766	0.785	0.743
ε	22.747	20.106	14.594	36.123	1.533	3.306	29.143

comprehensive and detailed investigation of migration trends using the NHSCR data. In this context, a research project is currently underway under the auspices of the IBG Limited Life Working Party on 'Migration'. The aim of the project is to describe the changes in internal migration flows, structure, patterns and behaviour and its affect on population within Britain, and to provide an interpretation of those changes. Researchers from different institutions have been invited to contribute national, regional or systematic perspectives. The project will provide a historical context for future studies of migration data from the 1991 Census.

Acknowledgements

The research reported in this paper has been undertaken in collaboration with Dr Peter Boden and Dr Philip Rees at the School of Geography, University of Leeds. We are very grateful to OPCS for the provision of the data.

References

Bates, J.J. and Bracken, I. (1987) Migration age profiles for local authority areas in England, 1971-1981, Environment and Planning A, 19, 521-535

Boden, P. (1989) The analysis of internal migration in the United Kingdom using Census and National Health Service Central Register data, Unpublished PhD Thesis, School of Geography, University of Leeds

Boden, P., Stillwell, J.C.H. and Rees, P.H. (1987) Migration data from the National Health Service Central register and the 1981 Census: further comparative analysis, Working Paper 495, School of Geography, University of Leeds

Boden, P., Stillwell, J.C.H. and Rees, P.H. (1988) Linking Census and NHSCR migration data, Working Paper 511, School of Geography, University of Leeds

Boden, P., Stillwell, J.C.H. and Rees, P.H. (1990) Internal migration projection in England and Wales: the OPCS/DOE model examined, Chapter 11 in Stillwell, J.C.H. and Congdon, P. (eds) Modelling Migration: Macro and Micro Perspectives, Belhaven Press, London (In preparation)

Bracken, I. and Bates, J.J. (1983) Analysis of gross migration profiles in England and Wales: some developments in classification, Environment and Planning A, 15, 343-355

Bulusu, L. (1988) The use of NHSCR data in migration statistics, Paper presented at the Institute of British Geographers Population Study Group Workshop on 'Problems and Solutions of Working with the NHSCR data', University of Birmingham, (28 April)

Courgeau, D. (1980) Analyse Quantitative des Migrations Humaines, Masson, Paris

Devis, T (1984) Population movements measured by the NHSCR, Population Trends, 36, 15-20

Devis, T. and Mills, I. (1986) A comparison of migration data from the National Health Service Central Register and the 1981 Census, OPCS Occasional Paper 16, OPCS, London

Ledent, J. (1980) Multistate life tables: movement versus transition perspectives, Environment and Planning A, 12, 533-562

Martin and Voorhees Associates and John Bates Services (1981) Developing the migration component of the official sub-national population projections, Final Report prepared for DPRP3, DOE, London

Nam, C.B., Serow, W.J. and Sly, D.F. (eds) (1990) International Handbook on Internal Migration, Greenwood Press, Connecticut

Newton, M.P. and Jeffery, J.R. (1951) Internal migration: some aspects of population movements within England and Wales, SMPS No. 5. HMSO, London

Ogilvy, A. A. (1980) Inter-regional migration since 1971: an appraisal of data from the National Health Service Central Register and Labour Force Surveys, OPCS Occasional Paper 16, OPCS, London

Ogilvy, A.A. (1982) Population migration between the regions of Great Britain, 1971-78, Regional Studies 16(1), 65-73

OPCS (1984) National grid references of the centres of population of local authority districts, counties and regions, England and Wales 1981, and movements of centres of population 1971-81, OPCS Users Guide 214, HMSO, London

Rees, P.H. (1984) Spatial population analysis using movement data and accounting methods: theory, models, the 'MOVE' program and examples, Working Paper 404/Computer Manual 23, School of Geography, University of Leeds (Also Working Paper No. 50, Netherlands Interuniversity Demographic institute, Voorburg)

Rees, P.H. (1990) Britain's Population: A Geographical Analysis, Cambridge University Press, Cambridge (In preparation)

Ricketts, J. (1988) Family Practitioner records and the NHSCR, Paper presented at the Institute of British Geographers Population Geography Study Group Workshop on 'Problems and Solutions of Working with the NHSCR data', University of Birmingham, (April 28)

Rogers, A.A. and Planck, F. (1984) MODEL: a general program for estimating parametrized model schedules of fertility, mortality, migration and marital and labour force status transitions, Working Paper 83-102, IIASA, Laxenburg

Rowntree, J.A. (1957) Internal migration: a study of the frequency of movement of migrants, SMPS No. 11, HMSO, London

Stillwell, J.C.H. (1984) IMP: a program for inter-area migration analysis and projection: users manual (revised), Computer Manual 12, School of Geography, University of Leeds

Stillwell, J.C.H. and Boden, P. (1989) Internal migration: The United Kingdom, Chapter 4.1 in Stillwell, J.C.H. and Scholten, H.J. (eds) Contemporary Research in Population Geography: A Comparison of the United Kingdom and the Netherlands, Kluwer, Dordrecht, pp 63-71

Stillwell, J.C.H., Boden, P. and Rees, P.H. (1987) Migration schedule construction using MODEL and GIMMS, Computer Manual 29, School of Geography, University of Leeds

Stillwell, J.C.H., Boden, P. and Rees, P.H. (1990) Trends in internal net migration in the UK: 1975 to 1986, Area, 22(1), 57-65

Thomson, P. (1984) A comparison of the 1981 Census and the NHSCR as sources of migration data, Paper presented at the Conference on 'Census Analysis and Applications', University of Sheffield, (20-22 September)

Willekens, F. and Baydar, N. (1986) Forecasting place-to-place migration with generalized linear models, Chapter 9 in Woods, R. and Rees, P.H. (eds) Population Structures and Models, Allen and Unwin, London, pp 203-244

Wilson, A.G. (1974) Urban and Regional Models in Geography and Planning, Wiley, London