A report on the statistics being collected under the exit checks programme

August 2016

Executive summary

This report provides the first introduction to the data collected on travellers departing and arriving in the UK as part of the Home Office exit checks programme with a particular focus on the requirements for statistical reporting rather than the operational use of the data.

The data collated by the Home Office will in due course provide a more comprehensive picture than has ever been available on those entering and leaving the United Kingdom. However, there are also a number of challenges if that data is to be used to produce statistical estimates. Those challenges include issues of coverage (i.e. whether systems are in place to collect data), completeness (i.e. here relating to data provided for particular flights) and the extent to which data on travel movements collected by airlines and other carriers can be matched to immigration system records.

A system of this kind, covering millions of travel movements and receiving data from hundreds of carriers, will inevitably have some gaps, and there are a number of ways in which a traveller's departure may legitimately not be recorded by the system. These include instances where outward travel is by the Common Travel Area, for example, or if a traveller unfortunately dies in the UK. There are also reasons why a person's record may not be matched (for example, if they are dual nationals and do not use a single document for travel or if different data systems have captured their name differently).

The comprehensive collection of exit check data only commenced on 8 April 2015 and there has not yet been time to record complete data on travel movements for longer-term migrants. This fact, together with the experimental nature of the procedures for analysis, and the above-noted needs to assess the coverage and completeness along with the processes for matching data, mean that any statistics will need to be interpreted with considerable care and any raw estimates on the numbers of visa-holders whose departure has not been recorded could be misleading. The matching of exit checks data with information provided by passengers prior to arrival in the UK is also an additional challenge, and at present the development of the analytical programme has focused only on those travellers from outside the European Economic Area (EEA) travellers who require a visa, although the source Semaphore system collects information on all travellers and every individual who crosses the border will be subject to security checks.

Notwithstanding these challenges, the collection of exit checks data will have the potential for monitoring the flows into and out of the UK for those people subject to immigration control, and in due course all travellers, as well as a wide range of other potential operational uses. The information being collected in relation to individual passengers is already being used for a range of operational enforcement purposes. Results from this statistical assessment of the data are promising with high levels of coverage and completeness now achieved.

The key findings in this report are:

- Estimated coverage and completeness of travel movements, using data supplied by carriers and port operators, have been increasing, with 100% coverage of outbound travel for in scope routes and very high completeness rates at voyage level for air travel. The data is already being used for operational purposes but there is more work to be done to generate regular, reliable and comprehensive statistics.

- Estimates of coverage for routes in scope had increased from 83% at the end of March 2015 to 100% at the end of June 2015 for outbound travel, with a step change between estimates for the quarters ending 31 March and 30 June 2015, with this step change

increase related to the formal introduction of exit checks on 8 April 2015 and actions taken then or shortly after by carriers and port operators to provide the information.

- By the end of June 2016 coverage for data provided by carriers had also reached 87% for inbound travel. For inbound travel the data provided by the carriers and port operators is further supplemented by the additional data captured as part of the regular immigration checks at the border which is used to mitigate the shortfall, in particular the juxtaposed controls that apply to rail travel and Calais and Dunkirk ferries.

- There are some routes which were not in scope for the introduction of exit checks, most notably travel within the Common Travel Area (CTA) to and from Ireland. Survey estimates for the International Passenger Survey set this in a statistical context and indicate that the proportion of non-EEA travellers whose departure would not be counted in this way is likely to be low.

- Coverage identifies the journeys which are in scope for the exit checks data collection and is not a measure of the completeness of data submitted by carriers. Estimates of voyage level completeness for outbound air travel have now become very high (averaging 98.7% for weeks in June 2016) though it should be understood that this relates to proportion of voyages where some data has been supplied and not completeness of data for each traveller.

- Systems for matching individual travel movements to data held on immigration or visa systems have been developed that generated low proportions of false matches in test training sets. Overall for the period April to December 2015 the percentage of travel movements that could be matched to immigration records for 'visa nationals'[1] were close to 90% inbound and outbound. Some degree of attrition in matching is to be expected, for example where individuals issued visas did not travel and individuals hold dual nationalities or replaced lost or expired passports.

- This level of matching can provide a good degree of operational confidence in checks made on cases of unrecorded departures for particular individuals or groups of travellers, but this can still leave considerable residual statistical uncertainty as there is a residue of unmatched cases which could result in identification of significant numbers of false-positive unrecorded departures.

---

[1] See the Glossary for definition of 'visa nationals' and 'non visa nationals'.

Introduction

Exit checks background

Before 1994, embarkation controls were operated by Immigration Officers at the majority of air and sea ports. Whilst this meant a visible presence in most locations, these controls were limited to only a subset of people crossing the border. The paper-based nature of the checks meant that there was only limited ability to match records and therefore no usable information about outbound passengers was published. In 1998 embarkation checks were abolished entirely.

From 2004, the UK moved towards a more sophisticated approach to checks by starting to collect Advance Passenger Information (API) for both inbound and outbound air passengers. There was, however, no process to match departures and arrivals, and no coverage of rail and maritime routes at this time. In 2010 the government subsequently altered the approach being taken by what was then known as the e-Borders programme, and committed to 'reintroduce exit checks'.

Although partly in place prior to that date as systems were being developed, exit checks were formally introduced on 8 April 2015, since when near-comprehensive data on departures from the UK has been flowing from ports and carriers into Home Office systems. Exit checks apply to over 100 million travellers a year, by air, rail and ferry, with a small number of routes 'out of scope' such as departures via the CTA.

Matching of data on entry, exit, length of permitted stay, type of visa is generating for Home Office decision-makers an 'Initial Status Analysis' relating to those individuals who have arrived since 8 April. At present the initial focus is on visa nationals. Non-visa nationals present greater challenges, in terms of volumes and the lack of detailed visa data in relation to visitors, but the lessons learned from matching data on visa nationals will be a significant benefit.

Exit checks data at aggregate and individual level offers detailed information for policy-makers and operations, in the Home Office and across government. The Home Office has already begun to use this information to identify non-compliant individuals who may have overstayed their visa. Over time, aggregate data will also provide information about specific groups with a higher propensity to overstay, informing policy and operational action to address this.

All this requires significant and continuing work on data improvement and assurance. Although legislation was introduced, through the Immigration Act 2014, to give carriers and port operators the powers to carry out embarkation checks on exit, and to give the government the power to compel them to do so if necessary, the Home Office has worked closely with the industry to deliver exit checks without the use of regulation. Since 8 April 2015, as a result of this collaboration, exit checks on outbound travel have covered all in-scope scheduled commercial international air, sea and rail routes.

Arrangements are in place with carriers to process API or Travel Document Information (TDI) for 100% of the outbound travel for modes of transport that are in scope. The government has been processing inbound travel data from carriers for some time and as of the end of June 2016 data was supplied on 87% of inbound routes, with 100% coverage for air routes and cruise ships. This data is further supplemented by additional data captured as part of regular immigration checks at the border.

Passenger data is submitted in advance of travel for most scheduled aviation journeys (as Advance Passenger Information or API) but only collected at the point of departure for other

modes of transport (as Travel Document Information or TDI).[2] For statistical purposes API and TDI provide essentially the same information on travel movements apart from TDI not being submitted in advance of travel. Both API and TDI are used to provide the information for exit checks on people leaving the UK and for the remainder of this report references made to API will also include TDI unless otherwise indicated.

Information provided by exit checks

Exit checks works by combining data on persons entering the UK, and the permission that they have to stay in the UK, with data on persons leaving the UK. Exit check data is collected from both API and TDI covering all in scope routes.

A detailed database, developed for monitoring movements of non-EEA nationals, termed the Initial Status Analysis (ISA) database, combines not just the API and TDI information required for exit checks, but also other sources of information including the main data on visas and other forms of permission granted for leave to remain in the UK such as extensions of leave or asylum claims.

The ISA is designed to allow data to be used operationally to identify individuals who may have overstayed their permission to stay in the UK, either on return to the UK or if in the UK illegally. The data also assists in the exercise of new powers under the Immigration Act 2014, which restrict the provision of various services to any person who does not have permission to reside in the UK. In addition the data is designed to give the government the ability to monitor the relative risks of non-compliance for particular immigration routes and visas.

While predominantly an immigration and data tool, the data can also improve security by helping the police and security services track the movements of known or suspected criminals and terrorists, supporting the wider work across government and law enforcement agencies.

In assessing the statistical use of exit checks data, it is important to be clear on what ISA and source Semaphore system are designed to do. Some have suggested that this information could potentially be adapted to provide direct measures of net migration. This is the estimate of long-term migration that adheres to the UN standard for the measurement of migration and currently provides the best estimate of the annual change in the normally resident population due to migration.

The potential for data from border systems to be used to measure migration has been considered in recent years by Home Office statisticians in collaboration with statisticians at the Office for National Statistics. The conclusions of the statistical community were reflected in the UK Statistics Authority's response[3] to the Public Administration Select Committee report on Migration Statistics. The Authority concluded that the information collected on travel movements cannot provide a direct statistical measurement of migration flows and therefore cannot replace the estimates for migration provided by the International Passenger Survey (IPS). The principal reason for this is that it is not possible to determine the very small fraction of long-term migrants among the very large number of travel movement per year (e.g. with 123 million admissions in 2015 but with estimated net migration for 2015 of 333,000 – being fairly close to the average number of admissions each day). It is also considerably slower to determine long-term movements through actual travel patterns than by asking individuals their intentions at the point of arrival (as the passenger survey does), because it is necessary to wait for a year after arrival

---

[2] Definitions of these terms are set out in the Glossary at the end of this report together with other technical terminology that has been used.

[3] This can be downloaded at https://www.statisticsauthority.gov.uk/correspondence/migration-statistics-10/

so the individual's stay can be recorded as long-term. It would be necessary to wait for at least a year after arrival so the individual's stay can be recorded as long-term. There may be some benefits from using actual travel movements in terms of detail but ISA will only hold the data relevant to the individual's travel or be linked to data held on other official records and datasets. This data would be available for legitimate visa holders but would not be available for UK and EEA nationals who are not currently subject to immigration control. Whilst real travel movements could potentially provide some migration data, this is dependent on the degree to which issues of coverage, completeness and data matching can be addressed.

Nevertheless, the ISA does have good potential for monitoring the flows into and out of the UK for those people subject to immigration control, as well as a wide range of other potential operational uses.

Visa and non-visa nationals

In this report reference will often be made to 'visa nationals' (i.e. nationals of countries, such as China, India or Russia, for which a visa is normally required to enter the UK) and 'non-visa nationals' (i.e. nationals of countries such as the United States, Australia or Japan, who are currently allowed to enter the UK without a visa as visitors for a period of up to six months and would only be required to obtain a visa[4] for longer-term activities, such as to work or study).

As set out below, the ability to assess matching procedures differs between these two groups as a lack of match to a visa for a non-visa national could indicate either a failure to match or that they had entered the UK as a visitor. The focus of this report is on visa nationals, but there is further consideration being given to how the Home Office can in due course assess the immigration compliance of non-visa national visitors.

Some 15.3 million non-EEA nationals were admitted to the UK in 2015, of whom 5.2[5] million were visa nationals. This means that even very small gaps in completeness of outbound data would have the potential to generate appreciable numbers of false positives, implying that a traveller has not departed the UK when that may not actually be true, a point that will be discussed further in the section that follows on completeness.

These large numbers of travel records also indicate the significant technical challenges in bringing together the various data sources as well as the complexity of obtaining information from large numbers of air, rail and maritime carriers.

The Initial Status Analysis (ISA) database

This report reviews only the data held on the ISA database covering non-EEA nationals, rather than the broader collection of information on travel movements, including EEA and UK nationals, which is available for operational and security purposes on what is currently called the Semaphore system. The ISA database has been designed as far as possible to create a simplified person-centric view of the key immigration and travel events for individuals within the immigration system.

The term *identity* is used to refer to a collection of records that relate to an individual. Each identity is allocated a unique identifier and consists of biographic details (i.e. name, passport, date of birth, nationality and gender) and links each identifier to associated events, such as a particular journey. These events when presented in a chronological order present the travel

---

[4] Referred to technically as 'entry clearance' for such nationals.
[5] Includes nationals of countries where visitors can obtain an electronic visa waiver in place of a visit visa.

history. This history is used in the assessment of an individual's immigration status at various stages of the immigration journey by comparing with records held on other systems relating to an individual, such as the visa they have been issued.

Through the provision of API data by carriers the Home Office is able to monitor people departing the UK. These checks allow operational assessment as to whether a person appears to have complied with the terms of their permission to stay in the UK; for example whether they left the country on or before the date of expiry of their visa and which may inform decisions on subsequent applications to re-enter. It also enables identification of those individuals who no longer have permission to stay in the UK but who have not been recorded as departing. The data also support national security by helping the police and security services track the movements of known or suspected criminals and terrorists.

The attributes of the data in terms of coverage, completeness and coherence (mainly in terms of data matching) have to be considered when assessing the confidence that can be placed on any statistical estimates that could be produced. The main aspects of data quality, and how it has changed, are considered in the remainder of this report and focus on the statistical uses, rather than operational functions, of the data.

Coverage

Coverage is a measure of the proportion of routes which are covered by the system for collecting data on border crossings; all carriers on a specified route are required to provide data for their flights, voyages or rail journeys. A calculation is then made of the estimated number of passengers who would have travelled on these routes. This is not the same as *completeness* of the API data provided in relation to flights by the carriers or the number of passengers for whom we have received API data (this is discussed in the next section, a route may be covered but data may not always be submitted by carriers).

For the purposes of the exit checks system both inbound and outbound API provide information on travel to and from the UK in order to build up information held in the ISA. Additionally for inbound travel movements Passport Scans and other information are collected from saved data when travel documents are checked, this mitigating the gaps in inbound API coverage.

Coverage can be estimated as a percentage of routes or as a percentage of volumes of travellers on routes. As some routes can have much higher volumes than others it is most meaningful to consider coverage weighted by numbers travelling. Using this measure it was estimated that at the end of 2009 the overall coverage of API/TDI collected by the Home Office for both inbound and outbound travel combined stood at 57%. Since then, the coverage has increased and by mid 2014 it had reached 81% and has continued to increase since then.

Figures are available from mid 2014 of coverage at the level of individual modes of travel (aviation, maritime, rail). There has already been some reporting of coverage rates in the National Audit Office (NAO) report *E-borders and successor programmes*[6] that states that by September 2015 rates were 100% outbound and 86% inbound but that the Home Office was not then receiving inbound data for the majority of ferry and rail passengers. Table 1 shows both how coverage has increased over time, reaching 100% for outbound routes by the end of June 2015 and 87% for all inbound routes by the end of June 2016, including 100% for arrivals by air and cruise liner passengers.

---

[6] This can be downloaded at https://www.nao.org.uk/report/home-office-e-borders-and-successor-programmes/

For outbound routes there was a step change between the quarters ending 31 March and 30 June 2015, with coverage rising from 83% to 100%. This resulted from the formal introduction of exit checks on 8 April 2015 and actions taken then or shortly after by carriers to provide the required information, including not only flights but also for maritime and rail. While the Immigration Act 2014 did create a statutory provision for the Secretary of State to direct carriers or operators of ports to examine embarking passengers, these powers have in the event not had to be used and exit checks data is instead collected and provided by carriers on a voluntary basis.

Table 1 indicates that progress in achieving coverage for API has progressed at a different pace across different modes of transport and between inbound and outbound travel. For *inbound* travel, coverage of ferry travel is currently partial (estimated at 32% coverage at the end of June 2016). Inbound coverage has not been introduced yet for rail (both Eurostar and Eurotunnel). It is important to note that although the API data provided by the carriers is not complete for all modes of transport, this does not affect the security checks that are carried out on all passengers on arrival at the UK border.

The prioritisation of outbound travel movements reflects the intention to use exit data to identify visa holders who have not complied with their conditions of entry by leaving the UK on time. As noted previously, for passengers leaving the UK there is now complete coverage in the provision of API data for scheduled international commercial aviation (other than on routes with the CTA).

The gaps in coverage prior to the full introduction of exit checks in April 2015 mean that analysis of historic data from the period before that date will be significantly incomplete and there is an appreciable chance that a traveller without a departure record may have in fact left the UK, for example by a route for which there was no coverage or incomplete coverage at the time of their departure. After the 8 April 2015, that is much less likely to be the case.

Although outbound coverage of in-scope routes has now reached 100%, there remain routes out of scope as discussed in the Public Accounts Select Committee's report *e-Borders and successor programmes*[7]. These include:

- Routes within the Common Travel Area (CTA) that includes routes between the UK and Ireland, and between the UK and the Crown Dependencies of the Channel Islands and the Isle of Man.

- General Aviation, that is non-scheduled flights (e.g. private airplanes) and General Maritime, that is non-commercial maritime traffic (e.g. private boats).

- Organised coach parties of school groups where students are aged 16 or under are also excluded.

---

[7] This can be downloaded at
http://www.publications.parliament.uk/pa/cm201516/cmselect/cmpubacc/643/64302.htm

Table 1: Percentage coverage[1], weighted by estimated traveller numbers, for aviation, maritime and rail journeys

Inbound travel[3]

	30-Jun-14	30-Sep-14	31-Dec-14	31-Jan-15	31-Mar-15	30-Jun-15	30-Sep-15	31-Dec-15	31-Mar-16	30-Jun-16
Aviation	96	96	96	97	97	97	100	100	100	100
Aviation EU	94	94	94	95	95	95	100	100	100	100
Aviation non-EU	100	100	100	100	100	100	100	100	100	100
Maritime	21	21	22	20	20	29	36	36	30[2]	32
Maritime ferry	16	13	13	13	13	16	23	23	22	22
Maritime cruise	53	65	71	59	59	94	100	100	100	100
Rail	0	0	0	0	0	0	0	0	0	0
All modes	81	81	81	82	82	83	86	86	87	87

Outbound travel[3]

	30-Jun-14	30-Sep-14	31-Dec-14	31-Jan-15	31-Mar-15	30-Jun-15	30-Sep-15	31-Dec-15	31-Mar-16	30-Jun-16
Aviation	96	96	96	97	98	100	100	100	100	100
Aviation EU	94	94	94	95	96	100	100	100	100	100
Aviation non-EU	100	100	100	100	100	100	100	100	100	100
Maritime	22	22	23	21	23	100	100	100	100	100
Maritime ferry	17	14	14	14	17	100	100	100	100	100
Maritime cruise	50	65	71	59	59	100	100	100	100	100
Rail	0	0	0	0	0	100	100	100	100	100
All modes	81	81	81	82	83	100	100	100	100	100

Notes:
1 Rounded to the nearest percentage point.
2 The fall in this percentage largely reflects a fall in the proportion of the maritime travellers arriving via cruises.
3 Coverage reflects only those routes in-scope for the programme introduction on 8th April 2015.

In understanding the implication of the gap in coverage for the CTA it is important to recognise that UK's unique and long-standing relationship with the Republic of Ireland and the Crown Dependencies. Although API is not routinely collected on travel in, out and through the CTA, there is strong governmental and operational co-operation over the sharing of information regarding movements across CTA borders including both for border control and security purposes. Continuing collaboration between the two countries is likely to further align the border and visa policy and processes in order to strengthen the external CTA border. A central focus of this work is on the development of electronic border management systems, including the collection and processing of API. The Republic of Ireland passed legislation at the end of March 2016 which allows the UK to ask carriers to provide API on UK-Republic of Ireland journeys where the carrier collects it. Sharing passenger data will further support the exit checks programme going forward.

Most flights from Ireland depart from Dublin, especially long haul routes that could be attractive for onward travel for non-EEA nationals who were previously in the UK. In total there were over 12.4 million passenger departures from Dublin in 2015[8], of which 4.5 million were to other airports in the CTA, 6.2 million to the rest of the EEA and 1.7 million to non-EEA. New York is the busiest non-EEA route from Dublin with close to 316,000 travellers, followed by Dubai at 202,000, Chicago at 171,000, Boston at 167,000, Abu Dhabi at 155,000, Toronto at 135,000 and Istanbul at 83,000. Other significant non-EEA routes tend to be to holiday resorts, and there are also flights via Gulf hub airports to destinations such as Australia.

---

[8] Information is derived from the website
http://www.cso.ie/px/pxeirestat/Statire/SelectVarVal/Define.asp?maintable=ctm01.

The International Passenger Survey estimates for 2015 that 354,000 non-EU nationals who were also non-EU residents travelled from Great Britain to the Republic of Ireland by air or sea[9]. Of these non-EU residents, about half of the total number were US citizens, a further quarter were citizens of Canada and Australia, and only very small numbers were visa nationals. The IPS figures do not distinguish between travellers to Ireland who are on return visits to Ireland from the UK, and those who are travelling one-way to Ireland, subsequently departing the CTA from Ireland. The latter are likely to constitute only a small proportion.

For General Aviation (small private planes), the current absence of a mandatory reporting requirement, and the disparate nature of the sector, mean that there are no definitive figures for the number of passengers who exit the UK via General Aviation. There are about 90,000 flights per annum, with on average 3 or 4 people per flight, and most of these will be UK or EEA citizens. The General Aviation flights represent a very small proportion of overall aviation numbers.

For General Maritime, passenger and crew data is not mandated from commercial general maritime or pleasure craft. Some data is submitted on a voluntarily basis that may be used for exit checks purposes.

Completeness of data

While coverage of exit checks data has now increased to 100% for the in scope outbound routes, carriers may not in all instances supply the data pertaining to a flight or an individual. Certain estimates of completeness of data are available for aviation routes and relate to the extent to carriers supply API data for voyages, though it should be noted that this does not mean that all data for every passenger was received. These completeness estimates are based on API information having been received for inbound and outbound flights.

These estimates for completeness stood at 90% overall in March 2014, rising to 94% in November 2014. These statistics can fluctuate but are now relatively stable at a high level. The 2015 NAO[6] report reported that completeness had reached 98% by September 2015. The latest position is that for the whole period September 2015 to June 2016 the average inbound completeness was 98.8% and for outbound it was 98.1%, and for the average for the weeks in June 2016 this was 98.8% and 98.7% respectively.

In February 2016 a new system was introduced to enable comparison of received data against airline schedules. For the first time, this provides the capability to identify completed commercial aviation flights for which no passenger data was received. The likelihood of this occurring is believed to be rare because non-provision of both passenger and crew data would require two independent systems to fail. Development of additional metrics is in progress to identify data quality issues at a more in-depth level for carrier resolution to further improve data quality.

Aviation accounts for around 83% of overall inbound and outbound travel movements. Rail travel (Eurostar and Eurotunnel) accounts for 8% and maritime travel for the remaining 9%. Completeness of maritime and rail data cannot be measured in this way, because only one message per passenger/service is received (i.e. no separate check in and departure confirmation). Measurement of completeness for these modes of transport is still in its early stages but the system is being developed to allow monitoring of receipt of data within expected volumes and time windows.

---

[9] These estimates do not include crossings of the land border with Northern Ireland. The total and the breakdown between nationalities were based on 368 survey respondents and so individual estimates are subject to considerable statistical uncertainty.

Although the available API completeness rates at flight level for aviation are high it should be recognised that they relate to whether some API data has been received for a flight. The impact of a failure to receive passenger details could have important impacts on any statistical estimates. However, if one assumed a non-completeness rate as low as 2%, this could still potentially generate a high number of false positives where no record of departure had been found (bearing in mind the 15.3 million non-EEA admissions overall in 2015, of which 5.2 million were visa nationals). Although some background level of unaccounted departures is inevitable, this needs to be borne in mind when interpreting figures.

Coherence: The processing of API data and matching to other sources

API data has been used for a number of years to inform operational assessments. Over time processes have been implemented to improve the quality of the data, such as addressing duplicate records and creating as far as possible a person-centric overview. In preparation for the generation of ISA data in April 2015, a number of further improvements were made. Additional de-duplication of data ensures as far as possible a person-centric view is produced, using links within immigration system data as well as matching processes for API data.

Processes for data linking

The ISA system combines immigration and travel events to an individual through a process of data linking which includes external linking based upon immigration system data and also data matching to create as far as possible a single person-centric dataset. Whilst linking various immigration systems data combines events that have common indexes, additional data matching combines events that have common biographic details. The data matching uses deterministic, probabilistic and event based matching methods. It is relatively easy to appreciate how each individual linking method operates in isolation; however, the combined effect of different methods working together is more complex. While, it is difficult to quantify, it is reasonable to assume that the quality of the linking process is directly related to the number of biographic details associated with an identity.

In order to test the accuracy of the data matching processes being applied in the ISA system, a sample consisting of pairs of individuals for each dataset combination was created. The resultant sample, described as the test training set, consisted of close to 40,000 candidate pairs. An overriding goal for the data linking process is to avoid false positives (i.e. combining events relating to two different individuals). Through visual inspection, each biographic pair in the training set was assigned a match status according to the similarity scores of the relevant match variables, that is name, passport held and date of birth. The resultant dataset was analysed using logistic regression to determine which match variables were statistically significant. The process was then repeated for each dataset pair combination, producing a result indicating whether these were accurate matches. On this basis, few false attributions were found to have been generated.

Achieved matching rates to API data

In considering matching rates of API data to records of visas and other leave held on Immigration systems it is important to distinguish between visa nationals, who are normally required to have a visa to enter the UK and non-visa nationals who are currently allowed to enter the UK without a visa as visitors for a period of up to six months but not for other purposes such as work or study.

For visa nationals, during the period April to December 2015 the match rate was 89% for inbound journeys and 90% for outbound. In other words, for around 9 out of every 10 visa nationals there was a record of a visa being issued and being used to enter the UK and there was a corresponding and linkable travel event to or from the UK.

As one would expect for the non-visa nationals (many of whom would be visitors without visas where no match to visa records will be found) these rates are much lower, with respective overall rates for inbound and outbound of 18% and 20% respectively. This lower matching rate is very much to be expected as most non-visa nationals arriving as visitors will not have a record on visa systems, although in due course it should be feasible to match their departure with their inbound record using only API data for the majority of travellers.

In addition, non-matching may legitimately occur because

- a passport has been renewed following expiry or loss whilst its owner was in the UK;
- some travellers may use a different travel document, for example when initially booking their voyage or because they hold dual nationality;
- a traveller may have indefinite leave to remain in the UK captured in an old passport which is presented on arrival, but use a more recent document on departure;
- a traveller may hold a residency card issued under EU law by an EEA member state together with other documentation that allowed visa free entry;
- others may be exempt from visa requirements, for example diplomats or merchant seamen.

There could also be inconsistencies between the data entered in booking a voyage as a result of error (e.g. on passport number or date of birth) that result in no match.

As a result, matching rates for non-visa nationals would currently have limited meaning. They do, however, indicate an important attribute of the system that many travellers only have a records based on their travel data. The focus of initial work for statistical purposes has been on visa nationals, but further consideration is being given to how the Home Office can in future begin to assess compliance of non-visa national visitors using inbound and exit data.

Table 2: Percentage of individuals' API data matched to same individual in visa systems data for the period April to December 2015 (excluding transit cases) for top ten visa nationalities and for all other visa nationalities

Nationality	Percentage of inbound API matched to CID/CRS		Nationality	Percentage of outbound API matched to CID/CRS	
	Number matched	% of all matched		Number matched	% of all matched
China	650,300	86%	China	682,100	85%
India	607,600	94%	India	635,600	95%
Nigeria	205,700	88%	Nigeria	234,400	88%
Russia	210,200	92%	Russia	220,400	93%
South Africa	150,600	87%	South Africa	156,900	90%
Turkey	146,900	94%	Saudi Arabia	169,700	97%
Saudi Arabia	148,700	95%	Turkey	156,100	95%
Pakistan	137,800	93%	Pakistan	141,500	94%
Kuwait	97,200	95%	Kuwait	103,200	98%
Thailand	81,800	95%	Thailand	87,700	96%
Other nationalities	1,032,200	86%	Other nationalities	1,121,100	87%
All visa nationals	**3,469,200**	**89%**	**All visa nationals**	**3,708,700**	**90%**

Transit cases where a person arrives and departs the UK within 48 hours are excluded.

Further work is needed to clarify why the differences in matching rates occur between different nationalities. In some cases it may in part be due to linguistic complexity (such as for Chinese travellers), differences in the degree to which travellers use on-line application systems (electronic data being easier to match), consistency in presenting biographical information in different databases or other reasons.

In addition, the system also records Passport Scan (PS) data from swipes of passports. This data can be used to supplement inbound API data and provides additional information for matching to immigration systems. At the same point at which PS data is collected there would normally be collection of biometric data used to match to the Immigration and Asylum Biometrics System (IABS). This data is combined with the PS data to also assist in making a match with API data. Both PS and IABS are used to provide more complete coverage for inbound data.

When matching API to PS and IABS, data rates are a little lower for non-visa nationals than for visa nationals (81% compared with 86%). This indicates the majority of API can be matched to PS and IABS data, but that these sources will not fully replace API where it does not exist for coverage or completeness reasons.

The matching rates are lower when looked at from the perspective of matching PS and IABS to API, being 71% and 70 % for non-visa and visa nationals respectively. However, it should not come as a surprise that these rates are lower as API coverage rates inbound were 86% at the end of 2015. Rates were also artificially low during the transitional period as exit checks were formally implemented on 8 April 2015, which will have affected this measure over the course of 2015.

As this system further develops there will be further opportunity to look at outcomes for longer term cohorts. Notwithstanding any limitations on coverage for outbound travel prior to 8 April 2015, patterns of non-compliance will in due course be available. There will also be further work to consider how best to monitor compliance of non-visa nationals.

Glossary

Advance Passenger Information (API)

This refers to travel document and service information submitted by carriers in advance of departure for passengers and crew. Data is often self-declared by passengers to carriers at time of booking.

Common Travel Area (CTA)

The Common Travel Area (CTA) comprises the UK, Ireland and the Crown Dependencies of the Channel Islands and the Isle of Man. A person who has been examined for the purpose of immigration control at the point at which he or she entered the area does not normally require leave to enter any other part of it. However certain persons subject to the immigration control who enter the United Kingdom through the Republic of Ireland do require leave to enter.

Completeness

In this report, completeness of data refers to the completeness in the provision of the supply of data agreed and delivered by carriers in relation to voyages, that is individual flights or journeys. Figures for voyage level completeness are currently available for aviation and relate to the percentage of voyages where at least one passenger or crew message was received (that is, the flight took place). This measurement does not mean that all data for every passenger was received.

Coverage

Coverage in this context is a measure of the degree to which those routes that are in scope have systems in place that allow submission of data to the Semaphore system and hence to the Immigration Status Analysis system. It should not be confused with completeness (see 'completeness' above).

Entry Clearance

A document issued to a **non-visa national** that is required to be obtained in advance of travel to the UK for purposes other than as a visitor, such as to take employment, study, or settle. In this report this is for brevity referred to as a visa (in line with the terminology used for visa nationals).

European Economic Area (EEA)

The EEA consists of the European Union member states as well as Iceland, Liechtenstein and Norway. Switzerland is also included here as although it is not formally a member of the EEA it has analogous status with respect to immigration control.

Immigration and Asylum Biometrics System (IABS)

The system that captures and holds biometric details provided prior to the visa and extension applications

Initial Status Analysis (ISA) database

A detailed database developed for monitoring movements of non-EEA nationals that combines not just the API and TDI information required for exit checks, but also other sources of

information including the main data on visas and other forms of permission granted for leave to remain in the UK such as visa extensions or asylum claims.

International Passenger Survey (IPS)

The IPS has collected information about passengers entering and leaving the UK continuously since 1961. Survey data is collected on the IPS via face-to-face interviews with passengers passing through ports and on routes into and out of the UK.

Non compliant

People who have remained in the UK after their permission to stay (such as a visa) has expired

Non visa national

A citizen of a country where there is no requirement to obtain a visa prior to travelling to the UK as a visitor. An **entry clearance** may be required for travel for other purposes, this being referred to as a **visa** in statistical publications to be in line with the terminology for 'visa nationals'.

Route

A route relates to all the scheduled flights or other journeys that are made between specified ports. The report discusses the routes that are in scope for exit checks, with CTA routes not being in scope.

Semaphore

The system that holds API or TDI data received from carriers, including air, rail and maritime.

Travel Document Information (TDI)

This refers to passenger information submitted by carriers or port operators at or shortly after the point of departure.

Visa

A document issued to a **visa national** that is required to be obtained in advance of travel to the UK.

Visa national

A citizen of a country where there is a requirement that they obtain a visa prior to travelling to the UK including as a visitor. For the purpose of this report nationals of countries where visitors can obtain an electronic visa waiver in place of a visit visa are also viewed as visa nationals.

Voyage

A voyage is a flight in the context of aviation and a train journey for rail.

The Home Office Responsible Statistician for this report is Chris Kershaw. If you have any comments, suggestions or enquiries please email them to:

MigrationStatsEnquiries@homeoffice.gsi.gov.uk

The Home Office Migration Statistics mission statement is:

We produce timely, accurate and objective statistics on immigration to support effective delivery of Home Office objectives and to inform government, Parliament and the public.

Press enquiries should be made to:

Home Office Press Office
Peel Building
2 Marsham Street
London
SW1P 4DF
Tel: 020 7035 3535