



## The Use of Mobile Phone Data in Tourism Statistics

Ossi Nurmi<sup>1</sup>; Pasi Piela<sup>1</sup>

<sup>1</sup> Statistics Finland

### Abstract:

Survey response rates are declining at an alarming rate globally. Statisticians have traditionally used imputing and recalibration of weights to deal with nonresponse. In case survey response rates are well below 50 %, these methods may often result in little more than guesswork. Alternative data sources need to be used to improve the accuracy of statistical estimates.

In the context of outbound tourism, mobile positioning data can be considered as such an alternative data source as it registers ‘traces’ of tourism trips. These traces are CDRs (call detail records) and DDRs (data detail records) and they are generated by the activities of mobile devices. Since 2016 Statistics Finland has worked closely with Finnish national mobile network operators (MNOs) to translate the CDRs and DDRs into tourism specific monthly aggregates such as number and duration of trips by destination country. Statistics Finland has learned that it helps to be very specific when approaching MNOs with data needs.

This paper provides a summary of the methodological process that the Finnish MNOs have followed to compile tourism statistics. A similar process may be used by National Statistics Institutes or other organizations who are approaching their national MNOs with the intent of obtaining data. The paper then presents the 2017 outbound tourism data from Finnish MNOs and highlights the shortcomings of the Finnish national tourism survey in light of mobile positioning data. Based on this analysis, the paper proposes a method to enrich the tourism survey using mobile positioning data.

One of the main needs for outbound tourism data comes from the Balance of Payments (BoP) statistics. The debit side of BoP requires quarterly data by destination country and purpose of trip. More accurate data is needed to reliably estimate the expenditure of resident tourists abroad. The paper proposes a method where mobile positioning data is first used to estimate the number and duration of outbound trips by country and month. The role of tourism survey data is then to provide ratios such as purpose of trip, means of transport and average expenditure. In this method the tourism survey is no longer needed for estimating the absolute number of trips.

### Keywords:

Mobile Positioning; Tourism; Geospatial data;

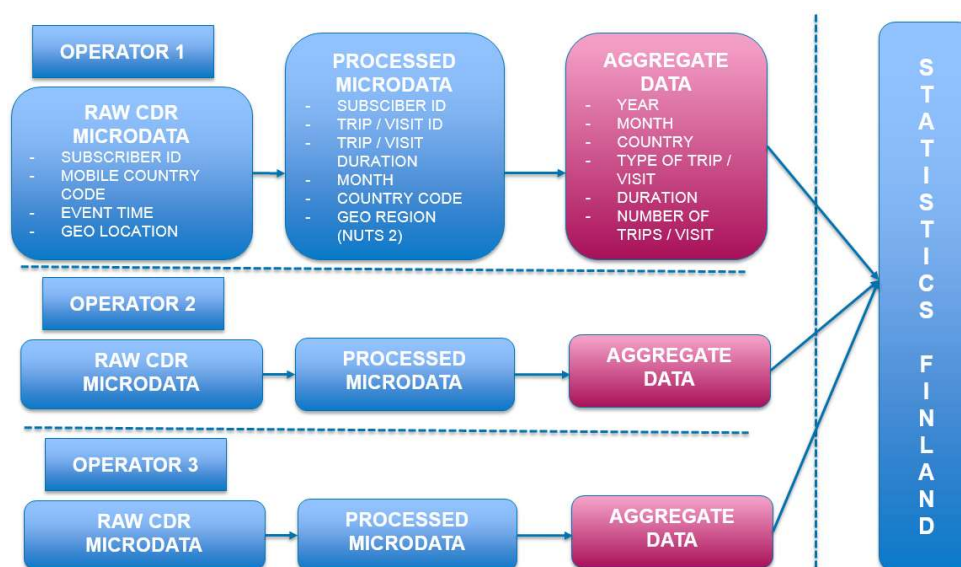
### 1. Introduction:

During 2016 up to 2018, Statistics Finland carried out the work in two phases within the context of Eurostat’s ESSNet Big Data –project. The focus in these projects was on three statistical domains: inbound tourism, outbound tourism and seasonal population.

The first phase in 2017 and 2017 focused on negotiations with national authorities and MNOs in order to set up such a process that is feasible from a legislative and technical point of view. The second phase was to carry out the process, collect data from each MNO and analyze the results.

The chosen approach relies on the operators to process the data and aggregate it for Statistics Finland. In the current Finnish legislation, only the operator is allowed to process the raw data using automatic

means. The size of the raw data is also massive, with annual data consisting of several billions of events per operator.



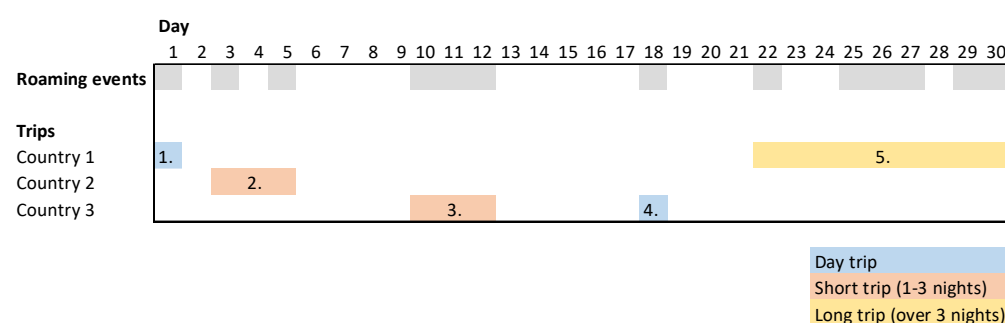
*Figure 1 Process from raw microdata to aggregated trips*

The starting point for each operator are the raw data of each of their subscribers. The subscriptions are associated with sim cards found on mobile devices. Machine-to-machine sim cards are excluded from the data as they do not represent the movement of people.

## 2. Methodology for raw data: processing roaming events to tourism trips

In case of outbound trips, the raw data consists of roaming events (calls, sms, mobile data) taking place outside of the subscriber's home network, in other words, a foreign country where the event took place. Based on the time gaps between these events, individual trips of each subscriber can be recognized.

The following example presents an imaginary case of the roaming events and outbound trips deduced for a single subscriber during a 30 days period. This subscriber is a particularly active traveller and five outbound trips are registered during the 30 day period.



*Figure 2 – Roaming events and outbound trips of a subscriber*

Roaming events are registered whenever the subscriber is making calls, sending sms or using data in a mobile network abroad. The first event in a country indicates the beginning of the trip to that particular country. A total of five outbound trips were registered for this subscriber.

## 3. Methodology for estimating outbound trips

The target population consists of all outbound tourism trips made by Finnish residents. The reference data is collected by Statistics Finland's Finnish Travel –survey with a sample size of 28,200 persons out of which 14,700 were interviewed by phone concerning trips that ended during 2017.

The data provided by the Finnish MNOs should be treated as samples of all outbound trips made by Finnish residents in 2017. Two out of three operators have provided the data and the market share of each operator is roughly one third of all subscribers. The source data thus contains the outbound trips of two thirds of the Finnish residents and one third are missing.

As with every sample, the MNO data is of limited use by itself and coefficient weights are needed to estimate the target population: all outbound tourism trips. According to the Finnish Travel –survey, Finnish residents made 10.5 million outbound trips in 2017. This is annually a relatively stable figure based on 2,700 trips reported by the respondents.

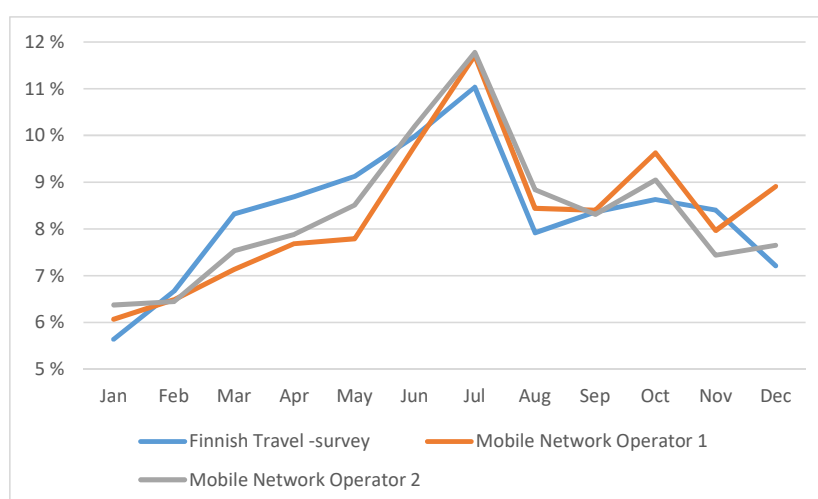
In order to avoid many of the known pitfalls related to over- or underestimation in MNO data, this annual figure of 10.5 million outbound trips, is used as a frame and MNO data is first simply weighted using the following formula:

$$\text{Weight coefficient} = \frac{\text{Annual outbound trips in tourism survey}}{\sum(\text{MNO annual trips})_n}$$

This simple weight coefficient is thus obtained by dividing the annual outbound tourism trips by sum of trips made by the subscribers of each MNO. In the case of Finnish data from two (out of three) operators, the coefficient for year 2017 is less than 1.3.

#### 4. Results: simple weight coefficient for total number of outbound trips

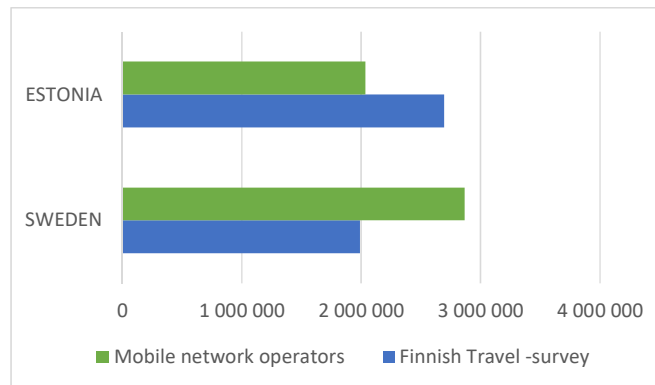
The monthly distribution of total outbound trips is shown below separately for the survey as well as the data provided by both MNOs.



**Figure 3 – Monthly seasonality of outbound trips using different sources in 2017**

The monthly seasonality of outbound trips is strikingly similar for both MNOs. Both register the summer holiday months of July and June as the peak months of outbound tourism. The third biggest month for both operators is October, which is the month of autumn holidays in Finnish schools. The Finnish Travel –survey registers more outbound trips than the MNOs during the months of March through May.

Concerning tourism by country, neighboring Estonia and Sweden are by far the most important outbound tourism destination countries. 45 per cent of outbound trips are made to these two countries.



**Figure 4 – Trips to Estonia and Sweden in 2017**

There were 2.6 million trips to Estonia in 2017 according to Finnish travel –survey. Using the top-down approach for mobile positioning data, 24% less trips appear in MNO data. This indicates that this method underestimates tourism to Estonia, as the 2.6 million trips from the survey is a very stable figure with only small changes annually.

In contrast, the outbound tourism to Sweden comprises of only 2.0 million trips in the survey, while the top-down approach estimates 44% more trips to Sweden in MNO data. The neighboring countries with open land borders have many possible sources of overestimation such as frequent non-tourism trips, border noise etc.

There is yet another reliable data source for providing the monthly seasonality of outbound trips to Estonia, the main outbound tourism destination. Nearly all passengers to Estonia use one of the ferries that operate between the capital cities of Helsinki and Tallinn. The Finnish Transport Agency compiles statistics on the total monthly passengers departing to Estonia, including passengers of all nationalities.



**Figure 5 – Monthly seasonality of trips to Estonia in 2017**

The seasonality of ferry passengers is in line with the outbound tourism data provided by the MNOs. In contrast, the seasonality of the Finnish Travel –survey seems to be affected by randomness. Some months in the survey are exceptionally high (March, November) and some too low (April, August) when compared to ferry passengers and MNO data.

Although the MNO data slightly underestimates the total number of trips to Estonia, it appears to be a better data source for estimating the monthly seasonality. Unlike the survey, it's not affected by the

randomness caused by a small sample size. This suggests, that MNO data should be used to adjust the monthly seasonality estimates produced the survey.

The main findings from these results are the following:

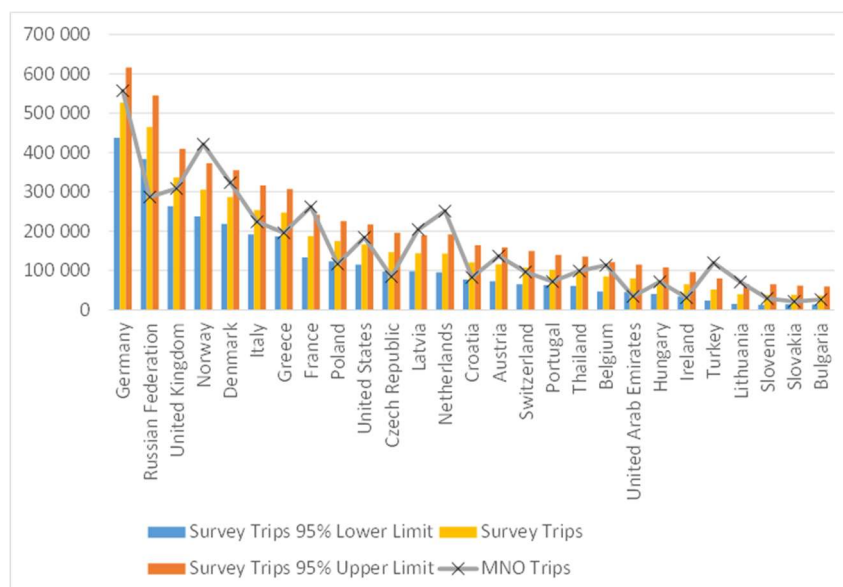
1. The data from two different operators are highly correlated. The monthly seasonality is nearly identical and in line with outbound tourism based on Finnish Travel –survey.
2. Depending on country of destination, the top-down approach dramatically over- or underestimates the total number of outbound tourism trips to that country. There are many known sources of bias in mobile data: non-tourism trips, border noise, devices switched off, multiple devices, transit corridors, conceptual differences etc.
3. Mobile positioning data provides a better estimate on the monthly seasonality of outbound tourism. The monthly estimates of Finnish Travel -survey are affected by randomness due to small sample size.

## 5. Results: country specific estimation

How should the survey data be enriched or recalibrated using MNO data in order to improve the accuracy? The proposed method for recalibrating the outbound trips data has to provide at least the following estimations:

1. Annual number of all outbound trips
2. Monthly seasonality of outbound trips to each country
3. Annual number of outbound trips to each country
4. Year-on-year change in the number of outbound trips

For annual number of outbound tourism trips (1.) the Finnish Travel -survey provides a solid estimate as shown earlier. For monthly seasonality (2.) the MNO data is more robust as it is not affected by survey randomness. For trips to each country (3.), the best data source depends on the country of destination. At present, it's not possible to evaluate the year-on-year change as MNO data is available only for 2017.



**Figure 6** – 95 % confidence intervals for top 30 destination countries (excluding top 3)

The figure presents the number of outbound trips to each country based on the Finnish Travel –survey as well as the upper and lower bounds of the 95% confidence interval for the survey estimates. The number of trips based on MNO data is plotted against them as a line graph.

Using this kind of method, the MNO data provides a ‘second opinion’ to the survey confidence intervals for each country. In case the MNO trips are outside of the confidence interval for a certain country, the MNO data most likely includes serious sources of over- or underestimation for that country. On the other hand, the survey estimates also become rather useless if the 95% confidence interval is too large. At present, countries with less than 170,000 trips have 95% confidence interval limits of plus or minus 30 per cent. For most of such small destination countries, the MNO trips can still provide a better estimate, given that the MNO trips are within the confidence interval.

There are currently only 24 destination countries where the annual estimates are considered reliable. On a monthly level, only trips to Estonia and Sweden are mostly reliable. In total there are 9 million outbound trips to these countries. The MNO data can potentially provide trips to 129 more smaller destination countries with 1,5 million trips in total.

## 6. Discussion and Conclusion:

In the context of outbound tourism, the strengths and weaknesses of mobile positioning and survey data can be summarized as follows:

*Table 1 – Strengths and weaknesses of survey and mobile positioning data*

	Finnish Travel -survey	Mobile positioning data
<b>Strengths</b>	<p><b>Scope</b> is clean: only tourism trips are included</p> <p><b>Provides supporting information</b> of the trip (ie. purpose of trip, expenditure, means of transport and accommodation)</p>	<p><b>Granularity:</b> millions of observations covering nearly all destination countries</p> <p><b>Monthly seasonality</b> of tourism is more accurate.</p>
<b>Weaknesses</b>	<p><b>Granularity;</b> very few observations per year, covering only a few destination countries</p> <p><b>Monthly seasonality</b> estimates are affected by randomness</p>	<p><b>Scope</b> is not clean, there are many sources of over- or underestimation</p> <p><b>No supporting information</b> of the trip</p>

By recalibrating the existing survey data using data from MNOs, it's possible to combine the strengths and alleviate the weaknesses in both data sources. This results in significant improvement in geographical as well as temporal granularity of outbound tourism statistics. Geographical granularity refers to the number of countries that can be reported, while the temporal granularity is the monthly tourism seasonality to each of those countries.

This study was conducted by using data from two MNOs consisting of roughly two thirds of all Finnish mobile subscribers. The data from both operators are highly correlated. This suggests that similar results could be obtained by using data only from a single MNO, given that its subscribers are not geographically or socio-demographically biased.

## References:

Eurostat (2018). ESSNET Big Data Project. Available at:

[https://webgate.ec.europa.eu/fpfis/mwikis/essnetbigdata/index.php/ESSnet\\_Big\\_Data](https://webgate.ec.europa.eu/fpfis/mwikis/essnetbigdata/index.php/ESSnet_Big_Data)

Statistics Finland. Finnish travel statistics. Available at:

[http://www.stat.fi/til/smat/index\\_en.html](http://www.stat.fi/til/smat/index_en.html)

Eurostat (2014). Feasibility Study on the Use of Mobile Positioning Data for Tourism Statistics.

Consolidated Report Eurostat Contract No 30501.2012.001 – 2012.452. Available at:

<https://ec.europa.eu/eurostat/documents/747990/6225717/MP-Consolidated-report.pdf>