Estimates of internal migration flows for the UK, 2000–2007

Adam Dennett and Phil Rees
University of Leeds

Abstract

Statistics on migration flows year by year within the UK are produced by the Office for National Statistics, the General Register Office for Scotland and the Northern Ireland Statistics Research Agency for migration within England and Wales, within Scotland and within Northern Ireland respectively. However, these flow statistics are not integrated across the UK. As there was a need for such integrated flow statistics at a sub-national scale known as NUTS2 for an EU sponsored project, the authors developed a synthetic estimate of migration flows for the calendar years 2000 to 2007 and the mid-year to mid-year intervals 1999–2000 to 2006–07. The estimates were controlled by the migration flows published at NUTS1 scale from the UK wide NHS Central Register to which country specific flows between NUTS2 regions from the various patient registers were fitted. The gaps, flows between regions in different devolved territories, were filled by adjusting comprehensive flow data from the 2001 Census to the published NHSCR flows. Age detail was added only to the total out-migration and in-migration flows from each region using a fixed national profile of migration rates by age from the 2001 Census. The resulting time series of flow data provide, therefore, the three matrix faces of an origin-destination-age array for each time period. The paper describes the details of the estimation process and reports on some of the trends that the time series show.

Contents

Abstract	82
Introduction	85
The spatial system.....	86
Data sources	87
Estimating migration flows for the UK.....	89
The estimation scheme	89
Limitations	99
Results.....	100
Conclusions	103
Key findings	104
Acknowledgements	104
References	105

List of figures

Figure 1	The estimation scheme	90
Figure 2	Examples of the migration, ancillary matrices and synthetic estimates	94
Figure 3	Migration probability by age for men and women.....	95
Figure 4	Variation in age-specific intra-country migration rates by UK constituent country, 2001 Census	100

List of tables

Table 1	NUTS region population thresholds and areas in the UK.....	86
Table 2	Data sources used in the estimation process.....	88
Table 3	Variables and indexes used in the estimation procedures	91
Table 4	Diagnostic statistics for different estimation schemes	93
Table 5	Example of single age to broad age proportions.....	96
Table 6	In-migration to NUTS2 regions by broad age group, 2000.....	96
Table 7	Net migration rates across NUTS2 regions, 2000–2007 (ranked by average).....	101
Table 8	Turnover rates across NUTS2 regions, 2000–2007 (ranked by average).....	102

Introduction

The European Commission and its statistical agency, Eurostat, has responsibility for monitoring, projecting and assessing demographic developments across Europe at member state and various regional scales. At regional scales it is essential to monitor internal migration between regions. The smaller the regions are, the more important is internal migration as a component of population change. The European Commission and Eurostat need robust time series of internal migration flows for the United Kingdom. Such a time series is not available for the UK for several reasons. There is a division of responsibility for many population statistics between three agencies¹. Migration data are produced from several different sources: UK-wide censuses, UK-wide registers, home country specific registers and surveys of households and individuals, which do not necessarily connect in a seamless fashion. This paper uses available published data sets on migration to create a synthetic estimate of migration within the UK in the current decade. This estimate is needed in an EU sponsored project and by Eurostat for future regional population projections.

In 2007 the European commission approved the new programme for the European Spatial Planning Observation Network (ESPON). The ESPON 13 programme has a broad remit of supporting policy development in relation to the harmonious and cohesive development of a European territory. ESPON 13 funds a number of applied research projects which aim to build a European-wide information and evidence base on a range of areas likely to influence this development, such as climate change, energy, transport, agriculture and population. DEMIFER – **DE**mographic and **MI**gratory **F**lows affecting **E**uropean **R**egions and **C**ities² – is one of these applied research projects with an aim to examine current demographic trends across European regions and forecast how population growth and changes will affect these areas; framing these forecasts within different scenarios which could affect demographic trends moving towards 2050.

One of the first objectives of DEMIFER was to build a database of fertility, mortality, inter-regional migration and international migration data for the 27 member states of the European Union, plus 4 other European states, for the NUTS2 geography. UK demographic data are not available at this geographical level, so new estimates were needed to fulfil the requirements of the project. One of the more challenging datasets to produce was the internal migration (inter-NUTS 2) data for the UK. The project requires that a time series of NUTS2 level inter-regional aggregate flow matrices be produced for calendar years from 1999 to 2007, with accompanying NUTS2 region inflow, outflow and net flow counts, disaggregated by single year of age. With these data not being available, a new methodology needed to be developed to produce a set of estimates for use in the project and this is the main motivation for this paper. An additional motivation was to develop a new methodological framework which could be adopted to produce similar national estimates for alternative geographies. For example, work has recently been undertaken to estimate national flows at the local authority district geography.

This paper details the estimation of these new NUTS2 data, with a plan as follows. The second section describes the spatial system, linking the UK's administrative geography with the Eurostat scheme for harmonized regions across the European Union. The third section describes the data sources available for the estimation. The fourth section discusses the conceptual framework for the estimation, the three "faces" (matrices) of an origin-destination-age array of migration flows. The fifth section describes the details of the estimation as a series of six steps. In the sixth section of the paper the method used to convert from mid-year to mid-year time intervals to calendar years is

outlined. The seventh section discusses the limitations of our estimates. The eighth section briefly describes some trends that characterize the new time series, while in the final section we draw some general conclusions.

The spatial system

In the 1970s, Eurostat proposed a single spatial system for dividing up the territory of Europe for statistical purposes³. The Nomenclature of Statistical Territorial Units (NUTS) system was introduced and remains in use today, albeit with a number of amendments to the initial incarnation. European regulation has given the NUTS system legal status, and with it certain defining characteristics. A hierarchical division of areas, the NUTS system has levels 1, 2 and 3 (previously 4 and 5 existed as well, although these have now been replaced by Local Administrative Unit – LAU – levels 1 and 2 since July 2003), each level being defined by certain maximum and minimum population thresholds. For ease of implementation, NUTS regions link to the existing administrative divisions in European states. For the UK, this means that the NUTS geography fits into the administrative geography described in **Table 1** –although Northern Ireland, whilst occupying its own NUTS1 level, does not fit the prescribed population limits.

The DEMIFER project is concerned with data at the NUTS2 level; the NUTS2 regions in the UK as currently defined by Eurostat are listed in **Table 4** (Eurostat provides an online boundary map³). It has been necessary, however, to slightly alter the boundaries of these UK NUTS2 regions for this project, resulting in a modified NUTS2 spatial system. The reason for this is that internal migration data from the 2001 Census at NUTS2 level will be important in the estimation of internal migration flows. These data are readily available for the whole of the UK from the Centre for Interaction Data Estimation and Research (CIDER)⁴ at the district (local authority/unitary authority/council area) scale, so an aggregation of these census data from district to NUTS2 is necessary.

Table 1 NUTS region population thresholds and areas in the UK

Level	Minimum Population	Maximum Population	England	Wales	Scotland	N Ireland	Total in UK	Example
NUTS 1	3 million	7 million	GOR	Wales	Scotland	N Ireland	12	UK1 – London
NUTS 2	800,000	3 million	Counties/Groups of Counties	Groups of Unitary Authorities	Combinations of council areas, LECs and parts thereof	N Ireland	37	UK11 – Inner London
NUTS 3	150,000	800,000	Counties/Groups of Unitary Authorities	Groups of Unitary Authorities	Combinations of council areas, LECs and parts thereof	Groups of District Council Areas	133	UK111 – Inner London – West
LAU 1	n/a	n/a	Districts/UAs	Unitary Authorities	Combinations of council areas, LECs and parts thereof	District Council areas	433	UK11101 – City of London
LAU 2	n/a	n/a	Electoral Wards/ Divisions	Electoral Divisions	Electoral wards or, rarely, parts thereof	Electoral Wards	10,664	UK11101001 – Aldersgate

Note: NUTS = *Nomenclature d'Unités Territoriales Statistiques* or Nomenclature of Territorial Units for Statistics

Source: http://epp.eurostat.ec.europa.eu/portal/page/portal/region_cities/regional_statistics/nuts_classification and www.statistics.gov.uk/geography/nuts.asp

This aggregation is most easily achieved through the use of a look-up table identifying the NUTS2 region each district in the UK falls within. In almost all cases the districts in the UK fit inside a single NUTS2 region, one district however, is split between two NUTS2 regions causing a problem. The district in question is the Scottish Council Area of North Ayrshire. The Isle of Arran is defined as part of NUTS2 region UKM6, whereas the mainland section of North Ayrshire is defined as part of UKM3. This presents the problem of how to allocate data from one district into two NUTS2 regions appropriately.

A number of potential solutions present themselves. One solution would be to take internal migration data from a lower geographical level, such as the ward level. As well as requiring a much larger data file (10,000+ × 10,000+ wards versus 400 × 400 districts), taking data from the ward level is likely to reduce the accuracy of flow data, especially where the flows are between Scotland and England and Wales thanks to the effect of small cell adjustment (SCAM)⁵, consequently this solution was ruled out. Alternatively, flows at the district level could be re-allocated using flow ratios from ward level data, but for an accurate estimate these proportions would need to be known for every ward level origin/destination pair – another large undertaking, thus also ruling this solution out. Analysis of the district's population falling into each NUTS2 region produced a third, simpler solution to the problem. Data from the 2001 Census reveals that the Isle of Arran contains just 3.72 per cent of the population of North Ayrshire, with the mainland section of the Council Area containing 96.28 per cent of the population. With such a large majority of the population of the district falling into the UKM3 region, it was likely that most flows for the district would occur between the mainland section and other districts. The simplest solution to the problem, therefore, was to allocate the whole of North Ayrshire to UKM3, in effect redefining the boundary of the NUTS2 zone for the purposes of this project.

Data sources

The estimates produced make use of a number of different internal migration and population data sources. These are outlined in **Table 2**.

The 2001 Census Special Migration Statistics (SMS) contain information about the inter- and intra-district origins and destinations of migrants in the year preceding the 2001 Census. Aggregated to NUTS2 regions, the SMS provide the only complete inter-NUTS2 migration matrix for the whole of the UK.

Patient register migration (PRDS) data for England and Wales are derived from the extracts of the patient registers held by Health Authorities^{6,7}. Where patients change residential address in any particular reference year, this transition is captured, with the results collated, adjusted to National Health Service Central Register (NHSCR) movement data internal migration estimates and disseminated at the district level by ONS for mid-year periods. The Centre for Interaction Data Estimation and Research provides online access to these data, along with the facility of aggregate to the NUTS2 level. Similar to the Patient registers in England and Wales, the Community Health Index (CHI) in Scotland also captures the movements of NHS patients. GRO Scotland provides collated mid-year moves between Scottish Council areas. Unfortunately neither the patient register data produced by ONS nor the CHI data from GRO Scotland records cross-border moves between districts in England and Wales and council areas in Scotland.

Inter-Government Office Region (NUTS1) movement data from the NHSCR are published by the ONS for rolling annual periods every three months. As such these include both mid-year

Table 2 Data sources used in the estimation process

Data set	Spatial scale and coverage	Time period(s)	Variables	Source
MIGRATION INFORMATION				
2001 Census Special Migration Statistics	LAU1 – District (aggregated to NUTS2) UK coverage	Mid 2000–01	Total migrants	http://cider.census.ac.uk
Patient register migration data	District (aggregated to NUTS2) England and Wales coverage	Annual data from Mid 1999 – 2007. 9 datasets.	Total migrants and by broad age group	http://cider.census.ac.uk
Community Health Index migration data	District (aggregated to NUTS2) Scotland coverage	Annual data from Mid 2002 to 2007 – 6 datasets	Total migrants and by broad age group	www.gro-scotland.gov.uk/
NHSCR migration data	Government Office Region or NUTS1. UK coverage	Mid 1999–2007 – 9 datasets Calendar year 2002–2007 – 6 datasets	Total migrants	www.statistics.gov.uk/
Single year of age migration rates – 2001 Census table C0527	UK, national rates	2000–01	Proportion of Male and Female population as migrants, by single year of age	www.ccsr.ac.uk/popgroup/about/fmm.html
POPULATION INFORMATION				
Single year of age population estimates	England and Wales	1961–2081	Male and Female count 0–85+	www.statistics.gov.uk/population/estimates/svg_pyramid/default.htm
		2002–2007	90–105+	http://www.statistics.gov.uk/statbase/Product.asp?vlnk=15003
		1981–2007	85–89	www.statistics.gov.uk/statbase/Product.asp?vlnk=15106
Single year of age population estimates	Scotland	1981–2031	Male and Female count 0–90+	www.gro-scotland.gov.uk/statistics/publications-and-data/popproj/projected-population-of-scotland-(2006-based)/additional-data.html
Single year of age population estimates	Northern Ireland	1961–2007	Male and Female count 0–85+	www.nisra.gov.uk/archive/demography/population/midyear/NI_Home_Pop_sya(1961-2007).xls
GEOGRAPHICAL INFORMATION				
District to NUTS2 lookup table	UK	N/A	N/A	Produced for the project

and calendar year migration estimates, something which will be important for adjusting mid-year estimates to calendar year estimates, explained later in this paper. Small variations exist between the figures published for the whole UK by ONS, and in-migration data from the rest of the UK published by GRO Scotland and NISRA. In the case of Scotland, the differences are due to very small rounding errors. In the case of Northern Ireland, the differences are large enough not to be caused by rounding, but, for consistency, we chose to use the NHSCR Southport data, published by ONS in preference.

In addition to the migration data described, where estimates of in, out and net migration for NUTS2 regions by single year of age and sex are required for the project, use has been made of single year of age population estimates for England and Wales, Scotland and Northern Ireland. These estimates are used in conjunction with a single year of age by sex migration schedule for the UK obtained from the 2001 Census.

Estimating migration flows for the UK

Conceptually, producing a new set of inter-NUTS2 migration estimates is relatively straightforward. It is a matter of using known, complete, data values and distributions to constrain other, incomplete data, to produce estimates which adhere to the known values. In practice this can be more challenging to implement, especially where data availability varies and the method needs adjusting according to this availability. For example, one assumption we chose to adhere to is that known, published data are likely to be more accurate than new estimated data, so where possible known data should be included in the final estimates in preference to estimated data. As is shown in Table 2, this means that as different data are available for different time periods and geographical areas, the estimation methodology changes slightly.

The estimation process is carried out in a series of stages – the results from the previous stage providing the data for the subsequent stage. Estimates of migration by single year of age for different years would not be possible without inflow and outflow totals. Therefore the first stage of the estimation process was to create a set of annual migration flow matrices with marginal inflow and outflow totals which could then be used for the age-specific estimates.

The estimation scheme

Figure 1 sets out the framework for the estimation. The input data are set out in the first two columns of tables (Census, PRDS/CHI and NHSCR data). The outputs of the estimation are set out in the third and fourth columns for mid-year to mid-year intervals and calendar years respectively. A matrix of inter-districts migration flows from the 2001 Census of Population were aggregated to NUTS2 region flows (shaded mid-green). These census data are used as first estimates where data are missing and adjusted to any known constraints. In effect, we assume the interaction structure of the migration flows for 2001 applies to the other years in the time series. Below the single census table for 2000–01 (the year prior to census date on 29 April), are the series of PRDS/CHI tables of migration flows captured from the Patient Registers of England, Wales, Scotland and Northern Ireland, which have missing sets of flows as described above. In the second column are tables for the same mid-year to mid-year intervals derived from the integrated NHS Central Register held at Southport. These data are used to constrain the first census estimates to deliver a set of synthetic estimates contained in the third column of tables. In the fourth column are represented the calendar year equivalents of the mid-year to mid-year synthetic estimates which are obtained by averaging. We now provide more details of the procedures that generate the synthetic estimates, stage by stage.

To help specify these methods as precisely as possible, we adopt the algebraic notation set out in **Table 3**.

Stage 1: Aggregating district level flows to NUTS2 regions

The first task in the estimation process is to aggregate known data into NUTS2 regions. As has been explained in detail elsewhere⁸ aggregating flow data from one geography into another needs to be carried out with care, and is not simply the case of summing the inflow and outflow totals. This is because where areas at one level aggregate into another, flows between those areas at the original level (inter-area flows) become flows within the area (intra-area flows) at the new aggregate level. With this in mind, flow data between districts from the 2001 Census, Patient

Figure 1 The estimation scheme

registers and CHI were aggregated into NUTS2 regions (with the minor re-allocation of the Isle of Arran) using a look-up table created with data available from GeoConvert⁹.

Stage 2: Estimating missing NUTS2 flows using existing data sources

In order to estimate complete mid-year NUTS2 flow matrices for the missing flows in Figure 1, the following estimating equation is used:

$$M_{ij}^t = N_{+j}^t [C_{+j}/C_{+j}] [C_{ij}/C_{+j}] \quad (1)$$

In this destination constraint method, the known information for in-migrations at NUTS1 scale (region J) is multiplied by two ratios. The first ratio C_{+j}/C_{+j} is the proportion census migration flows to NUTS1 region J from all NUTS2 regions i that come to NUTS2 region j. This is the marginal ratio. The second ratio, for internal values, C_{ij}/C_{+j} is the share of in-migration to NUTS2 region j that comes from NUTS2 region i. We use equation (1) for ease of computation in the relevant spreadsheet but note that the equation simplifies to:

$$M_{ij}^t = N_{+j}^t [C_{ij}/C_{+j}] = C_{ij} [N_{+j}^t/C_{+j}] \quad (2)$$

Table 3 Variables and indexes used in the estimation procedures

Variable, Index	Description	Migration concept used, ranges of indexes
<i>C</i>	Census migration flows (NUTS2 regions)	Transitions
<i>T</i>	Patient register migration flows (NUTS2 regions)	Transitions
<i>N</i>	NHS Central Register flows (NUTS1 regions)	Moves
<i>M</i>	Target migration flow (NUTS2 regions)	Moves
<i>i</i>	Index for NUTS2 origin regions	Range of values: $i = 1, 37$
<i>j</i>	Index for NUTS2 destination regions	Range of values: $j = 1, 37$
<i>l</i>	Index for NUTS1 origin regions	Range of values: $l = 1, 12$
<i>J</i>	Index for NUTS1 destination regions	Range of values: $J = 1, 12$
+	Indicates summation over index replaced	e.g. $C_{+j} = \sum_i C_{ij}$
1,2	Index used to label successive versions of the same <i>M</i> variable	
<i>t</i>	Index for mid-year to mid-year interval in a time series	The time interval extends from mid-year <i>t</i> in calendar year <i>y</i> to mid-year <i>t</i> + 1 in calendar year <i>y</i> + 1. We link mid-year time intervals by using a combined notation t_y , meaning <i>t</i> is the mid-point of year <i>y</i> . Range of values: $t = 1999-2000, 2006-2007$
<i>y</i>	Index for a calendar year in a time series	Range of values: $y = 2000, 2006$
C_{ij}	Census migration flow from region <i>i</i> to region <i>j</i> in year before 2001 Census	Transition between 29 April 2000 and 29 April 2001
T_{ij}^t	Patient register migration flow from region <i>i</i> to region <i>j</i> in year <i>t</i>	Count of reported transitions between 31 July in year <i>t</i> and 31 July in year <i>t</i> + 1 that cross an area boundary
N_{IJ}^t, N_{IJ}^y	NHS Central Register migration flow from region <i>I</i> to region <i>J</i> in mid-year <i>t</i> , and in calendar year <i>y</i>	Count of all moves reported to the NHSCR that cross a region boundary in the time interval indicated
M_{ij}^t, M_{ij}^y	Target migration flow from region <i>i</i> to region <i>j</i> in mid-year <i>t</i> , and in calendar year <i>y</i>	As target migration estimates are adjusted to NHSCR moves, these estimates can be regarded as move data

This version shows that we are adjusting the census migration flow by the ratio of the NHSCR total inflow to the NUTS1 region to the census total inflow to the NUTS1 region.

An alternative to equation (2) would be to use the equivalent out-migration (or origin constraint) terms rather than the in-migration:

$$M_{ij}^t = N_{i+}^t [C_{ij}/C_{i+}] = C_{ij} [N_{+J}^t/C_{+J}] \quad (3)$$

Another alternative would be to adjust the NUTS1 region to region flow sums using a ratio of the origin-destination flows:

$$M_{ij}^t = N_{IJ}^t [C_{ij}/\sum_{l \in I, j \in J} C_{lj}] = C_{ij} [N_{IJ}^t/\sum_{l \in I, j \in J} C_{lj}] \quad (4)$$

The advantage of equation (4) over equations (2) and (3) is that by constraining to C_{ij} cell values, the estimated internal matrix values sum to both the N_{i+} and N_{+j} marginal values as opposed to one or the other. The drawback, however, is that where N_{jj} values on the diagonal of the matrix do not exist, M_{ij} cannot be estimated, thus an additional estimation method would need to be implemented. Addressing both drawbacks, a fourth alternative could be to iterate between equations two and three using an iterative proportional fitting method^{10,11,12} which continually

adjusts M_{ij}^t cell values to N_{i+} and N_{+j} thus:

$$\begin{aligned}
 M_{ij}^1 &= N_{+j}^t [C_{ij}/C_{+j}] \\
 M_{ij}^2 &= N_{i+}^t [M_{ij}^1/M_{i+}^1] \\
 M_{ij}^3 &= N_{+j}^t [M_{ij}^2/M_{+j}^2] \\
 M_{ij}^4 &= N_{i+}^t [M_{ij}^3/M_{i+}^3]
 \end{aligned} \tag{5}$$

This process continues until convergence is reached and all internal cell values satisfy the marginal constraints for both origins and destination within a small error tolerance. In the long run it would be preferable to generate the full UK migration flow time series from a UK wide patient register. In the near term future the NHS's Personal Demographics Service is likely be tasked to develop a comprehensive demographic report from the UK NHS Register, the Spine, which will include population counts by age and sex and migration flows at all geographic levels.

In these estimates we chose the approach of equation (1) where the focus is on the destination. We chose the destination constraint, as it generally performed better in diagnostic tests (**Table 4**) than the origin constraint when recreating the known England and Wales PRDS data between 1999 and 2006; where the origin-destination constrained model, despite performing better than other methods was unable to provide estimates of flows within Scotland, and the IPF method shown in equation (5) failed to reach convergence after 20 iterations in 2003 and 2004 (almost certainly as a consequence of NUTS1 NHSCR in-flow and out-flow totals not being equal – a problem with using rounded data). To get around this difficulty, we could have adjusted origin totals so that their sum was the same as the sum of destination totals, and *vice versa*, but this introduces additional error. Of course, both of the latter methods might prove preferable were these issues addressed at a later date, but with our main focus on creating a consistent and immediately usable time series and our secondary focus on demonstrating a new estimation framework, equation (1) presented the best way forward. The model we have chosen is, in effect, a *pool* model. We estimate the total arrivals in a destination region and then allocate them to origins. Pool models are popular in projection work when direct origin to destination flow information is absent, but they do ignore some of the interaction structure in the flows matrix.

An example of the arithmetic involved in implementing equation (1) is presented in **Figure 2** for flows between two regions in northern England and Scotland and Northern Ireland. We have made the assumption that these are the only regions involved; in the actual estimation all regions are included. Figure 2a assembles the census migration flows and shows their NUTS2 and NUTS1 marginal totals. Note that in this case NUTS2 marginal totals add to NUTS1 totals. Figure 2b shows the census migration flows at NUTS1 scale. Figure 2c shows the NUTS2/NUTS1 in-migration and out-migration ratios in the row and column marginal cells, while the interior of the table shows the shares of the NUTS2 origin regions in the NUTS2 in-migration totals. Figure 2d shows the new information for 2006–7 that we have for migration between NUTS1 regions, which are used in equation (1) with the two set of ratios in Figure 2c to produce the estimates in Figure 2e. The grand total of estimates at NUTS1 scale agree exactly with the NHSCR totals. Note, however, that the NUTS2 out-migration totals do not sum exactly to the NUTS1 constraints for reasons already discussed. The equation (1) model was applied to all missing flows in a similar manner to the Figure 2 example, for all time intervals from 1999–2000 to 2006–2007. As shown in Figure 1, entire estimated matrices are not required, only migration flows to and from NUTS2 regions in Scotland and Northern Ireland. The estimates of these flows are added to existing

Table 4 Diagnostic statistics for different estimation schemes

Estimates for England and Wales compared with PRDS data									
Constraint	Diagnostic	1999	2000	2001	2002	2003	2004	2005	2006
Origin	Root mean square error	257	364	338	549	555	618	336	267
Destination		253	231	209	298	449	464	361	345
Cell		165	155	138	165	175	176	165	170
IPF		220	241	258	309	no conv	no conv	313	313
Origin	Average absolute difference between PRDS ij flow and estimated ij flow	134	134	117	150	163	163	135	133
Destination		138	122	105	135	155	160	144	147
Cell		103	99	87	106	110	109	105	107
IPF		125	123	109	129	no conv	no conv	135	138
Origin	Average ratio PRDS ij flow and estimated ij flow	1.01	1.02	1.01	1.02	1.03	1.02	1.00	1.00
Destination		1.03	1.02	1.01	1.00	0.99	0.99	0.99	0.99
Cell		1.01	1.01	1.01	1.01	1.00	1.01	1.00	1.00
IPF		1.01	1.01	1.01	1.01	no conv	no conv	1.00	0.99
		Rank	Rank	Rank	Rank	Rank	Rank	Rank	Rank
Origin	Root mean square error	4	4	4	4	3	3	3	2
Destination		3	2	2	2	2	2	4	4
Cell		1	1	1	1	1	1	1	1
IPF		2	3	3	3	n/a	n/a	2	3
Origin	Average absolute difference between PRDS ij flow and estimated ij flow	3	4	4	4	3	3	3	2
Destination		4	2	2	3	2	2	4	4
Cell		1	1	1	1	1	1	1	1
IPF		2	3	3	2	n/a	n/a	2	3
Origin	Average ratio PRDS ij flow and estimated ij flow	3	4	4	4	3	3	4	3
Destination		4	3	3	1	1	1	1	1
Cell		1	2	1	3	2	2	3	4
IPF		2	1	2	2	n/a	n/a	2	2

Patient Register migration flows and the totals for in-migration and out-migration adjusted to take account of the new estimated data.

Stage 3: Single year of age inflows and outflows

Once a set of flow matrices are created, complete with adjusted flows, the next step is to use the new data to estimate an accompanying set of in/out and net flow totals by age for each NUTS2 region. Single year of age in, out and net flow data are not available for districts (or NUTS2 regions) in the UK. Consequently these data are estimated using a combination of (1) national single year of age migration profiles taken from the 2001 Census, (2) single year of age population estimates produced by the different UK statistical agencies for constituent UK countries for each year of interest and, where available, (3) migration estimates by broad age group from patient

Figure 2 Examples of the migration, ancillary matrices and synthetic estimates

a. Example of Census NUTS2 flows

d. Example of Census NS-02 flows

Census NUTS2	O/D	UKM (Scotland)				UKN (N.Ireland)	C_i	C_j
		UKM2	UKM3	UKM5	UKM6	UKN0		
UKC (North East)	UKC1	561	330	114	280	132	1417	4079
	UKC2	1494	595	180	226	167	2662	
UKD (North West)	UKD1	480	517	61	111	66	1235	7168
	UKD2	461	272	90	117	78	1018	
	UKD3	803	601	136	217	371	2128	
	UKD4	634	601	94	195	228	1752	
	UKD5	310	268	104	133	220	1035	
	C_j	4743	3184	779	1279	1262		
	C_j	9985				1262		

b. Example of Census NUTS1 flows

Census NUTS1	UKM	UKN	C_i
UKC (North East)	3780	299	4079
UKD (North West)	6205	963	7168
C_j	9985	1262	11247

c. Ratios for estimation

	C_{ij}/C_j	UKM (Scotland)				UKN (N.Ireland)	C_i/C_i
		O/D	UKM2	UKM3	UKM5	UKM6	
UKC (North East)	UKC1	0.118	0.104	0.146	0.219	0.105	0.347
	UKC2	0.315	0.187	0.231	0.177	0.132	0.653
UKD (North West)	UKD1	0.101	0.162	0.078	0.087	0.052	0.172
	UKD2	0.097	0.085	0.116	0.091	0.062	0.142
	UKD3	0.169	0.189	0.175	0.170	0.294	0.297
	UKD4	0.134	0.189	0.121	0.152	0.181	0.244
	UKD5	0.065	0.084	0.134	0.104	0.174	0.144
	C_i/C_j	0.475	0.319	0.078	0.128	1.000	

d. Example of NHSCR NUTS1 flows for 2006-7

NHSCR NUTS1	UKM	UKN	C_i
UKC (North East)	4300	600	4900
UKD (North West)	7800	2100	9900
C_j	12100	2700	14800

e. Synthetic estimates of migration flows for 2006-7

Estimates NUTS2		UKM (Scotland)				UKN (N.Ireland)	C _i	C _j
	O/D	UKM2	UKM3	UKM5	UKM6	UKN0		
UKC (North East)	UKC1	680	400	138	339	282	1840	4900
	UKC2	1810	721	218	274	357	3381	
UKD (North West)	UKD1	582	627	74	135	141	1558	9900
	UKD2	559	330	109	142	167	1306	
	UKD3	973	728	165	263	794	2923	
	UKD4	768	728	114	236	488	2335	
	UKD5	376	325	126	161	471	1458	
	C _j	5748	3858	944	1550	2700	14800	14800
	C _J	12100				2700		

12100

Figure 3 Migration probability by age for men and women

Note: Source: Popgroup. See: www.ccsr.ac.uk/popgroup/about/fmm.html

register (PRDS) data. A 2001 single year of age migration schedule for males and females in the UK (**Figure 3**) is available from Census commissioned table C0527¹³. Population data for each of the constituent countries of the UK has been obtained from the different UK national statistical agencies, but the availability of different data for different countries in the UK means that the estimation process is carried out in different stages.

Stage 4: Estimating total migrants in each age group for England and Wales

The first stage in the process involves estimating total migrants for each age group, for each year of interest (2000–07) in England and Wales, using the published population estimates and the 2001 migration schedule (Figure 3). Applying the migration schedule to the population estimate produces an estimate of the total number of migrants in each age group. These single year estimates are aggregated into broad age groups to match the broad age groups (0–15, 16–19, 20–24 etc) in the published PRDS data. It is then possible to compute the proportion of each broad age group associated with each single year of age (**Table 5**), before using these proportions to disaggregate broad-age group PRDS data into single year of age data for each NUTS2 region in England and Wales. The 2001 Census migration schedule is disaggregated by males and females, so by applying the male and female ratios from the schedule to the PRDS data it is possible to disaggregate the estimated total flows by sex.

Stage 5: Adjusting data to totals

In producing these estimates, an issue was discovered with the PRDS data. To illustrate the issue, consider **Table 6**. The table shows an extract of the PRDS data for England and Wales for the year 2000. The rows show the age categories as published in the original data – eight broad age groups and an ‘all ages’ category. As is clear from the table, the sum of the age categories does not equal the ‘all ages’ total also published. In all cases, the sum of the age groups is considerably less than the all ages total given. Even at the district level (the level for which the PRDS data was originally

Table 5 Example of single age to broad age proportions

Estimated Migrants, Mid-2000				Single year proportion		
Age	Total	Male	Female	Total	Male	Female
0–15	1,026,170	523,284	502,886	1.000	1.000	1.000
0	101,777	52,022	49,756	0.099	0.099	0.099
1	90,657	46,774	43,883	0.088	0.089	0.087
2	81,456	41,572	39,884	0.079	0.079	0.079
3	76,221	38,926	37,294	0.074	0.074	0.074
4	68,553	34,984	33,569	0.067	0.067	0.067
5	64,981	33,553	31,428	0.063	0.064	0.062
6	63,864	32,902	30,962	0.062	0.063	0.062
7	62,079	31,685	30,395	0.060	0.061	0.060
8	60,659	30,844	29,815	0.059	0.059	0.059
9	58,852	29,885	28,967	0.057	0.057	0.058
10	54,986	27,802	27,184	0.054	0.053	0.054
11	52,658	26,761	25,898	0.051	0.051	0.051
12	50,320	25,994	24,326	0.049	0.050	0.048
13	48,611	24,934	23,678	0.047	0.048	0.047
14	42,982	21,412	21,571	0.042	0.041	0.043
15	47,513	23,234	24,279	0.046	0.044	0.048
etc						

Table 6 In-migration to NUTS2 regions by broad age group, 2000

Age group	UKC1	UKC2	UKD1	UKD2	UKD3
	Tees Valley and Durham	Northumberland and Tyne and Wear	Cumbria	Cheshire	Greater Manchester
All ages	21,092	23,091	12,320	31,209	50,358
0–15	3,682	3,334	2,127	5,517	7,138
16–19	2,965	3,490	646	1,731	6,318
20–24	3,663	5,216	1,604	4,757	12,250
25–29	2,364	2,864	1,256	4,387	7,987
30–44	4,448	4,500	2,695	8,405	10,126
45–59	1,795	1,693	1,825	2,953	3,161
60–74	836	866	1,057	1,445	1,430
75+	533	443	493	1,022	941
Sum	20,286	22,406	11,703	30,217	49,351
Difference: all ages less sum	806	685	617	992	1,007

Note Source: NHS Patient Registration Data System

published), when examining the in-flows and out-flows for each district, for each of the 376 districts in England and Wales, the 'all ages' total is higher than the sum of the age groups by an average of 122 people. This indicates why the discrepancies are so large when the districts are aggregated to the NUTS2 level.

The reasons for this discrepancy are likely to be due to these data being not being fully 'unrounded'. ONS produces three versions of the PRDS data:

- a) Normal rounded published PRDS tables – rounded to 10s
- b) Unrounded PRDS – rounded to integers
- c) Decimal unrounded PRDS – totally unrounded, containing long decimals – a by-product of constraining PRDS data to NHSCR flows – only used internally.

The decimal unrounded data are the most accurate, but are not available for these estimates, where only unrounded PRDS data are used. Where rounding is having an effect on data, it would be sensible to assume the 'all ages' total is a more accurate estimate of the total migrants than the sum of the age groups. This, then, presents a problem. Any single year of age estimates based on the broad age groups in the PRDS data are likely to be under-estimates. As such all single year of age estimates need to be re-scaled to conform to the higher, 'all ages' total in the PRDS data. The second stage in the process, therefore, involves re-scaling the in and out migration estimates for each year from 2000–07 to the PRDS 'all ages' totals for the corresponding years.

Stage 6: Inclusion of flows to and from Scotland and Northern Ireland

Published PRDS internal migration flows only cover England and Wales; therefore in the third stage of producing single year of age estimates, additional data from Scotland and Northern Ireland are added to both complete the set of NUTS2 regions, and adjust the England and Wales estimates to accommodate additional flows from Scotland and Northern Ireland. This is achieved through using the flow matrices described earlier in this paper, as well as the population data described in Table 2. Re-scaled England and Wales in and out migration data are adjusted to the UK flow matrix totals for each year. These matrix totals are then used to estimate the single year of age in and out flows for Scotland and Northern Ireland using single year of age population data for these countries along with the 2001 Census single year of age migration schedule. Male and Female ratios from these population and migration data are then used to split the total estimates into male and female estimates. The outcome of this estimation process is that single year of age in and out migration estimates for the UK for each year 2000–07 are produced which are consistent with the aggregate origin-destination flow matrices also produced.

Stage 7: Adjusting estimates from mid-year to calendar year

One of the issues with the initial migration estimates is that the data used, and therefore the estimates themselves, are for mid-year populations (June to June) rather than end of year populations (December to December). The DEMIFER project requires that the data cover calendar years rather than mid-years, partially because fertility and mortality estimates for the UK used in the project cover calendar years, but more importantly because demographic estimates for other European countries covered by the DEMIFER project are for calendar years. Consequently the mid-year estimates require conversion to calendar years. The conversion of these data was achieved through a several stage process using published NHSCR calendar year in and out

migration estimates at the NUTS1 level to constrain the average of two consecutive mid-year estimates. The exact process for estimating a new set of total flow matrices is as follows:

- Step a Take two consecutive mid-year to mid-year NUTS2 flow matrices and create a new matrix of the average values of the two, where y now refers to calendar year and t_{y-1} and t_{y+1} are the mid-years that start successive mid-year intervals (e.g. calendar year 2005 estimates are an average of 2004–2005 and 2005–2006 mid-year to mid-year interval estimates):

$$M_{ij1}^y = \frac{1}{2}[M_{ij}^{t_{y-1}} + M_{ij}^{t_y}] \quad (6)$$

The subscript 1 attached to the left hand side migration variable indicates that this is our first estimate of its value. We need now to make sure that these averaged flows agree with independently published calendar year migration flow matrices at NUTS1 scale. Again we adopted the in-migration approach rather than the out-migration or the interaction approach (see above).

- Step b Aggregate the estimated migration flows estimates at NUTS2 scale to migration flows at NUTS1 setting intra-region flows to zero and computing in-migration totals:

$$M_{IJ1}^y = \sum_{i \in I, j \in J, i \neq j} M_{ij1}^y \quad (7)$$

$$M_{II1}^y = 0 \quad (8)$$

$$M_{+J1}^y = \sum_I M_{IJ1}^y \quad (9)$$

- Step c Compute the in-migration totals from the averaged migration flows for calendar years

$$M_{+J}^y = \sum_I N_{IJ}^y \quad (10)$$

- Step d Compute the revised estimates of NUTS2 migration flows for calendar years as follows:

$$M_{ij2}^y = N_{+J}^y [M_{+j1}^y / M_{+J1}^y] [M_{ij1}^y / M_{+j1}^y] \quad (11)$$

To the constraint of the NHSCR in-migration total in NUTS1 region J , we apply two ratios derived from the initial averages. The first ratio is that of the NUTS2 total in-migration flow to the NUTS1 total in-migration, for all NUTS2 regions fitting within NUTS1 regions. The second ratio is the share of the total in-migration to NUTS2 region j which is allocated to origin region i . Equation (11) can be simplified by cancelling common terms in the ratios, to give:

$$M_{ij2}^y = N_{+J}^y [M_{ij1}^y / M_{+J1}^y] \quad (12)$$

This can be re-arranged to show the NHSCR constraints are applied to the initial migration estimates:

$$M_{ij2}^y = N_{ij1}^y [N_{+J}^y / M_{+J1}^y] \quad (13)$$

By following the methodology described above a new set of estimated calendar year flow matrices can be estimated. As before, we have chosen a destination/in-migration formulation for convenience. The other choices were an origin/out-migration approach or an origin-destination approach. These matrices can then be used to re-estimate in/out/net flow totals by single year of age for calendar years.

Limitations

This paper has described a new methodology for producing sets of annual internal migration estimates for NUTS2 regions in the UK. Whilst we believe this methodology offers the current best solution given the data available, there are a number of limitations which need to be acknowledged. Addressing these issues at a later date should improve the quality of the estimates.

Using 2001 Census NUTS2 flows to distribute NHSCR NUTS1 data

Flows between origins and destinations are unlikely to remain completely constant across different years. We have assumed that the relative NUTS2 shares of an origin to destination flow of the corresponding in-migration total, based on census data, will remain constant, whereas in reality the shares may vary. However, we believe the variation is likely to be small because studies of observed time series of migration flows have found that the interaction structures have remained relatively constant, even when origin and destination totals have varied. Of more impact may be the assumption that the ratio of in-migration in NUTS2 regions to in-migration in the containing NUTS1 region is constant over time. The attractiveness of regions to migrants does vary over time. This variation is quite slow: the north–south differences in in-migration shares in England have remained quite stable over the past four decades although the other home countries (Wales, Scotland and Northern Ireland) have experienced net in-migration balances. These assumptions only impact the migrations between England and Wales on the one hand and Scotland and Northern Ireland on the other. These flows are small relative to the migration volumes within each home country.

Using a UK migration schedule to estimate single year of age migrant flows

In estimating the numbers of in and out migrants by single year of age, a national, UK single year of age schedule was used. Applying a UK schedule is not a perfect solution as the rates of migration by age vary across the constituent countries of the UK. To demonstrate this, consider **Figure 4**. This shows the variation in the age specific intra-country migration rates (for age groups of varying sizes) across the constituent countries of the UK. There is considerable variation within the countries, especially at the ages of peak migration. For example, there is a difference of 13 per cent in the migration rate between England and Northern Ireland at ages 20–24. This kind of variation means that any estimates based upon an aggregate, national schedule at the peak migration ages, are likely to underestimate flows in England and overestimate flows in Northern Ireland. With these estimates, we recognise this drawback, and so include a ‘health warning’. An improvement to the methodology could involve using the national single year schedule to ‘smooth’ the age group schedule in Figure 4, before using the resulting smoothed schedule to estimate migrants at the sub-UK, home country level. Better still, these single year estimates could be improved substantially if single year of age migration schedules for England, Wales, Scotland and Northern Ireland (not publicly available, but in existence) were used instead.

Rounded data

As discussed earlier in this paper, there are a number of inconsistencies in the published PRDS data caused by rounding. This means that for the estimation of single year of age in/out migration flows, estimates are constrained to the totals – the assumption being that totals will be less susceptible to rounding errors and thus more accurate. Constraining the data in this way is likely to

Figure 4 **Variation in age-specific intra-country migration rates by UK constituent country, 2001 Census**

Note Source: 2001 SMS level 1, table 1.

introduce some error into the estimates; it is therefore far more desirable to avoid introducing this – something which will only be achieved with access to the decimal unrounded PRDS data.

Results

Taking the calendar year flow estimates from 2000–2007 along with population estimates for the same years, it is possible to examine some of the trends in migration for the UK over this eight year period. The first point of note is that from 2000 to 2007, the relative volume and direction of flows remained very stable. **Table 7** shows the net migration rate for each NUTS 2 region, ranked according to the average net migration rate per 1000 population across all eight years. As is clear from the table, the relative level and direction of net migration change varied little over the period. London has the highest rates of net out-migration throughout the period; Cornwall and Lincolnshire consistently have the highest rates of net in-migration. Whilst there is some variation across years, there is very little change in the rank order of regions over the period. Only Eastern Scotland shows some noticeable deviation from the mean, with Surrey, East and West Sussex exhibiting an increase in net in-migration towards the end of the period.

Of course, net-migration figures hide the volume of flows, so as advocated by Dennett and Stillwell¹⁴, we use here a measure of turnover (in-migration plus out-migration) to examine the relative intensities of migration across our chosen time scale. **Table 8** presents turnover rates per 1000 population¹⁵. As with net-migration, there is very little variation in turnover rates between regions across years. London exhibits the highest rates of turnover, with Northern Ireland and South Western Scotland the lowest. The spatial distribution of the regions with the highest rates of turnover is generally in the South of the UK, particularly around London and towards the South West (with the exception of the Scottish Highlands and Islands); rates decline as you move further

Table 7 **Net migration rates across NUTS2 regions, 2000–2007
(ranked by average)**

	Net migration rate 2000	Net migration rate 2001	Net migration rate 2002	Net migration rate 2003	Net migration rate 2004	Net migration rate 2005	Net migration rate 2006	Net migration rate 2007	
Area name	2000	2001	2002	2003	2004	2005	2006	2007	Average 2000–2007
Cornwall and Isles of Scilly	9.71	12.74	14.76	14.25	12.41	10.02	10.21	9.81	11.74
Lincolnshire	12.55	13.06	14.31	13.43	10.99	7.70	8.20	8.89	11.14
Devon	9.13	9.41	9.54	9.69	8.63	6.61	7.12	7.23	8.42
Dorset and Somerset	8.05	9.22	9.71	8.71	7.84	6.40	7.35	8.06	8.17
East Anglia	6.32	6.65	7.25	6.86	6.07	4.20	4.54	4.62	5.81
Herefordshire, Worcestershire and Warwickshire	3.58	4.47	5.95	5.24	4.65	3.92	3.22	3.42	4.31
Essex	4.83	4.65	4.55	3.71	4.13	3.51	4.36	4.53	4.28
Cumbria	2.96	3.65	5.37	6.58	5.04	3.16	3.26	3.31	4.17
Surrey, East and West Sussex	3.66	3.29	3.81	3.27	3.63	3.60	4.97	4.70	3.87
Kent	3.25	3.81	4.54	3.81	3.99	3.30	3.90	3.78	3.80
West Wales and The Valleys	2.36	3.27	5.18	5.91	4.43	2.84	3.21	2.93	3.77
North Yorkshire	4.74	3.43	3.70	4.47	4.01	2.90	2.75	2.28	3.54
Eastern Scotland	1.87	3.80	2.76	4.90	4.18	4.78	2.14	3.78	3.53
Cheshire	2.27	2.49	3.28	3.50	2.91	2.85	2.77	2.73	2.85
East Wales	3.13	2.74	4.30	3.83	2.70	1.48	1.74	1.46	2.67
Gloucestershire, Wiltshire and Bristol/Bath area	2.29	2.52	2.81	2.71	2.78	2.76	2.82	2.55	2.66
Lancashire	1.88	3.10	4.16	4.49	3.21	2.27	1.42	0.55	2.63
East Yorkshire and Northern Lincolnshire	1.97	3.20	4.73	4.84	3.02	1.19	0.86	0.86	2.58
Shropshire and Staffordshire	2.05	2.71	3.33	3.32	2.85	2.09	1.75	1.93	2.50
Leicestershire, Rutland and Northamptonshire	3.10	3.46	3.62	2.60	2.14	1.51	1.06	0.96	2.30
Derbyshire and Nottinghamshire	1.76	2.82	3.84	3.03	2.24	1.12	0.76	0.94	2.06
Hampshire and Isle of Wight	1.86	1.21	1.40	1.44	2.78	2.29	2.71	2.65	2.04
Tees Valley and Durham	−0.93	−0.27	1.33	1.75	1.46	1.16	1.18	1.04	0.84
South Yorkshire	−1.31	−1.02	0.41	0.99	0.77	0.07	−1.05	−1.62	−0.35
Northumberland and Tyne and Wear	−1.58	−1.07	0.12	0.09	−0.20	−0.14	0.02	−0.38	−0.39
Bedfordshire and Hertfordshire	−0.44	−0.91	−1.10	−2.03	−0.80	0.05	0.70	0.65	−0.49
Northern Ireland	−1.07	−0.44	−1.61	−0.92	−0.36	−0.56	0.14	−0.43	−0.66
North Eastern Scotland	−2.43	−0.87	−1.83	−0.07	−0.84	0.16	−2.13	−0.31	−1.04
South Western Scotland	−1.99	−1.15	−1.64	−0.66	−0.97	−0.56	−1.79	−0.91	−1.21
Highlands and Islands	−3.21	−1.04	−2.49	−0.03	−1.04	0.51	−2.67	−0.09	−1.26
West Yorkshire	−1.38	−1.72	−1.68	−1.25	−0.66	−0.77	−1.27	−1.52	−1.28
Berkshire, Buckinghamshire and Oxfordshire	−2.19	−3.73	−3.86	−3.34	−1.91	−1.34	−0.31	−0.53	−2.15
Greater Manchester	−2.84	−3.15	−2.78	−1.73	−1.60	−1.41	−1.85	−2.18	−2.19
Merseyside	−2.37	−2.22	−1.98	−2.03	−2.49	−2.21	−2.32	−2.69	−2.29
West Midlands	−5.54	−6.15	−6.43	−6.75	−5.98	−4.82	−5.29	−5.44	−5.80
Outer London	−6.60	−8.31	−10.40	−10.85	−9.91	−7.63	−7.20	−7.47	−8.54
Inner London	−13.69	−16.43	−20.49	−21.48	−19.64	−16.06	−15.22	−15.00	−17.25

Note

Net migration rate = $1,000 \times \frac{[\text{In-migration} - \text{Out-migration}]}{\text{Population at mid-year}}$

Table 8 **Turnover rates across NUTS2 regions, 2000-2007 (ranked by average)**

	Turnover Rate 2000	Turnover Rate 2001	Turnover Rate 2002	Turnover Rate 2003	Turnover Rate 2004	Turnover Rate 2005	Turnover Rate 2006	Turnover Rate 2007	
Area name	2000	2001	2002	2003	2004	2005	2006	2007	Average 2000-2007
Inner London	99.2	100.8	102.1	100.4	102.5	101.9	103.7	100.6	101.4
Outer London	80.0	82.0	84.2	82.6	83.5	82.2	84.1	82.3	82.6
Berkshire, Buckinghamshire and Oxfordshire	81.2	82.4	83.7	80.6	80.6	77.3	78.5	76.2	80.1
North Yorkshire	82.7	81.6	81.3	79.6	77.9	73.8	74.9	75.0	78.4
Lincolnshire	80.0	79.0	79.4	75.8	72.8	68.4	70.5	70.5	74.6
Highlands and Islands	68.1	76.2	72.8	79.7	76.5	77.3	67.6	72.5	73.8
Cornwall and Isles of Scilly	74.5	75.3	75.3	72.3	70.3	66.7	68.5	67.5	71.3
Bedfordshire and Hertfordshire	70.7	71.5	72.9	70.1	70.0	67.3	69.0	68.1	69.9
Surrey, East and West Sussex	70.5	71.7	72.5	70.0	69.9	67.0	68.5	66.9	69.6
Dorset and Somerset	71.8	72.6	72.6	69.5	68.5	66.2	68.2	67.4	69.6
Devon	72.8	73.3	73.5	70.9	69.1	65.4	66.3	64.6	69.5
Hampshire and Isle of Wight	68.3	68.6	69.0	66.5	66.5	63.7	64.9	63.0	66.3
East Wales	67.6	68.5	68.4	65.1	63.9	62.3	63.5	61.4	65.1
Herefordshire, Worcestershire and Warwickshire	66.3	66.4	67.2	64.5	63.4	62.3	63.1	62.8	64.5
Gloucestershire, Wiltshire and Bristol/Bath area	61.3	62.8	63.3	61.6	61.4	59.1	60.3	58.9	61.1
Leicestershire, Rutland and Northamptonshire	61.7	62.9	64.0	61.6	60.7	58.5	59.0	57.8	60.8
Cheshire	63.2	62.4	61.9	60.7	58.7	56.9	57.7	57.5	59.9
Essex	57.4	58.3	59.6	58.3	58.4	55.0	56.5	55.6	57.4
Kent	57.1	58.1	59.0	56.6	56.3	53.8	55.3	54.4	56.3
East Anglia	56.1	57.0	57.2	55.4	55.2	52.1	53.2	52.2	54.8
North Eastern Scotland	48.2	54.0	51.7	56.7	54.6	55.3	48.2	51.4	52.5
Shropshire and Staffordshire	53.6	54.1	54.9	52.9	51.8	50.1	50.8	50.5	52.3
Lancashire	53.3	53.9	54.4	53.3	50.8	48.8	48.8	47.9	51.4
Derbyshire and Nottinghamshire	49.7	51.1	52.4	50.7	50.3	48.3	48.7	48.3	49.9
Cumbria	51.9	51.6	52.4	52.3	49.4	47.0	47.3	46.5	49.8
East Yorkshire and Northern Lincolnshire	50.5	51.7	52.5	50.8	48.8	45.6	46.1	46.2	49.0
West Yorkshire	46.5	47.2	48.0	47.5	47.4	45.7	45.8	45.3	46.7
South Yorkshire	46.0	47.3	48.2	47.7	47.1	45.9	45.8	44.9	46.6
West Midlands	44.8	46.5	47.9	46.7	46.6	46.2	46.5	45.8	46.4
Greater Manchester	43.8	44.4	44.9	44.5	43.6	43.2	43.3	42.5	43.8
West Wales and The Valleys	44.9	45.2	45.6	43.9	42.9	40.7	41.3	40.4	43.1
Eastern Scotland	39.7	44.2	42.0	45.9	43.8	44.6	39.0	41.9	42.7
Tees Valley and Durham	40.6	42.0	43.4	42.3	41.5	40.7	40.6	40.4	41.5
Merseyside	39.3	40.0	40.2	39.5	38.2	37.7	37.8	37.1	38.7
Northumberland and Tyne and Wear	38.8	39.6	40.2	39.0	38.3	37.9	38.0	37.8	38.7
South Western Scotland	25.5	28.4	27.1	29.7	28.5	29.1	25.7	27.7	27.7
Northern Ireland	14.3	15.5	14.3	15.0	15.2	14.7	14.8	14.2	14.8
Net turnover, average rate	57.9	59.1	59.5	58.4	57.4	55.6	55.7	55.2	57.4

Note

Turnover rate = $1,000 \times \frac{[\text{In-migration} + \text{Out-migration}]}{\text{Population estimate at mid-year}}$

North, with the highest rates of turnover some seven times those of the lowest. The average turnover rates for each year reveal that from a peak in 2002, there has been a steady decrease in the overall levels of population turnover across the country, indicating less internal migration activity taking place. This is not the case everywhere, however, with Inner London actually increasing turnover rates over this period, and Outer London remaining quite stable.

One of the benefits of estimating a time series of UK data is that year-on-year change can be examined to see which regions are more dynamic and which more stable in respect of their migration flows. In addition, a key advantage of creating full flow matrices rather than just in and out flow estimates is that the relationship between areas can also be assessed. Which areas are showing increases in their migration exchanges and which are becoming less linked through migration over time? An analysis of the changing flow patterns over the time period showed that population exchanges between inner and outer London increased more than for any other pair of areas, revealing that not only was migration activity for London higher than for anywhere else, but the volume of activity also increased between 2000–2007. Exchanges between Outer London and Bedfordshire and Hertfordshire also increased. Despite some relatively low average overall turnover rates in parts of Scotland, population exchanges between almost all Scottish regions increased over the period. The Highlands and Islands region had high average turnover rates over the eight years but was also increasing its population exchanges with all other Scottish regions. Other pairs of areas where population exchanges are increasing are between Devon and Cornwall, the West Midlands and Herefordshire, Worcestershire and Warwickshire, and Northumberland, Tyne and Wear and Tees Valley and Durham.

Looking at pairs of areas where population exchanges decreased over the period of study, the largest decrease was between the two Welsh regions. Devon also saw a decrease in its exchanges with other regions in the South of England whilst maintaining a relatively high average turnover rate. Another area which shows a large reduction in exchanges with neighbouring regions is West Yorkshire with the two regions to the North. West Yorkshire exhibits low average turnover between 2000 and 2007, and experienced noticeable decreases in exchanges with other regions.

Conclusions

This paper has outlined in detail a new methodology for estimating a complete set of inter-region migration flows for the UK from 1999–2007 at NUTS2 scale. While the available data limits the accuracy of the results produced, the methods employed allow for new data which could improve the final output to be incorporated into the estimates subsequently. Furthermore, whilst these estimates use data from two particular levels of a geographical hierarchy, it is entirely feasible that the technique be applied to migration estimation at any other level where data from another level in the spatial hierarchy is present. For example, this methodology could be used in the estimation of a complete LAU1 (local authority district) level matrix for the whole UK. It is also the case that the particular version of the methodology employed to produce these estimates (destination constrained) might not represent the absolute optimum solution. Origin/destination flow constraints with an additional method employed to estimate flows between NUTS2 regions that fall within the same NUTS1 region might offer a more optimum solution, as would an iterative proportional fitting method which constrains to both origin and destination totals, as long as non-convergence problems could be addressed.

In producing a new consistent set of migration estimates, it has been possible to examine migration patterns over a number of consecutive years for the whole UK – something which until now has not been possible. Preliminary analysis has produced a number of interesting findings. Firstly, rates of net-migration for NUTS2 regions between 2000 and 2007 remain very consistent, with more-or-less the same regions appearing as net gainers or net losers of population. Levels of population turnover remain similarly consistent. Although, nationally, average rates of turnover reduce between 2002 and 2007, the relative ranks of these regions in the population turnover hierarchy fluctuate very little. Examining the population exchanges between regions over this period has shown that for some areas, despite a national background of declining migration activity, exchanges have been increasing – perhaps most noticeably in London.

The principal objective of this analysis at the outset was to produce a time-series of migration estimates for the whole UK to fill a gap in a Europe wide data set. We have developed a methodology for producing these estimates using data which are available in the public domain and, working in consultation with the Office for National Statistics, have made these data available to European users through Eurostat. These data now form part of a much larger, European database of flows at the NUTS2 level across 31 states. The DEMIFER project will use these data, along with complementary data on births, deaths and international migration to analyse past demographic trends before producing a suite of European population projections linked to different potential development scenarios.

Key findings

- A major gap in European migration statistics at NUTS2 scale for United Kingdom can be filled using published migration data with a new estimation model. To overcome a data roadblock we adjusted the definition of one NUTS2 region in Scotland by transferring the Isle of Arran back to the NUTS 2 region to which its containing Scottish Council Area belongs.
- The paper showed how migration flows between NUTS2 regions in England and Wales and NUTS2 regions in Scotland together with Northern Ireland could be estimated by adjusting the corresponding census flows, which are available, to information on migration flows at NUTS1 scale, which are also published from the NHS Central Register.
- Estimation models were formally defined that handled the inter-relationships between migration flows at different regional scales. There were three alternative forms of the estimation models: dependent on destination in-migration constraints at NUTS1 scale, origin out-migration constraints and origin-destination flow constraints. The in-migration constraint model was implemented.
- Analysis of the resulting time series of migration flows between NUTS2 regions revealed remarkable overall stability together with some lesser subtle shifts in the interactions between small sets of regions.

Acknowledgements

This research has been funded as part of the DEMIFER Applied Research Project under the ESPON 2013 programme. The authors are grateful for the comments and suggestions given by three anonymous reviewers.

References

- 1 The agencies are: (1) the Office for National Statistics (ONS) responsible for migration statistics in England and Wales and the UK as a whole, (2) the General Register Office Scotland (GROS) responsible for migration statistics for Scotland and (3) the Northern Ireland Statistics and Research agency (NISRA), responsible for statistics in Northern Ireland.
- 2 ESPON (2009) See: www.espon.eu/main/Menu_Projects/Menu_AppliedResearch/demifer.html
- 3 Eurostat (2009) 'The NUTS classification.' Available at: http://epp.eurostat.ec.europa.eu/portal/page/portal/region_cities/regional_statistics/nuts_classification. The UK NUTS2 map: <http://ec.europa.eu/eurostat/ramon/nuts/pngmaps/uk2.png>
- 4 See <http://cider.census.ac.uk>
- 5 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 – Statistics in Society* 170(2): 425–445.
- 6 Scott, A and Kilbey, T (1999) 'Can patient registers give an improved measure of internal migration in England and Wales?' *Population Trends* 96: 44–55.
- 7 Chappell, R, Vickers, L and Evans, H (2000) 'The use of patient registers to estimate migration.' *Population Trends* 101: 19–24.
- 8 Dennett, A and Stillwell, J (2008a) 'Internal migration in Great Britain – a district level analysis using 2001 Census data.' Working Paper 01/08, School of Geography, University of Leeds. Available at: www.geog.leeds.ac.uk/fileadmin/downloads/school/research/wpapers/internal.pdf
- 9 GeoConvert, see: <http://geoconvert.mimas.ac.uk/>
- 10 Deming, W and Stephan, F (1940) 'On a least squares adjustment of a sampled frequency table when the expected marginal totals are known.' *Annals of Mathematical Statistics* 11: 427–444.
- 11 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 Series A – Statistics in Society* 170 (4): 891–908.
- 12 Raymer, J, Smith, P W F and Giulietti, C (2009) 'Combining census and registration data to analyse ethnic migration patterns in England from 1991 to 2007.' *Population, Space and Place*. DOI: 10.1002/psp.565.
- 13 These data are provided on the Cathie Marsh Centre for Census and Survey Research POPGROUP site at: www.ccsr.ac.uk/popgroup/about/fmm.html
- 14 Dennett, A and Stillwell, J (2008b) 'Population turnover and churn – enhancing understanding of internal migration in Britain through measures of stability.' *Population Trends* 134: 24–41.
- 15 The population estimate data used were those provided by ESPON (the European Spatial Planning Observatory Network) for NUTS2 regions in the UK. These are data made available to ESPON by Eurostat, which in turn receive them from ONS.