Office for
National Statistics

2021
Census

ONS Census
Transformation Programme

Summary of feedback on
the 2015 administrative
data research outputs

May 2016

Table of contents

Office for National Statistics

1. Introduction

ONS would like to thank everybody who provided feedback on the Administrative Data Research Outputs that were published in October 2015.

Following publication, users were invited to provide feedback on:

- the level and quality of information contained in the main report

- how the methodology could be improved, including suggestions of additional administrative data sources to achieve this

- possible explanations for the differences between the Research Outputs and the official estimates

- the usefulness of the interactive content

- any general comments

The purpose of this report is to provide an overview of the feedback received including plans to make improvements as a result. Section 2 focuses on the feedback that was provided, Section 3 summarises the actions that ONS plans to take in response to this feedback and Section 4 outlines the plans for the next release of Research Outputs due to be published in autumn 2016.

2. Summary of feedback

Feedback was sought through a set of structured questions via a [feedback form](#) that was published alongside the Research Outputs. This section provides a summary of the comments received on each of the topics that respondents were asked to consider, including actions that ONS plans to take in response to the feedback. Feedback was received from a number of local authorities and Government Departments.

2.1 Level and quality of information presented in the research report

A theme throughout the majority of responses was that the main report was comprehensive and generally provided a clear account of how the Statistical Population Dataset[1] (SPD) had been constructed. The methodology used was well covered and the case studies provided useful explanations of the limitations of the current method. Some respondents commented that it would be more beneficial to select generic issues that were common across areas similar to their own for future case studies. This would help to better understand the discrepancies between the Research Outputs and the official estimates.

Respondents suggested that a more concise methodology section would be a positive step in the future as it would reduce the size of the report. However it was recognised that this level of detail was necessary for the first release to explain a rather complex new approach to estimating the population.

Action 1: ONS is currently reviewing the approach for disseminating future Research Outputs. Options under consideration are:

- reducing the size of the report

- enabling users to navigate through the report more easily

- making use of blogs to communicate key quality issues or present case study-type analyses, and

- making more use of visual applications

2.2 Improvements to the methodology used to produce administrative data population estimates

The feedback highlighted the need for a Population Coverage Survey to assess and adjust for coverage error in the administrative data. Reference was made to the need for 'activity data'[2] to strengthen the SPD exclusion rules and reduce the over-coverage on the SPD, especially for working age males. It was also suggested that such data might help to reduce the proportion of half-weights (where records are weighted 0.5 at conflicting administrative data source locations). Respondents also encouraged further research into the rules that are used to determine which records on the administrative sources should be included in the SPD.

---

[1] *Once individual records have been matched across data sources using the approach outlined in [Estimating the Population: In Practice](#), the information is pulled together into a single, coherent dataset that forms the basis for estimating the population. This is called a Statistical Population Dataset (SPD).*

[2] *Information from administrative data sources about when individuals have interacted with systems or services, such as the National Insurance, tax or benefits systems, or a hospital visit through the NHS system.*

Action 2: ONS is currently exploring how activity data can be used to improve administrative data population estimates (using data from interactions with a DWP/HMRC administrative system and from patients interacting with NHS systems through new registrations or to update their address and other details).

Action 3: ONS is considering a test of the PCS in 2018. Data collected through this exercise, data collected in the 2017 Census test and, where suitable and possible, existing survey data will be used to demonstrate how administrative and survey data can be combined to produce high quality estimates about the size (and characteristics) of the population.

Another suggestion was to publish measures or levels of uncertainty around the Research Output estimates, which would enable an independent assessment of the quality of the SPDs.

Action 4: ONS plans to review the current quality standards used to assess the quality of the SPDs. This work includes exploring how an independent quality measure (such as the production of confidence intervals around the SPD) might be produced. In the longer term, one use for the PCS will be to provide confidence intervals around the estimates.

Some respondents expressed concerns around the stability of administrative data over time as changes to the administrative data could potentially affect the methodology and the estimates. Similarly, changes to policy or processes for collecting the data could have an impact on the continued quality and accuracy of the data. The SPDs have been designed to make use of a range of data sources, both to maximise quality and to avoid reliance on a single source of data.

The Government recently consulted on a range of proposals for new legislation to enable Better Use of Data in Government, including for ONS's role in producing National and Official Statistics and statistical research. Proposed new legislation would include:

> *"An obligation to consult the National Statistician before changes to data collection are made in order to protect the security of data supply, and the accuracy and reliability of statistical outputs derived from these data".* This would help to improve the resilience of the administrative data, and therefore the SPDs, over time. Proposed new legislation is discussed in more detail in Section 6.1 of the 'Annual Assessment of ONS's progress towards an Administrative Data Census post-2021' published alongside this paper.

Action 5: ONS will continue to work closely with data suppliers to understand and improve statistical quality issues identified in the data. A Data Suppliers Group has been set up which includes members from key Government Departments across Whitehall. This group is a forum for discussing issues relating to data sharing, sharing research and building stronger relationships. If new data sharing legislation goes ahead, this group will be vital to the success of unlocking data across Government as work will need to be done to put in place practical arrangements for:

- data supply
- protecting that supply in the future, and
- improve quality

2.3   Other potential administrative data sources

Various administrative data sources were suggested as potential additions to improve the current methodology. The main ones are listed below:

- Council Tax data could be used for type of accommodation, number of rooms, vacant properties, HMO (Houses in Multiple Occupation) flag and special types of occupier (for example all student households)

- Lifetime Labour Market Data (L2) active National Insurance payments could be used as 'activity data' to indicate whether an individual is part of the usually resident population

- Individual Electoral Register (IER) could be used to provide information to indicate HMOs and large households, nationality and to improve coverage for older ages

- Child Benefits data could improve the coverage of 0-4 year olds

- DVLA could provide 'activity data'1 on applications for vehicle registration and road tax to help improve the quality of population estimates and information about the number of cars and vans

- Education data could include independent schools and boarding schools not covered by the School Census data, and also provide information about qualifications

- School Census data (England and Wales) was suggested to improve the coverage for school age children in the SPD

- Universal Credit data could be used as 'activity data'[3]

Suggestions were also made about sources that may help to improve estimates for special population groups such as information about armed forces from the Ministry of Defence, and prison data from the Ministry of Justice.

Annex A of the Annual Assessment report provides an update on the data that ONS has obtained access to, and data that we are focusing on next.

Action 6: ONS aim to show developments to the methods for producing estimates about the size of the population (including the use of School Census and DWP/HMRC activity data) in the next release of Research Outputs in autumn 2016.

2.4   Possible explanations for the differences between the Research Outputs and the official estimates

A number of possible factors were suggested by respondents to explain the differences between the administrative data based estimates and the official estimates for their particular area, including:

- areas with high population churn (in particular, for young adults)

- areas with high volumes of international migration

- student populations

---

[3] *Using 'activity data' (for example, when individuals have interacted with systems or services, such as the National Insurance, tax or benefits systems, or a hospital visit through the NHS system) could help reduce levels of over-coverage that are seen for particular age groups.*

- prevalence of armed forces including foreign armed forces and their dependents

- high proportions of second home ownership

- high proportions of part-time residency

- areas where private education and uptake of private healthcare is more prevalent (as these services are not covered on the administrative data used in the current methodology)

- communities that are reluctant to interact with services and public bodies

Some of these explanations were covered in the main report and are known issues with the methodology.

A key concern for respondents was how lags on administrative data (for example, through delays in registering for services or updating address information) might affect the quality of the SPDs and their ability to spot change over time. This can vary across population groups and areas. For example, younger males (who are generally recognised as being slow to update their records when they change address), or areas that have high levels of population churn. Similarly, areas with high proportions of special population groups, such as prisoners or armed forces personnel may be impacted as these groups may interact differently with administrative services.

Action 7: ONS will publish administrative data population estimates down to Lower layer Super Output Area (LSOA) and by single year of age (at LA level) in autumn 2016. Analysing data at these levels will help identify age groups and areas that have the biggest discrepancy from official population estimates, and will help to focus further refinement of the methodology and the data sources used.

Action 8: ONS is working closely with three groups of local authorities to better understand discrepancies between the SPDs and official population estimates. Where possible, a collaborative approach will be taken on future case studies.

2.5 Usefulness of the interactive content

The [interactive content](#) was generally well-received by users. Feedback mentioned that the data visualisation tool was easy to use and helpful in undertaking quick analysis of an area against the corresponding official estimate for the reference years. Some enhancements were suggested as follows:

- include total populations alongside males and females

- include a facility to include comparisons by neighbouring area or by larger area (for example, London, Counties and Districts) and nationally

- include additional quality measures as outlined in the main report and used in the case studies, for example, match rates and FP69 flags

Action 9: ONS will make improvements to the data visualisation tool for the next release, noting the comments received.

2.6   General comments on how to improve the Research Outputs

Concerns were expressed that the current methodology for the Research Outputs is stock based and refers to a snapshot of the population at a point in time. This is different from the flows based methodology for the MYEs which incorporates internal migration from GP patient register to track movement between years.

Action 10: ONS plans to investigate how the SPD methodology can be used to identify population flows, including migration.

Respondents were also very keen to see the SPD estimates on the size of the population extended to small area level. Action 7 sets out ONS's plans for this.

Respondents flagged a concerned that the methodology focuses on producing estimates for England and Wales, and does not yet extend to producing UK estimates. Scotland and Northern Ireland are aware of the current research being conducted by ONS and work will be done to understand how coherent UK estimates can be produced.

Action 11: ONS will continue to work closely with NRS (National Records of Scotland) and NISRA (Northern Ireland Statistics and Research Agency) to produce UK statistics from the census.

Currently population estimates are used to provide population projections, population estimates of the very old, and population estimates for small geographical areas. They are also used for weighting survey estimates and other social surveys to ensure that they are representative of the total population. Concerns have been raised about whether the new methodology will be sufficiently robust to provide this role.

Action 12: An assessment of ONS's ability to move towards an Administrative Data Census will be published each year. The first of these assessments is published alongside this paper. This assessment will be based on a range of criteria, including whether the methodology has the ability to meet the information needs of key users.

Other comments relating specifically to census methodology are not commented on here. ONS will set out detailed methodology for the 2021 Census in due course.

3. Plans for autumn 2016 Research Outputs Release

ONS plans to publish the next set of Research Outputs in autumn 2016. The aim is to expand the accuracy and/or detail and/or breadth (range of topics) of the outputs each year. The 2016 release will continue with the SPD time series released last year. ONS is also aiming to produce:

- population estimates down to Lower layer Super Output Area (LSOA) and by single year of age

- improvements to the methods used to produce administrative data population estimates

- outputs on the number of households

- research on income from combined PAYE and benefits data

Further details will be announced in due course. Users will continue to have the opportunity to provide feedback on these outputs.

Annex A Table summarising the actions that ONS plans to take in response to the feedback

Action number	Action
1	ONS is currently reviewing the approach for disseminating future Research Outputs. Options under consideration are: • reducing the size of the report • enabling users to navigate through the report more easily • making use of blogs to communicate key quality issues or present case study-type analyses, and • making more use of visual applications
2	ONS is currently exploring how activity data can be used to improve administrative data population estimates (using data from interactions with a DWP/HMRC administrative system and from patients interacting with NHS systems through new registrations or to update their address and other details).
3	ONS is considering a test of the PCS in 2018. Data collected through this exercise, data collected in the 2017 Census test and, where suitable and possible, existing survey data will be used to demonstrate how administrative and survey data can be combined to produce high quality estimates about the size (and characteristics) of the population.
4	ONS plans to review the current quality standards used to assess the quality of the SPDs. This work includes exploring how an independent quality measure (such as the production of confidence intervals around the SPD) might be produced. In the longer term, one use for the PCS will be to provide confidence intervals around the estimates.
5	ONS will continue to work closely with data suppliers to understand and improve statistical quality issues identified in the data. A Data Suppliers Group has been set up which includes members from key Government Departments across Whitehall. This group is a forum for discussing issues relating to data sharing, sharing research and building stronger relationships. If new data sharing legislation goes ahead, this group will be vital to the success of unlocking data across Government as work will need to be done to put in place practical arrangements for: • data supply • protecting that supply in the future, and • improve quality
6	ONS aims to show developments to the methods for producing estimates about the size of the population (including the use of School Census and DWP/HMRC activity data) in the next release of Research Outputs in autumn 2016.

Action number	Action
7	ONS will publish administrative data population estimates down to Lower layer Super Output Area (LSOA) and by single year of age (at LA level) in autumn 2016. Analysing data at these levels will help identify age groups and areas that have the biggest discrepancy from official population estimates, and will help to focus further refinement of the methodology and the data sources used.
8	ONS is working closely with three groups of local authorities to better understand discrepancies between the SPDs and official population estimates. Where possible, a collaborative approach will be taken on future case studies.
9	ONS will make improvements to the data visualisation tool for the next release, noting the comments received.
10	ONS plans to investigate how the SPD methodology can be used to identify population flows, including migration.
11	ONS will continue to work closely with NRS (National Records of Scotland) and NISRA (Northern Ireland Statistics and Research Agency) to produce UK statistics from the census.
12	An assessment of ONS's ability to move towards an Administrative Data Census will be published each year. The first of these assessments is published alongside this paper. This assessment will be based on a range of criteria, including whether the methodology has the ability to meet the information needs of key users.

RPT 02 CAD