

# Estimating British National long-term international migration using DWP Registration and Population Interaction Database (RAPID) data for the period YE Mar 2014 to YE Mar 2023

## EXECUTIVE SUMMARY

### Background

This paper focuses on a new method for estimating long-term international migration of British nationals using data from the Department for Work and Pensions' (DWP) Registration and Population Interaction Database (RAPID). The current methodology relies on the International Passenger Survey (IPS), which measures intentions rather than actual activity. Recognizing the limitations of IPS, particularly in underestimating immigration and emigration, the Office for National Statistics (ONS) is transitioning to an administrative data-based system.

### Introduction

RAPID contains data on National Insurance Numbers (NINo) and interactions with tax and benefits systems since 2008. Immigration and emigration is counted by assigning a residency flag to an individual record. A person is resident or not. If they were active in the data and now are not and do not have a death registration we infer they have emigrated. If they were not resident and now are and haven't just been born we infer they have returned from abroad.

### Quality Considerations

1. The data used in RAPID supplied by HMRC and DWP is collected for their business need. ONS's use of RAPID is stretching the data to its limits and putting it to a use that it was not originally collected for. That will inevitably introduce limitations for ONS purposes and we make a clear distinction between the quality of the raw data and ONS's use of that data for population and migration purposes.
2. Data Completeness: Some groups, such as children may not be fully captured in the RAPID dataset.
3. Timeliness and Lag: There is often a lag between data collection and its availability for analysis, affecting the timeliness of the data.
4. Estimation of Activity: Assumptions are made about activity continuity.

### Methodology

Several rules adapted from DWP's systems are used to flag changes in residency status:

- Single Tax Year Absence: Assumes continuity of UK residence if activity is found in adjacent tax years.
- Pension Rule: Assumes residency for those claiming state pensions, unless indicated otherwise.
- Address Change Rule: Uses address changes as indicators of activity.
- Child Benefit Rules: Assumes residency for children and their parents based on child benefit records.
- Student Rule: Identifies residency based on student loan repayments.
- Penultimate Year Rule: Adjusts for activity evidence in the penultimate year of the latest RAPID data to mitigate exclusion errors in the final year of available data.

### Findings

Comparison with IPS data indicates that RAPID-based estimates generally show higher immigration and emigration counts but more closely aligned net migration estimates. The discrepancies during the pandemic years highlight potential limitations of IPS and the greater coverage of RAPID in capturing actual migration.

### Conclusion

While RAPID has some limitations, for its use in the estimation of long-term international migration of British nationals, its coverage and accuracy present significant advantages over IPS. The report seeks feedback from the Methodological Assurance and Review Panel (MARF) on the standard of methods developed and their suitability for publication as official statistics in development.

### Questions for MARF

1. Does the panel think we have provided sufficient evidence that RAPID is more suitable data?
2. Does the panel feel there is sufficient evidence to support using the proposed method for the next migration statistics in development publication?

3. What feedback can you offer on future methods research?

## BACKGROUND

Legally there are 6 different types of British nationality<sup>1</sup>. These are:

- [British citizenship](#)
- [British overseas territories citizen](#)
- [British overseas citizen](#)
- [British subject](#)
- [British national \(overseas\)](#)
- [British protected person](#)

Our research into British nationals is ongoing. Currently the ONS uses International Passenger Survey (IPS) data. The IPS conducts face-to-face interviews with a sample of passengers as they pass through major UK airports, sea routes and the Channel Tunnel. More information on the survey methodology can be found in the ONS [IPS Quality and Methodology Information report](#). It is important to note that these data are intentions-based, and our research has shown that intentions do not always reflect actual behaviour. Further details can be found in our article [Estimating UK International Migration: 2012 to 2021](#)<sup>2</sup>. The IPS arrivals survey data collection stopped in the field from 1st July 2024 and once estimates from prior periods have come through the system, we will be changing to an alternative approach.

ONS have long acknowledged that the IPS has been stretched beyond its intended purpose. In 2019 we announced that we were moving away from the IPS and moving towards an administrative data-based migration statistics system. The latter measuring actual activity, while the former measures intentions, which can change. While we have made progress for measuring migration from EU and non-EU nationals, migration patterns for British nationals continue to use the IPS as the primary source. We have been investigating how we can use admin data sources instead, but this is more complex for British nationals as there are fewer admin data sources available to measure this group. We [revised our IPS-based migration estimates](#) in November 2023 following new information from the 2021 Census suggesting that the IPS was underestimating emigration of British nationals. This, coupled with our planned [change in collection method for the IPS](#), which includes no longer having an IPS interview for arrivals from June 2024 has accelerated our plans to use an admin-based method for British national migration.

## INTRODUCTION

The Registration and Population Interactions Database (RAPID) contains data for every National Insurance Number (NINo) and the interactions with the tax and benefits system since 2008. We then assume each NINo relates to a unique individual. We consider this a potential population stock dataset from which we can use activity starting or ending to infer a migration flow into and out of the UK. When we see evidence of these interactions stopping for a sustained period we could assume that this indicates an emigration. These data are based on an individual's recorded activity rather than their intention which is the case for the IPS.

### Quality Considerations in the RAPID database

1. The data used in RAPID supplied by HMRC and DWP is collected for their business need. ONS's use of RAPID is stretching the data to its limits and putting it to a use that it was not originally collected for. That will inevitably introduce limitations for ONS purposes and we make a clear distinction between the quality of the raw data and ONS's use of that data for population and migration purposes.
2. Data Completeness: Certain groups, such as children who arrive as migrants, may not be fully captured in the RAPID dataset because they do not appear in the Migrant Worker Scan (MWS) dataset unless they apply for a National Insurance number (NINo).

---

<sup>1</sup> <https://www.gov.uk/types-of-british-nationality>

<sup>2</sup> <https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/internationalmigration/articles/estimatingukinternationalmigration2012to2021/2023-11-23>

3. **Timeliness and Lag:** Data in the RAPID system often have a lag due to the time required for processing and validation. Migrants may not immediately register for services upon arrival, leading to delays in their appearance in the data<sup>3</sup>. The system is continually updated to incorporate new data sources and improve accuracy, but there can still be a significant delay from the time of data collection to its availability for analysis.
4. **Estimation of Activity:** The methodology used to estimate the activity and interactions of individuals in the RAPID database involves assumptions for example, the data might not show continuous activity, but assumptions are made about the total activity being sufficient to indicate long-term presence in the UK<sup>4</sup>

### Data Quality Initiatives

1. **Cross-Government Data Sharing:** DWP Digital is working to improve data sharing across government departments to enhance the accuracy and timeliness of the data captured in the RAPID system. This includes efforts to integrate data from various administrative sources and enhance the overall data quality<sup>5</sup>.
2. **Methodological Adjustments:** The Office for National Statistics (ONS) and DWP are continuously developing methodologies to better estimate international migration and address coverage gaps in the RAPID data. This involves making broad initial adjustments and refining these adjustments as new data sources become available<sup>6</sup>.
3. **Complexity and Categorization:** The RAPID system uses multiple categories to account for the complexity of individuals' lives and interactions. These categories help in distinguishing between long-term and short-term migrants based on their activity profiles within the administrative data<sup>7</sup>

Overall, while the RAPID database is a valuable tool for capturing interactions and estimating migration, it is being used for a project beyond its original intention which need to be carefully managed through ongoing methodological improvements and cross-departmental collaboration.

In November 2023, we updated users on how we plan to use RAPID for measuring British national migration patterns<sup>8</sup>. We acknowledge that the rules we use for EU-based migration using RAPID, as shown in our [long-term international migration user guide](#), may not be suitable for British nationals, because of the number of reasons a person may be inactive on this source, without actually having migrated. We have been working to develop some additional rules to account for some of these reasons, with many of these being based on rules, which were previously introduced in the [Department for Work and Pension's Longitudinal Lifetime Labour Market Database \(L2\)](#).

### **METHOD**

The rules we use apply a flag to a person's record, it does not remove or add people. These flag rules flag someone who is not resident to resident. Our overarching assumption is people are resident until we have reasonable grounds to assume they are not. This represents a potential bias within the rules. This is an area for future research. A person may receive more than one flag, but we only count that the person has a flag not the number of flags.

The flag rules we have adapted from DWP and the flag rules we are developing

1. Single tax year absence (Gap year) – adapted from DWP,
2. Pension rule (occupational and state) - adapted from DWP,

---

<sup>3</sup><https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/internationalmigration/articles/internationalmigrationdevelopingourapproachforproducingadminbasedmigrationestimates/2021-04-16>

<sup>4</sup><https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/internationalmigration/methodologies/methodsformeasuringinternationalmigrationusingrapidadministratedata>

<sup>5</sup><https://dwpdigital.blog.gov.uk/2022/02/17/dwp-digital-is-improving-cross-government-data-sharing/>

<sup>6</sup><https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/internationalmigration/articles/internationalmigrationdevelopingourapproachforproducingadminbasedmigrationestimates/2021-04-16>

<sup>7</sup><https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/internationalmigration/methodologies/methodsformeasuringinternationalmigrationusingrapidadministratedata>

<sup>8</sup><https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/internationalmigration/articles/improvinginternationalmigrationstatisticsusingadministratedata/2023-11-23>

3. Address change rule - developed by ONS for this research,
4. Child benefit (school age rule) grace period - adapted from DWP,
5. Child benefit for parent's grace period - adapted from DWP,
6. Child parent residency – developed by ONS for this research,
7. Students - developed by ONS for this research,
8. Penultimate year rule - adapted from DWP.

1. Single tax year absence (Gap year) - adapted from DWP,

The gap year rule looks for activity in both tax years either side of the tax year where there is no activity. If activity is found, in both tax years either side of the gap year i.e. year with no activity then it marks the person as "resident". For example, if a person has activity in year 1, no activity in year 2, and activity in year 3, with no abroad address, they would be considered resident in year 2. This is useful for classifying those, particularly of younger ages, who may take a year out between education and entering employment, or they may be in higher education and not in employment. For example:

- a. Students who leave education, fail to secure work and who do not claim JSA
- b. Students who only work part-year (holiday periods) or do no work at all in a particular year
- c. People who are unemployed but who do not claim JSA

However, there will still be a small number of people who don't have a change of address, but they are still absent, and we intend to develop a during the month indicator to more accurately count the time periods rather than just using the annual beginning and end dates.

2. Pension rule (occupational and state) - adapted from DWP

Owing to the eligibility requirements<sup>9</sup> for claiming a state pension, we can assume that someone who does not have any activity until they claim their state pension has remained resident in the UK for the duration of their inactivity. This rule changes a person's UK residency status to "yes" for those aged 50+ before claiming their state pension, therefore removing them from any migration estimates.

For British nationals living in a number of countries (excluding EEA countries and countries with a social security agreement), state pension uprating is not applied. A flag is applied in RAPID, which we can use to exclude some people from the pension age rule, as they have indicated to DWP that they are resident in another country and have migrated.

3. Address change rule - developed by ONS for this research

This rule considers anyone with no activity across benefits or earnings, but who has changed their address and their latest address is UK, as activity within the tax year. This is used as a flag to support the other rules and will be subject to further research.

4. Child benefit (school age flag rule) grace period - adapted from DWP

16-19 year olds are assumed to be resident if there was child benefit in payment for them at age 15. This is, in part, because many 16-19 year olds are in full time education and would not necessarily interact with national insurance, tax or benefit systems.

5. Child benefit for parents - adapted from DWP

Once the youngest child of a Child Benefit (CHB) recipient reaches the age of 15, they are considered to be resident even if no current activity is shown. This is because [longitudinal data](#) has shown that some people who have been in receipt of child benefit have remained inactive from date of benefit ceasing to the date where they begin claiming a state pension. If they subsequently interact with any of the systems within RAPID this grace period ceases to be applied. Data indicates the vast majority are female and likely caring for family.

6. Child parent residency – developed by ONS for this research

For those children who reach adulthood and appeared on child benefit, they may become inactive, and therefore potentially classified as non-resident if they do not have activity on RAPID. This rule looks at their linked parent through child benefit and assumes that their residence status is the same as that of their parents.

7. Students - developed by ONS for this research

---

<sup>9</sup> <https://www.gov.uk/new-state-pension#:~:text=You'll%20need%2010%20qualifying,or%20a%20parent%20or%20carer>

We identify people who have evidence of repaying a student loan. We use this information to look back and see if they were non-resident due to lack of activity, during the ages most commonly associated with students (16 – 22). We set them as resident as they were likely students. This assumes students have a student loan and we only capture people who have since met the repayment threshold. This could produce an under coverage.

#### 8. Penultimate year flag rule – adapted from DWP

Some of the rules that are applied throughout the time series cannot be applied in the latest year, for example the 'gap' year rule in the latest tax year because there is no tax year to follow. As a consequence, some people are incorrectly excluded in the penultimate year because there is no data in the latest year, or because the 'gap' year rule cannot be applied. In order to mitigate this, the penultimate year rule is applied to adjust the resident population in the latest year of the data. The rule works for those with evidence of activity in March of the penultimate year. If this criterion is met, then individuals with no activity in the final year are classed as resident. For example, our latest year of data available in RAPID is 2022/23, if someone has evidence of activity in 2021/22 but none in 2022/23, they are kept resident in 2022/23, this has the effect of reducing the number of emigration events in 2021/22.

## DISCUSSION

### Comparison with the previous move from IPS to RAPID for EU nationals.

Research from when ONS switched from IPS to RAPID to measure EU estimates for international migration found similar results though of a different magnitude. RAPID based estimates saw considerably more immigration and emigration<sup>10</sup>. This comparison is useful because it is a similar piece of work to this current research into switching from IPS to RAPID based international migration estimation.

Figure 1 – RAPID data show consistently higher estimation of long-term EU arrivals compared with LTIM data.

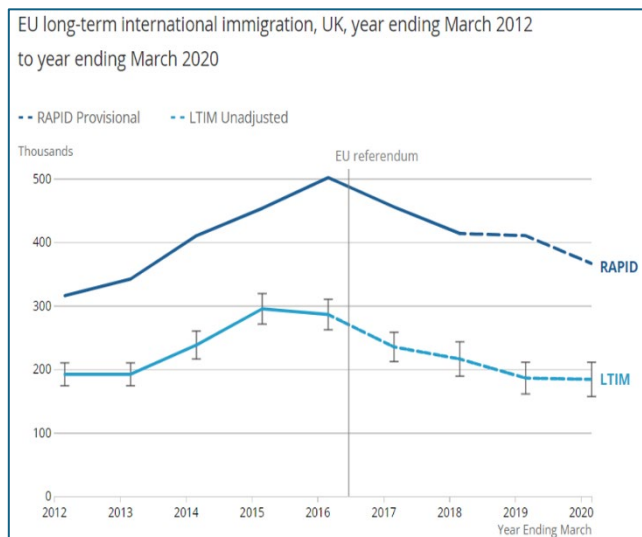
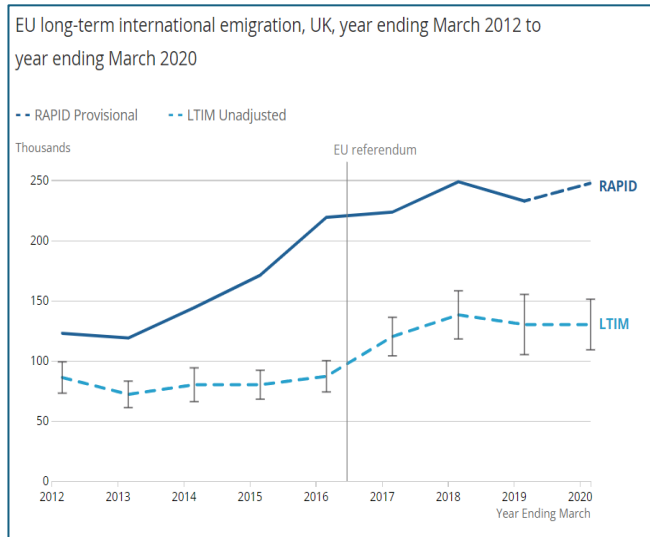


Figure 2 – RAPID estimates a higher number of long-term EU departures compared with LTIM data.



Following arrival to the UK, some people changed their intentions from visiting to staying or from short-term stay to long-term stay. IPS wouldn't capture this but RAPID would. Figs 1 & 2 show the differences between RAPID and IPS. This shows that for EU migrants the differences fell across all years meaning RAPID did not overestimating a particular year only.

### Dependency between the rules

We believe there is no dependency between the flag rules because the rules are run independently of each other. None of the flag rules build on a previous flag rule. The flag rules are only applying a flag not removing or adding people. A person may receive more than one flag, but we only count that the person has a flag not the number of flags.

### Timeliness

<sup>10</sup><https://www.ons.gov.uk/peoplepopulationandcommunity/> see section 4.

IPS has a quarter lag between the reference period and when ONS are able to process the data and report on it. However, it needs to be remembered that IPS reports on intentions to stay not actual stay. The IPS provisional estimate of intentions to stay is finalised in August the following year. RAPID has a potential lag of three quarters for provisional estimates because we receive the data once a year. The finalised estimate could be up to one year and 6 months from the reference period but it would be an actual finalised figure rather than an intentions figure as with IPS.

### Uncertainty measures

Provision of uncertainty measures for international migration are crucial for informing users about the quality of the estimates. ONS have been progressing with the provision of measures of uncertainty for administrative based migration estimates<sup>11</sup>.

The current research proposal is to quantify uncertainty based on the inflow adjustment and outflow adjustment. This is calculated by varying the assumptions behind how the rules are applied to generate a range of possible values that may have been used if we had chosen different assumptions thereby generating a range of values to give a possible indication of error. The unquantified uncertainty is planned for research later in the year. This will be an indication of the range of possible values based on variation caused the data itself such as duplicates and missingness.

**Inflow adjustment** - The aim of this adjustment is to account for the time needed in administrative data to assess whether an arrival is long-term. As RAPID is based on actual observed patterns of behaviour, it takes at least 12 months of activity to have taken place before a long-term measure of duration can be identified. For this current research we have applied an adjustment for arrivals in the most recent tax year of approximately 80%. This is based on the proportions of people who have arrived and remained long term in previous year. We will conduct further research to understand the accuracy of this adjustment.

**Outflow adjustment** - For this adjustment, for each year between the tax years ending 2012 and 2018 the proportion of long-term stock that go on to outflow is calculated using the long-term departures figure. Using the historical proportions calculated we calculate the predicted proportion for the final year.

## **FUTURE RESEARCH**

### Dual Nationalities

RAPID does not currently adjust for dual nationalities. There is an assumption that someone would use their British documents if they were a dual national on arrival to the UK due to the implications of recourse to public funds. However, someone who comes to the UK and subsequently gains British Citizenship will be recorded as their initial nationality in RAPID. Our current focus is on the common travel area between the UK and the Republic of Ireland as approximately 40% of Northern Ireland residents have Irish nationality. We are working with the Irish statistical offices on this issue.

### Prisoners

There were [46,519](#) sentenced admissions to prison in 2023. Approximately half (53%) of sentenced admissions during 2023 were for sentences of 12 months or more. People in prison are typically economically inactive and our rules will flag them as emigrated. We need a method to add prisoners of over a year sentence into our resident stock as they won't appear active in RAPID. Also prisoners on release could be classed as immigrants when they reappear as active in RAPID. Currently we can't account for this and further research and collaboration with the Ministry of Justice is needed to match the demographics accurately.

### Under coverage reduction

RAPID 8 due for release later in 2024 is expected to have extra data to plug the under coverage gaps reported above in the method section. This will reduce the reliance on many of the flag rules e.g. for students and activity for children missed. However because the data changes year on year, there is a risk to sustainability.

---

<sup>11</sup> our two working series papers:

[Measuring uncertainty in international migration estimates - Office for National Statistics \(ons.gov.uk\)](#)  
[Quantifying uncertainty in headline international migration estimates - Office for National Statistics \(ons.gov.uk\)](#)



### Self-employed

RAPID doesn't have complete self-employment data pre-2014. It only becomes reliable from 2014/15 tax year. DWP have created an adjustment which we used but need to research further to properly understand.

### Pandemic tax year YE March 2021

We are planning a case study for pandemic tax year YE March 2021. This is because of the irregular activity outside of the usual seasonality and trends in the data due to lockdowns. Also the pandemic introduced furlough and other economic payments and activity which previously didn't exist and we need to explore the impact of this on our flag rules in more detail.

### Naturalisation

A naturalisation adjustment needs to be made for people swapping their nationality after arrival and subsequently not been counted as UK stock in RAPID because RAPID keeps the previous nationality. When ONS initially published estimates from early research using RAPID we made an adjustment to non-EU outflow, to account for the fact a proportion would have become British Citizens. This was an adjustment to the flow and not to the stock. We essentially looked at outflow by time resident in the UK prior to outflow. For those who had lived in the UK over 10 years prior to outflow we assumed that a percentage would now be British Citizens so didn't count them in the estimates. Future research could use the HO data to show how many non-EU and EU nationals are becoming British Citizens and we can develop an adjustment for naturalisation<sup>12</sup>. This could overlap with dual nationalities.

### Deterministic vs Probabilistic

Our method is a deterministic method whereas a probabilistic method may be more appropriate. For example, rather than a flag of 1 we could estimate a probability as each of the rules has an assumption. If we could quantify the impact of that assumption, we could assign a probability e.g. the probability that someone who changed their address is therefore resident is not 1 because of lags.

## **CONCLUSION**

While RAPID has certain issues in measuring UK nationals' immigration and emigration, it offers advantages over the IPS in terms of coverage and accuracy. Here are the key points:

### Key Advantages

1. Actual Activity Measurement: RAPID captures real interactions with the tax and benefits systems, providing a potentially more accurate reflection of migration behaviours compared to the intentions based IPS. It is not a sample, so there is no sampling error.
2. Wide demographic coverage: RAPID includes data from multiple sources of interactions with DWP, offering a detailed view of individual activities over time.
3. Methodological Flexibility: The ability to apply various rules and flags allows for more nuanced categorization of migrants.

### Issues Overcome

1. Underestimation of Emigration: By moving away from the IPS, which underestimated emigration, especially of British nationals, RAPID offers the potential for more accurate estimates.
2. Incorporation of Administrative Data: Transitioning to administrative data has addressed the limitations of the IPS in capturing actual behaviors rather than intentions.
3. Improved Timeliness: Despite inherent lags, RAPID's ongoing updates and integration of new data sources help provide more timely and accurate data.

### Limitations and Future Work

1. Data Gaps: Certain groups, such as children and inactive individuals, may still be underrepresented in RAPID. Future versions (e.g., RAPID 8) aim to plug these gaps. However there is a risk that changes to the data may mean we have to continually review each rule and adjustment, risking stability and sustainability.

2. Lag and Timeliness: While RAPID is more timely than IPS, there are still delays. Future research will focus on reducing these lags through better data integration and processing.

---

<sup>12</sup> <https://www.gov.uk/government/statistics/immigration-system-statistics-year-ending-march-2024/how-many-people-are-granted-settlement-or-citizenship>

3. Complexity in Rule Application: The various flags and rules used to estimate residency status introduce complexity. Continuous methodological improvements could simplify these processes.

4. Dual Nationalities and Naturalization: Current methods do not fully account for dual nationalities or individuals who gain British citizenship. Future research will refine these adjustments in collaboration with relevant authorities.

In conclusion, while RAPID has its limitations for its use in the estimation of long-term international migration of British nationals, the system's wide coverage and methodological flexibility provide a viable framework for estimating migration. Ongoing research and development will continue to refine these methods, ensuring that future statistics are accurate, timely, and reflective of actual migration patterns. The feedback from MARP will be crucial in guiding these improvements and achieving the standards required for official statistics.