Combining available migration data in England to study economic activity flows over time

Peter W.F. Smith, James Raymer[1] and Corrado Giulietti
Southampton Statistical Sciences Research Institute, University of Southampton

11 March 2009

ABSTRACT

Information about internal migration in England may come from decennial censuses, surveys or population (health) registers. In this paper, we propose a methodology that allows us to combine aspects from multiple data sources to provide a time series of detailed migration flows. By detailed, we refer to a migration flow table cross-classified by origin, destination, age, sex and economic activity (e.g., employees, retirees or students). Our results can be used to analyse the movements of various population groups between counties in England over time.

Keywords: combining data, economic activity, England, internal migration, log-linear models

Acknowledgements**:** This paper was prepared for presentation at the 2009 Annual Meeting of the Population Association of America, Detroit, Michigan, 30 April - 2 May. Support for this research came from the Economic and Social Research Council (RES-000-22-2501). The data were provided by the Office for National Statistics and are Crown Copyright. The authors take full responsibility for the analyses and interpretations.

---

[1] Corresponding author: James Raymer, Southampton Statistical Sciences Research Institute, School of Social Sciences, University of Southampton, Southampton, SO17 1BJ, United Kingdom; telephone: +44 (0)23 8059 2935; email: raymer@soton.ac.uk

1. INTRODUCTION

In order to understand how a population responds to different economic situations, we need to have current and detailed information on migration patterns. For example, will employees and unemployed persons change their movements in response to an economic recession? Will there be more or less people migrating to become students in universities? Where will these students come from? What about the migration patterns of current retirees? Will they also change? These questions are relevant, particularly in times such as currently in the late 2000s when the economic situation in England is very uncertain. This paper provides a methodology to estimate migration flows that could help address these questions. Furthermore, it allows us to examine migration patterns of people at various stages in their life course.

In this paper, we develop a model that allows the combination of multiple sources of internal migration data for the purpose of studying how the internal migration patterns of six economic activity groups in England have changed over time. More specifically, we use three sources of internal migration data: the Patient Registry Data System (PRDS), which provides annual flows of migration by origin, destination, age and sex from 1999 to 2007; the 2001 Census, which provides the economic activity detail; and the Labour Force Survey (LFS), which contains quarterly information about the age, sex and economic activity of migrants but at very low geographic levels. The result is a synthetic data set of migration flows, cross-classified by origin, destination, age, sex and economic activity from 1999 to 2007.

The methodology described in this paper is a major extension of the framework provided by Raymer *et al.* (2007), which combined census and registration data to study elderly retirement and return migration over time across and twelve area groups in England and Wales. Our analysis combines three sources of migration data to study economic activity migration over time and across 47 counties in England for twelve age groups and by sex. The resulting data set can be used by local governments to improve their planning policies directed at supplying particular social services or at influencing levels of migration. This is important because migration is currently (and increasingly) the major factor contributing to population change at sub-national levels in England. Furthermore, our understanding of how or why populations change requires detailed information about migrants. Without these, the ability to predict, control or understand that change is limited. Finally, it is more economical and efficient to make the best use of available data rather than collect new data, which is costly and time-consuming.

To combine migration data from different sources, one first needs to account for the differences in measurement (see Bell *et al*, 2002; Long and Boertlein, 1990; Morrison *et al.*, 2004; Rogers *et al.*, 2003; Rogerson, 1990). For example, migration events, which can occur multiple times within a one year time period, are captured by population registration systems while changes in residential status (or transitions) from one point in time to another are captured by censuses (and surveys). These two data collection systems capture two different types of migration data, i.e., 'migrations' and 'migrants' (Rees and Willekens, 1986). Despite these conceptual differences, however, Boden *et al.* (1992) found high levels of correlation between in-migration, out-migration and net migration totals for England and Wales in the National Health Service Central Register (NHSCR) and census data. More recently, Raymer *et al.* (2007), in analysing elderly internal migration, found that that the main differences between the 2000-2001 NHSCR flows and the 2001 Census flows were in the levels of migration. The spatial patterns, on the other hand, were very similar when the levels were controlled for.

In the United Kingdom, there have been many studies that have examined or modelled internal migration flows (e.g., Bates and Bracken, 1982, 1987; Bell and Rees, 2006; Champion, 1996; Dixon, 2003; Kalogirou, 2005; Stillwell, 1994). Other studies have examined the determinants of internal migration (Fotheringham *et al.*, 2004) and the description of social change caused by international migration (Dorling and Rees, 2003; Rees and Butt, 2004), including the linkages between immigration and internal migration (Hatton and Tani, 2005; Stillwell and Duke-Williams, 2005). These studies have all relied on available data. They have not attempted to combine the various internal migration data sources available in the United Kingdom. Our research does.

2. AVAILABLE DATA

In England, internal migration data are available in several sources such as the decennial censuses, the PRDS and the LFS. Censuses contain much of the detail needed for analyses, but are only available every ten years and have problems with compatibility over time for certain variables (Stillwell and Duke-Williams, 2007). Migration data from the PRDS are available annually, but with minimal information on migrant behaviour (i.e., only origin, destination, age and sex are available) and with a tendency to miss important population groups, such as young adult males, who are known to be less inclined to register (Fotheringham *et al.*, 2004). However, the registration data constitute a good up-to-date source of internal migration as nearly all residents in England are patients of a general practitioner employed by the National Health Service, including those who may also have private healthcare provision. Furthermore, the average delay between moving house and registering with a new general practitioner is about one month (Office for National Statistics (ONS), 2002). Also, the patient register data are combined with information from the NHSCR to account for some missingness in the patterns. The LFS provides quarterly migration data with a rich detail of socioeconomic and demographic characteristics. These include, for example, ethnic group, country of birth, occupation and wages. The major disadvantage for the purpose of studying internal migration is that spatial information is only at regional level.

For this study, we estimate the 1999 to 2007 annual migration flows between the 47 counties of England as defined in the 1991 Census (see Figure 1) and for twelve five-year age groups (15-19, 20-24, ..., 70-74 years), two sexes and six economic activity groups (self-employed, employee, retired, inactive, unemployed and student). The Census and PRDS data were obtained from ONS by request. The 2001 Census migration data are contained in 'Table MG105'. The 1999-2007 annual PRDS data are included in tables 'Tab2aLA' (origin by destination) and 'Tab1LAQ' (origin by age by sex and destination by age by sex). Both the Census and PRDS data were tabulated at the local authority level. We used the '2001OA to 1991 Wards Lookup v1.0', provided by ONS on a CD-ROM, to aggregate these data to county level.

The 1999-2007 LFS data were downloaded from the Economic and Social Data Service website (http://esds.ac.uk/). The quarterly LFS migration data represent persons who changed their address in the past year, mixing residential movers and migrants. For our analysis, we assume that the age and sex profiles of these 'migrants' are the same as those moving between counties. We believe this assumption is not unreasonable based on a comparison of the age and sex profiles of persons changing address and those migrating between regions, obtained from the Sample of Anonymised Records from the 2001 Census. Finally, the LFS data were pooled to form yearly data, where each year begins on April, the same month as the Census.

Source: boundaries extracted from http://edina.ac.uk/ukborders/

Figure 1. Map of counties in England, 1991 definition

3. LOG-LINEAR MODELS FOR COMBINING DATA

3.1 Identifying Key Structures

Cross-classified tables are denoted by letters for the remainder of this paper. For example, OD is a two-way (origin by destination) table of migration flows, OAS is a three-way (origin by age by sex) table of migration flows and ODSE is a four-way (origin by destination by sex by economic activity) table of migration flows. Our aim is to estimate an ODASE table of migration flows by using information contained in lower order tables (e.g., OD and AS). Note, in the discussion below, we use the same notation to refer to lower order cross-classified tables and structures. For example, an OD table of migration can also be thought of as a partial table or a structure within the ODASE table.

The first step when combining migration data from multiple sources is to identify an overall model that could accurately predict the migration flows. This was undertaken by comparing various unsaturated log-linear model fits of two four-way migration flow tables, i.e., ODAS and ODSE, with the corresponding observed data, representing flows obtained from the 2001 Census. Note, the complete five-way table ODASE was not publicly available for disclosure reasons. This meant that we were unable to directly test the importance of age and economic activity.

Raymer *et al.* (2008), for interregional migration in England, found that a hierarchical log-linear model with the OD, OA, DA and AS structures provided a good fit to the four-way ODAS Census table. We found this also to be true for inter-county flows. For the ODSE table, we compared different unsaturated models and concluded that a hierarchical log-linear model with the ODE structure and the SE structure provided a good fit. This means that the various economic activity migrants have different spatial patterns and vary by sex, but do not have important differences represented by other structures in the ODSE table, such as OS, DS or OES. Although we were unable to formally analyse the significance of AE, we believe that this structure should be included in the overall model. For example, we know that the age profiles of retired migrants are very different to those of student migrants (for obvious reasons). To capture these important differences, we have to borrow this information from another source, such as the LFS. Therefore, we conclude that a good model for the five-way ODASE table would contain the following key structures: ODE, OA, DA, AS, AE and SE.

3.2 Model Specification

Our objective is to estimate migration flows for an ODASE table for each year from 1999 to 2007 with the diagonals of the OD partial tables (i.e., the within-region flows) excluded. The basic idea is to supplement information from the PRDS with more detailed information from the Census and LFS. The log-linear models developed by Raymer *et al.* (2007) and Raymer *et al.* (2008) are used as a starting point. These models combine marginal information available in the incomplete registration data with complete (but outdated) census data. In essence, the association structure of the census (auxiliary) data is imposed on the registration (incomplete) data. We extend this model to also include structures from a second auxiliary data source (i.e., the LFS).

Log-linear models can be thought of as spatial interaction models (Willekens, 1983). Spatial interaction models are commonly used to model origin-destination-specific migration flow data. Overviews of these models and frameworks can be found in Fotheringham *et al.* (2000, pages 213-235), Stillwell (2009) and Willekens (1999). A simplistic version of the spatial interaction model to estimate the number of migrations, n_{ij}^{OD}, in an incomplete data set, from origin i to destination j during a unit interval may be applied as in Willekens (1999):

$$\mu_{ij}^{OD} = \tau_i^O \tau_j^D m_{ij}^{OD}, \qquad (1)$$

where μ_{ij}^{OD} is the expected number of migration flows from origin *i* to destination *j* during the respective time interval and *i*, *j* = 1, 2,..., *R* for *R* origins and destinations. The τ_i^O and τ_j^D parameters represent background factors related to the characteristics of the origin and destination, respectively. The m_{ij}^{OD} factor is the auxiliary information on migration flows. This is additional data relating to migration between the same origins and destinations as in the incomplete data but is not a parameter in the model. As a result, the associations between origins and destinations in the auxiliary data are replicated in the estimated table of flows.

The above model focuses on estimating migration flows between two dimensions, origin and destination. Raymer *et al.* (2007) extended this model to include a third variable of interest not available in the incomplete migration data. For example, an origin by destination by economic activity table, with counts n_{ijz}^{ODE} can be modelled by using the following log-linear with offset form of the spatial interaction model:

$$\log \mu_{ijz}^{ODE} = \lambda_i^O + \lambda_j^D + \log m_{ijz}^{ODE}, \qquad (2)$$

where μ_{ijz}^{ODE} is the expected flows from origin *i* to destination *j* for level *z* of the third variable. The λ_i^O and λ_j^D parameters are related to the characteristics of the origin and destination, respectively, and m_{ijz}^{ODE} is the auxiliary information on migration flows. Note, there are no parameters corresponding to the dimension indexed by *z*. Here, we rely on the auxiliary data to provide the missing margin and association structures not contained in the incomplete data.

To fit our overall model identified in Section 3.1, we use the OD, OA, DA and AS tables from the PRDS data and impose the three-way ODE associations from the Census and the AE and SE associations from the LFS. (Note, as described later, the SE associations are taken from the LFS to allow these associations to vary over time.) This is achieved by fitting the following two log-linear models for n_{ijxyz}^{ODASE} , the counts in the five-way ODASE table:

$$\log\left(\mu_{ijxyz}^{*}\right) = \lambda_{xz}^{AE} + \lambda_{yz}^{SE} + \log\left(m_{ijz}^{ODE}\right) \tag{3}$$

$$\log\left(\mu_{ijxyz}^{ODASE}\right) = \lambda_{ij}^{OD} + \lambda_{ix}^{OA} + \lambda_{jx}^{DA} + \lambda_{xy}^{AS} + \log\left(\hat{\mu}_{ijxyz}^{*}\right). \tag{4}$$

Model (3) combines the LFS and Census data and provides an estimate of the ODASE table, $\hat{\mu}_{ijxyz}^{*}$, that has the same AE and SE associations as the LFS and the same ODE associations as the Census. Model (4) combines these association structures with the two-way OD, OA, DA and AS association structures from the PRDS. Should a different model for the flows be thought appropriate, then Models (3) and (4) can be modified by adding or removing interaction parameters, or by changing the offset terms, provided the pertinent information is available in one of the data sources.

Models (1) to (4) can be fitted using maximum likelihood estimation. It is straightforward to derive and solve, using an iterative procedure, the likelihood equations for these models to obtain estimates of the λ-parameters and flows. Raymer *et al.* (2007) did this for Models (1) and (2). However, since our interest is primarily in the estimation of the flows, we follow Raymer *et al.* (2008) and apply an iterative proportional fitting algorithm to obtain them instead. For Model (3), the initial values are given by the counts in the ODE table from the census: $\mu_{ijxyz}^{*(0)} = m_{ijz}^{ODE}$ for all *x* and *y*. They are then successively multiplied by adjustment factors so that the marginal tables match the counts in the LFS AE table and then the LFS SE table. This is repeated until the marginal tables of estimated flows simultaneously match all of the counts contained in the two

LFS tables. Model (4) repeats this process, using as initial values the $\hat{\mu}^{*}_{ijxyz}$ resulting from Model (3) and matching with the four PRDS two-way tables.

Raymer *et al.* (2007) assumed the three-way auxiliary interaction structure remained constant over time. Since Raymer *et al.* (2008) used information from two censuses, they allowed this structure to vary over time from 1991 to 2007 by geometrically interpolating the counts from 1992 to 2000 and by geometrically extrapolating forward from 2001. We, on the other hand, use the annual LFS data, to allow the association structure to evolve over time. This is achieved by fitting Models (3) and (4) for each year with the annual LFS data and PRDS data, respectively.

3.3 Model Implementation

The algorithm to fit Model (4) requires consistency in the marginal distributions of the incomplete data, namely of the OD, OA, DA and AS tables. Ideally, these would have come from a single four-way table. However, when we extracted the one-way margins from the OD, OAS and DAS tables provided by ONS, they did not match because the OAS and DAS tables included within county migration and migration to and from Wales, and the OD table included migrants aged 0-14 years and migrants aged 75+ years. To make them consistent, we used the following procedure. We started with the OD table and removed the diagonal elements and the rows and columns corresponding to areas in Wales. We then scaled the OA table so that its O margin matched that in the reduced OD table. Likewise, we scaled the DA table so that the D margin matched that in the reduced OD table. Next, we subtracted four age groups (i.e., 0-4, 5-9, 10-14 and 75+) from the scaled OA and DA tables, which were then used to scale the reduced OD and AS tables so that their O, D and A margins matched. Hence, all four tables required for modelling had the same one-way margins as required and had the necessary age groups for the study.

The PRDS migration data have a problem relating to the undercounting of young males, as seen in Figure 2 for 1999 and 2007. However, the 1991 and 2001 censuses indicated that the proportions of young adult male and female migrants were approximately equal. The reason for this difference has primarily to do with males being less likely to register, particularly in their young adult years (see Fotheringham *et al.*, 2004, pages 1637-1640). Note, this was not an issue in Raymer *et al.* (2007) because they only examined migration patterns of elderly persons, a group less likely to be missed in a health service population register.

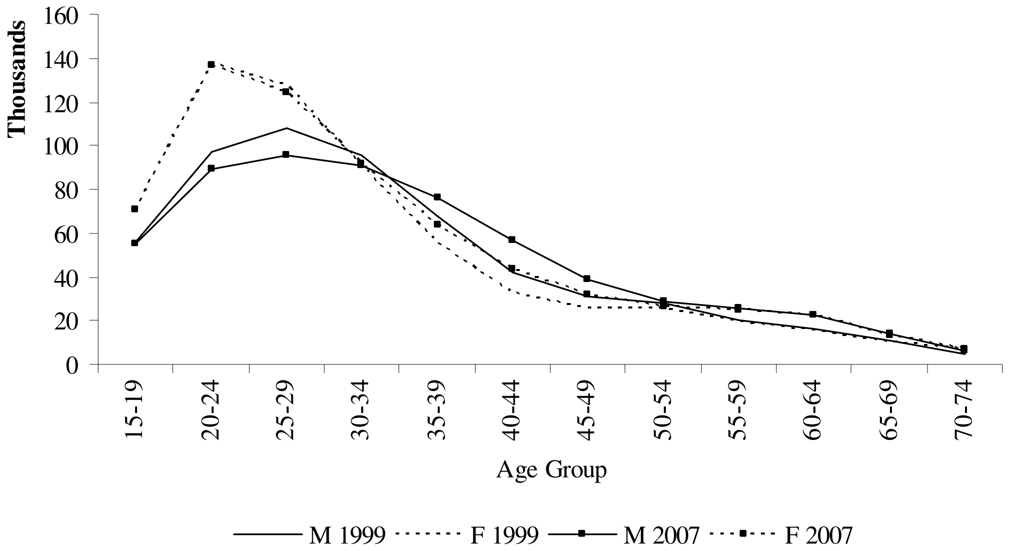

Figure 2. Age patterns of PRDS inter-county migration in England by sex, 1999 and 2007

As illustrated in Figure 2, nearly all the differences in the age patterns of male and female migration as reported in the PRDS data occur in the 15-19 year, 20-24 year and 25-29 year age groups. To correct for the differences in the age-sex patterns, we follow the procedure used in

Raymer *et al.* (2008), which assumes that females are counted accurately in the PRDS data and weights the estimates from Model (4) to account for the age-sex differences. The weights represent ratios of female to male migration for the 15-19, 20-24 and 25-29 age groups, marginalising over origin, destination and economic activity. This approach maintains all of the associations implied by Model (4). The weights applied to the male migrants in the three age groups are set out in Table 1, along with the resulting adjustment ratios for all males (a 13 to 15 percent increase) and both sexes together (a 6 to 7 percent increase).

Table 1. Adjustment ratios for PRDS migration data, 1999-2007

	Age Group (Males)			All	Both
Year	15-19	20-24	25-29	Males	Sexes
1999	1.264	1.422	1.182	1.131	1.063
2000	1.263	1.430	1.193	1.131	1.063
2001	1.277	1.441	1.187	1.134	1.064
2002	1.270	1.454	1.197	1.133	1.063
2003	1.289	1.468	1.211	1.137	1.066
2004	1.295	1.479	1.233	1.140	1.067
2005	1.287	1.454	1.231	1.140	1.067
2006	1.311	1.491	1.259	1.150	1.071
2007	1.289	1.523	1.295	1.153	1.072

4. PATTERNS OF ECONOMIC ACTIVITY MIGRATION

In this section, the estimated interregional migration flows by age, sex and economic activity group are presented. These estimates have been adjusted to correct for the male undercount in three age groups as discussed in the previous section. The full set of estimates produced by Models (3) and (4) consists of more than 2.8 million cells. The evolution of patterns over time is first described, followed by a focus on the age, sex and spatial patterns of employee, retired, inactive and student migrants from the Greater Manchester and Hampshire counties. These two counties and four economic activity groups were selected out of the set of estimates for illustration purposes only.

It is important to point out that the patterns described below are of economic activity groups, as measured at the time of the census or survey. There is no information on the economic activity status of these migrants prior to the move. Therefore, it is not possible to attach any causal relationships between economic activity and migration. For example, we do not know if an employed migrant was, say, employed, unemployed or a student prior to migrating. However, the patterns that we describe are informative about the movements of these persons prior to their current economic activity status and could be used to forecast future trends in economic activity by origin, destination, age and sex.

4.1 The Evolution of Economic Activity Over Time

The overall levels of inter-county migration by economic activity are set out in Figure 3 for the years 1999 to 2007. During this time, the total levels of migration increased from about 1.20 million to around 1.26 million. Employees make up the largest share of the total migrants. In 1999, this group represented about 50 percent of the flows. In 2007, this share was about 52 percent. Similarly, the other economic activity groups exhibited relatively small changes in their shares over time. Students and inactive persons represent the two other large flows. Students are counted at their term time address, rather than their home address in the 1991 Census. The inactive population consists of all individuals out of the labour force, such as those who take care of household members or those that are permanently sick or disabled. Flows of self-employed migrants are relatively small because this group only makes up a small portion of the population and possibly because their economic activity makes them less mobile than other groups. The share of unemployed as a total of the working age population (i.e., excluding male migrants aged over 64 and female migrants over 60) is on average five percent, in line with the national average unemployment rate (source: www.nomisweb.co.uk). Finally, retired migrants represent the smallest group. However, note that persons over 75 years are not included.

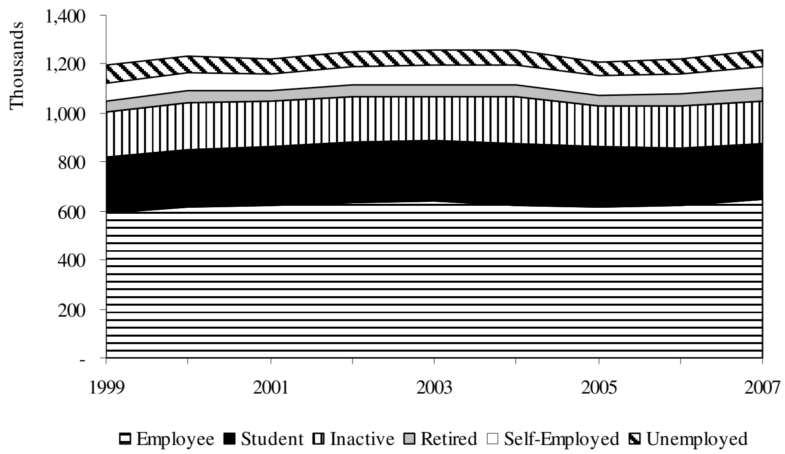

Figure 3. The levels of inter-county migration in England by economic activity group, 1999-2007

4.2 Spatial Patterns of Migrants from Greater Manchester and Hampshire Counties

The flows from Greater Manchester and Hampshire to the top ten destinations for employees, retired, inactive and students in 1999 are presented in Figures 4 and 5, respectively. The size of each arrow is proportional to the share of migrants (within each group) that move to a particular destination (refer to Figure 1 for names of counties).

The top left panel in Figure 4 presents the top ten flows of employee migrants from Greater Manchester. The two main destinations of Lancashire and Cheshire receive roughly 30 percent of all employee migrants. The other nearby destinations are Merseyside, West Yorkshire and Derbyshire. Not all employee migrants move to neighbouring counties, as demonstrated by the relatively sizeable flows towards Inner and Outer London and by the flows to Berkshire and Hampshire. The second panel illustrates the out-migration of retired migrants. Here, the largest share is to Lancashire, a well-known destination for retirees in the North of England. The remaining flows are distributed across several destinations in the North (e.g., Cumbria and North Yorkshire) and in the South (Cornwall & Isles of Scilly and Devon). The bottom left panel depicts migration of inactive migrants. The rather diverse type of destinations is indicative of the heterogeneity of this group; although the top four receiving counties (including London) are the same as those for employee migrants, other flows are directed to destinations that are more common for retired migrants, such as Cumbria and Somerset. The last panel presents student migrants. A large share of students had migrated to neighbouring counties that include university towns, such as West Yorkshire, Lancashire and Merseyside. The remaining flows are distributed across other counties where large institutions are located, such as Tyne and Wear, London and West Midlands.

Figure 5 presents the migration from Hampshire to the top ten destinations for the same four economic activity groups above. Here, employee migrants are more focused in terms of destination choice than those from Greater Manchester. This can be explained partly by Hampshire's relative proximity to London, which attracts about 30 percent of the flows of employee migrants. A relatively large share of employee migrants also moves to Surrey. The remaining flows are to neighbouring counties, such as Berkshire and West Sussex in the South East and Dorset and Wiltshire in the South West. The pattern of retired migrants is presented in the top right panel. The majority of retired migrants have moved to counties that are well-known tourist destinations in the South West region (Dorset, Devon and Cornwall & Isles of Scilly). The remaining top destinations are neighbouring counties, with the exception of Lincolnshire and Norfolk, which are relatively distant locations. Inactive migrants are represented in the bottom left panel. Similarly to the case of Greater Manchester, inactive migrants tend to move partly to destinations that are similar to employees (e.g., Surrey and Outer London) and partly to counties where retired migrants live (e.g., Dorset and Devon). The large majority of student migrants from Hampshire have moved to neighbouring counties, such as Surrey and Inner and Outer London, due to the presence of many university towns in these locations. A minor share of student migrants also moves to relatively distant locations, such as the West Midlands and Devon.

It is useful to investigate how the above spatial patterns evolved over time. In Figures 6 and 7, the migration patterns to the top five destinations for employee, retired, inactive and student migrants are presented for the 1999 to 2007 years for flows from Greater Manchester and Hampshire, respectively. The flow of employee migrants from Greater Manchester increased in all the receiving counties, except Outer London, where the levels remained constant over time. The patterns for retired migrants were irregular for the top two destinations (Lancashire and Merseyside), whereas they were pretty constant for the remaining counties. The patterns for inactive and student migrants were also fairly constant over time.

All of the top ten flows of employee migrants from Hampshire increased over time, except the one towards Surrey. The top ten flows of retired migrants, interestingly, exhibited decreases from 2000 to 2005 (except to West Sussex) but had similar levels in 1999 and 2007. The flows of inactive migrants from Hampshire were relatively constant over time, as were the flows of student migrants.

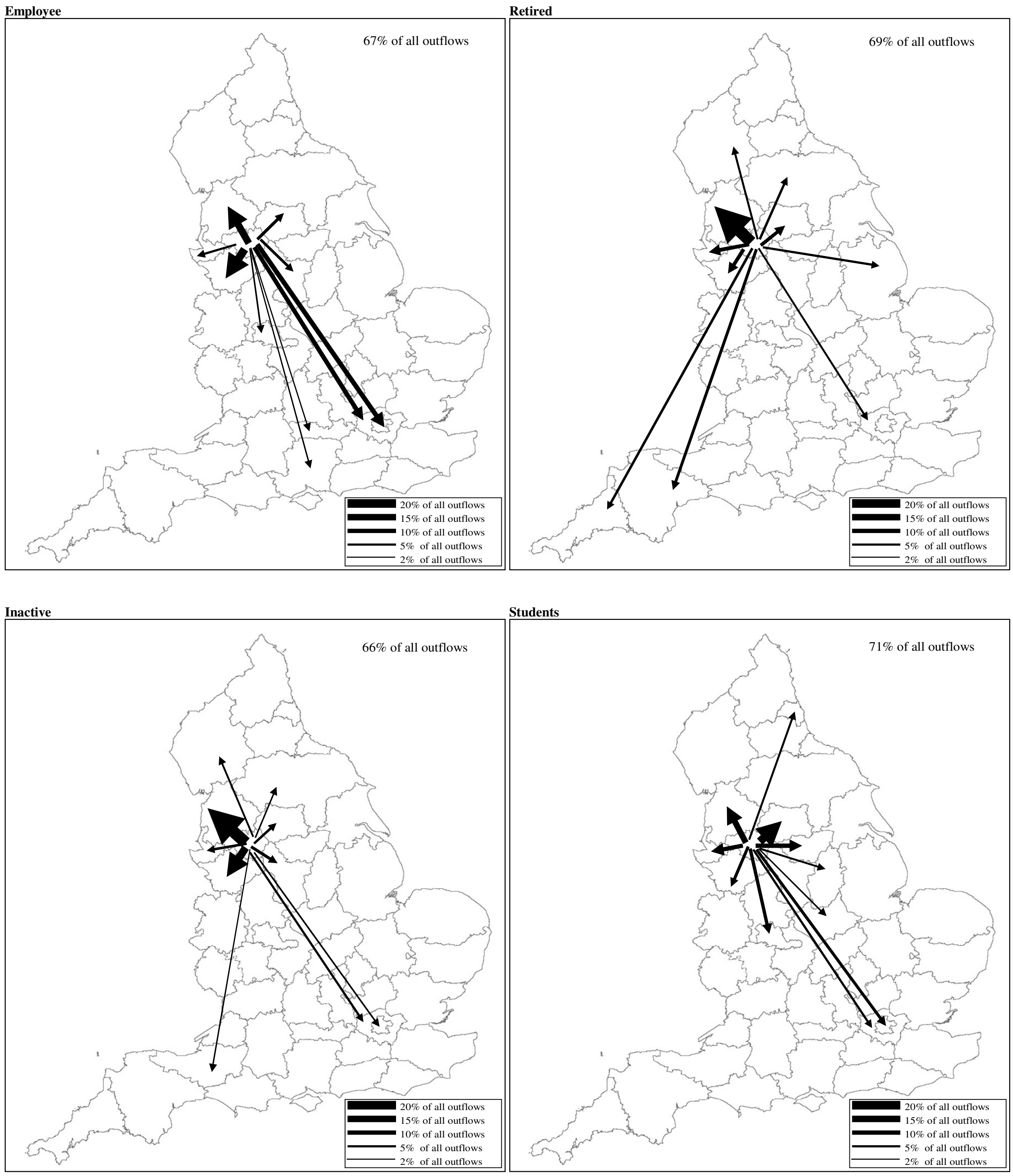

Figure 4. Top 10 destinations from Greater Manchester for employee, retired, inactive and student migrants, 1999

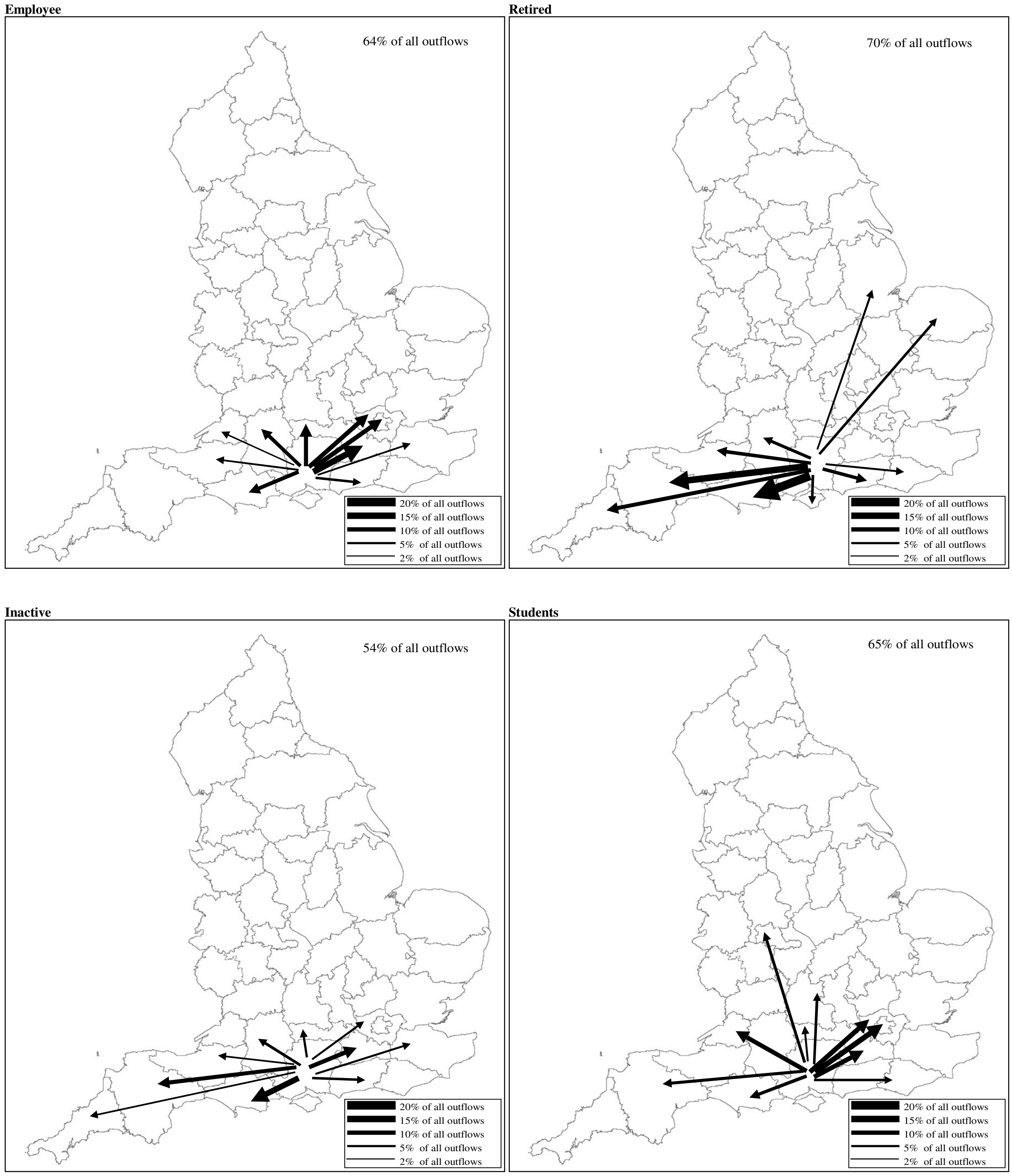

Figure 5. Top 10 destinations from Hampshire for employee, retired, inactive and student migrants, 1999

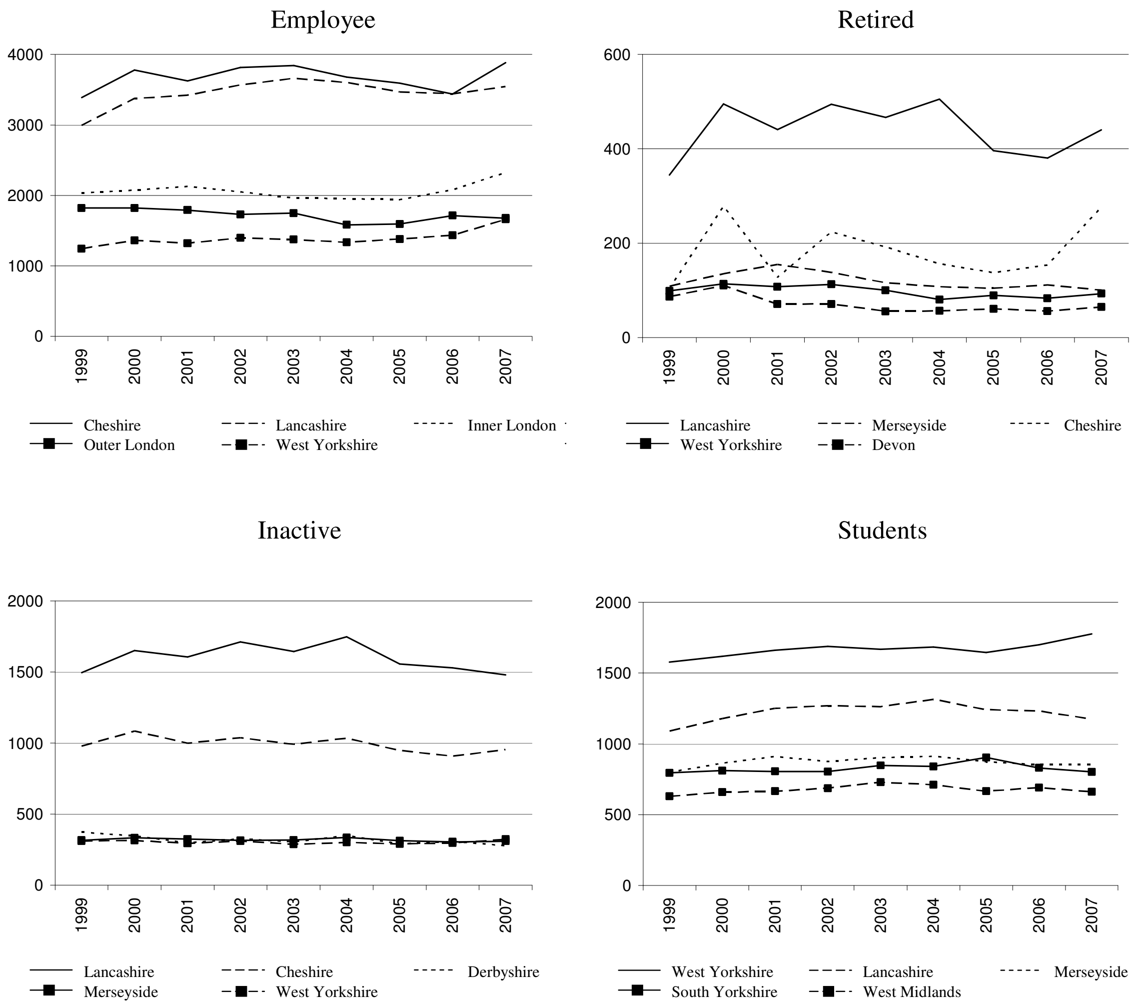

Figure 6. Top five destinations from Greater Manchester for employee, retired, inactive and student migrants, 1999-2007

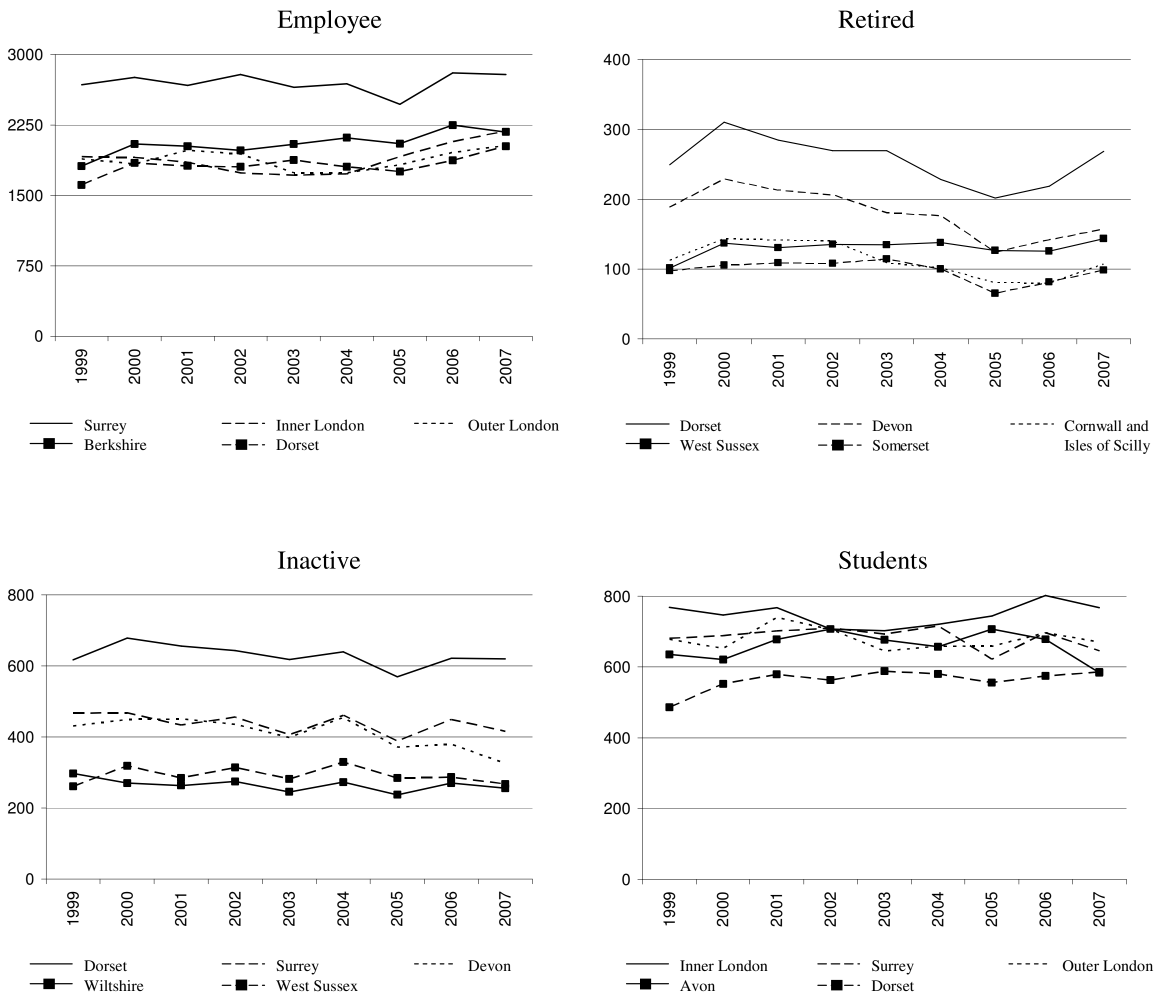

Figure 7. Top five destinations from Hampshire for employee, retired, inactive and student migrants, 1999-2007

4.2 Age- and Sex-Specific Patterns of Migrants from Greater Manchester and Hampshire Counties

In this section, the age- and sex-specific inter-county migration flows for employee, retired, inactive and student migrants are presented. For each of these economic activity groups, we first present the age profiles of male and female migrants to the top destination in 1999. Second, we present the corresponding female age profiles for 1999, 2003 and 2007 to examine changes over time.

The shapes of the employee migrant age profiles from Greater Manchester to Cheshire are very similar for males and females (see Figure 8), albeit with males exhibiting larger numbers in all age groups (except the oldest one). We expect male employee migrants to have higher levels of migration as they generally have higher levels of labour participation. The age profiles of these migrants exhibit a peak in the 25-29 year old age group, followed by a smooth decline towards the older age groups. The flows of retired migrants to Lancashire do not differ between sexes, except for the last age group, where there is a larger number of females, related to their lower mortality and higher population numbers. Flows of inactive female migrants to Lancashire are higher than for males. This is consistent with females being less likely to be in the labour force and more likely to be at home taking care of children. Flows of student migrants to West Yorkshire are represented in the bottom right panel. Levels for males and females (slightly higher) are very similar, with a peak in the 'college' years. The corresponding age profiles of migrants from Hampshire to top destinations are set out in Figure 9. These age-specific patterns are very similar to those from Greater Manchester (Figure 8), with the exception of student migrants to Inner London, who were estimated to have relatively lower numbers of migrants in the 15-19 year old age group.

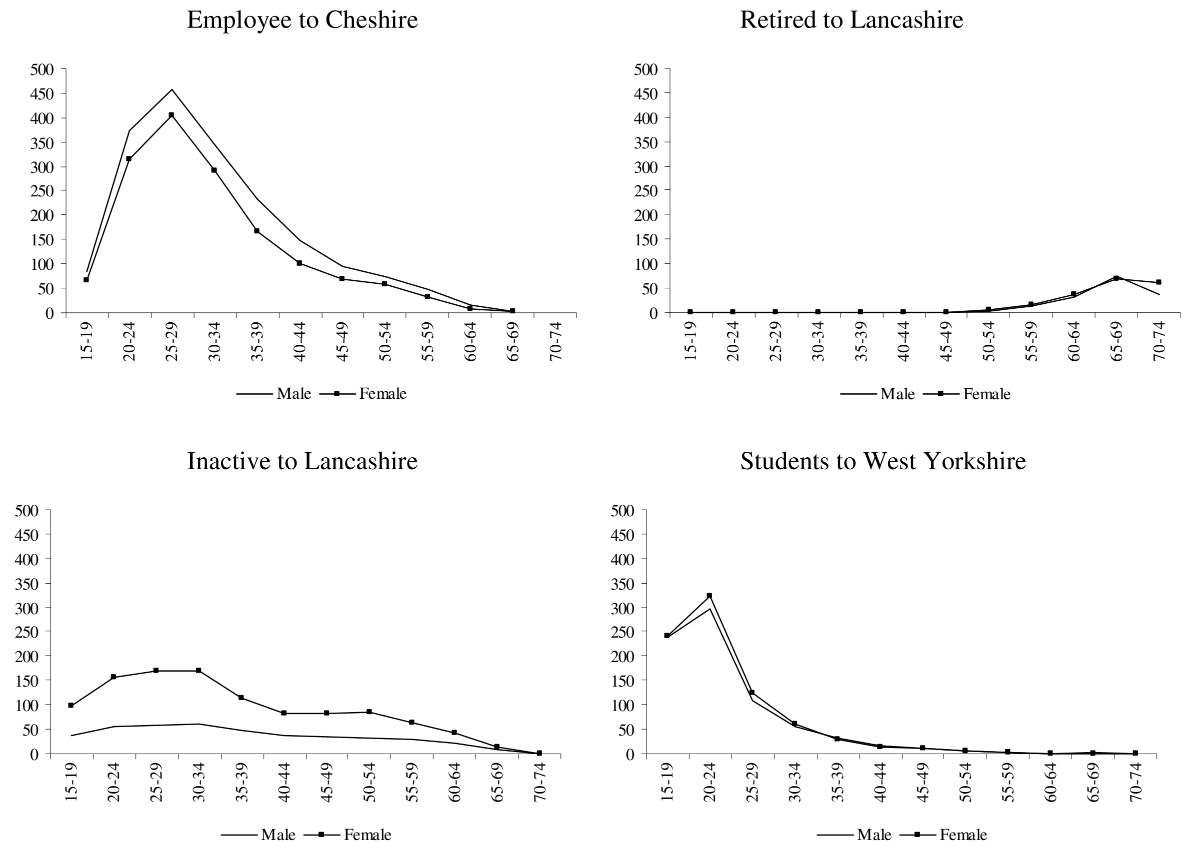

Figure 8. Age and sex patterns of migration from Greater Manchester: The top destinations for employee, retired, inactive and student migrants, 1999

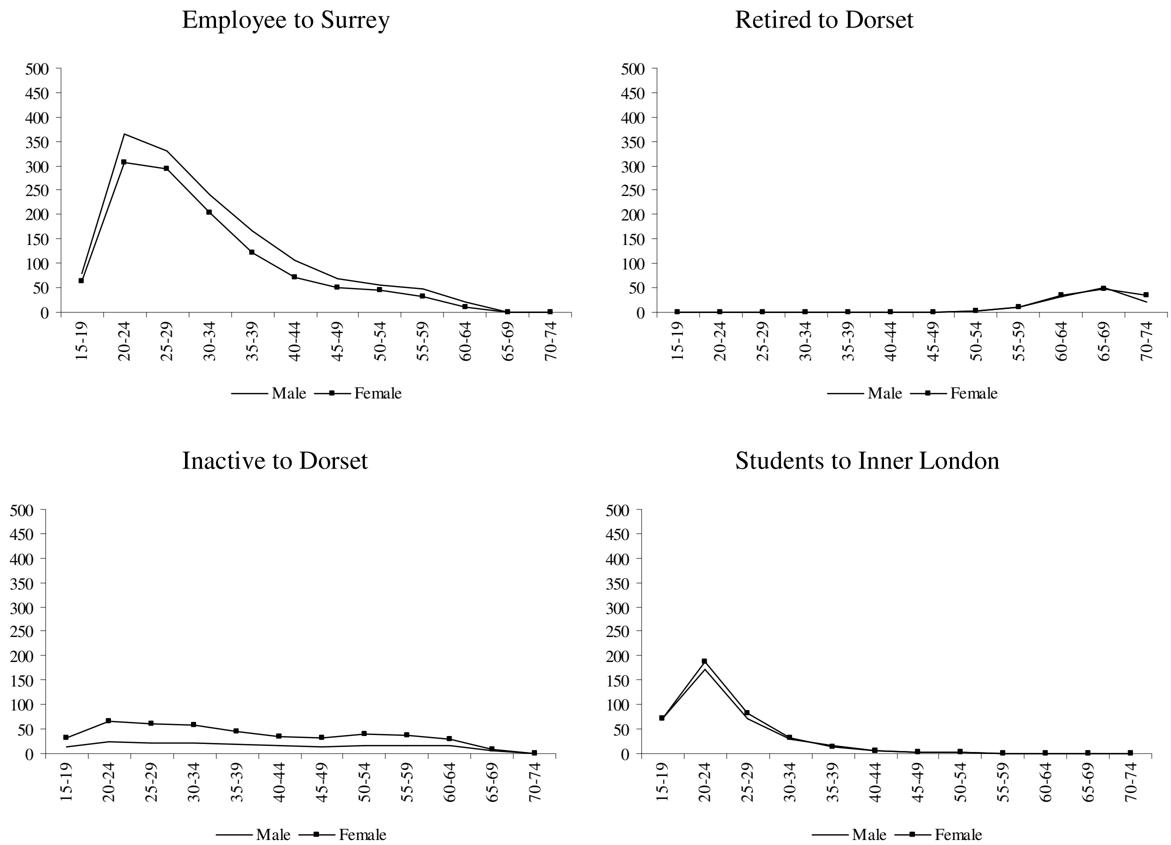

Figure 9. Age and sex patterns of migration from Hampshire: The top destinations for employee, retired, inactive and student migrants, 1999

To see how the predicted age patterns of migration changed over time, consider the female flows from Greater Manchester and Hampshire for the years 1999, 2003 and 2007 set out in Figures 10 and 11, respectively. In Figure 10, we see that the flow of employee migrants to Cheshire increased slightly over time, especially for age groups older than 30 years. The levels of retired migrants to Lancashire increased in the 60-64 year old age group, suggesting that females started retiring earlier. The pattern of inactive migrants to Lancashire is somewhat irregular over time, with some age groups gaining and others decreasing in numbers. On the other hand, there was a uniform increase in all the age groups for female student migrants to West Yorkshire. For the patterns set out in Figure 11, we see in the top left panel that female employee migrants to Surrey decreased in the age groups 20-24 and 25-29 years, but increased in the 35+ year old age groups. The migration of retired and inactive migrants to Dorset did not exhibit any major changes over time. Finally, there was a very slight increase in the flows of 25-29 year old student migrants to London in 2007 relative to the other two years.

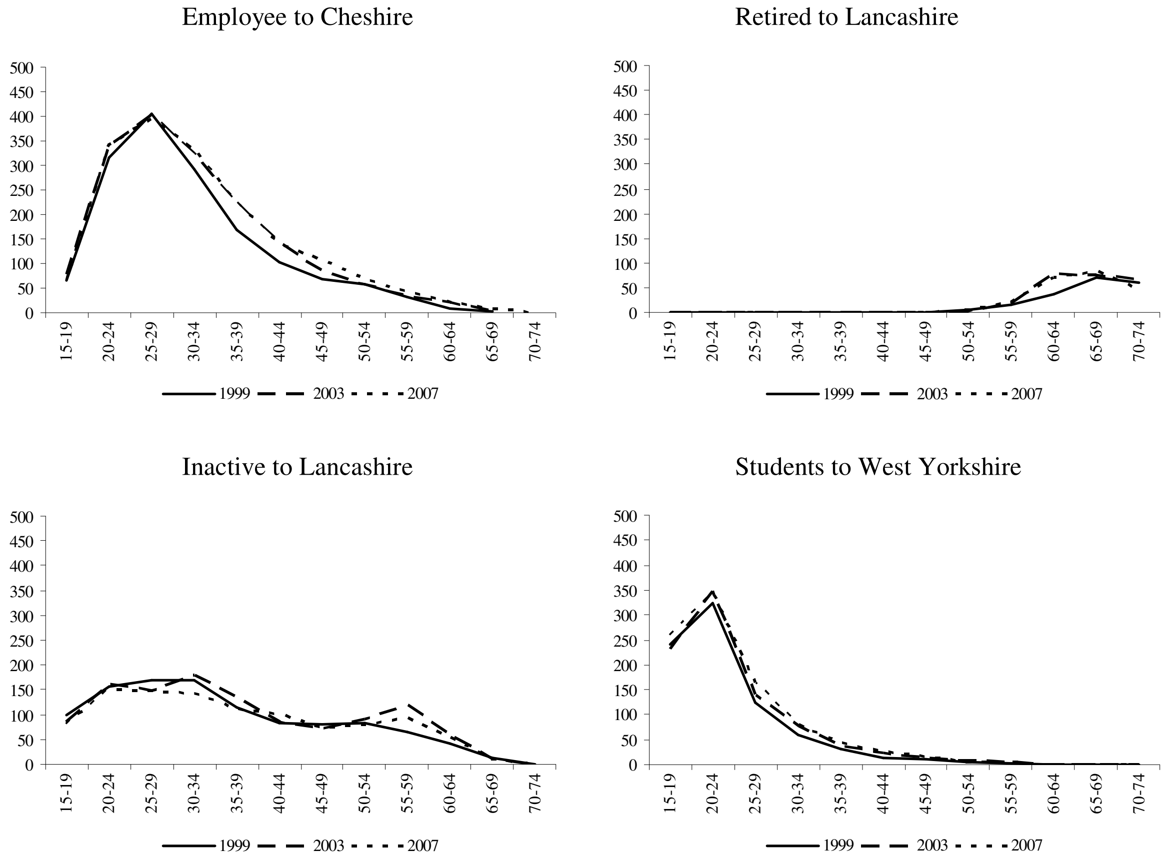

Figure 10. Age patterns of migration 1999, 2003 and 2007 from Greater Manchester: The top destinations for female employee, retired, inactive and student migrants

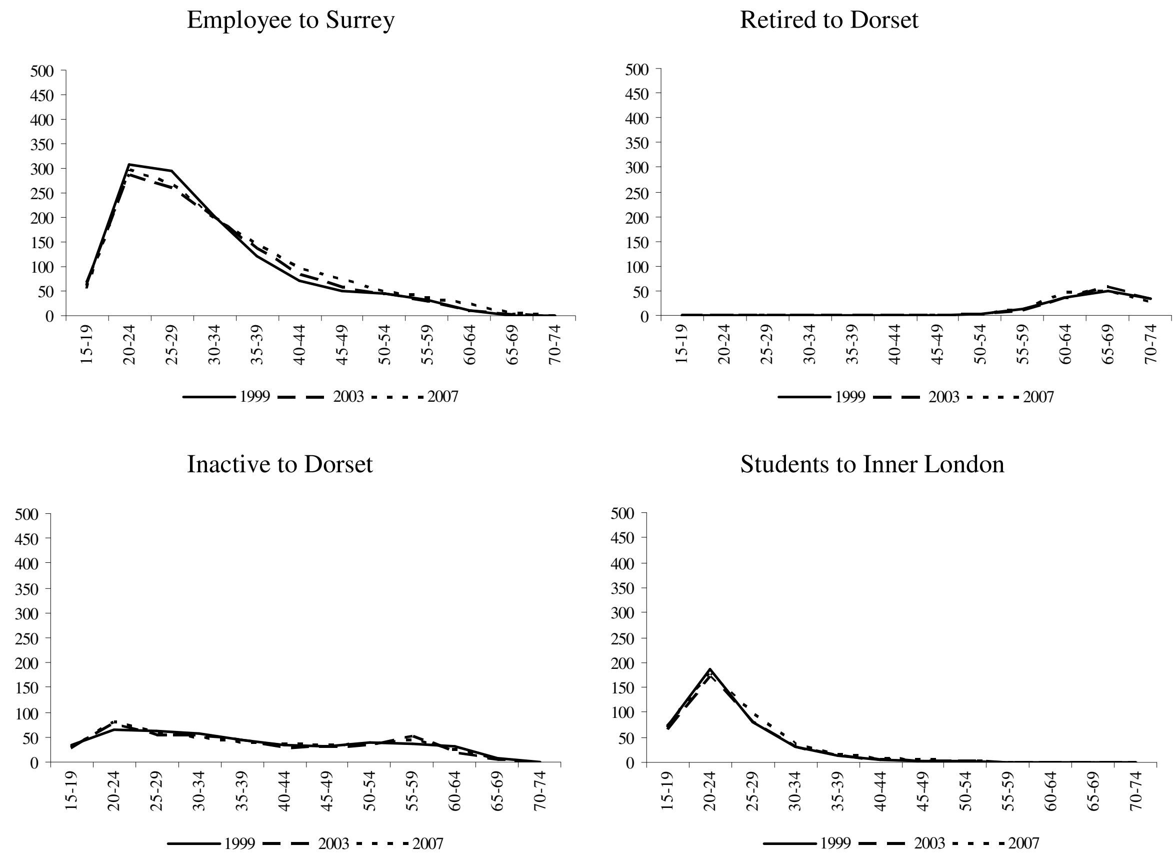

Figure 11. Age patterns of migration 1999, 2003 and 2007 from Hampshire: The top destinations for female employee, retired, inactive and student migrants

5. CONCLUSION

The methods described in this paper will help migration researchers and population planners make the best use of the data that are available to them, whether it comes from registrations, censuses or surveys. The above analysis has demonstrated the type of results that can be obtained from the estimated time series of economic activity migration flows in England by age and sex. It has allowed us to examine migration at various stages in the life course, demonstrating both the different age-specific shapes of migration according to economic activity, as well as the differences exhibited by males and females. For the examples described in this paper, we found that the estimated age, sex and spatial patterns of migration differed greatly across economic activity groups and demonstrated some interesting changes between 1999 and 2007. Note, we plan to make the entire estimated data set available in the near future so that others may analyse patterns of particular interest to them.

To conclude, population and migration analysts require detailed and up-to-date information to inform policy and planning. This information, however, is often not available. To overcome this limitation, we have introduced a methodology for combining migration data based on available registration, census and survey data. While we have used observed proportions and association structures from the census and LFS to estimate detailed migration flows, the methodology could also be used to create various scenarios, based on hypothetical proportions and association structures. For example, one could increase the proportion of student migrants over time and assess how this affects the overall migration patterns, assuming the other structures remain unchanged. This could be used to explore the consequences of more people investing in their education during times of recession. Alternatively, we could allow the age distribution of migrants to evolve in line with a projected ageing population (Little and Rogers, 2007).

REFERENCES

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

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

Bell, M., Blake, M., Boyle, P., Duke-Williams, O., Rees, P., Stillwell, J. and Hugo, G. (2002) Cross-national comparison of internal migration: Issues and measures. *Journal of the Royal Statistical Society Series A*, 165, 435-464.

Bell, M. and Rees P. (2006) Comparing migration in Britain and Australia: Harmonisation through use of age-time plans. *Environment and Planning A*, 38, 959-988.

Boden, P., Stillwell, J. and Rees, P. (1992) How good are the NHSCR data? In *Migration Processes & Patterns: Population Redistribution in the United Kingdom* (eds J. Stillwell, P. Rees, and P. Boden), pp. 13-27. London: Belhaven Press.

Champion, T. (1996) Population review: (3) Migration to, from and within the United Kingdom. *Population Trends*, 83, 5-16.

Dixon, S. (2003) Migration within Britain for job reasons. *Labor Market Trends*, April, 191-201.

Dorling, D. and Rees, P. (2003) A nation still dividing: The British census and social polarisation 1971-2001. *Environment and Planning A*, 35, 1287-1313.

Fotheringham, A.S., Brunsdon, C. and Charlton, M. (2000) *Quantitative Geography: Perspectives on Spatial Data Analysis*. London: Sage.

Fotheringham, A.S., Rees, P., Champion, T., Kalogirou, S. and Tremayne, A.R. (2004) The development of a migration model for England and Wales: Overview and modelling out-migration. *Environment and Planning A*, 36, 1633-1672.

Hatton, T.J. and Tani, M. (2005) Immigration and inter-regional mobility in the U.K., 1982-2000. *The Economic Journal*, 115, F342-F358.

Kalogirou, S. (2005) Examining and presenting trends of internal migration flows within England and Wales. *Population, Space and Place*, 11, 283-297.

Little, J.S. and Rogers, A. (2007) What can the age composition of a population tell us about the age composition of its out-migrants? *Population, Space and Place*, 13, 23-39.

Long, J.F. and Boertlein, C.G. (1990) Comparing migration measures having different intervals. Current Population Reports No. 166, U.S. Census Bureau, Washington, DC.

Morrison, P.A., Bryan, T.M. and Swanson, D.A. (2004) Internal migration and short-distance mobility. In *The Methods and Materials of Demography* (eds J.S. Siegel and D.A. Swanson), pp. 493-521. San Diego: Elsevier Academic Press.

Office for National Statistics (2002) Using patient registers to estimate internal migration: Customer guidance notes. Migration Statistics Unit, Office for National Statistics, Titchfield. (Available from http://www.statistics.gov.uk/statbase/expodata/commentary/WebGuidanceNotes.htm)

Raymer, J., Abel, G. and Smith, P.W.F. (2007) Combining census and registration data to estimate detailed elderly migration flows in England and Wales. *Journal of the Royal Statistical Society A*, 170, 891-908.

Raymer, J., Smith, P.W.F. and Giulietti, C. (2008) Combining census and registration data to analyse ethnic migration patterns in England from 1991 to 2007. Methodological Working Paper M08/09, Southampton Statistical Sciences Research Institute, University of Southampton.

Rees, P. and Butt, F. (2004) Ethnic change and diversity in England, 1981-2001. *Area*, 36, 174-186.

Rees, P. and Willekens, F. (1986) Data and accounts. In *Migration and Settlement: A Multiregional Comparative Study* (eds A. Rogers and F. Willekens), pp. 19-58. Dordrecht: D. Reidel.

Rogers, A., Raymer, J. and Newbold, K.B. (2003) Reconciling and translating migration data collected over time intervals of differing widths. *The Annals of Regional Science*, 37, 581-601.

Rogerson, P.A. (1990) Migration analysis using data with time intervals of differing widths. *Papers of the Regional Science Association*, 68, 97-106.

Stillwell, J. (1994) Monitoring intercensal migration in the United Kingdom. *Environment and Planning A*, 26, 1711-1730.

Stillwell, J. (2009) Inter-regional migration modelling: A review. In *Migration and Human Capital: Regional and Global Perspectives* (eds J. Poot, B. Waldorf and L. van Wissen), Chapter 2. Cheltenham: Edward Elgar.

Stillwell, J. and Duke-Williams, O. (2005) Ethnic population distribution, immigration and internal migration in Britain: What evidence of linkage at the district scale? Paper presented at the British Society for Population Studies Annual Conference, University of Kent, Canterbury, 12-14 September.

Stillwell, J. and Duke-Williams, O. (2007) Understanding the 2001 UK census migration and commuting data: The effect of small cell adjustment and problems of comparison with 1991. *Journal of the Royal Statistical Society Series A*, 170, 1-21.

Willekens, F. (1983) Log-linear modelling of spatial interaction. *Papers of the Regional Science Association*, 52, 187-205.

Willekens, F. (1999) Modeling approaches to the indirect estimation of migration flows: From entropy to EM. *Mathematical Population Studies*, 7, 239-278.