Study on New Mobility Patterns in European Cities

Task B: Targeted Survey on Urban Logistics

Annex 2 - Methodological approach of NMP survey

Written by Cartolano F., Vaghi C., Chiarilli S., Rodrigues M., Tharsis T., Borgato S., Maffii S., Mars KJ., Popovska T., Vincent V, Gayda S., Bogaert M.

July 2022

EUROPEAN COMMISSION

Directorate-General for Mobility and Transport (DG MOVE)
Directorate A – Policy coordination
Unit A.3 Economic analysis and Better regulation

Contact: Niccolò Pieri

E-mail: niccolo.pieri@ec.europa.eu

European Commission
B-1049 Brussels

Study on New Mobility Patterns in European Cities

Task B: Targeted Survey on Urban Logistics

Annex 2 - Methodological approach of NMP survey

Directorate-General for Mobility and Transport (DG MOVE)

Service Contract No MOVE/A3/SER/2019-401/SI2.814412

Europe Direct is a service to help you find answers to your questions about the European Union.

Freephone number (*)**:**

00 800 6 7 8 9 10 11

(*) The information given is free, as are most calls (though some operators, phone boxes or hotels may charge you).

Luxembourg: Publications Office of the European Union, 2021

ISBN 978-92-76-54831-7
doi: 10.2832/479946

1 Survey quality indicators

Survey quality indicators reported in the following subsections were calculated according to Eurostat manual and working papers as reported in the following.

1.1.1 Survey quality indicators for variables related to enterprise business and characterisation

These indicators were captured from questions included in parts 1 and 3 of the questionnaire, namely economic data of the enterprise and fleet size and future plans for low emission logistics. These indicators are:

- composition of fleet owned by the enterprise
- number of city deliveries by bikes
- number of city deliveries by powered two-wheelers
- share of companies who have a plan for low emission logistics
- share of companies who do not have a plan for low emission logistics

For the indicators related to this class, the reference population is the total number of companies performing urban logistics in major EU cities.

In order to identify such reference population, the ideal situation would be to have information on freight distribution of all enterprises performing delivery activities within a city, irrespective of where headquarters or local units are established. However, depending on the type of distribution service, some enterprises may be based in the city or elsewhere: in logistics it is possible to encounter enterprises with local business (this could be the case of couriers that often subcontract services to local enterprises) as well as larger enterprises performing inter-regional services ending with urban distribution (this is the case of operators fulfilling grocery stores).

As introduced chapter 2 of the report, the survey conducted used D&B database as sampling frame containing enterprises per the selected NACE codes and based in the related cities and this allowed to capture enough enterprises to complete the survey but introduced some limitations:

1. Enterprises based outside the selected metropolitan areas performing distribution services in the target cities were not captured.
2. Total number of trips, deliveries, parcels, volumes of freight moved by interviewed enterprises is unknown (because this data was not asked and, in any case, typically not disclosed).

Both limitations are somehow intrinsic to the methodology chosen: there is no register capturing all logistics operators doing distribution services in a city. This is a business-related variable not strictly dependent on the operational location of enterprises. As for the second one, enterprises tend to not disclose the total number of trips, parcels and volumes distributed for commercial sensitiveness reasons.

To deal with these limitations, a random selection was applied sequentially:

- to select candidate enterprises (among those that meet NACE and geographical criteria) from the database to be uploaded in the survey software,
- to select enterprises to be contacted,
- to select trips to be recorded among the total of daily trips performed by each company

This process ensured a strict randomisation in the selection of data to be collected and, given that variables analysed concern proportional distributions (e.g. type of vehicle), the Cochran's formula[1] was used to elaborate the sample size:

$$n_0 = \frac{Z^2 * p * (1-p)}{e^2}$$

Where:

- **Z** represents the z-score commonly set at 1.96 for 95% confidence level
- **p** represents the distribution shown for respondents for the entire population

The net sample is greater than 384, which is the minimum sample calculated to sit in the 95% of confidence interval and 5% or error. Therefore, results are largely within the range of the margin of error.

1.1.2 Survey quality indicators for variables related to trips

These indicators were captured from questions included in part 2 of the questionnaire, Activity and traffic data, concerning reference trips provided by respondents and are:

- number of deliveries
- number of parcels
- vehicle-km
- weight of parcels
- load capacity
- tonne-km
- average load factor
- pollutant emissions

For these indicators the reference population is the number of trips performed by a vehicle in a day. They are elaborated at city level, having each city peculiar characteristics with regards to urban distribution dynamics.

These indicators relate to continuous variables and the formula used to calculate the margin of error, given a sample size represented by the trips in each city, is the following[2]:

$$MOE_{95} = z_{0.95}\sqrt{\sigma^2/n}\P$$

[1] Eurostat guidelines on Passenger Mobility Statistics, December 2018 https://ec.europa.eu/eurostat/documents/29567/3217334/Guidelines_on_Passenger_Mobility_Statistics+%282018_edition%29.pdf/f15955e3-d7b4-353b-7530-34c6c94d2ec1?t=1611654879518

[2] https://ec.europa.eu/eurostat/documents/3859598/5927001/KS-RA-13-029-EN.PDF

Where:

- $Z_{0.95}$ represents the z-score commonly set at 1.96 for 95% confidence level
- **σ** represents the standard deviation of the variable
- **n** represents the net sample size (the reference trips collected in each city)

The following table reports the detail of interviews and related targets achieved, where:

- Reference trips are trips for delivering or collecting freight from the same addresses on a regular frequency
- Completed vehicle days are trips accounting for the net sample size in each city

Table 1 – Completed interviews per city

City	Completed interviews	Reference Trips	Completed Vehicle Days	Trips per Interview	Vehicle days per Trip	Vehicle days per Interview
London	101	155	711	1,5	4,6	7,0
Rotterdam	43	72	428	1,7	5,9	10,0
Amsterdam	27	49	279	1,8	5,7	10,3
Sofia	170	177	706	1,0	4,0	4,2
Prague	123	215	789	1,7	3,7	6,4
Lisbon	66	165	709	2,5	4,3	10,7
Bucharest	124	275	763	2,2	2,8	6,2
Barcelona	87	155	712	1,8	4,6	8,2
Madrid	88	167	781	1,9	4,7	8,9
Paris	122	171	809	1,4	4,7	6,6
Hamburg	52	87	399	1,7	4,6	7,7
Bremen	18	42	312	2,3	7,4	17,3
Budapest	158	293	737	1,9	2,5	4,7
Rome	105	176	712	1,7	4,0	6,8
Milan	88	161	705	1,8	4,4	8,0
Stockholm	38	98	597	2,6	6,1	15,7
Goteborg	16	28	111	1,8	4,0	6,9
Brussels	60	137	664	2,3	4,8	11,1
Antwerp	10	15	45	1,5	3,0	4,5
Berlin	53	93	467	1,8	5,0	8,8
Munich	41	69	267	1,7	3,9	6,5
Total	**1590**	**2800**	**11703**	**1,8**	**4,2**	**7,4**

Sources: DG MOVE, Targeted Survey on Urban Logistics, 2021

2 Approach to data cleaning/processing

This subsection summarises the approach undertaken for cleaning and processing data and then weighting results.

2.1.1 Data cleaning

Based on the collected responses, a data cleaning process was necessary for some interviews, in the following situations:

1. Correcting partial responses and errors in reporting, this situation occurred when:
    a. the respondent included the return trip as last leg: this leg was removed, and the trip considered valid
    b. the respondent included a large number of deliveries per trip but most of them have "don't know" in the field address; these fields have been removed and the trip considered valid.
2. Using Geographical Information Systems to select the urban leg from the urban boundaries selected; this situation occurred when there were found delivery addresses far from the urban areas. The trip was considered as inter-city and therefore removed.

2.1.2 Data processing

Data processing was carried out to correct some inconsistencies described in the following and to prepare data for elaboration of CO2 emissions. The following processes have been applied.

- **Incomplete trip data –** Some fields of the questionnaire, such as number of parcels and total weight, were not mandatory to be filled in and, in a limited number of interviews, data were missing. The approach for these fields was to discard the entire interview for CO2 emission computation only, in order to have more reliable results.
- **Incorrect addresses -** Some addresses were provided as zip code only or with topographic description that cannot be recovered from GIS tools: in this case pre-processing GIS algorithms were applied in order to retrieve geographical coordinates, when possible, and to assess distances driven for CO2 emission calculation.
- **Outliers -** Some fields were discovered to be outliers and were removed or replaced by average values. The incidence of this data is limited (about 3% in the worst case); following the list of data classified as outliers:
    - **Number of consolidation centres greater than 100**. Consolidation centres have several definitions: they may be intended as infrastructures serving single properties (an airport, a shopping centre) or dedicated to part of a city to facilitate deconsolidation and last mile delivery. In any case the number of such infrastructures per city is usually quite limited and a threshold of 100 is considered largely appropriate in literature[3]

[3] https://www.researchgate.net/publication/254324483_The_Role_of_Urban_Consolidation_Centres_in_Sustainable_Freight_Transport

- **Distances driven greater than 300Km.** Considering the maximum time spent for a delivery during a working day (8 hours) with a speed between 30 and 40 Km/h, the distance of 300 Km was chosen as threshold and greater values (some of them in the order of thousands of Km) considered outliers.
- **Number of daily trips per vehicle greater than 40.** Considering the maximum time spent for a delivery during a working day (8 hours) and 5 trips per hour, 40 trips per vehicle in a day was set as threshold and greater values (some of them in the order of hundreds) considered outliers.
- **Number of daily deliveries per bike and per powered two-wheeler greater than 40.** Considering the maximum time spent for a delivery during a working day (8 hours) and 5 deliveries per hour, 40 deliveries per vehicle in a day was set as threshold and greater values (some of them in the order of hundreds) considered outliers.

2.1.3 Data processing for pollutant emission computation

Once cleaned and pre-processed as described above, raw data collected (split per deliveries and pick-ups) from the survey included the following fields relevant for emission calculation:

- Respondent ID
- Number of reference trips
- Type of Vehicle
- Rigid or Articulated Vehicle
- Type of fuel
- EURO class
- Maximum allowed mass (kg)
- Total weight
- Distance travelled
- City

The pollutant emissions were then computed by using standard Copert data[4] applied per type of vehicle, Maximum allowed mass, Euro Class, Rigid/Articulated, Load factor (ratio between total weight and maximum allowed mass).
The following data processing has then been applied to further check accuracy of collected data:

- LGV (<= 3.5 ton) cannot be articulated, the attribute was removed, and the record considered valid
- Type of fuel for HGV cannot be LPG (Liquefied petroleum gas), the attribute was replaced with LNG and the record considered valid
- Type of fuel for LGV cannot be LNG, the attribute was replaced with CNG, and the record considered valid
- Copert emission for euro 0 vehicles are not available, the records was considered not measurable and then removed.
- For London, distances were converted from miles in Km

[4] https://www.emisia.com/utilities/copert-data/

3 Weighting procedure

Weighting procedures are used to accomplish the following objectives (Richardson, et al., 1995):

- To compensate for differential probabilities of selection among subgroups (in stratification procedures, geographical strata)
- To reduce the effects arising from non-response
- To compensate for inadequacies in sample frame
- To bring sample data up to the dimension of study target population

Generally, the weighting procedures rely on calibration of margins. Calibration consists in adjusting the original (sample) weights. Reweighted sample estimates conform to known population external totals. The method forces the estimates to equalise population parameters (e.g. the totals of population in territorial domains), which may be available from an auxiliary source such as administrative data, or from statistical sources such as Eurostat or other studies as mentioned in section 2. This stage is essential to ensure a representative sample and comparison with some other statistical sources (e.g. other national surveys). The calibration on margins must be implemented both on variables which explain (or are correlated with) transport behaviour, and also on the variables that explain the nonresponse mechanism for which the total is accurately known (Deville and Särndal 1994).

In the present study two weighting procedures were applied to either reduce the effects arising from non-response or to compensate for inadequacies in sample frame. These procedures were necessary for number of parcels and emissions calculation.

3.1.1 Number of parcels

In this dataset there was a large number of responses with "Don't know" value and the procedure in this case was to discard responses with such value and recalculate mean, standard deviation and margin of error using the reference trips with valid responses only.

3.1.2 Pollutant emissions

Pollutant emissions were calculated using the Copert data as weighted average from values resulting of applying the emission/km to distance driven and multiplied per the number of reference trips reported by the respondent.

Concerning Maximum allowed mass and Load factor, data are available in COPERT for a discrete number of classes and therefore the following approximations were applied:

- Load factor discretised at the following values:
  - 0: for ratios less than or equal to 0,25
  - 0,5 for values between 0,25 and 0,75
  - 1 for values greater than 0,75
- Maximum allowed mass discretised at the following values (in tonnes):

3,5;7,5;12;14;20;26;28;32;40;50;60.

Also, in this dataset there was a large number of non-valid responses with missing values for distances, weight, maximum allowed mass. Therefore, the emissions/km were computed for valid responses only and, in case of missing values for such dimensions, the entire trip was discarded.

Mean, standard deviation and margin of error were recalculated using the reference trips with valid responses only.

GETTING IN TOUCH WITH THE EU

In person

All over the European Union there are hundreds of Europe Direct information centres. You can find the address of the centre nearest you at: https://europa.eu/european-union/contact_en

On the phone or by email

Europe Direct is a service that answers your questions about the European Union. You can contact this service:

– by freephone: 00 800 6 7 8 9 10 11 (certain operators may charge for these calls),

– at the following standard number: +32 22999696, or

– by email via: https://europa.eu/european-union/contact_en

FINDING INFORMATION ABOUT THE EU

Online

Information about the European Union in all the official languages of the EU is available on the Europa website at: https://europa.eu/european-union/index_en

EU publications

You can download or order free and priced EU publications from: https://publications.europa.eu/en/publications. Multiple copies of free publications may be obtained by contacting Europe Direct or your local information centre (see https://europa.eu/european-union/contact_en).

EU law and related documents

For access to legal information from the EU, including all EU law since 1952 in all the official language versions, go to EUR-Lex at: http://eur-lex.europa.eu

Open data from the EU

The EU Open Data Portal (http://data.europa.eu/euodp/en) provides access to datasets from the EU. Data can be downloaded and reused for free, for both commercial and non-commercial purposes.

MI-09-22-351-EN-N

ISBN: 978-92-76-54831-7

doi: 10.2832/479946