Office for National Statistics

Methods to produce provisional long-term international migration estimates

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

Contact:
Brendan Georgeson
pop.info@ons.gov.uk
+44 1329 444661

Release date:
23 November 2023

Next release:
To be announced

Table of contents

1 . Purpose of this methodology

This methodology article summarises how we produced our provisional estimates to provide an early indication of migration in the year ending June 2023, as published in our Long-term international migration, provisional: year ending June 2023 bulletin.

The methods detailed in this paper are part of an ongoing programme of work to transform population and migration statistics. This work was accelerated in response to the coronavirus (COVID-19) pandemic, and we continue to build on these developments. More information can be found in our Improving international migration statistics using administrative data article.

2 . Improving how we measure international migration

The International Passenger Survey (IPS), which underpinned our estimates of migration until 2020, had been stretched beyond its original purpose. More information can be found in our Improving international migration statistics using administrative data article. To improve migration statistics and produce more regular and timely statistics without compromising on accuracy, we now focus on producing estimates using administrative data supported by statistical modelling.

As we make greater use of administrative data, we should be able to publish more timely estimates with a more consistent level of quality. The first, most timely, estimates will be provisional estimates, giving an early indication of population or migration based on early data and assumptions about people's migration status based on past behaviour. When more data relating to the reference period become available, these provisional estimates will be updated with new estimates, with reduced statistical uncertainty.

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

A common misconception is that it is easy to measure international migration 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. More information can be found in our Understanding international migration statistics methodology.

4 . The method for our latest estimates

Our provisional estimates for year ending June 2023 rely more on observed activity in administrative data, as explained in our International migration: developing our approach for producing admin-based migration estimates article.

Our year ending June 2023 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, and further nationality breakdowns for immigration into the UK.

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. A summary of the main concepts behind our estimates can be found in our Understanding international migration statistics methodology.

Non-EU nationals

Non-EU migration refers to the migration of people who do not hold British or EU nationality. We use Home Office Borders and Immigration data, which combine visa and travel information, to link an individual's travel movements into and out of the country. For more information, please see the Home Office statistics on exit checks: user guide.

Our first step is to identify which travellers meet the definition of a long-term migrant, filtering out those on long-term visit visas. Those on long-term visit visas are only eligible to stay in the UK for up to six months per visit. As these visas can be valid for up to 10 years, they would appear to be long-term migrants.

To estimate migration, we look for travel across an extended visa period. Visa periods are constructed by linking together any consecutive or concurrent visas held. If there is a gap of more than seven days between visas, then a new visa period is started.

For immigration, we look at the first arrival in a visa period and then 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 person will be considered a new long-term immigrant.

We use first arrival and last departure dates within a visa period as an approximation for length of stay in the UK. To estimate long-term international migration, this total length of stay must be over 365 days. The “first arrival, last departure” method allows us to exclude trips abroad over the course of an extended period of residence. If either (but not both) of these dates are missing and there is evidence of travel into or out of the UK during their visa, then visa start or end dates are used as a proxy.

To measure emigration, we identify previous long-term immigrants with a last departure from the UK during the reference period. We record them as a long-term emigrant if they do not return to the UK within 12 months, or if they only return for a short-term stay.

Adjustments to produce provisional estimates of migration

For individuals whose first arrival occurred within the 12 months before the end of the reference period (currently June 2023), there is not yet enough information to see a stay of 12 months or more. To provide estimates of immigration for this period, we use past behaviour showing the proportion of arrivals on long-term visas who actually stay short term. We use this to create an assumption of what proportion of arrivals on long-term visas will leave before staying a year.

This is known as an early leaver adjustment. We apply this adjustment to the most recent four year-ending (YE) periods (currently YE September 2022, YE December 2022, YE March 23, YE June 2023) by reason for migration because our research has shown that different groups exhibit different behaviours. These provisional estimates are updated as we get further travel data to identify actual behaviour during this time.

We also implement an early leavers adjustment for those arriving on the Ukraine schemes and British National (Overseas) arrivals where we have inadequate information to suggest how many stay for 12 months or more.

As these schemes opened in the last two years, we do not have historical information to inform this adjustment. We use insights provided by the Home Office showing the number of individuals on these schemes who have left and remained out of the UK for an extended period, suggesting they have not stayed 12 months.

To produce provisional estimates of emigration, we cannot exclusively use the “first arrival, last departure” approach as this requires a completed visa period within which a last departure can be identified.

For the most recent 12 months, not enough time has passed to see if an individual has left for 52 weeks or more with no re-arrival. We add an adjustment to include individuals who have left the UK and not subsequently returned and who have up to three months left on their visa.

For years other than the most recent 12 months, we include individuals in the emigration estimate who have been absent for 52 weeks or more and have an open visa period because these individuals would not be identified as an emigrant.

Our previous assumption was that people emigrated at the end of their visa, unless we saw evidence of travel on a new visa. For our latest estimates, we identified individuals, particularly students, who have moved to a new visa but did not travel on this. We have updated our methodology and implemented an adjustment to account for this.

Further information on each of these adjustments, and how we have improved our assumptions, can be found in our International migration research, progress update: November 2023 article.

EU nationals

Our latest methodology to estimate the migration of EU nationals is based on our Methods for measuring international migration using Registration and Population Interaction Database (RAPID) administrative data. 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.

In RAPID, we include two arrival categories to identify long-term migrants to the UK which most closely align with the United Nations (UN) definition of a long-term migrant.

RAPID is made available to the Office for National Statistics (ONS) on an annual basis in Quarter 3 (July to Sept) for the previous tax year. Currently we publish bi-annual international migration estimates, which requires forecasting RAPID for three or nine months.

Temporal disaggregation is used to perform the breakdown of the annual RAPID EU data to quarterly data and for forecasting of the RAPID EU estimates where the current international migration estimate is beyond the RAPID data we currently have. For the disaggregation, we use the Fernandez method which uses a regression approach to look for a relationship between two datasets. Forecasting is the process of generating figures beyond the timeframe of the RAPID data currently held. It is based on the signals and trends in the higher frequency time series.

We currently use the International Passenger Survey (IPS) as the input for the higher frequency time series. While we have acknowledged the long-standing issues with IPS measuring the levels of migration, the IPS seasonality and trend of migration flows are useful for the RAPID disaggregation.

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, and we adjust our EU estimates to account for this under-coverage as outlined below.

Inflow and outflow adjustment

International migrants that have recently arrived or departed the UK will not necessarily meet the UN definition of a long-term international migrant (LTIM). We adjust the estimates from the two most recent tax years to take account of this. This is done by calculating the proportion of migrants that have historically become LTIM, both immigrants and emigrants, and applying this to the number of arrivals and departures in the most recent tax years.

Students who may be excluded if they are not working or claiming benefits

To identify students immigrating into the UK long-term, we use the Higher Education Statistics Agency (HESA) dataset and HMRC Pay as You Earn Real Time Information (PAYE RTI) data as the best available data sources. Our latest method links these two sources using the Demographic Index to better understand how many international students are in employment alongside their studies. We are then able to adjust our estimates to include the proportion of students that do not show employment activity and are therefore not included in our early estimates.

Children aged under 16 years (U16s)

The adjustment for children aged under 16 years (U16s) uses an adult-to-child ratio derived from the IPS. Where IPS data are not available (2020), a three-year average ratio (2018, 2019 and 2021) is applied. This ratio is then used to calculate the number of U16s to add to the RAPID estimate. This ratio is calculated separately for both immigration and emigration.

During the coronavirus (COVID-19) 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 they will be missed in these estimates as they still appear active on RAPID.

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

Asylum applicants and resettlement scheme arrivals

We add the total number of asylum applicants and resettlement scheme arrivals to the international migration estimate not including arrivals from the Ukraine Schemes and British Nationals (Overseas). We link asylum applications data to visa data to identify who was already in the country before applying for asylum and link the application data to returns data to identify who left within a year of application.

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 2021 to December 2022. To cover the period when the IPS was suspended (March to December 2020), we use the state space model (SSM) time series analysis. This takes the available IPS and administrative data and uses the relationship between them to estimate the missing IPS data. The full technical details of the SSM can be found in our Using statistical modelling to estimate UK international migration methodology.

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. Further information on quality can be found in our Long-term international migration estimates: quality assuring of administrative data methodology.

For our most recent Long-term international migration, provisional year ending June 2023 bulletin we have included uncertainty intervals associated with some parts of the admin-based migration estimates (ABME) process: adjustments, modelling, and survey-based estimates. We quantified uncertainty across three nationality groups: EU, non-EU, and British nationals. These were drawn together by a composite measure to create total immigration, emigration, and net migration uncertainty intervals.

The uncertainty intervals are incomplete and should be interpreted with caution as our simulations do not measure all sources of uncertainty and likely underreport the true uncertainty. Some main sources of uncertainty, like uncertainty associated with the administrative data, are not included. More information can be found in our Measuring uncertainty in international migrations estimates working paper series.

6 . 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 as ensuring that 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:

- explore how administrative data could be used to produce estimates of migration for British nationals
- continue to investigate new visa requirements for EU nationals, set out in GOV.UK's New immigration system guidance, with the intention to incorporate associated data into our method
- explore how our methods can be improved to provide more timely and frequent measures of migration, and with breakdowns, including by reason for travel, geography, sex and age
- research how to integrate those who hold Indefinite Leave to Remain into our estimates
- work with other government departments to ensure published migration statistics are coherent
- implement a reproducible analytical pipeline (RAP) strategy to improve the quality of our analysis, increase trust in our analysis, and create more efficient processes

7 . Strengths and limitations

Strengths

The current approach, as published in November 2023:

- uses a wide range of data sources to observe the behaviour of migrants, greatly helped by the data-sharing powers of the Digital Economy Act 2017
- includes estimates increasingly derived by administrative data sources that capture actual, rather than intended, migration patterns (as was measured by the IPS)
- includes improved methods, taking account of new data sources, better 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
- had greater involvement and collaboration of experts and other government departments in the process of producing these migration statistics; this leads to enhanced trust in, and understanding of, these statistics, as well as more consistent methods across all government international migration data

Limitations

The current approach, as published in November 2023:

- can only produce headline figures for migration by direction of flow and broad nationality groups (British, EU, and non-EU), which does not meet all our users' needs
- our previous method counted someone as an emigrant at the end of their visa unless we saw evidence of travel on a new visa; we have applied an adjustment to account for this but will continue to improve the methodology to address this limitation in May 2024

8 . How our methods have changed over time

The way we measure international migration is evolving, taking account of new data sources, improved methods, and the changing needs of our users.

We summarise the different methods we have used to calculate international migration, developed since April 2021.

Version one

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

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

Version two

This version was published November 2021 and produced estimates up to December 2020.

The model in version two remained similar to version one, however, we updated one of our assumptions in the model by creating 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 nationals (as in version one).

Version three

This version was published in May 2022. This version relied more on administrative data and benchmarked the EU estimate to the Department of Work and Pensions (DWP) Registration and Population Interactions Database (RAPID) for the first time.

In version three, record-level Home Office Borders and Immigration data were not available for the final eight months of the time series (November 2020 to June 2021). This missing time period was estimated using an aggregated version of the data. We applied the pattern of change observed in the aggregate dataset with the Denton-Cholette method to predict the record level dataset for the missing period. This provided us with a rate of change for arrivals on a month-to-month basis.

For more information on the Denton-Cholette method, please see the European Commission's Temporal disaggregation, benchmarking and reconciliation guidelines (PDF, 2.3MB).

At the time, we did not have an equivalent method for measuring non-EU emigrants. Therefore, we calculated a ratio between emigration (numerator) and immigration (denominator) on a monthly basis from the aggregate dataset, which was applied to the calculated non-EU immigration estimates to estimate emigration. This assumed the trends in the aggregate dataset for the immigration series and emigration series are similar to the trends in the record-level data, as both are derived from the same source.

However, subsequent estimates of emigration using record-level data (introduced in version four) entailed average revisions of around 200,000 emigrants compared with the version two approach. These revisions are presented in Section 8: Revisions to migration statistics of our Long-term international migration, provisional: year ending June 2022 bulletin.

Version four

This version was published in November 2022 and produced quarterly estimates for year ending June 2020 to June 2022.

In version four, for EU nationals we introduced a more comprehensive student adjustment based on the linking of Higher Education Statistics Agency (HESA) and HM Revenue and Customs (HMRC) Pay as You Earn Real Time Information (PAYE RTI) datasets to understand how many international students were in employment during their studies. We also introduced an adjustment for those aged under 16 years, which until this iteration had been excluded from our estimates. This adjustment was based on an adult-to-child ratio derived from the International Passenger Survey (IPS).

The non-EU estimates in version four used the complete time series for record-level Home Office Borders and Immigration data, which had been unavailable in version three. The methodology of emigration for non-EU nationals was made comparable with the methodology of the immigration of non-EU nationals, and an adjustment for recent arrivals and departures was implemented to avoid overestimation.

Further detail on these changes can be found in our International migration research progress update: November 2022 article.

Version five

This version, published in May 2023, produced quarterly estimates for year ending December 2018 to December 2022.

Changes to EU estimation

The EU estimates of international migration continue to be based on DWP's RAPID. In our original research article into this administrative data source, we used four categories based on the amount of activity for the identification of long-term international migrants. For further information, please see our Methods for measuring international migration using RAPID administrative data.

The C1 and C2 categories most closely aligned with the UN definition of a long-term migrant. When we initially developed our method, we expanded on this definition of long-term activity to reflect the complexity of people's lives. This created two further categories: C3 and C4.

As part of the post-census reconciliation of the population estimates, we have reassessed the international migration component based on our new methodology using administrative data. Through this process, we have acknowledged that those categorised using the C3 and C4 classifications are likely to have pushed the definition of long-term migration too far, resulting in the overestimation of migration. Therefore, we have redefined our long-term international migration (LTIM) estimates to only include those people who meet the C1 and C2 classification criteria.

In addition to this, we have also brought our student emigration adjustment in line with the methodology introduced in version four for the student immigration adjustment.

Changes to non-EU estimation

We have continued to refine our methods using Home Office Borders and Immigration data, and we have updated several components, which are summarised as follows.

We now calculate the recent arrival adjustment based on reason for migration rather than at the headline figure (all immigration). This is to reflect that different types of migrants have different behaviours.

We have also implemented an adjustment for the Ukraine schemes and British National (Overseas) visa route arrivals where we do not have enough information to suggest how many people do not have a long-term stay of 12 months or more. Internal Home Office analysis shows that many people left the UK before reaching the 12-month period. We have used these data to implement an adjustment for these groups and removed these arrivals from our long-term immigration estimates where they have been outside the UK for eight weeks or more.

We added an emigration adjustment to account for those who have left the UK long-term but still have an open visa. For the most recent 12 months, we add an adjustment to include individuals who have left the UK and not subsequently returned, and who have up to three months left on their visa after the data extract. For other years, we add an adjustment to include individuals in the emigration estimate who have been absent for 52 weeks or more, and whose exits have not been identified because of the visa period still being valid.

Inclusion of asylum seeker applicants and returns, and resettlement scheme arrivals

Estimates included asylum seekers and resettled refugees, for the first time since moving to administrative data. An asylum applicant (also referred to as an "asylum seeker") is someone who makes a claim to be recognised as a refugee under the Refugee Convention.

We added the total number of asylum applicants and resettlement scheme arrivals to the international immigration estimate. We used internal Home Office analysis to apply an adjustment to account for people who applied for asylum after entering the country on a long-term visa (for example, on a student visa), by removing them from the count of asylum applications.

We added the total number of asylum returns to the international emigration estimate.

To include asylum seekers and resettled refugees in the long-term international migration (LTIM) estimation, we have made the assumption that:

- all asylum seekers are LTIM based on processing times in recent years, assumed deportation lag, and declared intent to stay in-country
- returns are people who have been in-country for more than 12 months
- all resettled persons are LTIM based on declared intent to stay in-country
- all asylum seekers arrived in the UK no more than 12 months before applying

All assumptions will be reviewed going forward.

Changes to British nationals estimation

Following the inclusion of British National (Overseas) visa route arrivals in the non-EU estimates based on Home Office Borders and Immigration data, we have made an adjustment to the IPS-based estimates to avoid double counting. We did this using a combination of country of birth and country of last residency, and removing these from the IPS estimates.

Version six

In this version, published in November 2023, we updated our assumption that people emigrated at the end of their visa, unless we saw evidence of travel on a new visa. We implemented a visa transitions adjustment to account for a change in behaviour, where individuals moved to a new visa, but did not travel on this, and have updated our year ending December 2022 estimates. More information can be found in our International migration research, progress update: November 2023 article.

For the disaggregation of our EU estimates, we transitioned from using the Denton-Cholette method to the Fernandez method.

We updated our estimates for the year ending June 2012 to June 2021. This allowed us to provide a coherent back series of EU, non-EU and British long-term international migration to and from the UK in line with evidence provided by the Census 2021 and other sources. More information can be found in our Estimating UK international migration: 2012 to 2021 article.

We improved the count of asylum applicants and returns by linking the asylum applications data to visa data to see who was already in the country before applying for asylum, and linking the applications data to returns data to see who left within a year of applying for asylum. Therefore, we no longer assume that all asylum seekers are long-term international migrants (LTIM) or that returns are people who have been in-country for more than 12 months.

We have published uncertainty estimates to accompany our LTIM estimates for year ending June 2023.

Changes to nationality breakdowns for immigration into the UK

The non-EU nationality breakdowns for immigration are calculated at the individual nationality level. To calculate total immigration for the individual nationality, unadjusted immigration counts are produced using Home Office Borders and Immigration data and are broken down by nationality. The early leavers immigration adjustment is applied, by reason for migration, for each of the nationalities and the count for each reason is summed to create the nationalities total immigration figure.

EU nationality breakdowns are calculated differently. They cover the year ending March 2023 and consist of country grouping level breakdowns (EU14, EU8, EU2 and EU other). The reason for the difference in time periods is that the EU estimates are calculated using the Registration and Population Interaction Database (RAPID). To produce EU nationality breakdowns, we take the inflow data from RAPID and disaggregate it to the country grouping level. We do not apply an under 16 years (U16s) or a student adjustment.

9 . Related links

Long-term international migration, provisional: year ending June 2023
Bulletin | 23 November 2023
Official statistics (in development) of UK international migration, year ending (YE) June 2012 to YE June 2023; estimates from YE December 2022 and YE June 2023 are provisional and will be updated when more complete data are available.

International migration research, progress update: November 2023
Article | 23 November 2023
An update on international migration methods and research.

Estimating UK international migration: 2012 to 2021
Article | 23 November 2023
The coherence of migration statistics is an important part of the work the Office for National Statistics (ONS) are doing to transform the way we produce population and migration statistics for our users. This article 1) explains how we arrived at the revised back series to be used for rebasing of mid-year population estimates, 2) sets out how migration has changed over the decade, 3) assures that new methods are robust.

Long-term international migration estimates: quality assuring administrative data
Article | 16 November 2023
Administrative data sources and quality assurance in the production of admin-based long-term international migration estimates published in bi-annual releases.

Improving international migration statistics using administrative data
Article | 23 November 2023
An article that explains to users why we have moved from a survey-based methodology for estimating international migration to using administrative data, why this is an improvement and then the remaining challenges. Its purpose is to bring all the information we have published over the years into one, easily digestible article.

10 . Cite this methodology

Office for National Statistics (ONS), released 23 November 2023, ONS website, methodology, Methods to produce provisional long-term international migration estimates