Office for National Statistics

Article

Population statistics research update: June 2017

A description of ONS research into improving standard population statistics outputs, including recent progress and plans for future research.

Contact:
Pete Large
pete.large@ons.gsi.gov.uk
+44 (0)1329 444236

Release date:
22 June 2017

Next release:
To be announced

Table of contents

1 . Introduction

This article describes the research and development planned for our standard population statistics outputs as at June 2017. To help users quickly find the information they need, the article includes a description of each project similar to that provided in the previous research update in January 2017, followed by an “Update” paragraph on progress since January.

Alternative methods of producing statistics following the 2021 Census are being investigated within a separate programme of work.

We welcome your comments on our proposed work and any suggestions for other research and development which would increase the value of our statistics to you.

The following sections will discuss and update on the progress of the various research and development planned for our standard population statistics outputs as at June 2017.

2 . Internal migration

The largest element of population change at the local authority level is internal migration – that is, people moving from one local authority in the UK to another. The methods for estimating internal migration were significantly improved in 2013 when the availability of new Higher Education Statistics Agency (HESA) data allowed a simpler and more reliable approach to be adopted for estimating migration of students. This further research has four main aims, which are detailed in this section.

Allowing the production of estimates for 2016 following the closure of the NHS Central Register (NHSCR) system (extracts from which are used in calibrating the estimates)

The fundamental approach to producing estimates of internal migration moves is comparing a file (produced by linking the Patient Register to the HESA dataset of people in higher education) of where people were living at the start of the year with the corresponding file of where they were living at the end of the year. Any record with a change of address must have moved (that is, migrated) during the year.

However, this approach misses some types of moves – where people have moved several times during the year or where they were not present at the start, or the end, of the year (for example, babies born during the year). Extracts from the NHSCR data are used to adjust the initial internal migration estimates to reflect these moves.

As the NHSCR data source was discontinued in February 2016, the production of the 2016 estimates requires an alternative method of estimating these additional moves. Our research indicates that using the previous year's adjustments is likely to be the best approach, both in terms of accuracy and simplicity, so we have adopted this method for the 2016 estimates. Estimates for later years should not require such adjustments.

Quality assuring the Personal Demographic Service (PDS) data source as an appropriate replacement for the Patient Register (PR) data source when the latter is discontinued

The PR data source currently used in producing the internal migration estimates is due to be closed in autumn 2017. However, the alternative PDS data source now available to us seems to have several advantages over the PR. We have started quality assuring the PDS data and checking that it is appropriate for use in producing the estimates, with a view to moving to using the PDS data in the mid-2017 population estimates.

Developing improved models for the destination of students after they graduate

The current methodology is thought to be an improvement on the previous method in that it is much more accurate in estimating migration of students to their place of study. However, there is still scope for improving methods for estimating the destination of students who move after leaving higher education.

At present, the assumption is that a student not updating their PR record upon leaving higher education will either stay in the local authority in which they lived while studying, or return to their previous address as recorded on the PR, with an increasing probability of moving to their previous address as time goes on.

We are developing an alternative approach of applying an Origin-Destination matrix to these students (so, for example, 10% of students in Southampton who don’t update their PR record when leaving higher education move to Portsmouth).

Improving estimation of migration moves within the year not identified by comparing addresses at the start and the end of the year

As described previously in this section, the current methodology uses NHSCR data to adjust the initial estimates to allow for moves of people who were only present at either the start or the end of the year, but not both. We are investigating an alternative approach of directly identifying these moves, for example, by linking the start-year and end-year population stocks file with registrations of births and deaths.

Methods developed in the first of these four strands are used in the 2016 mid-year estimates published in June 2017. Methods developed in the remaining three strands are planned to be implemented for the 2017 mid-year population estimates in June 2018. If possible, we will also use the new methods to produce a revised back-series of population estimates from 2011 to 2016.

Update: June 2017

This project is on track for delivery as planned. The methods for the “graduate adjustment” and direct identification of moves within a year are being quality assured. We are publishing details of the new methods as part of the release of the mid-year population estimates in June 2017, before implementing the methods in 2018 as described in this section.

3 . Local authority emigration

Estimates of emigration at the national level are derived from the International Passenger Survey (IPS), with adjustments made to reflect for asylum seekers. Disaggregating these estimates to the local authority level relies on complex regression models, which were first set up in 2011. This project aims to update these regression models to reflect as well as possible the current characteristics of migration at the local level.

The proposed model is broadly similar to the existing model: both being based on a Poisson Regression approach. However, there are three differences between the models.

Changing the explanatory variables

The proposed model uses a different set of explanatory variables to those used in the existing model. As with the existing model, the set of explanatory variables has been selected by a combination of manual selection and algorithmic stepwise selection, providing a compromise between strict statistical optimality (based on the estimation period) and intuitively plausible explanatory variables.

Using an “offset term”

The proposed model introduces an 'offset term' in the regression equation to reflect the size of the population of the LA. This term effectively changes the model from a direct model of counts to a model of rates (that is, modelling emigration as a proportion of the start population). Such a term is a standard feature of Poisson models where the 'population at risk' is different for different observations.

Removing new migration geographies for outflows from constraining process

The current model uses a non-standard geography called new migration geographies for outflows (NMGos). An NMGo is a collection of local authorities that is treated as a single source of emigration when constraining to the IPS figure. The proposed model eliminates the use of NMGos.

The removal of this constraint means we are making less use of the IPS data in allocating emigration, but also removes the practical problem that an over-estimate for one local authority could result in counter-balancing under-estimates in neighbouring local authorities. Removing the use of these non-standard geographies should also make the estimates more transparent and based on standard definitions.

Update: June 2017

This work is now complete and we plan to implement this new model alongside the new internal migration methods for the 2017 mid-year population estimates in June 2018. Again, we will also look to use the new methods in a revised back-series of population estimates from 2011 to 2016.

4 . Sex and age distribution of international migrants

The mid-year population estimates currently use 2011 Census data in applying a sex and age distribution to international migrants at the local authority level. We are investigating whether that is still the best approach or whether an alternative approach, for example, based on administrative data, should be adopted.

As with the previously-described projects, we plan to implement any new methods developed in this work in the 2017 mid-year population estimates in June 2018 and look to use the new methods in a revised back-series of population estimates from 2011 to 2016.

Update: June 2017

This work has concluded that the sex and age distribution of migrants has remained sufficiently stable since 2011 to retain the current approach based on 2011 Census data rather than switch to other datasets, which are more up-to-date but less directly related to our target concept of the usual resident population.

5 . Population characteristics

We have acknowledged user interest in a new set of population estimates by ethnic group. We launched a project in September 2016 looking at whether such estimates (consistent with the mid-year population estimates) can be derived primarily from the Annual Population Survey (APS). If this is successful we would look to extend these methods to cover country of birth and nationality.

For many years we have published tables of the population by country of birth and by nationality as part of our suite of migration-related statistics. These estimates are derived from the APS. Since that survey does not cover some people not living in households, these estimates are not consistent with the mid-year population estimates (which cover all usual residents in an area). This project is looking at combining APS data with other data sources (in particular, the 2011 Census) to produce estimates of the population by country of birth, nationality and ethnic group, which are fully consistent with the standard mid-year population estimates.

Update: June 2017

This project is on track for delivery as planned. We expect to publish a research paper, containing a description of the methods and with a set of “research estimates” in August 2017, seeking feedback on their usefulness and whether they should become a standard output.

6 . National population projections: uncertainty

Update: June 2017

Our work on this is in its concluding stages and we will provide a more detailed update later this year.

7 . National population projections: mortality assumptions

Assumptions on future mortality rates are one of the inputs into the national population projections (NPPs). We have been working with the University of Southampton to investigate whether a purely model-based approach can produce appropriate projected rates.

At present, mortality assumptions in the national projections are produced using a combination of a statistical model and expert advice from the ONS Demographic Analysis Unit and a representative of the Government Actuary's Department. This research has investigated modelling mortality improvement rates using a generalised additive model, with separate models for infant mortality and old-age mortality. The proposed approach presents a number of advantages when compared with the current method: for example, it makes fuller use of available data and is more transparent and efficient. However, there are also possible disadvantages – there is a risk of inconsistency with previous projections and the model may be more difficult to explain to users.

Update: June 2017

Following detailed testing and expert review, the NPP Committee has supported the use of the new model for the 2016-based NPPs. The current method will also be run, however, to allow validation and comparison of the results before a final decision is made. The projections are scheduled for publication in October and November 2017.

8 . Subnational population projections for England

We are working on improvements to the methodology of the subnational population projections (SNPPs), both to improve accuracy and to increase coherence with other population statistics. These will be incorporated in the methods used for the 2016-based subnational population projections for England, scheduled for May and June 2018.

Update: June 2017

In the January update, we noted completed work on improvements to estimating cross-border flows, as well as a new method for estimating the local authority distribution of asylum seekers flows by using the average of the last 5 years' data rather than only the most recent year. Since then we have progressed work on distributing UK armed forces returning from Germany and improving harmonisation of methods with the national projections. We will share more detailed information on methodological developments around the start of 2018.

In addition, the British Society for Population Studies kindly organised a meeting where we learned more about user needs for variant subnational projections. Whilst we won't include variants in our main publication, we are aiming to produce them later in 2018 and we will consult users on which variants would be most useful.

9 . Quality assurance of administrative data sources

Each of our regular statistical releases is accompanied by a Quality and Methodology Information document, which summarises information on the quality of the published statistics and the methods used to produce them. In addition, we are preparing reports on the quality assurance procedures adopted for each of the main administrative data sources used in producing our statistics.

This work is being conducted within the framework of the UK Statistics Authority's Quality Assurance of Administrative Data Sources Toolkit. All of the administrative data sources feeding in to our statistics have been assessed on the risk of quality issues having an impact on the statistics. This has allowed us to set out appropriate levels of documentation needed for each data source.

Update: June 2017

The full set of quality assurance of administrative data (QAADs) has now been published and reports are available through the following links:

- Births
- Deaths
- Patient Register
- NHS Central Register data
- International migration data for Scotland
- International migration data for Northern Ireland
- Migrant Worker Scan data
- National Insurance number (NINo) data
- UK Armed Forces data
- US Armed Forces data
- Prisoners data
- Northern Ireland internal migration and Northern Ireland Medical Card information
- Higher Education Statistics Agency data
- Home Office immigration data excluding asylum seekers
- Asylum support data
- Asylum seeker data and non-asylum enforced removals

10 . Household projections for England

We announced in January 2017 that we are taking responsibility, with immediate effect, for the production and publication of the household projections in England, previously produced by Department of Communities and Local Government (DCLG). We hope that this transfer of responsibility will further develop the consistency between the household projections and the national and subnational population projections, and allow us to improve efficiencies in their production. Following a discussion with DCLG on their experience of producing the household projections and a consultation with users, we are evaluating some proposed changes to the methods.

Update: June 2017

The consultation identified a number of areas of potential research as part of developing new methods and we have set up a Collaborative Working Group (with experts from within and outside ONS) to help take this work forward. We are exploring how we can incorporate improved methods in the 2016-based projections, planned for publication in Summer 2018.

11 . Adjustment for Foreign Armed Forces dependants

Whilst the methods for producing the mid-year population estimates work well for most local authorities in England, we recognise that they may not work so well for areas with a large number of Foreign Armed Forces, where their dependants are not reflected in available data sources in the same way as most international migrants. We are looking at whether we can use administrative data to reflect migration patterns of these dependants more accurately.

As with our research into internal migration and emigration described in this article, any methods developed through this work are planned to be implemented for the 2017 mid-year population estimates in June 2018. If possible, we will also use the new methods to produce a revised back-series of population estimates from 2011 to 2016.

Update: June 2017

We have developed a provisional method for dealing with this issue by treating these dependants as a “special population” (as we do the Armed Forces themselves) in the mid-year population estimates and the subnational population projections. Work continues on evaluating the provisional method before it can be formally adopted.

12 . Sexual identity estimates: local authorities

In October 2016, we published experimental official statistics on the sexual identity of the population of the UK and its constituent countries and regions. We are now investigating the feasibility of providing similar statistics, using information from the pooled Annual Population Survey dataset for 2013 to 2015, to the level of local authorities.

Update: June 2017

The subnational sexual identity estimates were published on 19 April 2017. Using the 3-year Annual Population Survey (APS) pooled dataset to produce subnational estimates of sexual identity does not seem to be a robust method, particularly for English counties and local authorities of Great Britain. The method means that for a few counties and a substantial number of local authorities, we cannot publish a robust estimate of their sexual identity distribution.

Increasing the sample size, either through a larger surveyed sample or through increasing the number of pooled years with a dataset, could improve sample sizes enough for us to produce and publish robust estimates for all counties and local authorities. However, rolling together more years will affect the timeliness of the estimates.

We will monitor the sample sizes for sexual identity at the subnational level within the next APS 3-year pooled dataset to assess whether these estimates are robust enough to publish.

13 . Gender identity

Working with Census, and on behalf of the office as a whole, we are taking forward work on gender identity.

The 2021 Census topic consultation identified a need among a number of data users for information about gender identity for policy development and service planning; especially in relation to the provision of health services. These requirements are strengthened by the need for information on those with the protected characteristic of gender reassignment as set out in the Equality Act 2010. The gender identity topic report on the consultation findings provides further information.

As we do not currently collect data on gender identity on any of our social surveys, research and testing work will inform our position on this topic. The gender identity research and testing plan sets out the work we will do to help us determine how to meet user needs for information on this topic.

On 13 January we published the gender identity update paper. This addresses our commitment to review the Trans Data Position Paper, which we published in 2009. The update outlines developments around the topic of gender identity. It covers:

- legislation
- the Women and Equalities Committee Transgender Equality inquiry
- data collection and question development worldwide
- our research, testing and findings so far
- the next steps and future work we will be undertaking

More information on our work on this topic is provided in our gender identity page.

Update: June 2017

Since the publication on 13 January 2017 of the gender identity update paper, in conjunction with the 2021 Census programme, two studies have been undertaken and a further one planned:

- a qualitative study to help gather public understanding and perception of collecting gender identity, both in the general population and amongst the trans community
- a quantitative study on the public perceptions around collecting gender identity (a similar exercise was carried out for sexual identity)
- a quantitative study to look at the quality of response and effect on response rates has been planned to take place in the summer

A further update on this work will be published later on in 2017 and another stakeholder engagement event on gender identity will take place in early summer 2017.