

Article

Population statistics research update: Jan 2017

An update from January 2017 on research into new and improved methods for producing our standard population statistics.

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

Release date:
24 January 2017

Next release:
To be announced

Table of contents

1. [Introduction](#)
2. [Research projects](#)

1 . Introduction

This note describes the research and development planned for our standard population statistics outputs as at January 2017. To help you quickly find the information you need, the note includes a description of each project similar to that provided in the previous research update in September 2016, followed by an update paragraph on progress since September. We also provide details of 4 projects not covered in the previous update.

A separate programme of work looking at [alternative methods of producing statistics following the 2021 Census](#) is also described.

We welcome your comments (pop.info@ons.gov.uk) on our proposed work and any suggestions for other research and development, which would increase the value of our statistics to you.

2 . Research projects

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 4 main aims.

Allowing the production of estimates for 2016 following the closure of the 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. We propose to adopt this method for the 2016 estimates. Estimates for later years should not require such adjustments (see the following aim).

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 ONS 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 leaving higher education and not updating their Patient Register record will either stay in the local authority in which they lived while studying or return to their previous address as recorded on the Patient Register, 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, 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 4 strands are planned for implementation for the 2016 Mid-Year Population Estimates in June 2017. Methods developed in the remaining 3 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: January 2017

This project is on track for delivery as planned. The interim method for calibrating moves has been agreed and we are working with the Census team to evaluate the PDS data source. A first version of the improved model for graduate destinations has been produced and we are using this to test the effect of different assumptions (for example, whether distance-learning students are included as "graduates" for this purpose). We have also succeeded in linking births and deaths registrations with the PDS as the first step in developing a more accurate alternative to the current calibration methods. We will publish a note with more detail on all this work in April 2017 and will invite comments on the work described.

Local authority emigration

Estimates of emigration at the national level are derived from the International Passenger Survey, 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. We have developed a proposed model to replace the existing model and are arranging for a further independent evaluation to ensure that this new model meets the quality standards required.

The proposed model is broadly similar to the existing model: both being based on a Poisson Regression approach. However, there are 3 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 local authority (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 NMGos from constraining process

The current model uses a non-standard geography called NMGo (New Migration Geographies for Outflows). An NMGo is a collection of local authorities that is treated as a single source of emigration when constraining to the International Passenger Survey 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 LAs. Removing the use of these non-standard geographies should also make the estimates more transparent and based on standard definitions.

If this research is successful 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.

Update: January 2017

This project is on track for delivery as planned following an independent evaluation of the proposed change by Southampton University. We will publish a note on the proposed update to the model, and its implications for the population estimates, in April.

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: January 2017

This project is on track for delivery as planned.

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) in order 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: January 2017

This project is on track for delivery as planned. If this research is successful we would look to publish a description of the methods, alongside a set of "research estimates" in summer 2017, seeking feedback on their usefulness and whether they should become a standard output.

National population projections: uncertainty

Update: January 2017

This project has evolved since the last research update in September.

We have traditionally expressed uncertainty in the projections by releasing variant projections reflecting different assumptions on future mortality, fertility and migration. While we stress that these are projections not forecasts, we recognise that a more quantitative expression of uncertainty might be useful – for example, statements like "our main projection for 2050 is x million, but there is an 80% likelihood of the value falling between y million and z million".

Developing a methodology to allow such probabilistic statements would be a very substantial investment and it will certainly not be feasible to introduce in time for our 2016-based projections, due in October or November 2017. However, this research project will investigate what such an approach could offer, what would be involved in developing it and whether it is something which we should take forward over the coming years. We plan to publish a more detailed update on this work in spring 2017.

National population projections: mortality assumptions

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

At present, mortality assumptions in the national projections are produced using a combination of a statistical model and expert demographic advice from the ONS Demographic Analysis Unit and a representative of the Government Actuary's Department. This research is investigating 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 compared with the current method: for example, it makes fuller use of available data; it is thought to be more flexible to allow improvement to the model over time and it is much quicker and cheaper. 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.

If this work is successful we would seek agreement from the National Population Projections (NPP) Committee (including the devolved administrations) before looking to implement the methods, if possible, in the 2016-based projections scheduled for publication in October 2017.

Update: January 2017

We are holding a meeting with expert users and representatives from the devolved administrations in late January 2017 to discuss our analysis of the new model and, on the basis of this, we will make a recommendation to the NPP Committee in February 2017 on the method to use for the 2016-based projections.

Subnational population projections for England

Following the release of the 2014-based Subnational Projections in May 2016, we plan to investigate a number of possible improvements in the methodology to improve the accuracy of the projections for some areas and to increase coherence with other population statistics.

The scope of this work will be determined by available resource. Proposed topics include:

- further harmonisation of methods with the national projections
- requirements for, and production of, variant projections as part of the subnational projections release
- enhancement of the methodology for estimating cross-border flows by treating flows to or from Wales, Scotland and Northern Ireland separately
- treating prisoners as a “special population” (not aged on with the rest of the population of an area)
- alternative trend estimation periods

We plan for the results of this work to be incorporated in the methods used for the 2016-based Subnational Projections, provisionally scheduled for spring 2018.

Update: January 2017

We have completed these planned improvements to the methods for estimating cross-border flows. We have also agreed a new method for estimating the LA distribution of asylum seekers' flows by using the average of the last 5 years' data rather than only the most recent year. This change is expected to improve the accuracy of the projections for this element of migration and also makes the approach consistent with that used for most other components of the projections.

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 these. 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: January 2017

We have moved back publication dates from the provisional timetable described in September and have combined some planned reports to make it easier to find the information relating to particular sources. The confirmed publication timetable for these reports is:

Table 1: Publication timetable

December 2016 (published)	January 2017	February 2017
Births	Scotland International Migration	National Insurance Health Card Information
Deaths	Northern Ireland International Migration	Higher Education Statistics Agency
Patient Register	Migrant Workers Scan	Certificates of Sponsorship/ Certificates of Acceptance/ Migrants Leave to Remain/ Visa Data
NHS Central Register	National Insurance numbers	"National Asylum Seeker Support Service"
	UK Armed Forces	Other Asylum Seekers and Non-Asylum Enforced Removals
	Foreign Armed Forces	
	Prisoners	

Source: Office for National Statistics

There are also 4 projects not covered in our previous research update.

Household projections for England

We announced in January 2017 that ONS is taking responsibility with immediate effect for the production and publication of the household projections in England, previously produced by the Department for 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 make some efficiencies in their production. Following a discussion with DCLG on their experience of producing the household projections, we are evaluating some proposed changes to the methods. We will shortly be launching a consultation inviting comments on the proposed changes and on the formats of the published statistics. Any changes to the methodology will be included in the 2016-based Household Projections due for publication in summer 2018.

Adjustment for Foreign Armed Force 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 earlier in this note, 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.

Sexual identity estimates: local authorities

In October 2016, we published experimental official statistics on 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. We aim to publish the outcome of this work by March 2017. If it is possible to produce LA level estimates, then we would expect to update these annually and publish them as part of our annual sexual identity release.

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 amongst 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 whether, and how best, to meet user needs for information on this topic.

On 13 January 2017, we published a [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](#).