

Article

# International migration statistical design: progress report

An update on the future of International Migration statistics.

Contact:  
Dominic webber  
pop.info@ons.gov.uk  
+44 1329 444661

Release date:  
26 November 2021

Next release:  
To be announced

## Table of contents

1. [Overview of migration statistics transformation](#)
2. [Progress in transforming migration statistics so far](#)
3. [Goals for transformed migration statistical system](#)
4. [Update on data sources for new migration estimates](#)
5. [Future developments](#)
6. [Related links](#)

# 1 . Overview of migration statistics transformation

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 lived in 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.

These modelled estimates will likely be revised at a later date using information provided by the ABMEs. The aim is that this future system will link different sources of admin data to ensure wider population coverage and provide granular detail and insight into migration and the wider population.

As part of a wider update on the population and migration statistics transformation, today we have also published:

- Admin-based population and migration estimates: research update
- Admin-based population estimates: data time series
- Understanding students across administrative data in England and Wales

## 2 . Progress in transforming migration statistics so far

We announced in August 2020 that we would not return to producing official migration statistics from the International Passenger Survey (IPS) following its suspension because of the coronavirus (COVID-19) pandemic, instead shifting to a new approach that will rely on administrative data first and foremost (admin-based migration estimates (ABMEs)). In the future we will be basing our estimates on actual patterns of migration, rather than relying on potential migrants to respond to survey questions about whether they are planning to remain in or out of the UK in the next 12 months. In our response to the Office for Statistics Regulation review of population estimates and projections, we also committed to keep users informed about our use of administrative data for producing estimates of international migration. Accordingly, this release provides an update on our progress and plans for reporting on the development of ABMEs supported by statistical modelling to produce estimates of international migration.

Our previous research [identified a range of data sources](#) held across government that can help us to measure migration including immigration, income, benefits and education data. So far, the Department for Work and Pensions (DWP) Registration and Population Interaction Database (RAPID) and Home Office (HO) Border Data are the two key sources with the greatest potential for measuring long-term international migration. We published our first piece of research using these data to measure international migration in [April 2021](#).

In the absence of any official statistics on international migration during the coronavirus pandemic, we published estimates using a modelled approach in [April 2021](#), covering the period up to June 2020. This was used to provide an indicative picture of what migration may have looked like in the first few months of the coronavirus pandemic by bringing together a range of data sources to model migration patterns. This was then fed into other components, such as the [population estimates](#). We updated these indicative estimates in November 2021, covering up to December 2020, providing a picture of international migration during the later stages of the coronavirus pandemic.

However, because of the nature and the assumptions of the model, the modelled estimates do not paint a complete picture of international migration. While providing a headline measure for the UK, broken down by broad category of citizenship, they do not provide information on country of origin or destination, or reason for migration, for example. This is where the additional insight of ABMEs can complement the more-timely modelled estimates.

Bringing this all together, our ambition is to integrate the modelled estimates with the ABMEs so that we can provide a coherent set of statistics that are timely and provide the level of granularity and insight that our users need. In this update we outline:

- progress in developing ABMEs since our previous update was published
- how we plan to integrate the modelling and ABME approaches, and how this is likely to affect official published estimates of UK migration in the future
- how we plan to use alternative definitions of migration alongside the long-term definition to be more reactive to user needs

### 3 . Goals for transformed migration statistical system

As we have outlined throughout our [Population and Migration Statistics Transformation](#) journey, our goal is to produce the best possible estimates of international migration. Through our engagement with users, we have built a good understanding of their most important considerations, in particular:

- **timeliness:** the length of time between the period that we are reporting, and the date we publish the estimates
- **accuracy:** the extent to which estimates reflect true patterns of migration
- **granularity:** the level of insight we can provide to users, including different sub-populations and geographies
- **relevance:** the extent to which we are able to provide the statistics needed to fill evidence gaps and inform decisions, including the use of clear and appropriate concepts and definitions

We have long acknowledged that there is no one single source that will tell the whole story of migration. Therefore, we are working with a range of data sources to help us tell this story. In [April](#) we introduced our two key sources: DWP's RAPID and Home Office border data. We outlined our progress in building an understanding of how we can use these sources to provide the best estimates of migration, and what we plan to do next to develop these sources more. We will discuss what the future international migration statistical system will look like post-coronavirus, including a three-phased approach of developing, refining and embedding the new migration statistical system:

- phase one: developing our understanding of the data sources
- phase two: improving our methods
- phase three: data integration and further refinements

These phases are described in more detail in the section "[Future developments](#)".

### 4 . Update on data sources for new migration estimates

## Registration and Population Interaction Database (RAPID)

In [April 2021](#), we provided an overview of the methods applied to the Registration and Population Interaction Database (RAPID) to estimate international migration. We outlined how we were able to identify a migrant through their interactions with earnings and benefit systems. We examined the length of interaction, together with their arrival information taken from the Department of Work and Pension's (DWP) Migrant Worker Scan (MWS), a database containing National Insurance registrations. Those that satisfied our 12-month residency rules were determined to be a long-term international migrant under the UN definition.

We acknowledge that RAPID does not cover the entire population, so adjustments were introduced to account [for those who may have migrated](#), but will not appear on RAPID, such as non-working students. There were also other populations that were not covered in the estimates, such as anyone under 16 years of age.

This first step in analysis considered only those who registered for a National Insurance number (NINo) with a non-UK nationality (an adjustment was made for those who may have taken on UK nationality since registering for a NINo). However, we know that non-UK nationals only form part of the picture when it comes to measuring migration and we have begun work exploring methods to extend use of RAPID data to measure migration of UK nationals. During our initial research, we have identified some challenges in measuring UK migration in the same way as non-UK migration in RAPID. It is clear that because of the more complex nature of interactions with the employment and benefits systems we cannot apply the same migration methodology to UK nationals as to non-UK nationals in RAPID. We are continuing to develop our understanding of how best to measure this group and will provide a further update in spring 2022.

### Home Office border data

In [April 2021](#) we also outlined a potential method using Home Office border data to produce separate preliminary estimates of long-term international immigration for non-EEA nationals. We also noted that further work is necessary to fully understand the complexities of the data and investigate what these data can tell us about particular groups of migrants. This includes bringing in data from EEA nationals as the new immigration system matures. Including information on EEA nationals will give us a more complete picture of migrant patterns and movements based on actual travel histories.

Further research with Home Office data has focused on the statistical modelling of international migration flows during 2020 and on using this data to improve our understanding of international students and transitions within the UK, for example movement from study to work visas. We will publish our [first results of this work](#) on 29 November 2021.

## 5 . Future developments

Looking ahead at developing a new migration statistics system, we have identified the following priority areas:

- linking administrative data sources together to provide a richer source of information and improve the accuracy of our migration estimates
- integrating modelling and admin-based migration estimates (ABMEs) to make the most of the strengths of each approach
- developing a system for revising migration estimates when we update a modelled estimate with one based on actual migrant behaviour
- our approach for 2022 and beyond when the Census will provide us with a rich source of data on population and migration
- how we are exploring alternative definitions of migration alongside the traditional long-term definition to deliver more timely estimates and give users more flexibility with the available data

More detail is provided on each of these areas in this section.

## Integrating modelling and ABMEs

Our [Admin-based migration estimates \(ABMEs\) report](#) highlighted how migration statistics could be produced using administrative data. The method is based on actual behaviours, which means waiting before a person has moved to the UK or left the UK for at least 12 months. Therefore, there will be time lags if we use these sources in isolation.

To mitigate this, earlier indicators through modelling are helpful in providing a timelier insight into international migration patterns. However, this approach does not provide the level of granularity that ABMEs can provide, such as breakdowns by local authority, single-year-of-age and sex. It is with this in mind that we are developing a system that integrates the modelled estimates and the ABME estimates, building on the strengths of each.

This will initially involve increasing the use of the main data sources from the ABMEs in the [model](#). While Home Office border data is already used as a key source, we are currently working towards adding the Registration and Population Interaction Database (RAPID) data into the model. This is expected to provide a more accurate basis for modelling migration behaviours for EU citizens and improve on the current approach that uses travel information of non-EU flows. In addition, we are exploring using administrative data to add the granularity that users require (by geography and age, for example) into the headline modelled estimates.

## Linking administrative sources together

As part of the development of the ABMEs, we need to continue to develop our understanding of the quality of the sources that feed into the estimates. This is important as none of these sources were designed for statistical purposes. Therefore, this work will help us identify further strengths and weaknesses of each source, which will then feed into how we refine our methods for estimating migration.

As we gain access to data at a more granular level, we intend to introduce data linkage as a means of addressing the coverage gaps in individual sources. By linking data such as RAPID and Home Office border data together, we can overcome challenges in quality and coverage of using the sources on their own. This will then build a more accurate estimate of international migration. However, the linkage process is complex, involving acquiring data with the required detail, carrying out the linkage, and quality assuring the linked dataset before conducting the relevant analysis. While it is not clear at this stage whether this data linkage will be viable, the potential benefits mean that it is important for us to explore its feasibility.

Another example of the benefits of data linkage could be to improve our understanding of the international migration of students. There are challenges with measuring the student population in administrative data, for example, those that do not receive earnings or benefits are likely not to appear in RAPID. In our [previous update on ABMEs](#) we outlined an adjustment to fill this coverage gap in RAPID data. We are continuing to develop this adjustment by linking data from the Higher Education Statistics Agency (HESA) to HM Revenue and Customs (HMRC) income data. This will provide a clearer picture of how many students are in employment while they study, therefore giving us a more reliable indication of how much of the student population is represented in RAPID.

The new changes in the immigration system mean that all international students are subject to immigration controls. Because of the recency of this change, data that might be derived from the immigration system is not available for use yet. We are working with the Home Office to assess how best to use this data when it becomes available.

## Revising migration estimates

The modelled estimates provide our initial best and most timely estimate of who is arriving and leaving the UK, which we will consider our "provisional" estimates in this new statistical design. These provisional estimates are likely to be revised as more data on the reference period, such as RAPID, become available. These could provide more insight and granularity than the provisional estimates and allow users to understand more about migration, and how it affects the UK.

Our [latest modelled estimates](#), covering up to December 2020, were published in November 2021. As well as providing the first official, albeit experimental, measures of international migration for Quarter 3 (Jul to Sept) 2020 and Quarter 4 (Oct to Dec) 2020, estimates for Quarter 2 (Apr to June) that were originally published in April 2020 were revised. This revision accounted for the fact that we had access to more complete information from Home Office border data for that period.

By March 2021 we expect to publish provisional estimates of international migration for Quarter 1 (Jan to Mar) 2021 and Quarter 2 (Apr to June) 2021. At the same time, we anticipate revising estimates for Quarter 2 2020, Quarter 3 2020, and Quarter 4 2020. Again, this will reflect the improvement in the completeness of the Home Office border data that we have access to but also we will be aiming to introduce RAPID data into the modelling for the first time in these estimates. Because RAPID covers tax years (6 April to 5 April), it will feed into the model up to Quarter 1 2021. We will update the model as data become available for the 2021 to 2022 tax year.

This approach to revisions reflects the need to build a migration and population statistical system using a range of sources. Our aim is to provide more information on the proposed framework for a migration statistics revision policy during spring 2022.

## Reviewing migration definitions

As a further means of improving timeliness and to meet a broader range of user needs, we are also considering how we define international migration in the future.

For a long time, we have used the UN definition of long-term international migration – that is "a person who moves to a country other than that of his or her usual residence for a period of at least a year (12 months), so that the country of destination effectively becomes his or her new country of usual residence". These feed into the mid-year population estimates.

We know that users of migration statistics need greater flexibility around how we define migration. In February 2020, we published initial findings on [how different definitions could be used to measure migration](#).

In a society where people are increasingly mobile, our statistics must reflect the needs of users and policy makers. Since the coronavirus pandemic there is more interest in who is in the country at any given time, whether in the short or long term. There is a unique opportunity to review the definitions of the population as part of our transformation to ensure we meet those changing needs. Therefore, we are exploring how we can provide additional and timely measures of the population, alongside the UN long-term definition, to meet those changing needs.

Alongside the user need to provide a more flexible set of migration definitions, another benefit stems from the way in which migration will be measured in the future. Moving away from an intentions-based system to one based on observed behaviors potentially involves waiting for the required period to elapse. Under the current definition of 12 months this means a lag of over a year to produce migration estimates, whereas a more flexible system allows the chance for more timely estimates.

Following the UK's exit from the EU and the changes to the immigration system, we will review how we classify different migrant groups. The migration patterns and behaviours in EU nationals may differ depending on their immigration status. In the future, we may provide further headline breakdowns of EU migration by those with pre-settled or settled status, and those who are migrating to the UK without pre-settled or settled status.

We will continue to engage with users on which definitions best suit their needs throughout 2022.

## Making use of the 2021 Census

Census 2021 results will provide high quality estimates of the population of both UK and non-UK born and nationals living in England, Wales and Northern Ireland on Census day<sup>1</sup>. Scotland's census is taking place in March 2022.

It will also provide information on migration through responses to questions such as country of birth, passports held, address one year ago and date of arrival in the UK.

Comparisons with the Census outputs will mainly help us understand the quality of our migrant stock estimates, but we will also consider how they may be used to improve published and future migration flow statistics.

## Our approach for 2022 and beyond

As we continue to develop both the modelling and ABMEs, we expect the transformation to be iterative with improvements implemented in phases throughout 2022 and 2023:

- phase one - spring 2020 to autumn 2021: developing our understanding of the data sources, which will feed into the ABMEs, as well as developing the fundamental methods for integrating them and measuring migration; modelled estimates will continue to be our key indicator of migration during this period
- phase two – autumn 2021 to summer 2022: improving our methods and building greater accuracy in our estimates, including addressing coverage gaps (some of which were highlighted in our April 2021 publication); this phase also allows us to be more robust in how we develop the migration estimates and explore the potential for greater flexibility in the population definitions used in our outputs
- phase three - summer 2022 and into 2023: bringing our administrative sources together through data linkage, providing even greater improvements and population coverage, and providing greater accuracy and granularity in our estimate

A lot of work is required to progress through our phases of development. To ensure the continued iterative development of these phases, the next steps that we will be focussing on in the next few months are:

- developing a method of measuring migration patterns of UK nationals using administrative data
- continuing to improve our adjustment for students through data linkage
- integrating the ABMEs and modelled estimates to form the concept of "provisional" and "revised" estimates; the first step is integrating RAPID data into the model
- developing methods to provide breakdowns of international migration required for population estimates, including local authority, single-year-of-age, and sex
- continuing to develop alternative definitions of migration to better align with user needs

We are committed to providing users with continual updates into our progress and will provide a further update throughout spring 2022.

Throughout all phases of development, we will be considering how we can integrate the transformed statistics system to ensure coherence between population stocks and flows. We are exploring Bayesian Demographic Accounts to achieve coherence across the different components of the population. We describe how the accounts work further in our Population and Migration Statistics Transformation Research Overview.

## Notes for: Future developments

1. The 2021 Census took place in England, Wales and Northern Ireland on 21 March 2021. Scotland's Census is due to take place on 20 March 2022.

## 6 . Related links

### [Long-term international migration, provisional: year ending December 2020](#)

Statistical Bulletin | Released 25 November 2021

Experimental statistics on UK international migration throughout 2020 including the impact of the coronavirus (COVID-19) pandemic.

### [Population of the UK by country of birth and nationality: year ending June 2021](#)

Statistical Bulletin | Released 25 November 2021

Latest population estimates for the UK by country of birth and nationality, covering the period from 2004 to the year ending June 2021.

[Admin-based population and migration estimates: research update]

Article | Released 26 November 2021

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

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

Article | Released 15 November 2021

Find out more about our population and migration statistics transformation journey.

### [International migration: developing our approach for producing admin-base 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.

### [Methods for measuring international migration using RAPID administrative data](#)

Methodology | Released 16 April 2021

A summary of the current methods in using the administrative data source the Registration and Population Interaction Database (RAPID) from the Department for Work and Pensions in estimating international migration.

### [Using statistical modelling to estimate UK international migration](#)

Methodology article | Released 16 April 2021

Estimating how levels of international migration have been affected by the coronavirus (COVID-19) pandemic by making innovative use of available data sources and methods.