Office for National Statistics

Internal migration estimates methodology document:

June 2016

1.   Introduction

1.1   Purpose

This document describes the methods we use at the Office for National Statistics (ONS) to calculate annual estimates of internal migration moves affecting local authorities in England and Wales. Internal migration is a challenging component of population change to estimate, as there is no compulsory system within the UK to record movements of the population. At present we use a combination of administrative data sources as a proxy for internal migration. Our current methods were introduced for the estimates for the year ending June 2012. This document also explains previous methodological changes.

Our estimates are used as a component of population change at the local authority level within annual population estimates. Subnational population projections use internal migration trends over a 5 year period to project how likely people are to move by age, sex and local authority.

1.2   Definitions

Internal migrant

We define an internal migrant as someone who moves home from one geographical area to another. This may be between local authorities, regions or countries within the UK. Unlike with international migration, there is no internationally agreed definition.

Cross-border moves

Cross-border moves are those where people move, in either direction, across the Scottish and Northern Irish borders. Cross-border inflows are where people move from Scotland and Northern Ireland to England and Wales. Cross-border outflows are where people move from England and Wales to Scotland and Northern Ireland.

We publish estimates of internal migration affecting England and Wales; National Records of Scotland (NRS) publishes estimates of internal migration affecting Scotland; and the Northern Ireland Statistics and Research Agency (NISRA) publishes estimates of internal migration affecting Northern Ireland. Our internal migration methodology page includes a comparison of the respective methods.

By consensus among the 3 statistical offices, the recipient nation estimates the total inflow (with ONS producing combined estimates for moves into England and Wales). The estimates are then shared between the offices to enable the calculation of estimates on a consistent basis across the UK.

1.3   Issues with measuring internal migration

As there is no single system to record population movements within the UK, we calculate internal migration estimates using proxy sources. We estimate moves between local authorities in England and Wales, and from Scotland and Northern Ireland into England and Wales, using a combination of 3 administrative datasets: the Patient Register (PR), the

National Health Service Central Register (NHSCR) and data from the Higher Education Statistics Agency (HESA). In addition we use information from the separate Scottish NHSCR and the Northern Irish medical cards register to improve estimates of cross-border flows.

We use this combination of sources because each helps overcome coverage or quality limitations in the others. However, there are still known issues which may cause differences between the estimates and the actual number of internal moves. We describe these later in this document.

2. Current internal migration methods

Before we discuss the methods in detail, here is a brief overview of our data sources and how we use them. We first used the current methods for the calculation of the estimates for the year ending June 2012 (which covers the period from 1 July 2011 to 30 June 2012).

2.1 The Patient Register (PR)

The Health and Social Care Information Centre (HSCIC) sends us a yearly extract from the PR, which lists all the people on the register in England and Wales at the end of July. We assume that people delay registering with a new National Health Service General Practitioner (NHS GP) by approximately a month after they move, meaning the data are taken to reflect people's location as at 30 June.

The PR contains information on patients registered with a NHS GP, and is collected by Clinical Commissioning Groups (the health authorities that replaced Primary Care Trusts). When a person changes GP their NHS number transfers with them so that the correct medical records can be sent to their new GP. By using the NHS number and postcode when comparing one year with another we are able to identify moves.

The strengths of the PR are that it includes almost everyone resident in England and Wales, and is sufficiently detailed to identify moves at local authority level. However, because the PR extracts are only taken annually, comparison of successive years' data will miss certain groups of internal movers including:

- children under 1 who have moved, as they were not born a year earlier

- people who were not in England and Wales a year earlier, but have already moved within England and Wales

- people who had moved since the previous extract was taken and subsequently died or left England and Wales (unless they have not been removed from the PR)

- interim moves of people who have moved more than once during the year

We address these limitations by using NHSCR data – more information on this is in the next section.

Another important limitation of the PR data is that because some people (mainly younger adults, especially males) are less likely to register with a new GP promptly after moving, their moves may not be picked up on the annual extract. We use data from the Higher Education Statistics Agency (HESA) to partially counter this.

A third limitation is that the PR data exclude moves for certain specific reasons, in particular moves either into or out of prison or the armed forces. Moves for these reasons are therefore excluded from our estimates.

2.2 The NHS Central Register (NHSCR)

The NHSCR provides details of patient record changes, also based on GP registrations, between former health authorities (FHAs). These are usually made up of groups of local authorities.

HSCIC send us a weekly update of changes allowing us to create annual totals of moves between FHAs in England and Wales, as well as moves into England and Wales from Scotland and Northern Ireland.

Like the PR, the NHSCR includes almost everyone resident in England and Wales and because we receive the NHSCR data weekly, multiple moves during the year are included. However, there are also limitations:

1. The data are only available for FHAs and not at local authority level. This means that moves occurring between local authorities within a single FHA are excluded. This exclusion is estimated, on average, to account for a third of all internal migration moves.

2. Like the PR data, the NHSCR data exclude moves of people who have not yet re-registered with a GP after moving. As explained, this limitation has most impact on the estimation of moves for young adults, especially males.

3. Also like the PR data, the NHSCR data cannot be used to identify moves for certain specific reasons, in particular moves either into or out of prison or the armed forces. This means that these types of moves are not included in the internal migration estimates at all. However, ONS's annual mid-year population estimates take account of these types of moves through separate processing.

2.3 The Higher Education Statistics Agency (HESA)

As mentioned, young adults, especially males, are less likely to register with a GP promptly following a move. A common reason for moving at these ages is to attend or leave university. To help improve internal migration estimates at these ages, we supplement the PR extract with data from HESA, which we regard as being of high quality.

The HESA data, which we receive annually, contain information on both home and term-time addresses of students registered in higher education in the UK during the data year August to July.

2.4 Information on cross-border moves

HSCIC provide us with yearly data, based on the PR, detailing cross-border moves between England and Wales (combined) and Scotland and Northern Ireland. We use these to create files called "flows from rest of UK" and "flows to rest of UK".

However, these yearly data do not include all moves so we also receive data from the Scottish NHSCR (provided by NRS) and the Northern Irish medical cards register (provided by NISRA).

2.5   Overview of the methods used to calculate internal migration estimates

At the start of the process, we remove duplicate records and records with missing variables for age, sex or location from the source data.

We match the previous year's stocks file for England and Wales (that is, the file detailing where everyone was last year) with the PR and HESA datasets to create a new stocks file for the current year.

Our rules for the new stocks file are:

- if people are no longer on the PR they are removed

- for students:

    o if the start of the HESA year (1 August) is more recent than the last PR update, the HESA location takes precedence (although we only do this if the HESA location is in England and Wales)

    o if the record has been updated since the start of the HESA year, the PR location takes precedence

- for lapsed students (that is, those who were on the HESA dataset in previous years but do not appear on it this year):

    o if their PR record has been updated then we use this

    o if their PR record has not been updated then they remain in their HESA term-time local authority, but each year we transfer a proportion to their PR local authority using post-study movement factors

The post-study movement factors determine whether we retain a lapsed student at their HESA term-time local authority or re-allocate them to the local authority in their PR record. If the latter then whether we move them depends on whether or not the HESA term-time local authority is the same as the local authority on the PR.

We calculated the post-study movement factors for each local authority in 2013 using data on those lapsed students who had registered with a new GP. The factor represents how many had registered in a different local authority from where they were studying.

Of all the "lapsed students" remaining in the dataset, we re-allocate a proportion corresponding to the post-study factor. So if the factor for a particular local authority is 0.6, we re-allocate a total of 60% of lapsed students in that area to their PR local authority (which, as explained, may or may not be different to their HESA local authority).

If we have not re-allocated a lapsed student to their PR local authority we repeat the process each year until they are re-allocated. After this we no longer class them as a lapsed student and instead treat them as a member of the general population in the stocks file.

We then add our annual "flows to rest of the UK" information to the new stocks file, and the "flows from rest of the UK" to the previous year's stocks file; this ensures that every mover appears somewhere in the UK on both files. We then identify movers by comparing the current and previous stocks files and creating initial totals for flows of people moving into and out of each local authority and FHA.

We can't identify moves of children born since 30 June of the previous year because they won't be present in last year's stocks file. Therefore we initially take PR data on moves of 1-year-olds into each local authority. We then scale these using PR counts of the number of children under 1 in that local authority who have moved in from outside the FHA during the year (we use moves from outside the FHA because the PR does not tell us whether they have moved into the specific local authority).

To take account of the fact that people may move multiple times during a year, we then calculate "movement factors" (also referred to as "scaling factors"). These are the ratio of NHSCR moves to PR moves for the origin and destination FHAs (as data from the NHSCR are not available at local authority level). We apply the movement factors to flows at local authority level, regardless of age and sex, between each pair of FHAs. Figure 1 provides an example of the calculation of these factors.

Figure **1: Movement factor example**

Let's say that there are 2 FHAs, X and Y.

According to the NHSCR, there were 20,000 moves into FHA X, and 30,000 into FHA Y. Note that this is total moves, not just moves between these two FHAs.

According to the Patient Register there were 15,000 moves into FHA X and 25,000 into FHA Y.

We calculate the movement factor for this pair of FHAs as:

$$\frac{\text{(Total NHSCR inflow to destination + total NHSCR outflow from origin)}}{\text{(Total PR inflow to destination + total PR outflow from origin)}}$$

$$= \frac{20{,}000 + 30{,}000}{15{,}000 + 25{,}000}$$

$$= 1.25$$

The result is that every internal migration flow between a local authority in FHA X and a local authority in FHA Y will be multiplied by 1.25.

We then multiply the number of moves of children under 1 within England and Wales by 1.5, to take account of moves between LAs within a single FHA. For flows to Scotland and Northern

Ireland, we calculate the flows of children under 1 by taking the flows of 1-year-olds and multiplying by 0.5. This is because the average age of a child under 1 in the dataset is approximately 6 months and as such they will have had half the time of a 1-year-old in which to move.

We then apply disclosure control to prevent the potential identification of an individual in the data. The method we use is perturbation, meaning we make changes to a small number of ages of movers and estimates of the numbers of movers. However, these small changes will have negligible impact on the overall quality of the data.

We then apply additional movement factors to the number of moves to Scotland and Northern Ireland to ensure that we meet the totals specified by NRS and NISRA, based on the Scottish NHSCR and Northern Irish medical card data respectively.

Overall the production of internal migration estimates can be split into 2 stages. The first of these stages creates the initial England and Wales stocks for the current data year (Figure 2) while the second stage creates the final estimates (Figure 3).
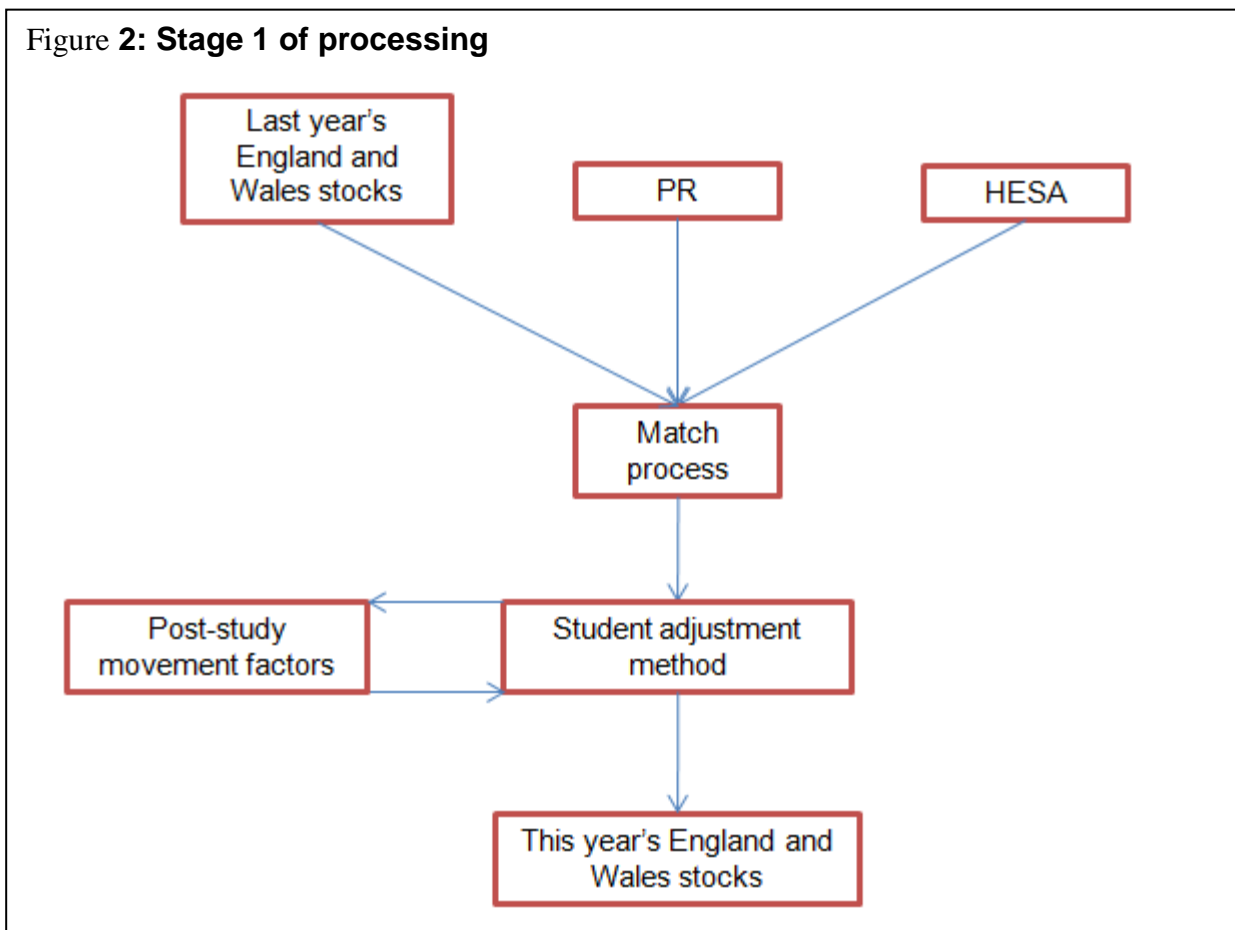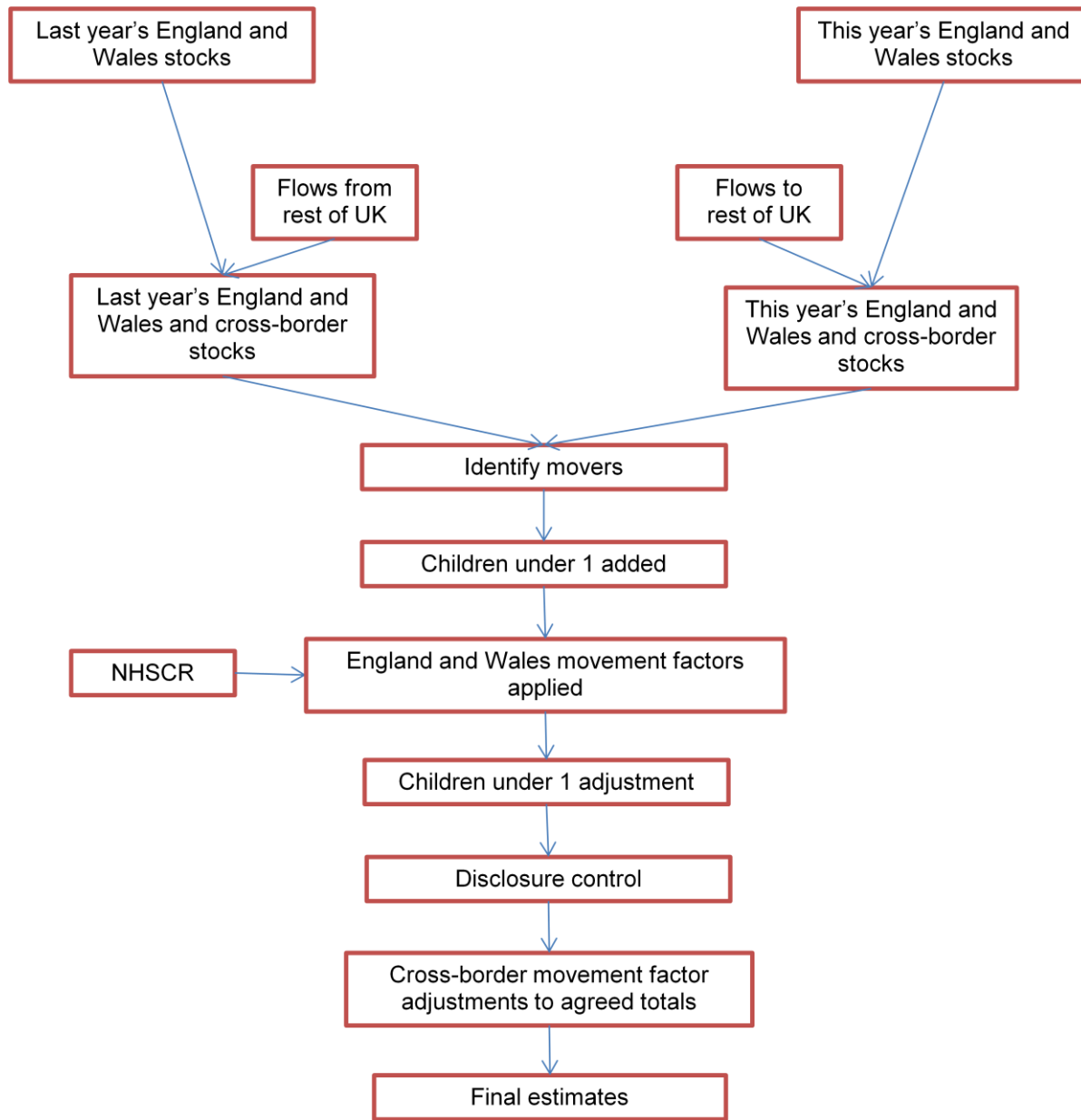

Figure **2: Stage 1 of processing**

Figure **3: Stage 2 of processing**

```
┌─────────────────────────┐                          ┌─────────────────────────┐
│ Last year's England and │                          │ This year's England and │
│      Wales stocks       │                          │      Wales stocks       │
└─────────────────────────┘                          └─────────────────────────┘
              ↘              ┌──────────────┐   ┌──────────────┐      ↙
                            │ Flows from   │   │  Flows to    │
                            │ rest of UK   │   │  rest of UK  │
                            └──────────────┘   └──────────────┘
   ┌─────────────────────────┐                    ┌─────────────────────────┐
   │ Last year's England and │                    │ This year's England and │
   │ Wales and cross-border  │                    │ Wales and cross-border  │
   │         stocks          │                    │         stocks          │
   └─────────────────────────┘                    └─────────────────────────┘
                      ↘                                   ↙
                          ┌─────────────────────┐
                          │   Identify movers   │
                          └─────────────────────┘
                                    ↓
                          ┌─────────────────────┐
                          │ Children under 1 added│
                          └─────────────────────┘
                                    ↓
   ┌──────────┐          ┌─────────────────────────────────┐
   │  NHSCR   │ ────────→│ England and Wales movement factors│
   └──────────┘          │            applied              │
                          └─────────────────────────────────┘
                                    ↓
                          ┌─────────────────────────┐
                          │ Children under 1 adjustment│
                          └─────────────────────────┘
                                    ↓
                          ┌─────────────────────┐
                          │  Disclosure control │
                          └─────────────────────┘
                                    ↓
                          ┌─────────────────────────────┐
                          │ Cross-border movement factor │
                          │  adjustments to agreed totals│
                          └─────────────────────────────┘
                                    ↓
                          ┌─────────────────────┐
                          │   Final estimates   │
                          └─────────────────────┘
```

2.6   Summary of data sources and methods

As explained, we derive the internal migration estimates from 3 main data sources (PR, HESA and NHSCR), with additional information from the Scottish NHSCR and the Northern Irish health card register to improve estimates of cross-border flows.

We use the annual PR extracts to identify moves at local authority level, but require the weekly NHSCR to cover within-year moves missed by the PR. We supplement these with HESA data in order to improve coverage of moves of young adults, who are less likely to re-register promptly with a GP after moving.

This combination of PR, NHSCR and HESA data will cover most moves. However, it is important to realise that it cannot pick up all moves. Limitations include:

1. The data exclude moves into and out of prison and the armed forces.

2. The additional information from the HESA dataset allows an adjustment for missed moves of those starting and finishing higher education. However, other groups of people may also not register promptly with a GP when they move, meaning their moves are still liable to be missed.

3. We don't apply the HESA adjustment to cross-border flows.

Despite these limitations, the estimates are the best currently available source of information on internal migration. However, over time data sources change and new data sources become available. We are therefore researching improvements to the methods and intend to introduce them in June 2017.

3.   Chronology of changes

In this section we outline the details of changes to methods over time, along with the reasoning for those changes and the years for which data are available based on those changes. Note that there have also been a number of geography code changes over the years; the recent changes can be found in Annex 1. We will announce future changes to the way in which internal migration estimates are calculated on the internal migration methodology webpage.

Table 1 summarises which methods underlie which years of the official internal migration estimates.

Table 1: Summary of methodology over time

Time period	Methodology	Lowest level of geography
Year ending June 2012 and onwards	New student adjustment method Combination of Patient Register (PR), NHS Central Register (NHSCR) and Higher Education Statistics Agency (HESA) data	348 local authorities
Year ending June 2009 to year ending June 2011	Old student adjustment method Combination of PR, NHSCR and HESA data	348 local authorities
Year ending June 2002 to year ending June 2008	Old student adjustment method Combination of PR, NHSCR and HESA data	376 local authorities
Year ending June 1999 to year ending June 2001	No student adjustment Combination of PR and NHSCR data	376 local authorities
Older estimates (back to 1975)	NHSCR data only	Health authorities

More information on each of the methods changes is as follows.

Internal migration estimates for the year ending June 2012

The 2012 estimates featured an updated student adjustment method and a change in the way children under 1 were calculated.

The changes to the student adjustment method are outlined in An Improved Method of Estimating Student Migration. The main change was a move from using aggregate Higher Education Statistics Agency (HESA) data to using record-level data in the processing.

We introduced the change to children under 1 because of a disproportionate number of these children being determined through the application of movement factors; the new method reduces the number of these moves being generated. We made a further minor amendment in the estimates for the year ending June 2013, to simplify estimation of moves between local authorities within each FHA.

We have also published a back series of data using the 2012 methodology back to the year ending June 2009. However, it is important to note that this was a research series only and does not replace the official estimates for those years.

Internal migration estimates for the year ending June 2009

In 2009, there were several counties where the districts were merged into 1 or 2 unitary authorities, reducing the total number of local authorities in England and Wales from 376 to 348. This meant that total levels of internal migration are reduced slightly from 2009 onwards compared with earlier years because moves between certain former districts were previously counted as a move between local authorities but are now counted as a move within a single local authority. For example, a move between Caradon and Penwith districts in 2008 would be an internal migration move, but in 2009 it would be a move within Cornwall unitary authority and so would not be included in the statistics. However, across the country the total reduction in moves is small; we compared data for the year ending June 2012 and found the reduction was around 35,000.

The full list of recent geographic changes is in Annex 1. In addition to the creation of the unitary authorities described, in 2009 there were also code changes to Powys and Merthyr Tydfil (because of a minor boundary change) and to the Isles of Scilly (to make the coding classification more consistent). There were also small numbers of local authorities that had their codes changed in 2012 and 2013 because of minor boundary changes.

Internal migration estimates for the year ending June 2002

We first introduced a student adjustment using HESA data in 2010, for the estimates for the year ending June 2009. However, we also published revised data back to the year ending June 2002, replacing the previously published estimates. As such, the estimates for the year ending June 2002 are the earliest for which we made a student adjustment.

Older estimates

Annual estimates at local authority level are available back as far as the year ending June 1999. However, before the year ending June 2002 we used a combination of PR and NHSCR data only, with no student adjustment.

Earlier statistics are available back to 1975 but only using the NHSCR as a data source, meaning they are only available at regional and health authority level. These statistics were produced quarterly.

4. Contact details

For more information on internal migration statistics and the methods used to create them, please contact:

Nicola White
Migration Statistics Unit
migstatsunit@ons.gsi.gov.uk
+44 (0)1329 444647

Annex 1: Geography changes

2009 changes

2008 local authority	2009 local authority	Old code	Updated code
Powys	Powys	W06000007	W06000023
Merthyr Tydfil	Merthyr Tydfil	W06000017	W06000024
Bedford	Bedford	E07000002	E06000055
Mid Bedfordshire	Central Bedfordshire	E07000001	E06000056
South Bedfordshire	Central Bedfordshire	E07000003	E06000056
Chester	Cheshire West and Chester	E07000013	E06000050
Ellesmere Port and Neston	Cheshire West and Chester	E07000016	E06000050
Vale Royal	Cheshire West and Chester	E07000018	E06000050
Congleton	Cheshire East	E07000014	E06000049
Crewe and Nantwich	Cheshire East	E07000015	E06000049
Macclesfield	Cheshire East	E07000017	E06000049
Caradon	Cornwall	E07000019	E06000052
Carrick	Cornwall	E07000020	E06000052
Kerrier	Cornwall	E07000021	E06000052
North Cornwall	Cornwall	E07000022	E06000052
Penwith	Cornwall	E07000023	E06000052
Restormel	Cornwall	E07000024	E06000052
Isles of Scilly	Isles of Scilly	E07000025	E06000053
Chester-le-Street	County Durham	E07000054	E06000047
Derwentside	County Durham	E07000055	E06000047
Durham	County Durham	E07000056	E06000047

Easington	County Durham	E07000057	E06000047
Sedgefield	County Durham	E07000058	E06000047
Teesdale	County Durham	E07000059	E06000047
Wear Valley	County Durham	E07000060	E06000047
Alnwick	Northumberland	E07000157	E06000048
Berwick-upon-Tweed	Northumberland	E07000158	E06000048
Blyth Valley	Northumberland	E07000159	E06000048
Castle Morpeth	Northumberland	E07000160	E06000048
Tynedale	Northumberland	E07000161	E06000048
Wansbeck	Northumberland	E07000162	E06000048
Bridgnorth	Shropshire	E07000182	E06000051
North Shropshire	Shropshire	E07000183	E06000051
Oswestry	Shropshire	E07000184	E06000051
Shrewsbury and Atcham	Shropshire	E07000185	E06000051
South Shropshire	Shropshire	E07000186	E06000051
Kennet	Wiltshire	E07000230	E06000054
North Wiltshire	Wiltshire	E07000231	E06000054
Salisbury	Wiltshire	E07000232	E06000054
West Wiltshire	Wiltshire	E07000233	E06000054

2012 changes

Local authority	2011 Code	2012 Code
St Albans	E07000100	E07000240
Welwyn Hatfield	E07000104	E07000241

2013 changes

Local authority	2012 code	2013 code
Gateshead	E08000020	E08000037
Northumberland	E06000048	E06000057
East Hertfordshire	E07000097	E07000242
Stevenage	E07000101	E07000243