Office for National Statistics

Methods to produce provisional long-term international migration estimates

An explanation of the methods used to produce the latest experimental and provisional experimental statistics on migration flows into and out of the UK.

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

Release date:
26 May 2022

Next release:
To be announced

Table of contents

1 . Purpose of this article

This article summarises the methods used to produce experimental and provisional estimates of long-term international migration flows, published in our Long-term international migration, provisional: year ending June 2021 bulletin. It also summarises previous methods implemented in April 2021; further details are available in our Using statistical modelling to estimate UK international migration methodology, updated November 2021. We explain how our methods have changed and why future revisions of the estimates will be an important part of producing migration estimates going forward.

The research detailed in this paper is carried out as part of an ongoing programme of work to transform population and migration statistics. This work has been accelerated in response to the coronavirus (COVID-19) pandemic. For more details, please see our How we are improving population and migration statistics article, published 15 November 2021.

2 . The need to change how we measure international migration

We have long acknowledged that the International Passenger Survey (IPS), which underpins our previous estimates of migration, has been stretched beyond its original purpose. We need to consider all available sources and methods to estimate international migration flows (from here on referred to as migration).

In March 2020, the IPS was suspended because of the coronavirus (COVID-19) pandemic. In response, we accelerated our approach for transforming migration statistics using new methods and administrative data, supported by statistical modelling. The IPS resumed operation in January 2021 but, because of its known limitations in measuring migration, we continue to focus on using administrative data supported by statistical modelling.

Our latest published estimates (26 May 2022) cover the year ending (YE) June 2020 and YE June 2021. We have not provided a back series because of:

- the experimental nature of the methods
- planned revisions following the release of Census 2021 results later this year
- subsequent rebasing of mid-year estimates

Going forward, the revision of long-term international migration statistics will be an important part of the production of these estimates. Provisional estimates are released with the expectation they may be revised as more complete data become available. In addition, our methods are still experimental and we will therefore revise the estimates as our methods mature. Further information on our approach to revisions can be found in our Population and international migration revisions policy (May 2022).

3 . Why we cannot count people in and out at the border

A common misconception is that it is easy to measure international migration simply by counting people in and out as they cross the border. There are many reasons why it is difficult to count migrants by monitoring cross-border travel data using passport scans at airports. For example:

- some people hold two passports and use different passports for incoming and outgoing journeys
- the UK and Ireland belong to a free travel zone called the Common Travel Area; people can travel freely between the two countries and movements across the land border between Northern Ireland and the Republic of Ireland are not tracked

Countries across the world have encountered similar measurement challenges and we work closely with many of them to learn and share best practice.

The Office for National statistics (ONS) is committed to exploring all sources of information to produce migration estimates. This includes advanced passenger information, which provides passenger information for a large proportion of inbound and outbound air passengers. We aim to publish more detail on these plans later in 2022.

4 . The method for our latest estimates

Our latest experimental estimates (for year ending (YE) June 2021 and YE June 2020) are produced using a new method that relies less on International Passenger Survey (IPS) data and statistical modelling, and makes greater use of administrative data. This is in line with our ambitions published in April 2021, moving from intentions based estimates from the IPS to estimates based on actual observed activity in administrative data.

Our latest estimates use different data sources and methods for each nationality grouping. We currently publish estimates on immigration, emigration and net migration for non-EU nationals, EU nationals and British nationals.

Non-EU Nationals

Non-EU migration refers to estimates of migration for people who do not hold British or EU nationality. We use data from the Home Office initial status analysis (ISA) system, which combines visa and travel information to link an individual's travel movements into and out of the country. This dataset is known as the exit checks dataset, with more information provided in the Home Office statistics on exit checks: user guide.

We continue to use the United Nations' (UN) definition of a long-term migrant: a person who moves to a country other than that of their usual residence for at least a year.So, our first step is to identify which travellers meet the definition of a long-term migrant, filtering out those on long-term visit visas.

We use arrival and last departure dates within a visa period as an approximation for length of stay in the UK. Short trips abroad over the course of an extended period of residence are excluded. If either the first arrival or last departure information is missing, then visa start or end dates are used as a proxy.

Visa periods are constructed by linking together any consecutive or concurrent visas held. If there is a gap between visas, then a new visa period is started. We look at any previous visa period to determine if this is a new long-term immigrant or one who has previously been in the country. If no presence is identified in the country during the 12 months preceding first arrival on a given visa, or the previous visa period had a length of stay of less than 12 months, then this pattern of travel will be considered as identifying a new long-term immigrant.

The ISA record level dataset is not yet available for the final eight months of the time series (November 2020 to June 2021). This missing time period has been estimated using an aggregated version of the ISA data to November 2021. We applied the pattern of change observed in the aggregate data set (high frequency series) with the Denton-Cholette method to predict the record level data set (low frequency) for the missing period. This provided us with a rate of change for arrivals on a month-to-month basis. See the Temporal disaggregation of time series guidelines (PDF, 2.3MB) for more information on the Denton-Cholette method).

The ISA aggregated data set provides similar information as the ISA record level (arrivals and departures of visa holders to and from the UK). However, as it is not person level data, it is not possible to identify multiple visas and /or travel events for a person. This limits its use in counting arrivals and departures of people, but provides suitable information on the trends of flows to and from the UK.

In our previous reports (Exploring international migration concepts and definitions with Home Office administrative data from February 2020 and April 2021), we set out our journey to better understand this complex data, including the caveats. For example, we acknowledge that it excludes a minority of non-EU migrants that do not require a visa. Since 2016, the Home Office has published an annual report on exit checks data.

Currently we do not have an equivalent method for measuring non-EU emigrants. Therefore, we calculate a ratio between emigration (numerator) and immigration (denominator) on a monthly basis from the aggregated ISA dataset, which is applied to the calculated non-EU immigration estimates to estimate emigration. This assumes the trends in the aggregated dataset for the immigration series and emigration series are similar to the trends in the exit checks data; this is a reasonable assumption, as both datasets are derived from the same source.

EU Nationals

We cannot apply this same method to estimate EU nationals. During the reference period, there was free movement between the EU and UK, where EU nationals were not required to hold a visa to live in the UK.

The latest methodology to estimate the migration of EU nationals is based on previous research to measure international migration using the Registration and Population Interaction Database (RAPID). RAPID currently provides the best insight into the migration of EU nationals.

RAPID is created by the Department for Work and Pensions (DWP) to provide a single coherent view of citizens' interactions across the breadth of systems in the DWP, HM Revenue and Customs (HMRC) and local authorities via Housing Benefit. RAPID covers everyone with a National Insurance number (NINo) and for each person, the number of weeks of "activity" within these systems is summarised in each tax year. Records are then categorised as either long-term or short-term by looking for patterns of interactions with the tax and benefits system.

RAPID is a financial YE March dataset, whereas estimates of migration published in May 2022 are provided on a YE June basis. Therefore, we need to disaggregate the financial year estimates into quarters, and then predict the final quarter (April to June 2021) not covered in RAPID.

The disaggregation and prediction are based on an alternative measure of EU migration derived from IPS data. Long-standing issues with the IPS for measuring migration make it a lower quality source for estimating levels, but its monthly frequency of reporting does enable estimates of within-year seasonal trends. We derive a monthly EU migration time series, taking IPS data collected before it's suspension, and predicting a more current series up to June 2021. This prediction is derived using state space modelling (SSM), and visa data on non-EU nationals.

Using the Denton-Cholette method we applied the predicted IPS series to the financial year RAPID estimates to both disaggregate it to a monthly series and then predict this RAPID-based measure for April to June 2021.

These EU migration estimates are limited by the population coverage of RAPID. Anyone arriving in the UK needs to apply for a NINo to work, claim benefits or apply for a student loan. The coverage is extensive for most migrants because of the wide range of data sources included. However, there are some populations who have less or no interaction with the source datasets, for example, those aged under sixteen years. Also, students are only included on RAPID if they are working or claiming benefits. Using findings from previous research linking HMRC's Real Time Information (RTI) data to Higher Education Statistics Agency (HESA) data, we have applied an adjustment that estimates the non-UK born students who would not normally appear on RAPID.

In addition, the NINo registration service was disrupted during the coronavirus (COVID-19) pandemic. Operational services were partially suspended in March 2020 for certain customers, with a phased return to normal operations completed by April 2021. Consequently, some of the trends seen in EU migration from RAPID may be affected, where some migrants may have arrived in the UK but were unable to register in the normal time frame for their NINo. Such records may not be included in these estimates from RAPID. More information on the suspension of the NINo application process is available in DWP's National Insurance numbers allocated to adult overseas nationals article. Further details of the noted coverage issues in RAPID can be found in our Methods for measuring international migration using RAPID administrative data article.

During the coronavirus pandemic, some EU workers may also have been furloughed under the Coronavirus Job Retention Scheme. It is likely that some of these workers will have left the country for more than 12 months (and therefore should be measured as a long-term emigrant) but will be missed in these estimates as they still appear active on RAPID.

Further research is required and planned for the methodology for determining migration estimates using the RAPID migration dataset.

British nationals

Our research into British nationals is ongoing, but the complexity associated with identifying these migrants in administrative data means we cannot use such data at this time. For our latest estimates of migrants with British nationality, the IPS data are still our main source of information.

The IPS was reinstated in January 2021 and we use these data as our estimates for January to June 2021. To cover the period when the IPS was suspended (March to December 2020), we use the SSM time series analysis. This takes the available IPS and administrative data and uses the relationship between them to estimate the missing IPS data. We assume that the pattern of British nationals' immigration to the UK is equivalent to non-EU nationals' emigration from the UK (measured using visa data as described above) and vice versa.

The full technical details of the SSM can be found in our Using statistical modelling to estimate UK international migration methodology, updated November 2021.

5 . Production of outputs and quality assurance

The estimates undergo quality assurance to ensure they are plausible and meet the standards of experimental statistics set by the UK Statistics Authority. This includes comparing the trends with other comparable data and statistics, and consulting with the Government Statistical Service Migration Steering Group and the Migration Statistics Expert Group.

For our most recent Long-term international migration, provisional year ending June 2021 bulletin, the experimental methodology used to produce our estimates means that we are currently not able to quantify this uncertainty. Our most recent estimates therefore have an unknown degree of uncertainty around them and are experimental and provisional.

Confidence intervals were calculated in our previous versions of the model to indicate the degree of uncertainty around each estimate. They assumed that the underlying assumptions in the model were true and made no account for the uncertainty in the assumptions themselves. Users reviewing previous outputs of the model should therefore exercise caution when interpreting estimates of uncertainty.

6 . How our methods have changed over time

The way we measure international migration is constantly developing, taking account of other new data sources, improved methods and the changing needs of our users. Detailed information about our previous methods is given in our Using statistical modelling to estimate UK international migration methodology.

We summarise two former different methods we have used to calculate international migration, both developed since the suspension of the International Passenger Survey (IPS).

Version one

The first version of the model was published April 2021 and provided migration estimates for March 2020 to June 2020.

Since 2019, our methodology has been exploring time series modelling as an approach to estimate migration using administrative data. Time series modelling was selected because of the strong seasonal trends that are evident in international migration over time.

The state space model (SSM) was introduced in the first version. We projected the trends and seasonality of the previous IPS data forward and then adjusted it by the structural shift seen in the Home Office visa data for non-EU citizens. This version includes assumptions on EU nationals having different travel options during lockdowns. When airports were closed, EU nationals were still more able to travel by ferries and Euro Tunnel.

Version two

This version was published November 2021 and produced estimates of migration for Quarter 3 (July to Sept) and Quarter 4 (Oct to Dec) of 2020. Estimates produced from version one of the model (March to June 2020) were also revised because of data updates and improved methodology since April 2021.

The model in version two remained relatively similar to version one, with the main change being that we created an EU proxy series for estimating migration of EU nationals. This new EU trend used in the SSM was based on historical movements of EU nationals, rather than an assumption that EU nationals behave the same as non-EU (as in version one).

7 . Strengths and limitations

Strengths

The current approach, as published in May 2022:

- provides a calculated estimate of migration flows at a time when traditional data collection through the International Passenger Survey (IPS) was absent and migration patterns were changing because of the coronavirus (COVID-19) pandemic and Brexit
- uses a wide range of existing and new data sources to observe the behaviour of migrants, greatly helped by the data-sharing powers of the Digital Economy Act 2017
- includes provisional estimates that will be routinely revised in response to more timely and relevant information from the data
- includes estimates increasingly derived by administrative data sources that capture actual rather than intended migration patterns (as was measured by the International Passenger Survey (IPS))
- includes constantly evolving methods, taking account of other new data sources, improved processing methods and the changing needs of our users; it will join up with other Office for National Statistics (ONS) transformation work on population and migration statistics
- involved developing our measures, which has supported and encouraged greater involvement and cooperation of experts and other government departments in the process of producing migration statistics; this leads to enhanced trust in, and an understanding of, these statistics

Limitations

The current approach, as published in May 2022:

- has discontinuity with migration flow estimates prior to 2020
- can only produce headline figures for migration, by direction of flow and broad nationality groups (British, EU and non-EU); this does not meet all our users' needs
- includes estimates that do not currently cover all types of migrants, as some migrants are harder to identify within administrative data sources
- includes estimates that are currently experimental; there is a degree of uncertainty around them, which we are unable to quantify at this point in time
- relies on administrative data, so these estimates are not timely (almost one year after the reference period); we will publish more information in summer 2022 on how we aim to improve the timeliness of these estimates
- requires users to adapt to the new practice of regular revisions of migration estimates

8 . Future developments

The Office for Statistics Regulation's (OSR's) review on migration statistics encouraged us to improve and broaden our user engagement as well ensure we have coherent plans across our transformation work. We recognise the need to continuously improve our methods with our users. If you would like to find out more or have any feedback, please email us at pop.info@ons.gov.uk.

We continue to explore other data sources as they become available and existing sources as they are updated. In the coming year, we plan to:

- compare our outputs to final Census 2021 data, which will provide the most robust and comprehensive picture of the England and Wales population
- explore insights available from advanced passenger information
- investigate new visa requirements for EU nationals, with the intention to incorporate associated data into our method
- investigate how other types of migrants, including refugees and asylum seekers, can be better incorporated into future versions of our method
- explore how these models can be improved to provide more timely and frequent measures of migration, and with breakdowns, including by reason of travel, geography, sex and age
- work with other government departments to ensure published migration statistics are coherent

9 . Related links

Using statistical modelling to estimate UK international migration
Methodology | Updated November 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.

Long-term international migration, provisional: year ending June 2021
Statistical bulletin | 26 May 2022
Experimental statistics on UK international migration throughout 2020 to 2021, including the effects of the coronavirus (COVID-19) pandemic.

Population and international migration statistics revisions policy
Methodology | 26 May 2022
This policy explains how we will deal with revisions specific to population and international migration statistics.

Irregular migration to the UK, year ending December 2021
Article | 24 February 2022
This release from the Home Office provides an overview of irregular migrants who come to the UK, including those arriving on a small boat across the English Channel (a "small boat arrival").

Guide to statistical revisions
Policy | Released 28 January 2016
A brief introduction to the reasons why statistics may need to be revised after their first publication or release.

How we are improving population and migration statistics
Article | Released 15 November 2021
Latest update on our population and migration statistics transformation journey.

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.

EU settlement scheme quarterly statistics, March 2022
Article | Released 26 May 2022
This report provides detailed statistics on applications made to the EU settlement scheme (EUSS) from 28 August 2018 to 31 March 2022, and applications concluded during the same time period.

Irregular migration to the UK, year ending March 2022
Article | Released 26 May 2022
Statistics on irregular migration to the UK, including small boats.

Immigration statistics, year ending March 2022
Article | Released 26 May 2022
Quarterly and annual statistics relating to those coming to the UK, extending their stay, gaining citizenship, applying for asylum, and being detained or removed, as well as immigration for work, study and family reasons, including new visa routes where these are operational.

Statistics relating to passenger arrivals since the COVID-19 outbreak, May 2022
Bulletin | Released 26 May 2022
A statistical report showing the effect of COVID-19 on the immigration system, up to April 2022.