Office for National Statistics

International migration data for Scotland: quality assurance of administrative data used in population statistics, Jan 2017

The quality assurance undertaken on administrative data for international migration data for Scotland used within Population Statistics Division (PSD) publications.

Contact:
Pete Large
pop.info@ons.gsi.gov.uk

Release date:
27 January 2017

Next release:
To be announced

Table of contents

1 . Introduction

Population Statistics Division (PSD) within Office for National Statistics (ONS) receives international migration data for the devolved administration of Scotland from National Records of Scotland (NRS).

International migration data are fundamental to a number of our publications including but not limited to National Population Projections (NPP), the Local Area Migration Indicator Suite (LAMIS) and the Mid-Year Population Estimates (MYE).

Although estimates of international migration for Scotland are created using a number of different data sources, in the interest of transparency and quality assurance the datasets provided to ONS from NRS are regarded as administrative data sources.

Within PSD, International Migration data for Scotland were assessed separately by relevant teams using the UK Statistics Authority's Quality Assurance of Administrative Data Toolkit, which assesses an administrative data source in terms of the risk to data quality and its onward use in statistics as well as the profile of the statistics produced from the source. The results of those assessments are an A1 rating.

The A1 rating means that a basic level of assurance is required for these sources and this report will provide information to meet this level of assurance. If you feel that this report does not adequately provide this assurance then please contact pop.info@ons.gov.uk with your concerns.

The toolkit outlines 4 areas for assurance; the rest of this report will be split into these areas. The areas for assurance are:

- operational context and administrative data collection

- communication with data supply partners

- quality assurance principles, standards and checks applied by data suppliers

- producer's quality assurance investigations and documentation

2 . Operational context and administrative data collection

The main source of data used to produce international migration estimates is the International Passenger Survey (IPS). In addition to the IPS, the Labour Force Survey (LFS), National Health Service Central Register (NHSCR) and the Community Health Index (CHI) are all used to derive international migration estimates for Scotland.

2.1 International Passenger Survey (IPS)

The IPS is a continuous sample survey carried out by us which collects information about passengers entering and leaving the UK at all principle airports, sea ports and the Channel Tunnel. Each year, between 700,000 and 800,000 people are sampled through the IPS. The survey is voluntary and the participant is not obliged to complete the survey.

IPS interviewers are stationed across the UK at all principal airports, sea ports and the Channel Tunnel. Using systematic sampling, interviewers approach passengers at each point of entry and carry out the survey if consent from the participant is obtained.

The data obtained from IPS is intention based and therefore carries the risk of producing data which may not reflect the true outcome of a migrants journey, for example, if a migrant states they intend to live in the UK for more than 12 months, they are counted as an international migrant, however it is entirely possible that this person changes their mind after only 3 months and leaves the UK.

To account for this, we publish a data series known as the Long-Term International Migration (LTIM) on a quarterly basis. LTIM is created by making adjustments to IPS data, accounting for change of intentions, asylum seekers and cross border migration flows. These are the data National Records Scotland (NRS) use when producing their international migration estimates.

2.2 Labour Force Survey (LFS)

While IPS data is used to provide an estimate of international migration into the UK, the allocation of these migrants to Scotland is based on the LFS.

LFS is the UK's largest household survey, collecting information on employment, unemployment, housing, ethnicity, religion, health and education. Although the IPS asks participants the question of where they intend to settle within the UK, since the IPS is based on intentions only, it is likely that migrants actually settle somewhere other than their intended destinations, which in turn could result in lower data quality. Data collected by the LFS are based on actual locations of migrants and are therefore seen to be more reliable than IPS data for the purpose of calculating Scotland's share of total UK migrants.

2.3 National Health Service Central Register (NHSCR)

The NHSCR system records registrations with NHS doctors, new registrations from migrants coming from overseas are used to allocate international migration flows to National Health Service (NHS) board areas. In addition, the age and sex distributions of international immigrants are obtained directly from the NHSCR.

A known issue with the NHSCR is the fact that emigrants often do not de-register with their doctor when moving overseas resulting in an undercount of migrants, to address this, NRS adopted methodology to allocate the international outflows from NHS board areas using:

- international inflows from 2 years prior to the reference date

- outflows to the rest of the UK

- the population share of each NHS board area

The age and sex distributions for international emigrants also cannot be reliably estimated from directly using NHSCR data due to the undercounting mentioned above. Current NRS methodology states that international emigrants' age and sex distributions in an NHS board area are estimated based on:

- out-migrants to the rest of the UK leaving from that NHS board area

- in-migrants from overseas 2 years prior to the reference date, aged on by 2 years

Please see the Information on the quality assurance arrangements for administrative data used in population estimates report for further information regarding NRS' use of NHSCR data.

2.4 Community Health Index (CHI)

The CHI is a register of all patients in NHS Scotland. Anonymised CHI data containing the postcodes of NHS patients are used to distribute migrants to council areas (CA) within Scotland. It is known that CHI data also suffers from an undercounting of international emigration like the NHSCR, so international emigrants are allocated using a combination of in-migrants to Scotland from overseas and migrants leaving Scotland for the rest of the UK.

Please see the [Information on the quality assurance arrangements for administrative data used in population estimates](#) report for further information regarding NRS' use of CHI data.

2.5 Home Office (HO) Data

Data from the HO is used to account for the number of asylum seekers, as well as used in the adjustments of IPS data for creating LTIM estimates. Separate quality assurance reports for data received from the Home Office are being published by Population Statistics Division (PSD) in 2017.

2.6 Further information

The data collection process through to the supply of data to PSD is outlined in this section.

There are 4 routes that could be taken for inclusion in the Scottish long-term international migration estimates: the IPS route, the LFS route, the medical information route and the asylum seeker route.

The IPS route is where a migrant enters the UK through an airport, a sea port or the Channel tunnel. Should they then be approached by an IPS interviewer and agree to take part then their information will be recorded.

The LFS route is where the household that the migrant is part of is chosen for sampling. Should a member of the household agree to take part in the survey then information relating to the migrant will be collected.

The medical information route is where a migrant registers with an NHS doctor in Scotland. If this is their first registration, then an indicator to denote that a first registration by an overseas national is used. General Practitioner (GP) registration data is then collected by the NHSCR and CHI. This indicator is used as the basis to extract data for age, sex and area distribution within Scotland.

The asylum seeker route contains information from obtained from applications for asylum, collated by the Home Office. A separate Quality Assurance Administrative Data report covering asylum seeker data is scheduled for publication in February 2017.

The data from the various routes are combined by NRS and used to determine long-term international migration figures for Scotland; for more information on NRS' methodology for producing international migration estimates please see [Mid-Year Population Estimates for Scotland: Methodology Guide 2015](#). The mid-year tables are then sent by NRS to PSD for use in statistical outputs.

3 . Communication with data supply partners

3.1 Population Estimates Unit (PEU)

The population estimates unit has no explicit and specific agreement with National Records of Scotland (NRS) for the supply of this data, but provision is implied by the accepted role of us in collating and publishing the UK population estimates. Regular contact between colleagues from both organisations is maintained throughout the year.

NRS publish their mid-year estimates on an annual basis in April. The data received by PEU are identical to published mid-year estimates available on the NRS website.

3.2 Population Projections Unit (PPU)

We are currently commissioned to prepare and produce principal and variant national population projections on behalf of the devolved administrations of Scotland, Northern Ireland and Wales.

There is currently an agreement in place whereby the devolved administrations, including Scotland, satisfy our principal data requirements for the production of National Population Projections (NPP), this includes the latest available migration data at the time of request. Formal requests for data are sent by PPU in years when NPP are to be published.

ONS and NRS engage in regular discussion on the progress of work and data requirements. Additionally, in a NPP publication year, an initial NPP committee meeting is held in January where ONS will outline their intended progress and provide NRS with a time table of events, data requirements are discussed in this meeting and a guide date for NRS to provide us with Scotland international migration data is set. A formal data request is sent after this initial meeting. In July of the same year, a second committee meeting takes place during which ONS and NRS discuss the agreed assumptions for international migration.

The data PPU receive from NRS is provided as a spreadsheet containing international migration estimates to and from Scotland by sex and single year of age.

3.3 Migration Statistics Unit (MSU)

There is no formal agreement for international migration data between MSU and NRS. A formal data request are sent by MSU when data are required. MSU provides NRS with a template for inputting international migration data as part of their formal data request. The template is a spreadsheet consisting of empty tables labelled by year. NRS supply international flows based on the years requested in the template.

3.4 Engagement with users

PSD continually engages with users, through a variety of means, to understand how our outputs are meeting their requirements. Feedback provided tends to relate to the overall statistical methodology and the impact on the final statistics, rather than to any individual data source. To date no specific feedback on the use of this data source has been provided.

4 . Quality assurance principles, standards and checks applied by data suppliers

The data received by Population Statistics Division (PSD) are identical to those published in National Records of Scotland (NRS) international migration estimates. As such, NRS deems the following quality assurance is necessary when supplying the data to us.

4.1 Population Estimates Unit (PEU)

No additional quality assurance is carried on the data set received by PEU: this is because no reformatting or alterations are made to the data. The spreadsheet supplied to PEU is identical to NRS' published mid-year estimates tables, in which international migration is listed as a component of change.

4.2 Population Projections Unit (PPU)

Two NRS statisticians collate the requested data into separate spreadsheets, the resulting tables are compared for consistency, if the resultant tables are identical, the data are supplied to PPU. If there are differences between the 2 tables, the cause of the differences will be investigated by NRS statisticians.

4.3 Migration Statistics Unit (MSU)

Two NRS statisticians collate the requested data into separate spreadsheets and the resulting tables are compared for consistency. If the resultant tables are identical, the data are supplied to MSU. If there are differences between the 2 tables, the cause of the differences will be investigated by NRS statisticians.

4.4 Auditing of data

Separate audits are not conducted for international migration data NRS provides to us. As described at the start of this section, the figures in the data received by us are identical to those published in NRS' international migration estimates.

5 . Producer's quality assurance investigations and documentation

5.1 Population Estimates Unit (PEU)

Population Estimates Unit (PEU) receives data for international migration flows to and from Scottish council areas. The data are supplied in a spreadsheet listing various components of change including international migration. This spreadsheet is identical to National Records of Scotland (NRS) published mid-year population estimates.

These data feed into the production of mid-year population estimates as components of change. The data are also included in the MYE3 and MYE3B tables published by PEU annually which present components of change. Simple checks are carried out to ensure the correct data has been supplied and that there are no causes for concern when comparing the latest available data against historical data from NRS.

5.2 Populations Projections Unit (PPU)

Populations Projections Unit (PPU) uses Scottish international migration data from NRS to apply the Scottish age and sex distribution of the National Population Projections (NPP).

Data are checked for consistency by comparing the latest received data to previously received data. Migration figures are also checked to ensure that there are no unexplainable trends. The nature of international migration data means significant differences are to be expected in relation to political developments. For example, a significant increase in working-aged migrants following the removal of work restrictions would not be deemed to be a cause for concern as there was a clear cause for the observed trends. However if significant increases were observed in only those over pension age, NRS would be contacted for an explanation and the data would be further scrutinised for potential errors.

Output data from the NPP are also checked for consistency against previous projections. The international migration component of change is checked against previous NPP output to ensure there are no unexplained trends in the projection. Should an anomaly be found, PPU will conduct further checks to identify the cause of the issue.

5.3 Migration Statistics Unit (MSU)

Migration Statistics Unit (MSU) receives international migration flows to and from Scottish council areas through a template provided by MSU in their formal data request. The latest available data is supplied alongside up to 10 years of historical data.

MSU follows a 2-person quality assurance process whereby 1 colleague verifies the latest data received is in fact the latest data and that no mistakes were made by NRS during their input of data into the template. For example data for the year 2014 would be checked against the 2014 data sent in a previous request to ensure all figures are identical. Should MSU observe any inconsistencies with the data, NRS would be immediately contacted regarding the issue and an investigation into the cause of the inconsistency is initiated.

Once the quality assurance process is complete, a second colleague from MSU carries out the same checks. NRS data are only used by MSU once both colleagues have agreed that no issues were found during the quality assurance.

6 . Conclusions

6.1 Assurance Ratings

Population Estimates Unit (PEU)

Population Estimates Unit (PEU) uses Scotland international migration estimates for the production of the UK mid-year population estimates (MYE) and the components of change tables.

The data are accredited National Statistics and fully quality assured at the time of supply. Given the low number of variables, the regular communication maintained between colleagues in Office for National Statistics (ONS) and National Records of Scotland (NRS) in regards to data needs and quality and the fact that PEU use the data for the same purpose for which they have been accredited as National Statistics, there is deemed to be a low risk of quality concerns.

MYEs are produced on an annual basis and are of great significance to local authorities for planning and the allocation of resources. There is also considerable interest from academics in demographic research, as a result, the MYEs are considered to have a higher public interest profile.

PEU judge that an A1 level of assurance was appropriate for this data source for the purpose it is used for.

Population Projections Unit (PPU)

Scotland international migration data is used by the population projections unit for the production of National Population Projections (NPP).

The data are accredited National Statistics and fully quality assured at the time of supply. Given the low number of variables, the regular communication maintained between colleagues in ONS and NRS in regards to data needs and quality and the fact that PPU use the data for the same purpose for which they have been accredited as National Statistics, there is deemed to be a low risk of quality concerns.

The NPP are used both within and outside government as the definitive set of population projection figures. Within government, they are used for important decision-making, such as planning the allocation of resources as well as being the basis for other official projections such as the Subnational population projections and Household projections. Outside government, NPP are widely used by academics in demographic research as well as other statistical agencies such as Eurostat for the production of their own population projections. As such, NPP are considered to have a higher public interest profile.

PPU judge that an A1 level of assurance was appropriate for this data source for the purpose they are used for.

Migration Statistics Unit (MSU)

Scotland international migration data are included in the Local Area Migration Indicators Suite (LAMIS).

The data are accredited National Statistics and fully quality assured at the time of supply. Given the low number of variables, the regular communication maintained between colleagues in ONS and NRS in regards to data needs and quality and the fact that MSU use the data for the same purpose for which they have been accredited as National Statistics, there is deemed to be a low risk of quality concerns.

LAMIS is published annually and the data is of interest to local authorities and health services in particular for the purposes of planning resource allocations. There is interest from academics working in the area of demographic research, however, LAMIS does not feed into central policy decisions and does not have receive significant media attention. As such MSU deem LAMIS to have a medium public interest profile.

MSU judge that an A1 level of assurance was appropriate for this data source for the purpose it is used for.

6.2 Strengths and limitations

The data are deemed to be the best measure for international migration into Scotland at both the national and subnational level by NRS and are accredited National Statistics for the purpose that Population Statistics Division (PSD) use them.

However, there are some recognised limitations to data quality:

1. There is a recognised deficiency in the number of registrations and de-registrations from young adult males.

2. There is a lag between the arrival of a person into Scotland and their registration with a General Practitioner (GP).

3. International emigration is estimated from GP registration data and there are known to be instances of people failing to de-register.

To account for this these issues adjustments are made in the Scottish international migration methodology, for more information please see Mid-Year Population Estimates for Scotland: Methodology Guide 2015.