

## **Travel and Tourism Reform Paper**

Boran Shenhuy (Methodology and Quality) and Katie Davies (Methodology and Quality)

### **1.0 Purpose**

This paper outlines the proposed methodology for the new and improved travel and tourism statistics for the UK. The panel members are invited to comment on:

1. Assurance that the proposed methodology is being supported by evidence both by carried out and ongoing explorations.
2. Advise on any room for improvement for tourism statistics.
3. Comment on the direction of any outstanding exploration.

### **2.0 Executive Summary**

The International Passenger Survey (IPS) has been running since 1961. Data collection involves collecting information at various ports, seaports and tunnels at both departure and arrival stages. For airport departure, this data is collected after security and for airport arrival, it is collected at the baggage reclaim.

The IPS has been the main source for travel and tourism statistics. These include data on expenditure reasons for travel.

Following a review, it was concluded that the IPS is not an efficient way to compile these statistics alone as there are other sources that could be used instead of or alongside the survey. The paper proposes a few changes:

1. Stop IPS arrival data collection and instead collect data from a household survey.
2. Move IPS departures to follow the Civil Aviation Authority (CAA) Departures Survey design with the aim of increasing sample size and thus improving standard errors.
3. Change data collection across ports and tunnels to more efficient methods.
4. Explore better ways of imputing missingness on key variables than current methods.
5. Weight modular survey data using composite calibration.

Some explorations have been done for these changes:

1. Two separate data collection pilots were carried out to assess (1) data quality and (2) practicalities of data collection. The first pilot was successful in helping us understand how to collect data and was used in creating data collection scenarios. The second pilot has shown that non-response for expenditure at gate-rooms is minimal.
2. Standard errors were compared on both CAA Departures Survey and IPS to explore in which scenarios would there be better precision. Standard errors produced on the CAA Departures Survey are lower than those produced from IPS.
3. The Great Britain Tourism Survey (GBTS) and the Great Britain Day Visits Survey (GBDVS) has been explored to show how it can be used to provide statistics on UK Residents.

Overall, the changes have two major benefits: (1) better accuracy on the data and (2) savings of around £1.5M per year.

Following internal review, there are a number of further avenues of research which we are exploring. These are in Appendix 1.

### **3.0 Introduction**

The International Passenger Survey (IPS) was first created in 1961. It currently collects information from both arriving and departing passengers from the United Kingdom at different airports, seaports, and tunnels. The Travel and Tourism estimates from the survey are used by various government departments (such as the Department for Transport, Home Office and more), and across ONS publications.

In October 2019, [the Office for Statistical Regulations \(OSR\) recommended](#) that the Travel and Tourism statistics be improved on two counts:

1. Better user engagement.
2. Smaller confidence intervals and more accurate estimates.

To be able to meet these requirements, the Travel and Tourism Reform Project was started. Although interconnected, this paper focusses largely on the latter point. However, users have been updated as to the progression of the project and there is ongoing engagement with them on the user needs to be able to provide fit for purpose statistics.

Following the OSR review, [the final report](#) on progress was published in May 2022. The paper concluded that while exploring the administrative data (transaction, mobile data, Exit Checks, etc.), there is also the need for statistics in the short and medium term. They concluded that the Civil Aviation Authority (CAA) Departures Survey has shown lower standard errors around the estimates, compared with the International Passenger Survey (the latter being used for the Travel and Tourism statistics).

In January 2023, further work was carried out to mimic the findings of the May 2022 report and for further exploration. Research was pursued to find ways of implementing the survey. This paper summarises the results of that work.

It is worth noting that unless specified otherwise, the changes proposed affect the data collected at ports (airports and seaports) and tunnels in Great Britain. There are ongoing explorations about using Northern Ireland Passenger Survey (NIPS) to estimate the travel and tourism statistics in Northern Ireland. However, more work is needed on this and will be reported when complete.

This report will follow a process order on explaining the proposed changes. Firstly, a proposed sampling strategy will be explained. Following this, new data collection field methods will be explained in more detail. In relation to this, findings from the pilots carried out (one by ONS and one by CAA) will be given and compared. Moreover, in relation to missingness, explorations on future imputation methods will be proposed. The paper will then explain proposed weighting methods, then compare CAA and IPS data in order to understand the magnitude of change. Finally, the paper will describe the method for producing UK Residents total expenditure.

The next section will outline the steps to be taken under the proposed design to sample journeys.

## **4.0 Discussion**

### **Definitions**

There are two major definitions used throughout this paper.

A journey refers to a single journey, for example London Gatwick to Larnaca Cyprus on a given day and time.

A route-carrier combination refers to journeys from port A to port B by a specific carrier.

### **Sampling Departures**

The CAA Departures Survey design involves: (1) creating an annual sample, (2) using the traffic data to allocate the annual sample to quarters, (3) selecting the journeys (a specific journey from an airport to another, on a given day and time). To select the journeys, CAA use one level of stratum (Carrier and Origin), and two levels of clusters (shifts, then flights separately).

IPS, currently, (1) create an annual sample, (2) allocate the annual sample to periods and ports and tunnels, (3) select the journeys. However, different to CAA Departures Survey, IPS use one level of stratum (Port or tunnel, Terminal), and a level of clusters (shifts).

After exploring the standard errors on both designs, we propose that IPS consider including flights as a sampling stage to be able to benefit from both a merged sample, and the benefits of lower standard error. Questions identified as necessary by both ONS, and CAA will be carried out by both surveys which will allow a combined bigger sample size. Questions that are identified as non-core will only be asked by ONS or CAA (depending on who is using the data).

Sampling will involve three distinct stages:

1. Creation of an annual sample.
2. For the new Travel & Tourism Survey, dividing the annual sample into quarterly and monthly samples (depending on the seasonality) and allocating the monthly sample to ports based on traffic.
3. Sampling the journey (journey in this context is an individual journey from a port to another on a given day and time).

Please note that currently we are exploring two different ways of sampling the journey. This is because stratification has previously been criticised for use in IPS. Systematic stratification on AM and PM shifts were seen as not randomised enough. Therefore, a Latin squares approach has been used. In this sense, we are exploring whether to continue using Latin squares for choosing the AM and PM shifts, and possibly on flights.

There are no plans to reduce the sample size, as we are wanting a bigger sample size with the combined surveys. The following sample sizes relate to the airport departures.

Table 1: Sample Sizes of Residence for CAA and IPS 2019 departures after cleaning the datasets.

Residence	CAA	IPS
Foreign	62,195	20,915
UK	91,442	15,791

Table 2: Sample Sizes of Residence by Purpose for CAA and IPS 2019 departures after cleaning the datasets.

Residence	Purpose	CAA	IPS
Foreign	Business	13,250	5,092
Foreign	Leisure	48,921	15,575
UK	Business	14,629	1,844
UK	Leisure	76,808	13,730

### IPS Data Collection

Currently, data across both departures and arrivals are collected on AM and PM shifts by:

- Selecting on average every 1 in 30 passengers crossing an imaginary counting line after security and border control, respectively at airports.
- Selecting every 10<sup>th</sup> passenger at seaports. This is done by working through vehicles in car lanes at all ports but Dover.
- Selecting every 10 passenger on Channel Tunnel trains.
- Selecting every 30<sup>th</sup> passenger passing an imaginary counting line at the exit to the platform at St Pancras.

It is proposed in the to-be design that the interviewers will collect data in AM (06:00 to 14:00) or PM (14:00 to 22:00) shifts. The shifts will be allocated across the year, quarters and months using the traffic data from previous years as described in the sampling sub-section. Each shift will consist of two interviewers. There will also be a shift manager to oversee the process.

The traffic during the two shifts is assumed to be similar to the flights running between the PM and AM shifts (22:00 to 06:00). Therefore, they are adjusted for in the out-of-hours weights now, along with for ports where IPS is not running (IPS is not and will not be running at ports with traffic less than 250,000 persons in a quarter).

Data will be collected only at the departures stage (not arrivals). In this design, information about foreign residents' trip in the UK will be collected in full. However, full expenditure of UK Residents cannot be collected at this point as they have not left the UK. Therefore, expenditure is proposed to be collected as part of a household survey. This is explained in the UK Post-departure expenditure sub-section.

There are seven different scenarios for data collection. They include:

1. Airport departures (defined gate-rooms, mixed-gate, and 'Virtual' gate).
2. Eurostar.
3. Eurotunnel Cars.
4. Foot-passengers.
5. Eurotunnel Freight.
6. Dover (and Portsmouth) Quayside.
7. Freight on ferry.

Selection of passengers at in these scenarios will vary based on practicalities associated with them.

### **Weighting**

Weighting will need to be run with separate processes for core questionnaire and non-core questionnaire. Of note is that for the core questionnaire, we will weight both the CAA and IPS departures surveys' data together, whereas the non-core weights will only be for the IPS dataset.

The weights to be calculated are:

1. Design weights: which are the inverse of the probability of inclusion
2. Calibration weights: which calibrates the weights to the population total of a route-carrier combination per period (which is administrative data collected by the CAA) using composite calibration which will weigh the modular data (both Core and Non-core) together.
3. Regional Weights: which calibrates the weights for regional population totals for sub-national estimates.

### **Imputation**

Imputation will only be carried out in the case of item non-response (where a person has responded to other questions but for whatever reason they have not to another question). Under the current IPS model, the following variables are imputed for those eligible for imputation (i.e., those with item non-response):

1. Number of nights stayed, using a Poisson regression.
2. Airfare, using donor mean imputation.
3. Expenditure, using multi-level linear regression.
4. Rail-fare, using donor mean (currently going through customer needs engagement).

### **Estimation**

In this paper estimation will be mentioned on two accounts: Precision and Accuracy.

#### ***Precision***

CAA Departures Survey shows lower standard errors than IPS departures. This is due to two factors (1) sample size is bigger on the CAA Departures Survey, and (2) the standard errors shrink with the introduction of the second level of clusters.

## **Accuracy**

Currently, we're working to understand more about the bias of the estimates produced from CAA. There's known bias on the current IPS design, as the number of passengers with certain countries of residence do not have equal estimated arrival and departure statistics. However, after exploring the CAA data, we have found that results are more balanced and around expected levels.

On the other hand, we have also discovered that CAA has different estimates on purpose of visit compared to the IPS. We are currently exploring why this is the case through (1) questionnaire design and harmonisation, (2) sampling, (3) weighting.

## **UK Post-Departure Expenditure**

As explained above, we plan to stop collecting data at the arrivals across all ports and tunnels. This means that we will not be able to collect information from passengers on how much they have spent after leaving the UK. To collect this information, we are planning to utilise the [Great Britain Tourism Survey](#) (GBTS) and the Great Britain Day Visits Survey (GBDVS). GBTS/GBDVS are online non-probability panel surveys, with an annual responding sample size of 95,000 across England, Scotland, and Wales.

From GBTS/GBDVS, we will be creating a linear regression model that will be able to predict how much a person has spent abroad, given their destination and how long they have stayed. We will then use this model to predict how much a person that has been sampled on the IPS or the CAA might have spent to obtain post-departure expenditure. The pre-departure expenditure and the post departure expenditure will then be used to create total expenditure.

Previously, we have explored recall abilities of participants on Opinion and Lifestyles Survey (OPN). There was no statistical significance of the difference between responses given eight weeks after their journey and twelve weeks after their journey.

We are also exploring the model fit and the correlation of variables with post-departure expenditure. We are planning to evaluate multiple models to ensure the best model is used for the data available. Some of these models are multivariate linear regression, multi-level linear regression, and neural networks.

There are couple of constraints with this approach. Firstly, while predicting values, we can only use variables that are shared across GBTS/GBDVS and IPS. Secondly, low level estimates will not be produced. However, the latter is being communicated with the users and their feedback is being sought.

## **5.0 Conclusion**

Considering the above, the travel and tourism is proposed to undergo many changes. Sampling will, in the to-be design, incorporate a new sampling stage with flights as clusters. Weighting will need to change to incorporate the changing probability of inclusion and the new population groups. Imputation will be explored to move to a more sophisticated methodology. These changes are expected to improve the precision of the estimates. We are currently exploring how the accuracy will change, and what bias, if any, will there be to minimise.

The panel members are invited to comment on the aspects of the methodology, any room for improvement, and direction of any outstanding exploration.

## **Appendix 1 – Further Research**

Actions to take forward in response to internal review:

- Review of power calculations.
- Conduct further pilots to gain further understanding of the practicalities of data collection and outline data collection processes.
- Explore the best approach for the creation of regional weights and if they can be incorporated with calibration weights.
- Ensure documentation about weights produced is clear allowing users to make informed decisions on their use.
- Investigate best method for calculating standard errors across both IPS and CAA.
- Produce within and between cluster variances to understand their contribution to the overall standard error.
- Continue investigations into measuring bias in the estimates.
- Explore other methods of predicting post-departure expenditure.