

Article

Admin-based population and migration estimates: research update

Provides an overview of all the latest research into the transformation of population and migration statistics to an administrative data first methodology.

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

Release date:
26 November 2021

Next release:
To be announced

Table of contents

1. [Overview](#)
2. [Admin-based population estimates time-series analysis](#)
3. [Admin-based population estimates compared with mid-year population estimates](#)
4. [Measuring internal migration in admin-based population estimates](#)
5. [Understanding international migration based on administrative data](#)
6. [Specific population groups' interaction with administrative data](#)
7. [Exploring data sources](#)
8. [Bayesian demographic accounting](#)
9. [Future developments](#)
10. [Related links](#)

1 . Overview

- Following a [blog by Pete Benton](#), this article summarises the population and migration transformation research papers that we published on 26 November 2021.
- We are transforming the way we produce population and migration statistics to better meet the needs of our users and aim to produce statistics from the best-available data at any given point in time; we will keep users regularly informed on our progress.
- This article provides an update on recent research into putting administrative data at the core of population and migration statistics; this is an ongoing process - see our [overview of the population and migration statistics transformation system](#).
- By exploring data based on our current methods over a 5-year time-series we have developed our understanding of the characteristics of these methods for producing admin-based population and migration estimates.
- We have also explored population groups which may not interact with our admin sources in the same way as the rest of the population, such as students and those that live in communal establishments; this is important to ensure that our statistics are fully inclusive.
- We also signpost exciting new developments in Bayesian demographic accounting methods that we are exploring; we will publish more in the near future.

More about the population and migration

- Latest estimates on [migration into and out of the UK](#).
- View estimates of the [UK population by country of birth and nationality](#).
- Find out about the [future design of migration estimates](#).
- Latest research on [population estimates using administrative data](#).
- Research into [estimating the student population](#).

2 . Admin-based population estimates time-series analysis

We have two working methods for producing admin-based population estimates (ABPE), referred to as ABPE v2.0 and ABPE v3.0. Each method uses a different set of rules to estimate the usually resident population¹ from administrative data records, resulting in different estimates for each method.

[The ABPE v2.0 method](#), previously statistical population dataset (SPD) v2.0, includes records in the population if present on two of the administrative data sources in the relevant year. [Previous research shows that ABPE v2.0 typically over-estimates the population](#) compared with official estimates, particularly for working age males, because people remain present on administrative data sources after they have moved or left the country.

We published [our ABPE v3.0 method paper](#) in June 2019. The design aimed to incorporate more data sources into the methodology and use “activity-based” inclusion rules to generally address the over-coverage patterns observed in ABPE v2.0. Records are included if they have interacted with a single data source in the 12 months prior to the population reference year. [Research](#) published last year found ABPE v3.0 generally showed lower estimates than the 2011 Census, although they were higher for some populations.

In April 2021, we published modelled estimates for international migration during the coronavirus (COVID-19) pandemic by bringing together a range of sources. We also updated on the [development of an admin-based migration estimates \(ABMEs\) based on actual behaviours](#).

This article summarises our most recent research to further explore our admin-based methods. This ongoing research takes into account our [commitment to consider current methods and concerns when developing ABPEs](#), following the recommendations in the OSR's review of population estimates and projections.

Notes for: Admin-based population estimates time-series analysis

1. We are currently adopting the UN definition of “usually resident” – that is, the place at which a person has lived continuously for at least 12 months, not including temporary absences for holidays or work assignments, or intends to live for at least 12 months (United Nations, 2008).

3 . Admin-based population estimates compared with mid-year population estimates

For the first time, we have produced ABPEs for 2016 to 2020 using both admin-based population estimates (ABPE) methods. We compared both against official mid-year population estimates (MYE) to explore the stability and coverage of each method over time. We have published our detailed research and analysis into these ABPE time-series in an [accompanying paper](#). We recognise that as we move further away from the 2011 Census, the MYEs become less reliable, however these comparisons can help us identify the challenges remaining in the ABPE methods.

Both ABPEs are broadly in line with the MYEs over the time-series for England and Wales as a whole. ABPE v2.0 is consistently higher than the MYEs throughout the series (up to 2% higher in 2020) and ABPE v3.0 remains lower than the MYEs (up to 2.4% lower in 2018). This hides patterns of over- and under-coverage that are revealed when we look at results by age, sex and local authority. For ABPE v2.0 most local authorities are within positive or negative 5% of the MYEs, while for ABPE v3.0 most are within negative 10% and positive 5%. When we look at estimates over time there is clear evidence of some specific issues with each method.

Comparing ABPE v2.0 with the MYEs shows that in each year of the time-series, ABPE v2.0 population counts by year-of-birth grow considerably for young adults. While we would expect some net immigration in younger working ages, the growth seen in ABPE v2.0 was much higher than in the MYEs. This suggests the ABPE v2.0 method does not adequately remove those who have left England and Wales, despite the rules requiring a record to be found on at least two sources.

We also found that, for individuals aged between 35 and 50 years, ABPE v2.0 was introducing more males each year than females (relative to population sizes), which gradually moves the sex ratios in our ABPEs away from the official figures over time.

Although ABPE v3.0 generally underestimates the population in most age, sex and local authority groups, some over-coverage remains. However, we do not see the same issue as with ABPE v2.0. This suggests the activity-based rules are removing people who have left England and Wales more effectively.

We can use this understanding to improve the rules we use in future versions of the ABPE.

4 . Measuring internal migration in admin-based population estimates

Linking consecutive years of the time-series has enabled us to explore the flows into and out of our admin-based population estimates (ABPE) and estimate internal moves (moves within England and Wales) within the ABPE. This allows us to understand the stability of individuals' inclusion in the dataset over time and may highlight further areas where we need to strengthen our ABPE rules.

The analysis shows that the rules used in ABPE v3.0 result in more people leaving and joining the ABPE between years than ABPE v2.0. This reinforces the conclusion that ABPE v3.0 is better at removing people who have left England and Wales. However further research is needed to understand whether this is capturing real trends or is triggered by patterns in how people relate to different administrative sources at given points in their lives.

The internal moves within both ABPEs show a similar pattern by age and sex as the official internal migration estimates with a notable exception around ages 18 to 24 years where we know that both ABPE methods and the official internal migration estimates are affected by the way students interact with administrative data. We plan to investigate the internal moves around life transition ages in more depth to build a fuller picture of the strengths and limitations of each method.

5 . Understanding international migration based on administrative data

Until recently, international migration was measured based on a person's intentions through the International Passenger Survey (IPS). We are now exploring how to use actual behaviours recorded in administrative data to identify a long-term international migrant (someone who has arrived or left the UK for at least 12 months). These admin-based migration estimates (ABMEs) will use statistical models to bring data together and predict the likely outcome of long-term international migrants who have recently immigrated or emigrated. More timely statistics can be produced by this method.

The future system is expected to provide provisional estimates of international migration that are modelled and timely. These modelled estimates will likely be revised using information provided by ABMEs, which uses observed behaviours to determine whether a person has migrated to or from the UK. The aim is that this future system will link different sources of administrative data to ensure wider population coverage and provide granular detail and insight into migration and the wider population. A separate article, [International migration statistical design: progress report](#), outlines these plans more fully.

6 . Specific population groups' interaction with administrative data

There are some population groups that are difficult to measure at their usual residence using our admin-based population estimates (ABPE) rules, such as students and those living in communal establishments (for example, prisons and care homes).

Students are generally a young and mobile population whose interaction with administrative sources is influenced by many factors, such as where they are studying and whether they are working.

For example, of the 2018 to 2019 academic-year cohort, only around 57% of students were registered with the NHS Personal Demographic Service (PDS) in the term-time address provided on the Higher Education Statistics Authority data.

To ensure our admin-based methods are as accurate as possible when it comes to students, we have explored [how we can identify and measure this population in administrative data](#).

Communal establishment populations

We have also looked in more depth at those who are resident in communal establishments (accommodation that is under part- or full-time management). These include student halls of residence, military bases, care homes and prisons.

Measuring communal establishment populations correctly within ABPEs can be a challenge; these residents may not interact with some services in the same way as the rest of the population. Some may be present but recorded as residing elsewhere. For example, communal establishments for asylum seekers and foreign armed forces may contain people who do not interact with the data sources we are using, while other groups such as prisoners may remain on administrative data sources at their home address.

Further discussion of the coverage of communal establishments in our local authority-level ABPEs is included in the [time-series analysis](#).

Information on those in communal establishments is not currently collected in social surveys, which focus on collecting information about private households. This means that there is an opportunity to improve the inclusivity of our statistics by using administrative data to provide detailed insights on the characteristics for those in communal establishments in the intercensal period, as well as providing a basic count by age, sex and geography.

7 . Exploring data sources

We continue to explore a range of data sources to assess how they may improve our admin-based population estimates (ABPE) and admin-based migration estimates (ABME). For example, we have considered the use of RAPID¹ data to measure activity and migration patterns for British nationals. For the latter, we have identified some challenges in measuring British nationals' migration in the same way as non-British nationals' migration. We are also exploring the use of HM Revenue and Customs (HMRC) PAYE-RTI data to provide more regular activity to improve our rules. While these data sources show promise, the varied nature of different population groups' interactions with the employment and benefits systems requires further work to understand their complexity.

Measures introduced to reduce the spread of coronavirus (COVID-19) are likely to have changed the way individuals interact with administrative data. We are researching how individuals' interactions with our administrative data sources have changed since the coronavirus pandemic began to prepare us for possible effects on our ABPEs. There is also the opportunity to explore new data sources that are available as a result of the coronavirus pandemic, such as vaccinations data.

Initial analysis of the PDS over this period suggests people's patterns of movement changed during the coronavirus pandemic. There may be a variety of reasons for this. Further detail on this analysis can be found in the [ABPE time-series analysis](#).

For many of our key administrative data sources, updates that cover the coronavirus pandemic are not yet available. When this information becomes available for more of the data sources, we will carry out detailed analyses into the changing nature of data sources during coronavirus and the likely impacts on the ABPEs.

Notes for: Exploring data sources

1. Department for Work and Pensions (DWP) Registration and Population Interaction Database.

8 . Bayesian demographic accounting

The research outlined in this article has highlighted challenges that we need to address to produce population and migration estimates that meet user needs.

We are likely to need additional methods for improving the coverage, timeliness and coherence of these estimates.

We are exploring Bayesian demographic accounting¹ methods to achieve timely population estimates and coherent estimates across all population components over time.

A demographic account provides a framework where we can produce an internally consistent set of demographic stock and flow estimates. The method captures the regularities of the demographic system, allowing expert understanding of demographic trends to be used in the framework instead of ad-hoc adjustments, which are used in the current method for creating subnational population estimates.

Bayesian methods allow us to feed in our prior understanding of the population, such as admin-based population estimates (ABPE), admin-based migration estimates (ABME) and internal migration and combine these with evidence from different data sources that measure population change. Bayesian demographic accounts will allow us to use our estimates, information around the uncertainty relating to these data and estimates, and information about demographic trends to produce coherent estimates.

Notes for: Bayesian demographic accounting

1. Bryant, J., & Zhang, J. L. (2018). Bayesian demographic estimation and forecasting. Chapman and Hall /CRC. This is a collaborative research project with Peter Smith, Paul Smith, Jakub Bijak and Jason Hilton at the Southampton Statistical Sciences Research Institute and John Bryant of Bayesian Demography Limited.

9 . Future developments

Our research is evolving rapidly to explore how we will be able to meet the need for more frequent and timely population estimates. We are developing the Bayesian demographic accounting method to first allow us to produce a proof of concept set of estimates for England and Wales. As we build our understanding of the approach and develop the data sources, we plan to provide more granular estimates, including at local authority level. Our intention is to work towards releasing very timely, provisional indicators that will be revised over time to take into account increased availability of more sources of administrative data.

While the Bayesian demographic accounting methods are developed, we will continue to research and test the quality of new and current data sources and methods, which feed into the whole system, and explore the impact of the coronavirus (COVID-19) pandemic on these.

This includes continuing to refine our modelling approach to measuring international migration and introducing more sources of administrative data to improve their accuracy. These developments will be introduced in our migration estimates up to June 2021, published by March 2022.

In 2022 we will be developing a new version of ABPE, using the strengths of ABPE v2.0 and ABPE v3.0 as well as introducing new methods and data sources. Alongside the analysis outlined in this article, this will be informed by our deeper understanding of internal migration and special populations. This will be combined with research to explore how life transitions are captured in administrative data, as well as how the coronavirus pandemic has impacted this. These improved ABPEs and international migration estimates will be developed as an input to the demographic accounts system.

We will link the 2021 Census with administrative data sources, which will allow us to explore where the administrative data is inconsistent, define the coverage adjustment problem that we need to solve, and further refine our ABPE rules. We plan to use this information to develop adjustment methods that can be built into the demographic accounts system.

We will publish further updates on this work in 2022 to keep users informed of how this research can be used in a future population and migration statistics system and as part of a recommendation in 2023 of the future of the Census.

10 . Related links

[Developing admin-based population estimates, England and Wales: 2016 to 2020](#)

Article | Released 26 November 2021

Progress made to extend our understanding of the stability and behaviour of two methods for producing admin-based population estimates.

[Understanding students across administrative data in England and Wales](#)

Article | Released 26 November 2021

Update on our latest research into how the student population of England and Wales appears in administrative data.

[International migration statistical design: progress report](#)

Article | Released 26 November 2021

An update on the future of International Migration statistics.

[International migration: developing our approach for producing admin-based migration estimates](#)

Article | Released 16 April 2021

A summary of our current research into estimating international migration using administrative data focusing on the Registration and Population Interaction Database (RAPID), which includes data from the Department for Work and Pensions (DWP) and HM Revenue and Customs (HMRC); and Home Office Border data.

[How we are improving population and migration statistics](#)

Article | Release 15 November 2021

Report updating our users on our population and migration transformation journey.

[Update on our population and migration statistics transformation journey: a research engagement report](#)

Article | Released 30 January 2019

Report updating our users on our population and migration statistics transformation journey and seeking feedback on our future plans.

[International migration](#)

Web page | Updated as and when new releases are published

People moving into and out of the UK, long-term migration, short-term migration, and non-UK residents data providing a picture of those entering and leaving the UK, covering all lengths of stay. All data are published from the [Centre for International Migration](#).

[Population estimates](#)

Web page | Updated as and when new releases are published

Annual population estimates. Figures are available for various administrative and electoral geographies and for different population sub-groups, for example, estimates of the very old and estimates by marital status.

[Population estimates for the UK, England and Wales, Scotland and Northern Ireland](#)

Statistical Bulletin Series | Updated as and when new releases are published

National and subnational mid-year population estimates for the UK and its constituent countries.

[GSS harmonisation guidance](#)

Web page | Updated as and when new releases are published

Information and guidance on the main concepts and definitions used for migration, country of birth and citizenship across the Government Statistical Service (GSS).