Office for National Statistics

Article

Population statistics research update: September 2016

An update from September 2016 on research into new and improved methods for producing our standard population statistics.

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

Release date:
21 September 2016

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 September 2016. A separate programme of work looking at [alternative methods of producing statistics following the 2021 Census](#) is described on our website.

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. Please get in touch using the contact details accompanying this article.

2 . Executive summary

Internal migration summary

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 a new data source 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 NHS Central Register (NHSCR) system (extracts from which are used in calibrating the estimates)

- quality assuring the Personal Demographic Service (PDS) data source as an appropriate replacement for the Patient Register data source when the latter is discontinued (this quality assurance work will also cover use of this data source in the local authority distributions of immigration in the mid-year population estimates , and in the small area population estimates)

- improving models for the destination of students after they graduate

- improving the estimation of migration moves which are not identified by comparing addresses at the start and the end of the year (for example, the move of someone born after the start of the year)

Find out more in the main internal migration section.

Local authority emigration summary

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 2010 and updated to reflect 2011 Census data in 2013. This project is to update these regression models to reflect (as well as possible) the current characteristics of emigration at the local authority level.

Find out more in the main local authority emigration section.

Sex and age distribution of international migrants summary

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 plan to investigate whether that is still the best approach or whether an alternative approach, for example based on administrative data, should be adopted.

New methods developed in these 3 projects are planned to be used in the 2017 mid-year population estimates planned for June 2018. If possible, we will also use the new methods to produce a back series of population estimates from 2011 to 2016.

Find out more in the main distribution of international migrants section.

Population characteristics summary

We have acknowledged user interest in a new set of population estimates by ethnic group. We plan to launch 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. If this is successful we would look to extend these methods to cover country of birth and nationality.

Find out more in the main population characteristics section.

National population projections: Uncertainty summary

This planned project is intended to extend the existing information on uncertainty around the national population projections, both in the provision of variant projections and the analyses published in the Accuracy Report, to help users interpret the projections and make appropriate use of them. Subject to discussions with the National Population Projections (NPP) Committee (including the devolved administrations), we provisionally plan for the results of this work to be released with the 2016-based projections in autumn 2017.

Find out more in the main population projections: uncertainty section.

National population projections: Mortality assumptions summary

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. If this work is successful we would seek agreement from the NPP Committee (including the devolved administrations) before implementing the methods, if possible, in the 2016-based projections scheduled for publication in Autumn 2017.

Find out more in the main population projections: mortality section.

Subnational population projections for England summary

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. 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.

Quality assurance of administrative data sources summary

Every one of our outputs 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. Publication of this series of reports is planned to be completed by the end of 2016.

Find out more in the main quality assurance of administrative data sources section.

3 . 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.

1. 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 NHS Central Register (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.

2. 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.

3. 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 (PR) 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).

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

As described above, 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.

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 2010. 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:

1. 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.

2. 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.

3. 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 which is treated as a single source of emigration when constraining to the International Passenger Survey (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 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.

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 plan to investigate 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.

Population characteristics

We have acknowledged user interest in a new set of population estimates by ethnic group. We plan to launch 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. If this is successful we would look to extend these methods to cover country of birth and nationality.

As this project has not yet started we can give only a very high level description of the planned work. 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 Annual Population Survey (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 expected to look 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.

If this research is successful we would look to publish the first estimates using these methods in summer 2017, seeking feedback on their usefulness and whether they should become a standard output.

National population projections: Uncertainty

This planned project is intended to extend the existing information on uncertainty around the national population projections, both in the provision of variant projections and the analyses published in the Accuracy Report, to help users interpret the projections and make appropriate use of them. As this project has not yet started, we can't yet describe the specific approaches that will be investigated.

Subject to discussions with the National Population Projection (NPP) Committee (including the devolved administrations), we provisionally plan for the results of this work to be released with the 2016-based projections in autumn 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 our 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 to the current method: for example, it makes fuller use of available data; it is thought to be more flexible to allow improvement 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 Projection (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.

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:

1. Further harmonisation of methods with the national projections.

2. Requirements for, and production of, variant projections as part of the subnational projections release.

3. Enhancement of the methodology for estimating cross-border flows by treating flows to and from Wales, Scotland and Northern Ireland separately.

4. Treating prisoners as a "special population" (not aged on with the rest of the population of an area).

5. 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.

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 key 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. Table 1 sets out the provisional timetable for publication of these reports.

Table 1: Provisional timetable for publication

September 2016	October 2016	November 2016
Births	Patient Register	Higher Education Statistics Agency (HESA)
Deaths	National Records of Scotland international migration	Northern Ireland health card information
NHS Central Register cross-border flows	Northern Ireland Statistics and Research Agency international migration	Other asylum seekers
NHS Central Register multiple moves	Migrant Workers Scan	Dependants' data from UK forces Germany
Certificates of sponsorship	National Insurance numbers (NINos)	UK regular forces and Gurkhas
Certificates of acceptance		National Records of Scotland internal migration
Scottish cross-border flows		Northern Ireland Statistics and Research Agency internal migration
Migrants leave to remain (study)		National Asylum Seeker Support Service
Migrants leave to remain (employment)		Non-asylum enforced removal
		Prisoners
		US Air Force
		Visa data
		International Passenger Survey (IPS)

Source: Office for National Statistics