INSTITUTO NACIONAL DE ESTADISTICA

Mobility study based on mobile telephony during the July-December 2020 period (EM-3)

Technical Project

Sub-Directorate General of Socio-Demographic Statistics

November 2020

Experimental

Index

Mobility study based on mobile telephony during the July-December 2020 period (EM-3)

1 Introduction: mobility studies 2019-2020

During the years 2019 and 2020, the INE has undertaken what can be considered a single mobility measurement project based on mobile telephony, but which has had three phases due to the contingencies of the Covid-19 pandemic. We are therefore talking about three different "studies":

- The original EM-1 study[1], which was designed to measure daily mobility taking as a reference a specific week in November 2019 and four days, also in 2019, to measure how the population is located outside its usual place of residence on designated dates (two days of summer, one normal weekend, Christmas day).
- The EM-2 study during the state of alarm. The enormous impact that the coronavirus outbreak in March 2020 had on the mobility of the population and the consequent promulgation of the state of alarm led to the urgent consideration of a second mobility study, which provided daily information on movements among the 3,214 areas that were designed for the project. Due to the exceptional circumstances in which this study was carried out (telephone operators provided free daily information during the state of alarm), the dates are limited to the state of alarm (March 16 to June 20).
- The EM-3 study, after the March-June state of alarm. Once the aforementioned state of alarm has ended, the need arises to continue with the series of mobility data in order to monitor what is then called "return to normality". At the time of writing this project (end of October 2020) there are abundant restrictions on mobility in many parts of Spain (perimeter confinements) that even increase the interest in having this information.

Since study EM-3 fully shares the methodological principles with the previous two, the reader is referred to the original study EM-1 for more details. Only the specific questions of this third phase of the mobility project are presented here.

2 Objectives

The main objective of the EM-3 is to measure mobility between areas during the second half of 2020, more specifically, data are collected for the period June 24 to December 30, 2020.

The original project EM-1 includes three types of mobility matrices:

- Matrix 1: daily mobility
- Matrix 2. Day and Night Population
- Matrix 3: seasonal population

For this study, EM-3 dispenses with the type 2 matrix, which does not yield very useful information for these purposes; the type 1 and 3 mobility matrices are constructed.

[1]For more information https://www.ine.es/experimental/movilidad/experimental_em.htm

3 Scope of the investigation

3.1 POPULATION SCOPE

The population scope consists of the mobile phones of the resident population in Spain with service from one of three above-mentioned operators; that is, mobile phones with national dialling codes. Foreign-numbered telephones on roaming -normally used by tourists- are excluded.

Therefore, we can implicitly affirm that the population scope is the population resident in Spain.

The information requested refers only to mobile phones, not to all devices that may have a SIM card.

3.2 GEOGRAPHICAL SCOPE

The geographical scope is the entire national territory. The territory is divided into 3,214 “mobility areas” and information is provided for each of them.

3.3 TEMPORAL SCOPE

The collection period foreseen in type 1 matrices is just over six months. Data are taken from every Wednesday and Sunday for the period from June 24 to December 30, 2020, both inclusive.

Therefore, the chosen days are: Wednesday, June 24, Sunday, June 28, Wednesday, July 1, Sunday, July 5, and so on until the end of Wednesday, December 30, 2020.

For matrix type 3, four specific days are taken: July 18, August 15, November 22 and December 25, all of them from 2020. It really is about the early mornings of those days, that is, the data for August 15 are those for the night of August 14 to 15.

It is intended with them to have two representative days of summer - a time of maximum seasonal population such as August 15 and another not as intense as a Saturday in July - Christmas day, and a "valley" weekend. It is intended to be able to compare mobility with similar days in 2019 that were taken in the original EM-1 study.

4 Characteristics of the investigation

TABLE Nº 1: DAILY MOBILITY MATRIX

It is an origin-destination matrix formed by N (3,214) areas x N areas. For each day observed, the box "row A-column B" would show the number of terminals moving from the residence area (A) to the destination area (B).

For the construction of this matrix, the area of residence of each of the mobile phones of each MPO must first be determined, using this method:

The area of residence of each mobile phone is the area where the mobile phone is found for the longest time between 10:00 p.m. the previous day and 06:00 a.m. during the observed day.

Assigned for each mobile phone the area of residence, an attempt is made to determine the area of daily destination. The method is similar to that of assigning the residence area but it will not always give a value: all the areas (including the residence area) in which the terminal is located must be taken during the hours of 10:00 a.m. to 4:00 p.m. observed. If there is more than one area, the one that is most frequent (in which it is the longest), other than the residence area, will be taken. If only the residence area meets this condition, then the residence area will also be the destination area.

As mentioned above, this method will not always be able to obtain a destination area, because in some cases it will not be possible to determine that there is an area that is visited on a recurring basis. Therefore, it is not necessary to assign a destination area to all mobile phones, but it is nevertheless necessary that all mobile phones have an associated residence area.

TABLE Nº 2: THEY ARE NOT CALCULATED FOR THIS STUDY.

TABLE Nº 3. SEASONAL POPULATION MATRIX

It is about building an N x N x 4 matrix (or four N x N matrices). Each of the four corresponds to a specific day. The days of July 18, August 15, November 22 and December 25 have been considered, all of them in 2020.

For each day chosen, the row A-column B box shows the number of people who, residing in area A, spend the night in area B during that day, according to the following specifications:

For the chosen day, the area where it stays overnight is determined for each terminal (where that mobile phone is most frequently found during the period from 10:00 p.m. the day before to 06:00 on that date). Therefore, more precisely, the data for July 18 correspond to the period between 10:00 p.m. on July 17 and 6:00 a.m. on July 18, and thus respectively with the other three dates.

This method provides an overnight area for all telephones present in Spanish territory on those dates.

For each of the mobile phones, their area of residence is found as the one where the mobile phone is found for the longest time between 00:01 and 06:00 in the last two or three months (the number of days is different according to the operator but at least they take 60 days). In other words, to determine the area of residence necessary to prepare the December 25 matrix, the area of residence where that telephone is most frequently found between 00:01 and 06:00 hours during the previous period considered (between two and three months).

5 Data collection and information processing

The INE receives the already aggregated matrices prepared by each operator. Therefore, it does not have individual information of any kind at any time, but only matrices of final results prepared by each of the OTMs that offer counts at the mobility area level.

It is important to note that the source used for the estimated location of mobile phones is from the telephone antennas, so the location of a phone cannot be determined with total precision. The margin of error is from tens of meters in urban areas to hundreds of meters, even kilometers, in rural areas

In addition, movements observed between bordering areas should be taken with caution because they may be caused by an error made when setting each terminal's position. This can result in a terminal being counted that is immobile in two different adjacent areas at different times.

This is an extract with simulated data of the information of the files type "matrix 1" sent by one of the operators:

```
54Mu;51MU;26
54Mu;52MU;30
54Mu;53MU;118
54Mu;54Mu;1797
54Mu;55MU;48
54Mu;56MU;#
54Mu;57MU;27
54Mu;58MU;130
54Mu;59MU;33
54Mu;60MU;56
54Mu;61MU;21
54Mu;62MU;#
54Mu;63MU;52
54Mu;64MU;20
```

That is, it would indicate that from area 54MU 26 telephones move to area 51MU, or that from area 54MU a number less than 10 or 15 terminals move to area 56MU (depending on the operator).

Whenever the number of observations is less than 10 or 15 in an area, for a given operator, the result will be censored at origin. Therefore, origin-destination mobility data could only be offered for areas in which at least one of the three OTMs counts more than that threshold.

The OTMs also offer the total flows for each origin and destination area, that is, they do not identify flows smaller than 10-15 but they do provide the data of the total number of telephones that leave or enter an area, as long as those same are respected thresholds, which allows estimating flows lower than 10-15, obviously always with a certain error.

The three complete matrices are published in numerical form and also in the form of maps. As for the numerical information, the important thing about the project is to know the people who move between areas, so it has been chosen to offer information, not on the number of telephones that move but elevated to population figures.

5.1 PREPARATION OF MATRIX 1 (DAILY POPULATION)

The first section of the publication consists of six tables:

1.1 Summary of daily mobility by Autonomous Communities, provinces, islands and main cities (over 50,000 inhabitants or provincial capitals).

1.2 Summary of daily mobility by areas of residence

1.3 Origin-destination flows between areas (more than15 people)

1.4 Destination-origin flows between areas (more than15 people e)

1.5 Origin-destination flows between main cities with more than 50,000 inhabitants or provincial capitals (more than15 people)

1.6 Destination-origin flows between main cities with more than 50,000 inhabitants or provincial capitals (more than15 people).

The flows of the three OTMs are added, provided that you have numerical data for at least one of them. These values are raised to the resident population aged 12-89 years for each area of residence. The choice of the age range is somewhat arbitrary. However, the user could undo this elevation and adjust it to their own since the population data 12-89 for each area is also published.

This age interval has been chosen since to measure daily mobility it is assumed that for these age ranges the probability that the person carries a mobile phone is much higher.

The following data is provided for each mobility area:

- Resident population as of January 1, 2019 (according to Register) (A)
- Resident population that is located during the day in your area of residence. It is about counting the population whose area of residence coincides with the destination area (B).
- Resident population found during the day in another area (C)

- Non-resident population that is located during the day in the area (D)
- Total population located during the day in the area (E = B + D)
- Population balance entering and leaving the area (F = DC)
- Flows of more than 15 people between the area and others, originating from or destined for the area under consideration.
- Flows of more than 15 people between the main cities (greater than 50,000 inhabitants or provincial capitals) that have their origin or destination in the city considered.

Thus, the resident population in an area (A) can therefore be divided into three groups: those found in the same area (B), those found in different areas (C) and the rest, which may or may not be found. They have a destination area in which they are located most of the time (A- (B + C)).

In addition to publishing the details of all flows over 15 inhabitants, data on the number of different origins or destinations is also provided which, based on the data received from the OTMs, exceed the threshold of 10-15.

In addition, a map is provided that provides the following information:

- Population summary by areas
- Origin-destination flows: for each mobility area, the flows of residents in it are analysed.
- Destination-origin flows: for each mobility area, the arrival flows to it during the day from other areas are analysed.

5.2 PREPARATION OF MATRIX 3 (SEASONAL POPULATION)

The tables offered are:

3.1 Summary of seasonal population by dates, CCAA, provinces, islands and main cities (over 50,000 inhabitants or provincial capitals).

3.2 Summary of seasonal population by dates and areas of mobility

3.3 Flows of residence / overnight stay by dates and areas (more than15 people)

3.4 Overnight / residence flows by dates and areas (more than15 people)

3.5 Flows of residence / overnight stays by dates and main cities, greater than 50,000 inhabitants or provincial capitals (more than15 people)

3.6 Overnight / residence flows by dates and main cities, over 50,000 inhabitants or provincial capitals (more than15 people)

3.7 Population staying overnight in a different Autonomous Community than the one of residence, by dates

3.8 Population staying overnight in a different province than the one of residence, by dates

The following data is provided for each mobility area:

- Number of different overnight areas in which the population residing in the area is spending the night
- Number of different areas of residence in which the population that is spending the night in the area resides
- Resident population (A)
- Total resident population found on selected date (B = C + D))
  - Resident population that spends the night in their area of residence (C)
  - Resident population found that stays overnight in another area (D)
- Non-resident population that stays overnight in the area (E)
- Total population that stays overnight in the area (E)
- Population balance of the area (G = F - A)
- Flows of more than 15 people between the area and others, who reside or stay overnight in the considered area.
- Flows of more than 15 people between the main cities (over 50,000 inhabitants or provincial capitals) who reside or stay overnight in the city considered.

Likewise, a map is offered for each day, for the analysis of the seasonal population that offers the following information:

- Summary of seasonal population by areas of mobility
- Residence-overnight flows: for each area, it is analyzed where residents are spending the night in it.
- Overnight-residence flows: for each mobility area, the population that stays overnight is taken and the areas of residence of all of them are analyzed.