Economic and Social Council

Economic Commission for Europe

Inland Transport Committee

Working Party on Transport Statistics

Seventy-third session
Geneva, 15 – 17 June 2022
Item 7 (b) of the provisional agenda
**Data collection, methodological development and
harmonization of transport statistics: Common questionnaire**

Common questionnaire: quality improvement

Note by the secretariat

I. Background

1.   The common questionnaire allows countries to simultaneously provide inland transport statistics to three organizations – the Economic Commission for Europe (ECE), International Transport Forum (ITF) and the Statistical Office of the European Union (Eurostat). This exercise provides a rich collection of indicators, most of which are not available through other data collections at the international level. The principal methodological document for this collection is the Glossary for Transport Statistics, jointly published by the three organizations, the fifth version of which was finalized in 2019.

2   This document discusses some common quality issues with data collected through the common questionnaire, to identify areas where quality can be improved. This analysis is largely based on work led by Eurostat. The document is organized by theme, with a description of the issue and a proposal for actions to be taken. It covers only the main points in the Eurostat analysis; further detail will be provided at the annual session (both as an informal document and in presentations).

II. Load capacity of road vehicles

3.   In the road chapter of the questionnaire, there are 362 indicators requested, and vehicle equipment makes up 275 of these. Of these, 32 indicators relate to the load capacities of goods vehicles (based on both the total fleet number and new registrations in the reference year), where the unit is tonnes of load capacity. For these indicators, it seems that some countries have been consistently supplying the number of vehicles, rather than their carrying capacity, and there are unit problems for other countries. While the partner organizations all conduct quality checks on all common questionnaire indicators, likely errors in this set of indicators have gone unnoticed because checks typically focus on internal consistency and trends over time, rather than on values relative to other characteristics (i.e., number of vehicles in this case).

4.   The issue seems to exist for all 32 indicators, but to give an example of compliance the load capacity of all lorries (total fleet) is highlighted. For this indicator, seventeen ECE countries appear to give compliant data, while nineteen do not provide any data at all. Six countries (Germany, Ireland, Italy, North Macedonia, Russian Federation and the Netherlands) provide capacity data identical to their fleet number figures. Two other countries (Bosnia and Herzegovina, and Moldova) seem to provide the data in tonnes, and not thousand tonnes. A further two countries could probably estimate or calculate their total capacity as they seem to report the sub-components (broken down by weight class) of the total.

5.   Taking this analysis further, the ratio of load capacity per goods vehicle (by weight class) can be calculated to see the average capacity of each goods vehicle. This gives an average of 0.9 tonnes for the category of vehicles up to 3.5 tonnes of Permissible Maximum Gross Weight (PMGW); 3 tonnes average for vehicles of 3.5-7.5 tonnes PMGW; 5 tonnes average for vehicles of 7.5-12 tonnes PMGW; 12 tonnes average for vehicles of 12-40 tonnes PMGW; and 29 tonnes average capacity for vehicles over 40 tonnes PMGW.

6.   The partner organizations propose to delete the capacity data that are equal to fleet numbers and request that countries provide the correct capacity figures for these series. Further, any data that seem to be out by a factor of 1000 will be corrected as well, after confirmation by the country. To improve data quality, validation rules will be developed that can detect these issues, by using the number of goods vehicles as a benchmark for total load capacity.

III.   New vehicles registered versus new registrations

7.   The Glossary for Transport Statistics[1] defines the date of first registration of motor vehicle as follows (emphasis added):

"The date of first registration of a motor vehicle is the first-time registration of the vehicle **as new** in a motor vehicle register, irrespective of the nationality of the register. The dating of the registration is the date on which the registration was recorded at the motor vehicle registration office. The **registration of an imported second-hand vehicle is not a first-time registration** but should be regarded as a reregistration."

8.   Despite this definition, it is known that for some countries, new registrations do actually include second-hand imports. In the common questionnaire metadata for example, Greece, Hungary and Lithuania are all flagged as including used vehicles from abroad or re-registered vehicles as new registrations. It is worth noting though that many countries may have misunderstood the definition, and/or not provided metadata documenting their differences from the definition.

9.   New registrations of vehicles are an important economic indicator in their own right. In addition, the breakdown of these data by fuel are an important leading indicator to know the future development of the fleet. For example, the breakdown by fuel type allows an assessment of how quickly Battery Electric Vehicles (BEV) are populating the fleet, important to know for climate change reasons as well as local pollution, electricity demand and charging infrastructure planning.

10.   There is a good case for data collection covering both cases. All new registrations (including second-hand imports) show how the fleet is changing, while new vehicles registered shows what type of completely new vehicles consumers are buying from forecourts.

11.   Therefore, Eurostat analyzed the new vehicles registered data in the common questionnaire against two separate sources. The first was the European Alternative Fuels Observatory[2], which is a "European Commission funded initiative which provides open and free information to support Member States with the implementation of EU Directive 201/94

---

[1] unece.org%2FDAM%2Ftrans%2Fmain%2Fwp6%2Fpdfdocs%2FGlossary
    _for_Transport_Statistics_EN.pdf&clen=2764597&chunk=true.

[2] https://www.eafo.eu/.

on the deployment on of alternative fuels infrastructure". The second source was new passenger cars (registered) taken from the European Central Bank (ECB).

12. For new passenger cars, the Eurostat analysis (against the ECB data) shows that 24 countries (with data in the European Central Bank database) seem to be fully compliant with the Glossary definition. Eleven member States either deviate seriously from the Glossary definition or provide on their national statistical website multiple different notions of "new registrations" but have not selected the right one in the common questionnaire (Cyprus, Croatia and Malta). Ten other countries do not provide any data.

13. For BEVs, analysis compared common questionnaire data against the values from EAFO. For BEVs a challenge is that Glossary definitions are somewhat simplified between e.g., Battery-Only Electric Vehicle and Plug-in Hybrid Electric Vehicle, whereas some countries consider separately a Range-Extended Electric Vehicle, which can contain a small petrol motor. For 14 countries the data seemed to completely agree with the EAFO. Seven other countries provide figures in the common questionnaire differing significantly from EAFO, which is likely due to the concept of new vehicles including second-hand imports (as above). Five further countries have significant differences with the EAFO data yet to be explained, and 21 countries did not have common questionnaire data (nine of which were included in the EAFO database). 21 countries did not have data for this comparison.

14. The partner organizations propose to reaffirm the Glossary definition for new vehicle, only including vehicles registered for the first time and excluding second-hand imports. This can be achieved through outreach to the countries, as it should be possible for most countries to have a Glossary-compliant definition available. In addition, ensuring that the definition of BEVs is fully compliant with the Glossary and does not include certain types of hybrid vehicle, should further improve data quality.

15. A final possibility on improving common questionnaire data would concern using the EAFO data directly for relevant member States, with a full citation or footnote making it clear that these are sourced from another organization rather than a national statistics contact. This would only be undertaken when common questionnaire data are not available for a country, and with the EAFO data first shared with the national statistics contact for their verification and approval.

IV. Rail lines and tracks

16. This area of data quality principally concerns countries covered by European Union Agency for Railways (ERA), namely European Union countries, Norway and Switzerland. In the European context, it is not always clear which data source (common questionnaire versus ERA data, as well as data from the International Union for Railways (UIC)) should be used for reference in certain publications. Further, some data in the common questionnaire seem to disagree internally, in that the length of lines is greater than the total length of tracks (each line is made up of one or more tracks so this should not be possible).

17. Seven UNECE countries provided lengths of tracks equal to their lengths of lines in 2019, despite the fact that they also say that some lines are double tracked or greater. These countries are Bosnia and Herzegovina, Belarus, France, Republic of Moldova, Russian Federation, Slovakia and Ukraine. Conversely, North Macedonia and Montenegro report different lengths of tracks and lines and yet all of their lines are declared to be single track.

18. The partner organizations propose that a new plausibility check is conducted to ensure that track length is always greater than or equal to line length (in practice track length should always be strictly greater than line length unless all lines are singular).

V. Road tractive vehicles

19. Eurostat has noticed differences between the common questionnaire data and the "statistical population of tractive vehicles" information received every quarter through their regulated collection of Road Freight Transport Survey (RFTS) results (based on EU Regulation 70/2012). For the majority of countries, common questionnaire figures for fleet

populations are typically 50% higher than in the RFTS. These differences may be due to the sampling scheme of the RFTS, due to certain age ranges of vehicles for example. It may be that countries are deliberately not sampling older vehicles because it is known that they are driven less (and so this difference in statistical populations will not significantly affect vehicle-km and tonne-km figures).

20.   The partner organizations propose that countries with a very large difference between the two datasets (Bulgaria, Italy and Ireland) should check these datasets and verify the differences or rectify them as appropriate.

VI.   Road traffic of motor vehicles

21.   The road transport vehicle-kilometres chapter was developed several years ago as a pilot collection, from the 2008 year onwards. When this occurred, no attempt was made to streamline the existing road chapter, resulting in ten repeated indicators (despite slightly different language used between the two chapters).

22.   The partner organizations propose that given the overlap between the two datasets, when data are provided in one chapter but not the other, the data are to be repeated (as the different chapters of the questionnaire feed into different dissemination cubes).

VII.   Concluding remarks

23.   The partner organizations will continue their efforts to improve data quality, in continuous liaison with member States. Statistical contacts are reminded of the importance of where possible following definitions from the Glossary, and of including metadata descriptions in all cases where this is not possible. Further, many common questionnaire contacts are also data users themselves, and sharing any observations on quality with their respective organization allows more of these kinds of analyses to be conducted.

―――――――――