

International Migration Statistics, current position and plan

May 9, 2024

Dominic Webber, Laura Cheatham, Brendan Georgeson and Melissa Randall, Principal Researchers, Migration Statistics Division

Mark Edwards, Methodology and Quality Division

1. Executive Summary

The Office for National Statistics (ONS) produces regular estimates of international migration. Long-term international migration (LTIM) statistics estimate the flow (or movement) of migrants to and from the UK. This includes our headline estimate for net migration, which is the difference between the number of people who enter the UK (immigration) and the number of people who leave (emigration).

To produce these estimates, we use the [UN-recommended definition of a long-term international migrant \(PDF, 5.0MB\)](#), 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 estimates are important because migration statistics help describe how the number of people living in different parts of the country changes over time. This helps decision makers to plan and allocate resources effectively, including the funding of local services.

To ensure that we continue to provide accurate and timely statistics that meet users' needs, we have been improving our methods. This includes moving away from survey-based data where we would ask people if they were moving to or from the UK, to using administrative (admin) data. Admin data are data collected for the purposes of services such as tax, benefits, and health systems. This includes data collected at the border as people enter or leave the UK.

This paper presents an overview of the methods for long-term international migration statistics, consistent with the latest estimates published in November 2023 for [year ending June 2023](#). It describes the approach for measuring EU, non-EU and British migration. This is followed by a description of the assumption-setting process which enables provisional estimates 5-months after the reference period ends, despite not yet having 12-months of travel information for recent arrivals and departures. It concludes by highlighting the main methodological improvements current research is aimed towards.

2. Ask

Members of the Methodological Assurance Review Panel (MARP) are invited to:

- Provide feedback on our current methods for measuring international migration.

- Note our current research plans to explore new methods for EU and British migration, and to expect research papers on these during the Summer.
- Help direct these research projects, by providing their views on methods to for evaluating the appropriateness of sources.
- Provide views on how best to draw strength from multiple sources which have differing levels of quality for different sub-populations.
- Provide feedback on the appropriateness of the assumption setting process for producing timely estimates of international migration.

3. Short history of international migration methods

Until the early 2020s the International Passenger Survey was the primary source for measuring international migration. These data are based on the reported migratory intentions of UK residents departing the country and foreign residents arriving. Adjustments were then made to account for asylum seekers, resettlement schemes, changes in intentions, and flows to and from Northern Ireland.

This measurement method had longstanding challenges: mainly that some peoples' plans were transient and so their stated intentions didn't always reflect their actual behaviours. Indeed, research from 2019 ([Understanding different migration data sources article](#)) concluded that the IPS had been underestimating EU immigration and underestimating non-EU student emigration during the 2010s.

While adjustments were introduced for these groups, the ONS recognised that the IPS was being stretched beyond its original purpose. In August 2019, we asked the Office for Statistics Regulation (OSR) to support our [reclassification of international migration statistics to Experimental Statistics](#) (now referred to as official statistics in development). In August 2020, we announced in our [Population and migration statistics system transformation article](#) that we would be moving away from using the IPS to produce official international migration statistics. Instead, we would be moving to produce ABMEs. The suspension of the IPS during the coronavirus pandemic accelerated this ambitious programme of work. Since then, migration methods have transitioned from modelling IPS data using admin data.

The decision to move away from IPS data to admin data was more recently vindicated through analysis comparing estimates from these two sources against the changes between the two most recent Censuses. The [Estimating UK international migration: 2012 to 2021 article](#) demonstrates that estimates of migration derived from administrative are more accurate than previous IPS-based ones.

4. Methods for producing LTIM estimates.

The following section provides a high-level overview of the methods used to produce international migration estimates. The annex of this paper provides a more detailed technical description of the methods, as presented in our forthcoming Technical User Guide (link available on 3rd May).

Non-EU

Non-EU migration refers to the migration of people who do not hold British or EU nationality. We use Home Office Borders and Immigration (HOBI) 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.

EU

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.

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. To convert the tax-year dataset into a quarterly series, we use the [Fernandez method](#). This uses a regression approach to look for a relationship between two datasets.

We then use forecasting techniques to generate figures beyond the timeframe of the RAPID data currently held (i.e., for when the reporting period is beyond the end of the tax year for the latest version of RAPID we currently have access to). 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. As mentioned in the Areas of Current Research section, our ambition is to replace RAPID-based estimates of EU with methods using Home Office data. This will mean that we have a more timely and frequent data source, as well ensuring consistency with non-EU methods.

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 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** - The adjustment for children aged under 16 years 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.

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

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. Producing timely, provisional estimates of migration

Estimates of international migration are published just under 5-months after the reference period. That means, despite the move from an intention to actuals-based measurement approach, statistics on migration are as timely as they were before the switch away from IPS data.

In order to achieve this level of timeliness, a set of assumptions are applied, for instance on expected length of stay. This addresses the fact that while the administrative data allows a clear picture on the total number of arrivals and departures within the reference period, at the point of publication, we don't know the extent to which they will be long-term or not. This means that these first estimates are provisional, have greater uncertainty, and will be

updated as we gather a full 12- months of travel data to confirm people's long-term migration status.

The following describes the immigration early leaver adjustment, while more information on the full range of assumptions is provided in [International migration research, progress update: November 2023](#).

Immigration early leaver adjustment

For individuals whose first arrival occurred within the 12 months before the end of the reference period, we do not yet have enough information to see a stay of at least 12 months as not enough time has passed. Therefore, the early leaver adjustment uses past trends to estimate how many people arriving on long-term visas will stay in the UK for over a year.

The early leaver adjustment is challenging to implement, as trends are currently changeable following the pandemic. Using international students as an example, the extent to which students left the UK early has changed significantly since 2021 compared with the period before. This will reflect different factors affecting student behaviours over this time, such as the pandemic, student dependents coming to the UK long term, and the introduction of the new Graduate visa scheme.

The early leavers adjustment is applied by reason for migration as our research has shown that different groups exhibit different behaviours. Using study and work as examples, our assumption for data published in May 2023 were based on data for YE December periods 2018, 2019 and 2020.

We have improved our assumptions to give more weight to recent data that are likely to better reflect future trends. For estimates published in November 2023, we avoided using data that were affected by the pandemic, and therefore included YE June 2019 and YE June 2022. We have double weighted YE June 2022 as this is most likely to reflect future trends.

The proportion of total long-term study arrivals who did not stay for at least 12 months in the:

- YE December 2018 was 18%
- YE June 2019 was 21%
- YE December 2019 was 34%
- YE June 2020 was 32%
- YE December 2020 was 20%
- YE June 2021 was 21%
- YE December 2021 was 4%
- YE June 2022 was 14%

Our previous assumption used in estimates published in May 2023 was based on the YE December 2018, 2019, and 2020, which estimated an average of 24% of study arrivals did not become long-term. Our updated assumption used in estimates published in November 2023 is based on the YE June 2019 and 2022 (double weighted), which estimated an average of 16% of study arrivals did not become long term.

In previous iterations, we applied 100% of the adjustment to the most recent year ending period, and applied proportionally back 75%, 50% and 25% to the remaining three-year end periods. We have now applied this adjustment to immigration in individual quarters to reflect that arrivals are not constant throughout the year. For example, there were 1 million arrivals on long-term visas in the YE December 2022. Previously, we would have applied the adjustment to 50% of the total arrivals (covering the second half of the year, where we do not have complete travel information). 64% of arrivals occurred between June and December 2022, so we have applied the adjustment to this group.

6. Areas of current research

British Nationals

While long-term international migration estimates for non-British nationals are now derived from administrative data, estimates of immigration and emigration of British nationals remain based on the International Passenger Survey (IPS).

As part of our work to [update the back series of UK international migration statistics for the period 2012 to 2021](#), we revised our historic estimates for the international migration of British nationals. This is to take account of evidence from the 2011 Census and Census 2021 on the change in the size of the UK-born population over the decade. These changes have resulted in an increase to our estimates of emigration of British nationals from 2012 to 2021.

We do not currently have sufficient evidence to indicate whether the higher levels of British emigration seen prior to Census 2021 continued into 2022 and 2023. Therefore, from the year ending (YE) September 2021 onwards, we are continuing to use estimates from the IPS, while we undertake further research to identify an administrative data-based alternative.

There are currently two main alternatives in scope: a RAPID-based approach, following by research exploring the viability of Advance Passenger Information.

- Current research is examining the potential of RAPID to identify the migration of British nationals, using methods that are similar to those currently used to measure EU nationals. The main challenge is the uncertainty of migration events for British nationals. There are numerous ways someone can be inactive on the Department for Work and Pensions' and HM Revenue & Customs' systems as a British national and remain in the UK (for example, a student or someone not in employment, education or training (NEET)). The main thrust then of this research is to design a series of rules that aim to enable a robust classification for those who appear and disappear from RAPID so that they can be accurately reported as migrants or otherwise. We are testing our new method for using RAPID to measure British national migration throughout 2024 and will aim to bring a paper to MARP in August 2024.

- Advanced passenger information (API) is constructed of records generated by the process whereby passengers are required to provide certain information to the airline or travel company before travel. The information normally details passport number, the name on the passport, gender, and date of birth. In July 2022, a feasibility study was conducted on API with the aim of assessing whether it could provide an additional data source for migration estimates that improve the quality. The findings found API could be useful for assessing arrivals and departures of British nationals. We are in the process of acquiring a dataset of API from which we aim to conduct further research. We aim to have a first research output by November 2024.

Measuring EU

The Registration and Population Interactions Database (RAPID) remains the best available administrative data source for estimating the migration of EU nationals. However, since January 2021, EU nationals wanting to move to the UK either needed a visa or to apply for the EU settlement scheme (EUSS), which opens the possibility of using Home Office Borders and Immigration (HOBI) (the principal source for non-EU migration).

To date we have applied the same “first arrival, last departure” (FALD) method to EU visa holders as we currently use to produce migration estimates for non-EU nationals. [Early analysis](#) has compared these estimates of long-term immigration with data published by the Home Office showing how many long-term visas were granted to EU nationals in the same period. Initial analysis suggests that, of EU nationals that had a visa granted in the year ending (YE) June 2023, 56% arrived in the UK in that 12-month period.

We have also produced estimates of EU visa holders emigrating out of the UK up to the YE June 2023 (Figure 3), using HOBI data. We expect these estimates to be low because EU nationals have only required a visa to move to the UK since January 2021, and therefore not enough time has passed for many individuals to be coming to the end of their visas and leaving the UK. For example, if a person arrived in January 2021, they would need to remain in the UK for a period of 365 days (and be confirmed as a long-term immigrant), and then subsequently leave for another period of 365 days before being considered a long-term emigrant. Initial estimates of emigration of EU visa holders are very low until the YE December 2022, where we are starting to see the first EU visa holders coming to the end of their visas and leaving the UK.

In order to derive a new method for visa-based travel for EU nationals needs to account for the travel patterns of those with settled status. The EUSS enables EU, other European Economic Area (EEA) and Swiss citizens and their family members who are residents in the UK by the end of the EU exit transition period (31 December 2020) to obtain the status required to live and work in the UK.

Initial research using the “first arrival, last departure” (FALD) method to identify those who hold indefinite leave has shown it may not be suitable for estimating this group. Most people who are entitled to indefinite leave were residents in the UK before their application. This makes our FALD method (which assumes a visa is granted before arrival into the UK) unsuitable for this leave type, as most of those applying will already be residents in the UK and should not be included in our long-term immigration estimates. The last departure approach makes assumptions around an absence from the UK relative to the time remaining

on a period of leave. Those with indefinite leave do not have a time limit on the amount of time they can reside in the UK.

The method currently being researched is based around travel patterns and the amount of time spent in and out of the country. The method has been designed to be flexible in terms of defining travel periods and the maximum length of time required for someone to be categorised as a long-term migrant. For example, we do not want to count a weekend away or a week-long holiday as part of an emigration calculation.

Figure 1: Illustrative example of periods in and out of the UK

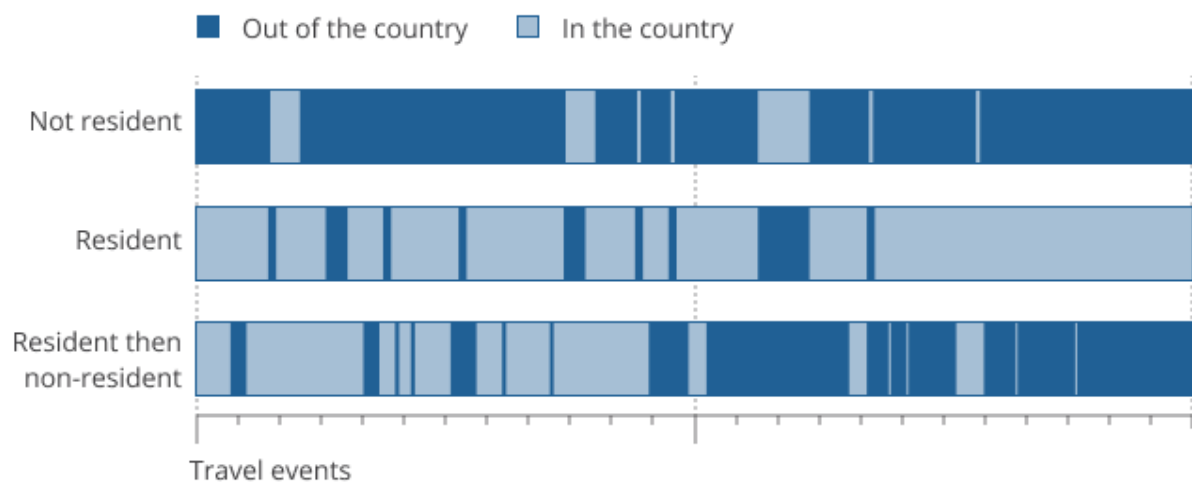


Figure 1 provides an illustrative example of three types of travel history we might see in the data. The first example shows evidence of a person who is not a resident in the UK but took several short trips to the UK. The second example shows someone who is a resident in the UK and took several trips outside of the UK. The third example shows someone that was a resident then transitioned to a non-resident. The third example also highlights one of the challenges that we face with categorising an emigration event and determining exactly when it occurred.

Owing to the requirement of residency to be entitled to settled status, the approach to classifying someone as a long-term immigrant or emigrant must be reassessed. We start from a position where someone with settled status is a resident in the UK, and therefore, the first event we look for is a long-term emigration. We have travel data going back to 2015, which means we can assess travel patterns over several years.

Using travel dates, we calculate the length of stay in and out of the UK. We identify extended trips out of the country and group them into a single travel period; we then assess the cumulative time in and out of the country. We look for two pieces of evidence that a person may have emigrated long-term. Firstly, we identify those who spend more time out of the UK than in. We then look at the total amount of time spent outside of the UK. If the total time out of the UK meets a set emigration criterion, we class them as a long-term emigrant. Once

someone has emigrated, we can look for immigration events using an inverse approach (identifying extended periods within the UK, which we link together to determine whether a person meets the immigration criteria).

Research into the use of Home Office Borders and Immigration (HOBI) data for people with EU nationals – both with a visa and settlement leave is ongoing. We are aiming to conclude a first comprehensive research paper in August which will present to MARP. That paper will aim to answer the fundamental question: do we obtain higher quality estimates of EU migration with HOBID than those available from RAPID.

Measuring uncertainty

Providing measures of uncertainty for statistical estimates is important to inform users of quality and follows best practice for communicating uncertainty with the estimates. The Year Ending June 2023 provisional LTIM estimates was the first to include measures of uncertainty to accompany the mainly administrative based international migration estimates. These uncertainty measures were in the form of uncertainty intervals and represented the quantification of doubt with the estimates.

Producing administrative based international migration estimates is a complex process with multiple steps in the estimation process. Our measures of uncertainty that accompany LTIM estimates are only a partial quantification of doubt with the statistical process and likely underestimate the associated uncertainty. It is therefore feasible that revised LTIM estimates will be outside of the uncertainty intervals until the quantification of uncertainty is more comprehensive. This will require quantifying doubt in more of the steps of the LTIM estimation process.

Currently, LTIM uncertainty intervals only include quantifying doubt associated with adjustments, modelling, and survey based-estimates from the IPS. Two main factors could cause revised LTIM estimates to be outside of uncertainty intervals published with provisional LTIM estimates. First, uncertainty associated with steps in the LTIM process prior to applying any adjustments. Second, there has been substantial behavioural change in migration patterns, as we use historical data to inform the degree of uncertainty with adjustments.

The ONS has published two working series to outline our current progress with quantifying uncertainty in LTIM estimates. A paper was publishing outlining our simulation based approach and a paper published outlining how we produce a composite measure of uncertainty for headline LTIM estimates. The Office for Statistical Regulation have recognised the ONS progress and commitment to providing users with a clear and comprehensive understanding of uncertainty in LTIM estimates.

We are continuing to research suitable approaches to provide more comprehensive measures of uncertainty for LTIM estimates. As a case study, we tried using expert judgement and published our findings. Our results from the case study suggest that expert judgement has potential for quantifying uncertainty but would require some further steps and research.

Two approaches are currently being researched to assist in quantifying more of the uncertainty in the LTIM estimation process.

First, we are exploring the feasibility of using structural equation modelling (SEM), with our approach building on a previous ONS example of using SEM to quantifying error in floor space measurement. One reason for exploring SEM is that it offers an approach for error estimating that does not necessarily require a gold standard for comparison purposes.

Second, we are exploring the feasibility of adapting the method used for producing mid-year estimates (MYE) and administrative based population estimates uncertainty intervals (ABPE). Both methods are based on the assumption that uncertainty intervals can be produced through comparing estimates with a benchmark source, which serves as the measure of truth – ABPE and MYE use the census as the benchmark. We propose adapting this method by using multiple benchmarks for comparison, rather than using one benchmark.

Both approaches outlined will go through the necessary quality assurance prior to being included to produce more comprehensive measures of uncertainty for LTIM estimates.

7. Conclusion

This paper provides an overview of our current approach for measuring international migration using administrative data. We set out the methods for measuring EU, non-EU and British migration separately, and set out the challenges of continuing to produce timely estimates of migration under an actuals, rather than intentions-based approach. We conclude by highlighting the research ongoing (and that will eventually be shared with MARP) that aims to continue to build the quality of these estimates, particular for measuring British and EU national migration.

We welcome a discussion with the panel on our current approach and the research that we have in scope. We also look forward to hearing views on appropriate methods for considering migration where there a number of imperfect sources, and how best to bring them together to provide the highest possible quality estimates. We also would appreciate the panel's view on the assumption setting process, and whether there are alternatives that will support the production of timely migration estimates.