

Article

Population and migration statistics system transformation – recent updates: evaluating coverage and quality in the admin-based population estimates

The latest update on our population and migration statistics transformation journey.

Contact:
Ann Blake
pop.info@ons.gov.uk
+44 (0)1329 444661

Release date:
27 July 2020

Next release:
To be announced

Table of contents

1. [Context](#)
2. [Conclusions](#)
3. [Next steps](#)

1 . Context

In June 2019, we published our latest [admin-based population estimates](#) (APBE). Instead of a "presence on two or more sources" approach, we moved to an "activity-based" approach using registration activity and other interactions with administrative sources as an indication of usual residency in our inclusion rules.

Our design aim for these research outputs was to reduce the over-coverage observed in our [previous versions](#) when compared with the official estimates. This aim is driven by our understanding that to produce ABPEs of acceptable quality for our users, we need to adjust them using a Population Coverage Survey and an estimation method. One estimation approach is to use a method that will require zero to minimal over-coverage in the admin data base.

Today (27 July 2020) we publish three reports assessing the quality of our admin-based population estimate Research Outputs so far and evaluating progress towards our design aim.

Measuring and adjusting for coverage patterns in the admin-based population estimates, England and Wales 2011

A condition of some of our traditional estimation methods is minimal over-coverage in our base ABPEs, which we have tried to address by using an activity-based approach. [This report](#) assesses the extent to which we have removed this over-coverage by linking our 2011 ABPEs to Census responses and accounting for missed links and Census non-response.

This article summarises our results showing that, despite achieving net under-coverage, there are records remaining in ABPE V3.0 that are not in the Census estimates, suggesting over-coverage remains. Some reasons for this might be the possible inclusion of emigrants and short-term residents and those records without real activity in the data sources, however these need to be fully explored. This has helped us to understand possible improvements for future iterations of the ABPEs.

Developing our approach for producing admin-based population estimates, subnational analysis: 2011

[This report](#) compares the latest ABPE V3.0 with our 2011 Census estimates down to output area (OA) level. The analysis shows similar patterns to the national comparisons outlined in the Measuring and adjusting for coverage patterns in the admin-based population estimates report (less over-coverage and more under-coverage) with some exceptions.

We use cluster analysis to group and analyse the local authority (LA) coverage patterns, and outlier analysis at the OA level to understand how the challenges we see at the national level have an impact at the local level. We have identified the types of areas where there is greater under-coverage, for example, for the working-age population in urban areas but also other areas where there is over-coverage, for example, females in student areas. This shows the need to ensure complete coverage from administrative data of communal establishments (CEs) and to continue to refine our methods.

Admin-based population estimates and statistical uncertainty

[This report](#) introduces exploratory research to measure uncertainty for the ABPEs based on our [previously published mid-year estimate approach](#). The report focuses on what the measures of statistical uncertainty can tell us about the ABPEs and provides further information about the coverage patterns at LA level discussed in the Developing our approach for producing sub national admin-based population estimates report. This approach considers what we know about the uncertainty in the different estimates, developing our understanding of how significant differences might be between our official population figures and the ABPEs.

We have also published [International immigration distribution at local authority level: research into potential improvements](#). Our analysis shows that our current methods cannot be meaningfully improved now through the inclusion of cross-year linkage of administrative data.

The [transformation overview](#) has also been updated to reflect this progress, together with information on the implications of the coronavirus (COVID-19).

2 . Conclusions

From the reports, we conclude that moving to an activity-based approach has removed some of the over-coverage we observed in our previous admin-based population estimates (ABPEs), however over-coverage remains. This will need to be reduced further to enable the production of robust estimates using existing coverage adjustment methods.

The analysis suggests that even after accounting for possible linkage error, we may still be including emigrants and short-term residents in our ABPEs. Similarly, there are still some gaps in the coverage of our ABPEs, particularly for age groups where we do not yet have access to a real source of activity, for example, the self-employed or the working-age population who are neither working nor receiving a benefit.

These coverage patterns are broadly reflected in the local area (LA) analysis, indicating some records may be missing or not captured in the right place. This is particularly evident when looking at small areas where there are communal establishments (CEs) and/or particular population groups. Our analysis shows some CE population groups are either missing completely (for example, foreign armed forces and some school boarders), or we may be including them at an incorrect address, perhaps because of poor quality address information or lags in the administrative sources (for example, care home residents and prisoners).

3 . Next steps

This evaluation of our admin-based population estimates (ABPEs) has taught us a lot about how our current methods work and how we might refine our rules and incorporate new data sources to produce better quality estimates in the future. A priority next step for us is developing our quality frameworks further and using the knowledge gained here to make the best use of the data. This is important for understanding how interactions with administrative sources change over time in both our existing and future data sources.

Given the current context of the coronavirus (COVID-19), we need these frameworks in place so we can understand any changes in the interactions with administrative sources, for example, students moving away from their university accommodation or deferring their courses. Now more than ever, we see the need for more timely information about the size and structure of the population and population change.

Our research so far suggests that that we are unlikely to be able to produce high quality population estimates without a robust estimation framework and some form of Population Coverage Survey. This analysis will help inform where we need to target our efforts for the development of these methods.

Our users have told us about the need for good quality information about the statistics we produce to help them understand how best to use these outputs. We will be building on the uncertainty measures and other quality indicators with the aim to provide information about the quality of the ABPEs in the future.