

Long-Term International Migration estimates methodology

In-depth look at methodology of both Long-Term International Migration (LTIM) and the International Passenger Survey (IPS), including assumptions made for variables, and changes over time.

Contact:
Nicola White
migstatsunit@ons.gov.uk
+44 (0)1329 444097

Release date:
24 May 2018

Next release:
To be announced

Table of contents

1. [Introduction](#)
2. [Current methodology](#)
3. [Assumptions made in order to produce LTIM](#)
4. [Provisional and final estimates of long-term international migration](#)
5. [Reliability of migration estimates](#)
6. [Sampling distribution and the normal distribution](#)
7. [Accuracy and confidence intervals](#)
8. [Sampling and non-sampling error](#)
9. [Measuring change – statistical significance](#)
10. [Summary and timeline of methodological changes](#)

1 . Introduction

1.1. Definition of a long-term migrant

Office for National Statistics (ONS) uses the United Nations (UN) recommended definition of a long-term international migrant:

“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.”

Page 18, [Recommendations on Statistics of International Migration \(PDF, 5.0MB\)](#)

This is the definition used to calculate net migration and is also used for the UK usually resident population estimate series. This definition does not necessarily coincide with those used by other organisations.

1.2. Issues with international migration

There is no single, all-inclusive system in place to measure all movements of people into and out of the UK. Therefore, it is necessary to use a combination of data from different sources, which have different characteristics and attributes, in order to produce estimates of international migration. These sources are described and discussed in [Comparing sources of international migration statistics](#).

Measuring illegal migrants presents another challenge. For more information on measuring illegal migration, please see the [International migration – terms, definitions and frequently asked questions](#).

2 . Current methodology

The methodology outlined in this section was first applied in 2009 for the calculation of the 2008 estimates and revisions were made to earlier years as appropriate.

As a consequence of these methodological improvements, the Long-Term International Migration (LTIM, formerly known as Total International Migration or TIM) back series was revised.

2.1 Data sources used to compile final estimates of Long-Term International Migration (LTIM)

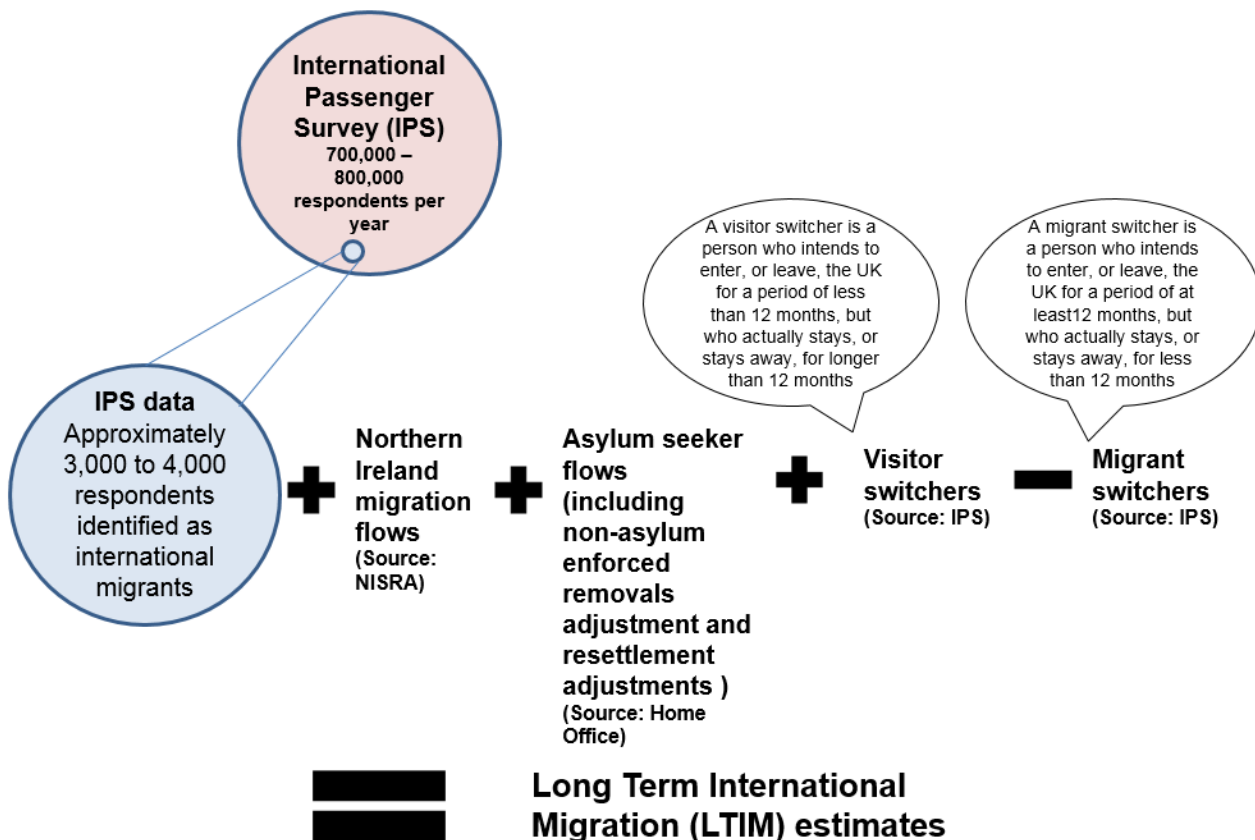
Estimates of LTIM are produced from these main data sources:

- International Passenger Survey (IPS)
- [Labour Force Survey \(LFS\)](#) – provides a geographical distribution of migrants for the calibration methodology
- Home Office data on asylum seeker flows and their dependants, on non-asylum enforced removals and on people resettled in the UK under various resettlement schemes
- forecasted LTIM estimates based on previous GP registrations from the Northern Ireland Statistics and Research Agency (NISRA) for estimating long-term international migration to and from Northern Ireland and the Rest of the World, from 2008 onwards; forecasted data are replaced with final data for LTIM final annual estimates

2.2 Components of Long-Term International Migration (LTIM)

LTIM comprises a number of components, which are described in Figure 1.

Figure 1: Estimating Long-Term International Migration



Source: Office for National Statistics

2.3 International Passenger Survey (IPS)

The [International Passenger Survey \(IPS\)](#) is a sample survey of passengers arriving at, and departing from, UK air and sea ports and the Channel Tunnel. It is carried out by Office for National Statistics (ONS) for a range of public and private sector organisations. In particular, it provides figures used for the travel account of the balance of payments, captures data on international tourism as well as providing data on the numbers and characteristics of short-term and long-term international migrants.

The long-term international migration data from the IPS is the largest component of the Long-Term International Migration (LTIM) estimates. It is important to note that these data are intentions-based, for example, the survey asks how long each migrant intends to remain in or out of the UK, as opposed to recording what they have done on their later journeys.

In 2009, adjustments were made to the methodology of the IPS, in terms of both sampling and data processing. This resulted in a sample design that is better optimised for collecting data on migrants.

A review into the [Quality of LTIM estimates \(PDF, 1.0MB\)](#) showed that the IPS missed migration flows, particularly of EU8 citizens, prior to improvements to the survey in 2009 that increased its coverage of regional airports. As a result, the net migration estimates for 2001 to 2011 were revised in light of the results of the 2011 Census, which showed that net migration was higher than implied by published migration estimates.

[International Passenger Survey: quality in relation to migration flows](#) provides further detail on the IPS, including its sample design; how the collected data are weighted to be representative of the total numbers travelling; information about the quality of the estimates; and details of the recent changes and their impacts. Further general information about the IPS can be obtained from the annual report [Travel trends – a report on the International Passenger Survey](#).

A copy of the IPS migrant trailer questionnaires [are available on the International migration methodology page](#).

The IPS has some limitations with respect to measuring immigration and emigration, as it:

- is a sample survey and so only a sample and not every migrant to or from the UK is interviewed; as a result, the estimates are subject to a degree of uncertainty
- does not capture all asylum seekers who may be entering or leaving the UK
- does not take into account the changing intentions of passengers (those who intended to remain in or out of the UK for 12 months, but actually spent less than a year and those who believed they would be staying or leaving for less than a year but actually spent longer)
- does not capture those who are crossing the land border between the UK (Northern Ireland) and the Republic of Ireland

2.4 Calibration of International Passenger Survey (IPS) data

The IPS asks long-term immigrants to state where they intend to move to within the UK. Our research, as part of the [Migration Statistics Improvement Programme \(MSIP\)](#), compared IPS data with the 2001 Census and the [Labour Force Survey \(LFS\)](#) (a sample survey of households living at private addresses in Great Britain). This revealed that there are some migrants who will live at their intended destination for only a short period of time before moving elsewhere. In particular, IPS data show a greater proportion of migrants stating London as their destination compared with either the LFS or census data. One explanation is that London is an international gateway to the UK and is therefore a transition point before they settle in other parts of the UK.

The geographical distribution of immigrants who were recorded entering the UK by the IPS can therefore be improved with the use of the LFS. This is because it asks respondents where they lived a year ago and this identifies recent migrants. The LFS can therefore provide more reliable data on where migrants actually live rather than on their intentions when they first arrive.

A methodology has been developed that adjusts the IPS data to the geographical distributions provided by the LFS (known as “calibration”) and is described in detail in the article [The use of calibration in estimating international in-migration to UK countries and the regions of England \(PDF, 43KB\)](#). The main steps are as follows:

- LFS data are used to identify the geographical distribution of recent immigrants (those that arrived in the UK within the last year) by UK constituent countries and regions
- these distributions are applied to IPS inflows to create a “control total” for each geographical area
- IPS data are calibrated to each control total

An IPS dataset is created, which has the same total inflows as the original, but the estimates by geographical area are consistent with LFS data on where recent migrants are living.

Calibration is applied to individual IPS contacts, potentially changing the weight of each contact so that regional proportions match those of the LFS. Calibration can therefore affect all IPS breakdowns (for example, citizenship) not only regional breakdowns.

Outflow data are not put through the process of calibration and remain unchanged. An assessment of the impact of these changes on the LTIM series can be seen in [Impact of revised methodologies on total international migration \(TIM\) estimates \(PDF, 77KB\)](#).

This improved methodology has been implemented back to 1999 because the scope of the original research only went back this far. Prior to this, the IPS alone was used to distribute migrants around the UK. Care therefore needs to be taken when examining detailed breakdowns of the IPS estimates before and after 1999, particularly when comparing regional or country estimates before and after this point.

2.5 Migrant and visitor switchers

One of the main features of the IPS estimates of long-term international migration is that they are based on passengers' intentions. The IPS classifies long-term international migrants as travellers who intend to change their country of residence for at least a year. This can be either overseas residents arriving to live in the UK, or UK residents leaving to live abroad.

Sometimes these intentions may not be realised. People who enter or leave the UK intending to be a visitor, that is, staying or being away for less than 12 months, may actually migrate for more than a year. These people are, in effect, visitors who subsequently become migrants and are referred to as “visitor switchers”. These migrants must therefore be added to the estimate of migration to make it comprehensive.

Alternatively, some people who enter or leave the UK intending to migrate (for 12 months or more), may actually stay in or leave for less than a year. These people are known as “migrant switchers” as they intended to be migrants, but were actually visitors. They need to be removed from IPS migrant flows. These adjustments improve the accuracy of the LTIM estimates.

These switchers are identified by the IPS as they complete their journey when subsequently entering or leaving the UK. The passenger is asked how long they intended to stay in the UK or overseas when they initially arrived or departed and for how long they actually remained in or out of the UK.

2.6 Visitor switcher methodology

In response to a need for more accurate estimates of visitor switchers, new IPS questions were introduced in 2004. These questions collect data on respondents who did not intend to stay in or leave the UK for longer than a year, but subsequently did. These data are then used to provide a more informed indication of how many visitors will change their intentions and become migrants. This is an improvement to the previous methodology, which estimated how many of the potential visitor switchers would become migrants, without the additional information from the IPS.

It is known that the likelihood of a visitor changing their intentions can vary depending on their citizenship and place of last or next residence. To take these differences into account, the visitor switchers are split into four groups before any calculations are carried out: those entering the UK who are EEA and non-EEA citizens, those leaving the UK who are EEA citizens going to the EU and all “other” citizens leaving the UK going to anywhere in the world. (The EEA refers to the European Economic Area, which is the EU plus Iceland, Liechtenstein and Norway.)

Figure 2: For each group the following calculation is made

$$\frac{\text{(respondents who did not intend to stay in or leave the UK for longer than a year, but subsequently did, over previous three years)}}{\text{(respondents who stated an intention to stay in their destination country for 6 to 12 months or possibly 12 months, over previous three years)}} \times \text{(respondents who stated an intention to stay in their destination country for 6 to 12 months or possibly 12 months, this year)}$$

Source: Office for National Statistics

For details of the proportions of components that make up LTIM estimates, please see [Table 1.01](#) for long-term international migrants.

2.7 Migrant switcher methodology

The new IPS questions introduced in 2004 also collect data that can help improve the estimation of the number of migrant switchers. As with visitor switchers, these questions gather information on a traveller’s completed journey, therefore allowing the estimate to be calculated using actual migrant switcher data, as opposed to just using data for those who originally intended to be migrants.

As with the calculation of visitor switchers, a fraction is produced that takes the number of migrant switchers (over the previous three years) and divides these by the number of migrants recorded by the IPS in the previous three years. This denominator is therefore the pool of travellers who could potentially become migrant switchers as they were initially recorded stating an intention to be migrants. It is produced separately for both immigration and emigration. Unlike visitor switchers, there is no distinction between citizenships or countries of last or next residence for migrant switcher calculations.

The number of migrant switchers is then removed from the estimate of Long-Term International Migration (LTIM) in the reference year as these people are not migrants. The proportion will change each year depending on the number of both migrants and migrant switchers captured by the IPS.

For details of the proportions of components that make up LTIM estimates, please see [Table 1.01](#) for long-term international migrants.

Due to the new IPS questions being introduced in 2004, a decision was made to apply the new methodology to the 2004 estimates onwards. Care therefore needs to be taken when comparing LTIM estimates before and after this year. The 2006 calendar year estimates were the first to use a full three years of data as required by the methodology as the new questions in the IPS were only introduced in 2004.

A comparison of how the fractions have changed using the previous and current methodologies are provided in Appendix C of [Impact of revised methodologies on total international migration \(TIM\) estimates \(PDF, 77KB\)](#).

2.8 Asylum seekers and non-asylum enforced removals

The IPS does not interview all asylum seekers entering or leaving the UK. In order to produce LTIM estimates, we obtain data from the Home Office (as they are responsible for immigration control and applications for settlement, citizenship and asylum) on principal applicant asylum seekers and their dependants. Details can be found on the [Home Office \(UK Visas and Immigration Agency\) website](#).

Data are provided for different types of asylum seekers. This includes the number of those who applied for asylum, were refused asylum, appealed against their asylum decision, asylum seekers who were returned home and those who withdrew. These different categories dictate whether the asylum seeker is leaving or entering the UK. An adjustment for asylum seekers returned, departing voluntarily, or withdrawing their application and leaving the UK within a year of the application, is made. This therefore excludes those who are not long-term international migrants from the LTIM estimates.

Asylum applications covered by the Home Office can be identified as either “port” or “in-country”. Port asylum seekers – the minority – are those who apply at port when entering the UK. Most port asylum seekers are not captured in the IPS because they are usually escorted over the IPS counting line. An allowance is made when estimating port asylum seekers for the small number of migrants in the IPS data who give “seeking asylum” as their reason for entry and will therefore be double-counted if kept in.

In-country asylum seekers are those who enter the UK and later apply for asylum while in the UK. It is assumed that information about planned duration of stay given to the IPS interviewer is the same as that given to the Immigration Officer and, therefore, that in-country asylum seekers are unlikely to be captured as migrants in the IPS.

In summary, asylum seeker immigration figures are based on the number of people applying for asylum. This data is used to adjust the IPS estimates in order to:

- exclude those asylum seekers counted by the IPS on arrival in the UK to remove the possibility of double-counting
- allow for the small numbers of those counted in both the principal applicant and dependant applications data
- exclude those who were returned within a year of their application

Asylum seeker emigration figures are based on:

- the number who were returned to their country of origin
- the number who withdrew their application and were known to have left the UK
- a small number of applicants who had been refused asylum in the previous year (and, if appropriate, had been unsuccessful at appeal) or who had withdrawn their application and were not known to have left the UK.

The Home Office also collects data on non-asylum enforced removals – these are people who have been removed from the UK and who have not claimed asylum at any stage. They would not be interviewed by the IPS upon leaving the UK as they would not cross the sampling line. No adjustment for these people is required for inflow estimates, as they were not asylum seekers and therefore would have crossed the IPS sampling line on entering the UK.

For 2013 data onwards, it is possible to identify long-term international migrants within the data on non-asylum enforced removals. Therefore for 2013 estimates onwards, an adjustment is made to include non-asylum enforced removals in asylum seekers' emigration estimates. The approximate impact of applying the adjustment is to increase emigration estimates by 2,000 to 3,000 per year and reduce net migration by around 1%.

For 2015 data onwards, adjustments using Home Office data on people resettled in the UK under four resettlement schemes have been included within LTIM estimates. This affects the inflow and, as a result, net balance estimates for LTIM.

2.9 Estimate of migration to and from Northern Ireland

The IPS does not sample those passengers who cross the land border between the UK (Northern Ireland) and the Republic of Ireland. In addition, no ports in Northern Ireland have historically been surveyed in the IPS, although this started at Belfast International Airport in 2009.

Family doctor registration data is the most complete source that can be used to estimate international immigration to Northern Ireland. This source gives information on an intention to stay for a period of time and covers all age groups.

The health card system records de-registrations with family doctors in Northern Ireland, while the Central Statistics Office (CSO) Ireland Quarterly National Household Survey provides the number of people moving from Northern Ireland to the Republic of Ireland. In combination, these sources are used to estimate emigration from Northern Ireland to all countries outside the UK.

These estimates are then incorporated into the LTIM estimates. A more detailed explanation of this methodology and the recent changes is available in [Improving estimates of international migration in Northern Ireland, and between the UK and the Republic of Ireland \(PDF, 57KB\)](#). Further information about international migration statistics for Northern Ireland is available at the [Northern Ireland Statistics and Research Agency \(NISRA\)](#).

It should be noted that since 2014, we use forecasted data provided by NISRA for quarterly LTIM estimates. This is to improve timeliness of statistical outputs. Forecasted data is replaced with final data for the annual final LTIM estimates.

3 . Assumptions made in order to produce LTIM

The published Long-Term International Migration (LTIM) figures are broken down to show estimates by variables such as citizenship and age and sex. To produce estimates for each of these variables, data from the sources that contribute to LTIM also need to be broken down by the same variables.

Migrant data from the International Passenger Survey (IPS) is available broken down by each variable. Data on Northern Ireland flows and asylum seeker data are not and need to be derived using a series of assumptions. In addition, the IPS data used to calculate the visitor switcher adjustments are based on a relatively small sample size each year, but still need to be broken down in the same way.

The following tables detail how the Northern Ireland flow data, asylum seekers and visitor switcher data are broken down for each variable. It is not necessary to do further processing to form assumptions for the migrant switcher data as it is applied as a direct proportion of the IPS migrant estimates.

Table 1: Assumptions made for citizenship

Source	Assumption made for citizenship
Northern Ireland immigration data	Uses IPS immigration data constrained to NISRA family doctor registration data by age, sex, country of last residence and reason for migration
Northern Ireland emigration data	Uses 75% of the immigration distribution from Northern Ireland data and 25% of the IPS emigration distribution, constrained to total Northern Ireland emigration
Asylum seeker, non-asylum enforced removals and resettlement data	None as Home Office provides data by citizenship
Visitor switcher data	A three-year average of the visitor switcher data by citizenship is used.

Source: Office for National Statistics

Table 2: Assumptions made for country of last or next residence

Source	Assumption made for country of last or next residence
Northern Ireland immigration data	Data come from family doctor registration
Northern Ireland emigration data	Uses 75% of the immigration distribution from Northern Ireland data and 25% of the IPS emigration distribution, constrained to total Northern Ireland emigration
Asylum seeker, non-asylum enforced removals and resettlement data	Assumed to be the same as citizenship
Visitor switcher data	A three-year average of the visitor switcher data by country of last or next residence is used

Source: Office for National Statistics

Table 3: Assumptions made for main reason for migration

Source	Assumption made for main reason for migration
Northern Ireland immigration data	Data come from family doctor registration
Northern Ireland emigration data	Uses 75% of the immigration distribution from Northern Ireland data and 25% of the IPS emigration distribution, constrained to total Northern Ireland emigration
Asylum seeker, non-asylum enforced removals and resettlement data	Is always included in the “other” reason for migration category
Visitor switcher data	A three-year average of the visitor switcher data by reason for migration is used

Source: Office for National Statistics

Table 4: Assumptions made for usual occupation (prior to migration)

Source	Assumption made for usual occupation (prior to migration)
Northern Ireland immigration data	Uses IPS immigration data constrained to NISRA family doctor registration data by age, sex, country of last residence and reason for migration
Northern Ireland emigration data	Uses 75% of the immigration distribution from Northern Ireland data and 25% of the IPS emigration distribution, constrained to total Northern Ireland emigration
Asylum seeker, non-asylum enforced removals, resettlement and visitor switcher data	20% of age 16 and over assumed Professional / Managerial 20% of age 16 and over assumed Manual / Clerical 20% of age 16 and over assumed Student 40% of age 16 and over assumed Other All aged 15 and under assumed Child

Source: Office for National Statistics

IPS data are adjusted using the LFS distributions to more reliably distribute immigrants throughout the UK, known as calibration. The following assumptions are made for people who are not covered by the IPS or whose intended length of stay changes:

Table 5: Assumptions made for origin or destination distribution within the UK

Source	Assumption made for origin or destination distribution within the UK
Northern Ireland flow data	Northern Ireland
Asylum seeker and non-asylum enforced removals data	Distribution calculated using asylum support data (a data-set containing the number of asylum seekers receiving support and their locations), which is supplied by the Home Office
Resettlement data	Syrian Vulnerable Persons Resettlement Scheme (SVPRS) refugees: Distribution calculated using SVPRS lead local authority data supplied by the Home Office Other scheme refugees: Distributed to reflect distribution of SVPRS refugees
Visitor switcher inflow data	Use the LFS distribution to assume the migrant's destination within the UK
Visitor switcher outflow data	Calculated by using the most recent three-year average of the IPS visitor data by origin

Source: Office for National Statistics

Table 6: Assumptions made for age and sex

Source	Assumption made for age and sex
Northern Ireland flow data	Data come from family doctor registration
Asylum seeker, non-asylum enforced removals and resettlement data	None as the Home Office provides the data by age and sex
Visitor switcher flow data	Calculated by using the most recent three-year average of the IPS visitor data by age and sex

Source: Office for National Statistics

Table 7: Assumptions made for sex and marital status

Source	Assumption made for sex and marital status
Northern Ireland immigration data	Data on sex come from family doctor registration. Data on marital status use IPS immigration data constrained to NISRA family doctor registration data by age, sex, country of last residence and reason for migration
Northern Ireland emigration data	Data on sex come from family doctor registration. Data on marital status use 75% of the immigration distribution from Northern Ireland data and 25% of the IPS emigration distribution, constrained to total Northern Ireland emigration
Asylum seeker, non-asylum enforced removals, resettlement and visitor switcher data	Estimated using a three-year average of the marital status distribution taken from the IPS. Carried out separately for males and females aged 15 and over

Source: Office for National Statistics

A distribution for intended length of stay from the IPS is used for those entering and leaving the UK for work or study.

Table 8: Assumptions made for intended length of stay

Source	Assumption made for intended length of stay
Northern Ireland immigration data	Uses IPS immigration data constrained to NISRA family doctor registration data by age, sex, country of last residence and reason for migration
Northern Ireland emigration data	Uses 75% of the immigration distribution from Northern Ireland data and 25% of the IPS emigration distribution, constrained to total Northern Ireland emigration
Asylum seeker, non-asylum enforced removals, resettlement (inflow only) and visitor switcher data	25% assumed intended length of one to two years 25% assumed intended length of three to four years 25% assumed intended length of more than four years 25% assumed unsure

Source: Office for National Statistics

4 . Provisional and final estimates of long-term international migration

Provisional figures allow for a timely comparison of recent migration patterns on a quarterly basis. However, these are subject to change as their calculation is based upon provisional data. The final Long-Term International Migration (LTIM) estimates are considered to provide a more reliable picture of migration and allow for annual comparisons over time.

4.1 Provisional LTIM (rolling year quarterly)

LTIM estimates offer the most comprehensive early indication of migration flows into and out of the UK. LTIM are about 90% based on data from the International Passenger Survey (IPS). In addition, they are adjusted to account for asylum seekers (including a non-asylum enforced removals adjustment and people resettled in the UK under four resettlement schemes), migration to and from Northern Ireland and people whose length of stay changes from their original intentions.

Releases of provisional LTIM data every quarter show overall estimates of immigration, emigration and net migration by citizenship and main reason for migration, and are available for the last 10 years.

4.2 Provisional IPS (rolling year quarterly)

These estimates are based solely on data collected by the IPS. They offer an early indication of changes in patterns of long-term international migration to and from the UK. The rolling year-based estimates are produced every quarter and provide up-to-date data on citizenship and main reason for migration and are available for the last 10 years.

4.3 Final LTIM calendar year (1-series and 2-series)

LTIM estimates provide the most comprehensive estimates of long-term international migration to and from the UK. The IPS provides the foundation of these estimates. Final LTIM estimates are published annually in November or December. Data are available from 1991 and include tables by calendar year, half year or mid-year.

1-series (methodology) contains tables showing the components and adjustments for LTIM and the confidence intervals and non-response associated with the IPS estimates. These are available from 1975 onwards.

2-series tables usually focus on one characteristic of migrants, from the following list:

- citizenship
- country of last or next residence
- country of birth
- age and sex
- sex and marital status
- usual occupation (prior to migration)
- main reason for migration
- origin or destination within the UK
- intended length of stay

4.4 Final IPS calendar year (3-series and 4-series)

These estimates are published annually in November or December. They provide detailed cross-tabulated data of all available variables. Since November 2012, the data within the 3-series tables have been available from 1975 onwards (superseding the 2-series tables previously published covering 1975 to 1990).

Characteristics of migrants that are in one or more of the tables are

- citizenship
- country of last or next residence
- country of birth
- age and sex
- sex and marital status
- usual occupation (prior to migration)
- main reason for migration
- previous main reason for migration
- origin or destination within the UK
- intended length of stay
- actual length of stay
- route

4-series tables are underlying datasets of final IPS data. These annual tables provide further breakdowns of migrant characteristics by individual country rather than country grouping. Underlying datasets are available from 2000 to the latest annual estimates (2016).

4.5 Comparison of provisional and final data

Estimates of LTIM are about 90% based on data from the IPS, supplemented by adjustments made for asylum seekers, non-asylum enforced removals, flows from Northern Ireland and visitor and migrant switchers. Therefore, the majority of the differences between provisional and final LTIM estimates can be accounted for by the replacement of provisional IPS data with final IPS data. Nonetheless, the differences between overall provisional and final estimates of long-term international migration from the IPS are minimal. Table 9 gives an example of the specific difference between provisional and final IPS data.

Table 9: Percentage differences between provisional and final 2016 IPS estimates, UK

Rolling year ending	Inflows (long-term migrants)			Outflows (long-term migrants)		
	Provisional IPS estimate	Final IPS estimate	Percentage (%) difference	Provisional IPS estimate	Final IPS estimate	Percentage (%) difference
March 2016	566,000	567,000	0.2	287,000	287,000	0
June 2016	581,000	582,000	0.2	293,000	293,000	0
September 2016	530,000	531,000	0.2	300,000	300,000	0
December 2016	525,000	527,000	0.4	315,000	316,000	0.3

Source: Office for National Statistics – International Passenger Survey

Once the final IPS data are incorporated in the LTIM estimates, final data for asylum seeker flows, non-asylum enforced removals and flows from Northern Ireland are incorporated to produce the final LTIM estimates. These adjustments are typically negligible (Table 10).

Table 10: Percentage differences between provisional LTIM estimates (including final IPS data) and final LTIM estimates, UK, 2016

Rolling year ending	Inflows (long-term migrants)			Outflows (long-term migrants)		
	Provisional LTIM estimate (including final IPS data)	Final LTIM estimate	Percentage (%) difference	Provisional LTIM estimate (including final IPS data)	Final LTIM estimate	Percentage (%) difference
March 2016	638,000	638,000	0	311,000	311,000	0
June 2016	653,000	652,000	-0.2	317,000	316,000	-0.3
September 2016	598,000	598,000	0	325,000	325,000	0
December 2016	589,000	589,000	0	340,000	340,000	0

Source: Office for National Statistics

5 . Reliability of migration estimates

Estimates of Long-Term International Migration (LTIM) are:

- the best estimates available at this time
- based on a consistent definition since 1991
- produced in accordance with the [Code of Practice for Statistics](#)

To see how these estimates of LTIM measure against the Code of Practice Pillar of Quality, please see the [Quality and Methodology Information Report for Long-Term International Migration](#).

The quality of international migration estimates improved following the introduction of fundamental changes to the International Passenger Survey (IPS). These changes included:

- redesigning the sample design at the beginning of 2009 to make the survey more focused on identifying migrants
- rebalancing the IPS interviewer resource away from Heathrow and towards other routes
- introducing a new IPS processing system early in 2009 to enable improvements to be made to the IPS weighting methodology

Further information on these changes can be found on the [Improving Migration and Population Statistics web pages](#). Further information is also available in [IPS: quality information in relation to migration flows](#).

We continue to keep abreast of available administrative sources and explore whether additional administrative data could be used to improve international migration and population estimates. Please see [International migration data and analysis: improving the evidence](#) for more details.

6 . Sampling distribution and the normal distribution

The different possible samples of passengers that could have been selected by the International Passenger Survey (IPS) can be used to produce a sampling distribution for the figure we are trying to estimate. For example, if we are estimating immigration within a particular year, one sample may produce an estimate of 500,000, another may have resulted in an estimate of 515,000 and another may have produced an estimate of 490,000. If we could take a lot of samples like this and plot the estimates from each sample, we would produce a chart of the sampling distribution of our estimate.

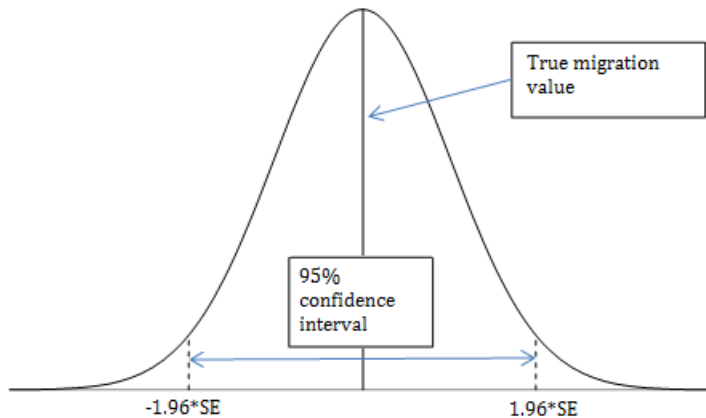
Assuming that the estimation method we use to produce the estimate for each sample is unbiased, the shape of the plot would follow the widely recognised normal distribution, where the most likely estimates of the true value are centered towards the middle and the least likely estimates are at the “tail ends”.

In practice, to estimate the true value for a specific population of, say, the number of immigrants, we take one sample and produce a single estimate. We assume that the sampling distribution of our estimate would follow approximately a normal distribution, centered on the true value and we can use a statistical formula to calculate the standard error around the estimate. This is a measure of the accuracy of the estimate.

At the 95% confidence level, which is a widely accepted level, we would expect the confidence interval to contain the true value 95 times out of 100. Equivalently, we can say that there would be a 1 in 20 chance that the true value would lie outside of the range of the 95% confidence interval.

As Figure 3 illustrates, 95% of the estimates would lie within 1.96 multiplied by the true standard deviation of the sampling distribution. This also works the other way round, so we can say that for 95% of random samples taken, our estimate will be no more than 1.96 multiplied by the standard error of that estimate away from the true value that we are trying to estimate. Using this knowledge, we can calculate a confidence interval around our estimate.

Figure 3: Normal distribution curve



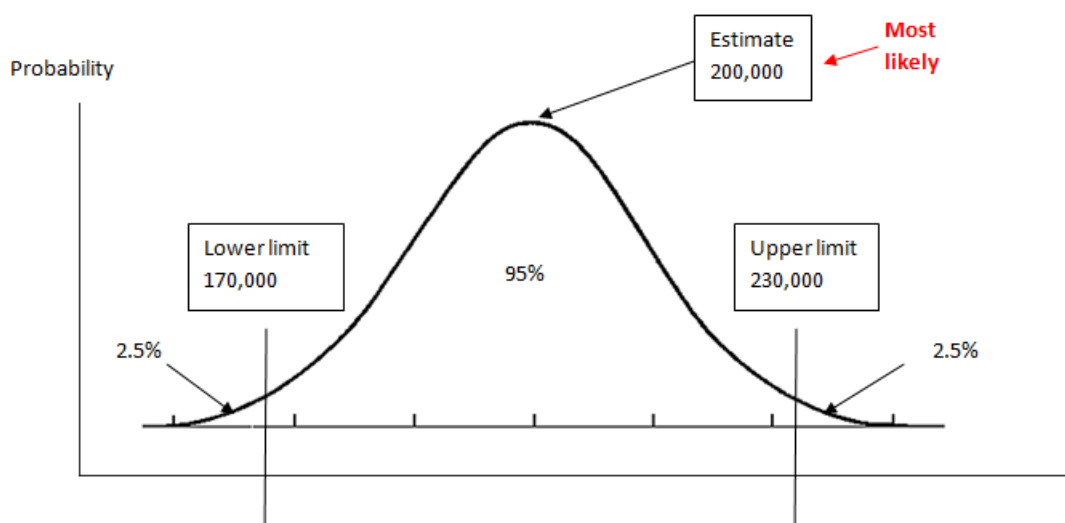
Source: Office for National Statistics

7 . Accuracy and confidence intervals

Confidence intervals are indicators of the extent to which the estimate may differ from the true value. The larger the confidence interval, the less precise the estimate is. The central value within the confidence interval is the best estimate of the true value. The confidence interval around the estimate captures the uncertainty of the estimate and gives an interval within which we can say that there is a high probability that the true value lies.

A confidence interval is not a range in which any figure is equally likely, but the most likely true figure is the estimate itself (Figure 4). However, when interpreting these confidence intervals and statistically significant changes, you should be aware that there is no method for quantifying the error associated with the non-survey components of Long-Term International Migration (LTIM), which are unlikely to be random.

Figure 4: Explaining uncertainty with confidence intervals



Source: Office for National Statistics

The published estimate 200,000 is the most likely figure in the range of +/- 30,000.

Users are advised to be cautious when making inferences from estimates with relatively large confidence intervals. Confidence intervals become larger (meaning there is more uncertainty) for more detailed estimates (such as citizenship by reason for migration). This is because the number of people in the sample who have these specific characteristics (for example, EU8 citizens arriving to study) is smaller than the number of people sampled in higher-level categories (such as the total number of EU citizens arriving to study). Where possible, it is better to use the highest level breakdown of data available. For immigration and emigration estimates where the lower confidence interval is below zero, users should assume the estimate is above zero.

8 . Sampling and non-sampling error

Details of the possible effects of sampling error on the migration estimates by various characteristics are given in [Table 1.02](#) of the Long-Term International Migration (LTIM) tables. Entries in this table show that estimates based on the sampling of passengers on certain routes have proportionately larger confidence intervals associated with them. Thus, generally speaking, the reliability of the estimate increases in proportion to the size of the estimate.

Sampling error arises due to the variability that occurs by chance because a sample, rather than an entire population, is surveyed; that is, sampling error results because not every migrant who enters or leaves the UK is interviewed. Sampling errors are determined both by the sample design and the sample size. Sampling error may sometimes present misleading changes as a result of the random selection of those included in the sample.

Non-sampling error is all error that is not sampling error. The challenge with non-sampling error is that it is difficult to directly calculate a numerical measure of its effect. This, therefore, makes it hard to incorporate when analysing results.

Non-sampling error is best understood by referring to examples that apply to the International Passenger Survey (IPS).

The first non-sampling error may be due to non-response. Bias will occur when passengers who do not respond to the survey have different characteristics to those who do respond.

A further source of bias may arise from contacts deliberately concealing their migration intentions from the interviewers. In addition, the question that determines whether the contact is a migrant or not and their length of stay, is based on intentions and not actual behaviour. Measurement errors could therefore be introduced if there is a discrepancy between those intending to migrate, but who subsequently stay less than a year and those not intending to migrate, but who stay for a year or more.

For those contacts identified by the IPS as migrants, the level of non-response is very low for most characteristics. Latest details of survey non-response can be found in [Table 1.03](#).

[International Passenger Survey: quality information in relation to migration flows](#) provides an overview of the quality and reliability of the IPS in relation to producing estimates of migration flows.

9 . Measuring change – statistical significance

Changes in the estimates from the International Passenger Survey (IPS) from one period to the next may occur simply by chance. In other words, the change may be due to which individuals were selected to answer the survey and may not represent any real-world change in migration.

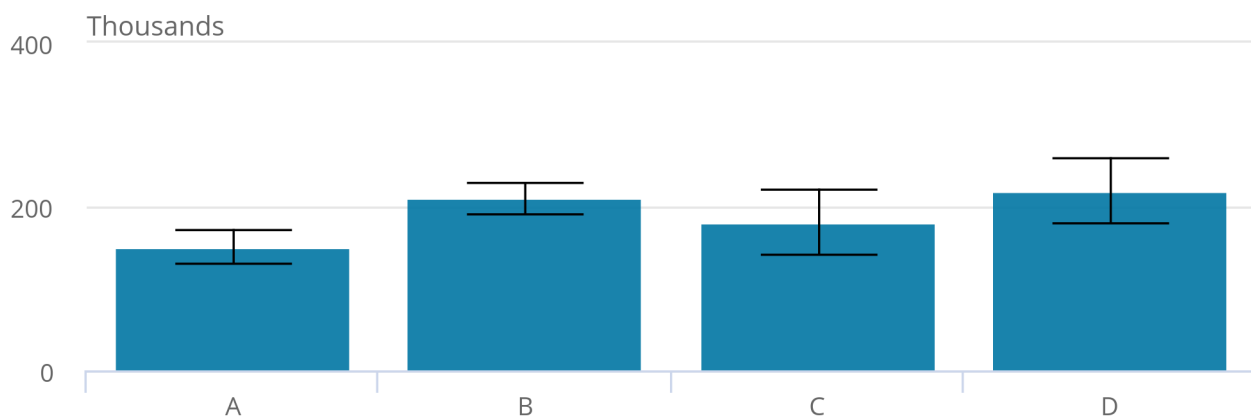
We are able to measure whether this is likely to be the case using standard statistical tests. These tests examine the difference between two estimates and calculate a confidence interval of the difference. The usual standard is to carry out these tests at the 5% level of "statistical significance". When we report on statistical significance, we provide an assessment of how likely it is that we would see results as unusual as these if the true value of the population remained unchanged. The phrase "statistically significant at the 5% level" indicates that, if the true value of the population remained unchanged, a result like this would occur less than 5% of the time.

When comparing two estimates, a t-test is performed, which results in the calculation of a 95% confidence interval for the difference between these estimates. If this interval excludes the value zero, then we can conclude that the difference is very likely to be a real difference in migration figures and not a result of sampling variation.

A quick method of identifying if the difference between two estimates is statistically significant is to determine if there is an overlap of their confidence intervals. If they do not overlap, then the differences can be described as statistically significant. However, if they do overlap, then a t-test should be performed to determine statistical significance. For example, there is a significant difference between datasets A and B, but may not be between C and D (Figure 5).

Figure 5: Is there a statistically significant difference?

Figure 5: Is there a statistically significant difference?



Source: Office for National Statistics

A t-test ascertains if the difference between two estimates is statistically significant, that is, if it were repeated with a different sample, the difference would occur 19 out of 20 times.

This test divides the difference of the estimates by the square root of the sum of the squared standard errors. The resulting t-value needs to be greater than 1.96 to be 95% certain that the estimates are different. It can also be used to create a confidence interval around the difference. It calculates the standard error of the difference directly from using the difference between the two individual standard errors.

All main statistical software packages have the functionality required to perform a t-test. If you need assistance with identifying whether the difference between two international migration estimates is statistically significant then please contact migstatsunit@ons.gov.uk.

10 . Summary and timeline of methodological changes

Over time, small adjustments have been made in the methodology to produce Long-Term International Migration (LTIM) estimates. These reflect improvements in the components or statistical techniques used to estimate the flow of international migrants. Table 11 summarises changes, and Table 12 identifies where discontinuities exist in the present time series of LTIM, from 1991 to the latest time period.

Table 11: Summary of methodological changes to components of final LTIM estimates, 1991 to 2016

Period of data on which change was implemented	Component of LTIM affected	Changes made	Impact of changes	Further information
2015	Asylum seekers	Including an adjustment for people resettling in the UK under various resettlement schemes.	Increases immigration by the precise number of Great Britain immigrants under the resettlement schemes.	
2013	Asylum seekers	Including an adjustment for non-asylum enforced removals within the asylum seeker adjustment.	Decreases emigration by the precise number of Great Britain non-enforced asylum removals.	
2009	International Passenger Survey (IPS)	Changes to sample design and data processing.	Survey more statistically robust, particularly for international migrants not entering the country through Heathrow.	International Passenger Survey Methodology
2008	IPS and Republic of Ireland component	IPS figures on the Republic of Ireland used in LTIM instead of data from the Central Statistics Office of Ireland.	Discontinuity between 2008 and earlier years, as revisions to the back series of final IPS data (used in final LTIM figures) have not been run. But it is thought that between 2001 and 2006, UK immigration would have increased by an average of 10,000 per year and UK emigration would have decreased by an average of 2,000 per year.	Improving estimates of international migration in Northern Ireland, and between the UK and Republic of Ireland
2008	IPS and Northern Irish component	Northern Irish data is calculated using data from family doctor registrations in Northern Ireland. Previously data from the IPS was used.		
2008 and back series to 1999	IPS – geographical distribution of immigrants	IPS inflow calibrated to Labour Force Survey (LFS) distribution of recent migrants.	Inflows – reduction of migrants into London and increase of migrants elsewhere. Calibration also introduced some secondary effects on other variables. Revised back to 1999 – small discontinuity with years prior to 1999.	Impact of revised methodologies on Total International Migration (TIM) estimates
2006 and back series to 2004	Visitor and migrant switchers	New IPS questions introduced to capture “actual migrants”. These data inform the visitor and migrant switcher calculations.	Applied to 2004 data onwards. Total inflows – negligible impact. Total outflows – reduction of 20,000 a year in 2004 and 2005 incorporated into revisions. Small discontinuity with years prior to 2004.	
2003 and back series to 1991	Visitor switchers	Previously estimates were based just on visa data supplied by the Home Office. Now calculated by Office for National Statistics (ONS) for the first time using fixed proportions.	Between 1992 and 2001, these changes led to a reduction in the net inflow of 97,000. Incorporated into 2003 revisions.	Impact of revised methodologies on Total International Migration (TIM) estimates

2003 to 1991	Migrant switchers	Assumed that 5% of those identified by the IPS as immigrants became migrant switchers. For outflows it was assumed that 1% of emigrants returned to the UK within a year, also became migrant switchers.	Same methodology used for years 1991 to 2003.	Impact of revised methodologies on Total International Migration (TIM) estimates
2002 backdated to 1991	Asylum seekers	Improvement in the estimation of the outflow of unsuccessful asylum seekers; and an allowance for the dependants of asylum seekers not captured by other data sources.	Negligible difference to net flow of asylum seekers.	
1999 backdated to 1992	Adjustments to IPS weightings	Improvements were made to the methods used to weight up the estimate of migrants from the IPS.	The net effect of these weighting changes in 1999 was to decrease the inflow by 4.2% and increase the outflow by 3.4%. The adjusted flows for each year remained within originally published error bands for the IPS estimates.	MN No.26 International Migration 1999
1992 to 1994, revised	Adjustments to Home Office data (visitor switchers and asylum seekers)	Revisions to the previously provided visitor switcher and asylum seeker data for 1992 to 1994.	These revisions reduced the total visitor switcher inflow by 26,000 and increased the total asylum seeker inflow by 21,000 over the three-year period. The net effect on LTIM estimates was small, as it only reduced the inflow by 5,000.	Impact of revised methodologies on Total International Migration (TIM) estimates

Source: Office for National Statistics

Table 12: Discontinuities in the component data of final LTIM estimates, 1991 to 2015

Period	International Passenger Survey (IPS)	Data on flows to and from Northern Ireland	Data on flows to and from the Republic of Ireland	Visitor switchers	Migrant switchers	Asylum seekers
--------	--------------------------------------	--	---	-------------------	-------------------	----------------

1991		Data from the IPS	Data from Central Statistics Office of Ireland.	Estimated using fixed proportions of European Economic Area (EEA) and non-EEA citizens to EU and Other countries.	Estimated that 5% of inflow of IPS migrants and 1% of outflow become Migrant switchers each year.	Calculated by ONS using Home Office Data.			
1992									
1993									
1994									
1995									
1996									
1997									
1998									
1999	IPS inflow calibrated to LFS distribution of recent migrants.								
2000									
2001									
2002									
2003									
2004							Visitor switcher calculations based on information from new IPS questions on actual length of stay.	IPS data used.	Migrant switcher calculations based on information from new IPS questions on actual length of stay.
2005									
2006									
2007	IPS inflow calibrated to LFS distribution of recent migrants. Changes to sample design and data processing.	Data supplied by NISRA using family doctor registrations.							
2008									
2009									
2010									
2011									
2012	Adjustment for non-asylum enforced removals included.								
2013									

2015		Adjustment for people resettling in the UK under various resettlement schemes.
------	--	--

Source: Office for National Statistics

Unfortunately it is not always possible to apply changes in methodology to the entire back series. This is because of problems regarding data availability in earlier years. This has resulted in unavoidable discontinuities in the time series. Overall these discontinuities are small and it is important not to confuse the size of the revisions with the real underlying trends in LTIM.