Economic and Social Council

Economic Commission for Europe

Inland Transport Committee

Working Party on Rail Transport

Seventy-first session
Geneva, 27-29 November 2017
Item 17 of the provisional agenda
Rail Statistics

Promotion of Data Quality in Rail Statistics

Note by the secretariat

I. Background

1.   Rail transport for both passengers and freight is a part of daily life for people across the ECE region and beyond. With a seemingly ever increasing availability of data, transport policymakers should be able to answer many of the vital questions facing the future of rail transport. For example, how does the utilization of rail networks differ between countries? How does the extensiveness of rail networks in countries correspond to the transport of goods and passengers?

2.   To answer these and other questions, users have many options, including big data collected by many private organizations. However, the most comprehensive source at the country level remains the traditional collection official statistics by national statistics offices or transport ministries. These statistics typically represent data from the entire country rather than only a subsection of the country (e.g. those using a particular rail company). Work remains to fill the still present gaps in these data as statisticians continue to work to improve the comprehensiveness and comparability of these data.

3.   To help to address these issues and encourage communication between official statisticians on data collection methods, in early June 2017, the United Nations Economic Commission for Europe (UNECE) held a workshop entitled Promoting Data Quality in Rail Statistics: Sharing Country Practices. The workshop was the first in the UNECE series on transport statistics and was convened in conjunction with the annual session of the UNECE Working Party on Transport Statistics. Presentations were given by national statistics offices, transport ministries and international rail organizations. A discussion between all workshop participants concluded the workshop.

II. Workshop on promoting data quality in rail statistics

4.   Challenges in data collection were a major point of discussion for many presenters. For many countries, statistics on rail are a combination of administrative data directly reported from rail operators and surveys sent to rail operators or sometimes railway users. Unfortunately, the collection of these data is not always consistent and may lack the comprehensiveness required to report the data received as official data. One can note this lack of data through the international databases that are maintained by several organizations. At the workshop, the European Rail Agency (ERA) in particular reported on the low data availability levels in some of the countries in its region (see Figure 1). While some countries have provided nearly all data required in the database of ERA, many countries lag behind.

Figure 1
Network and parameter data completeness in European Rail Agency database
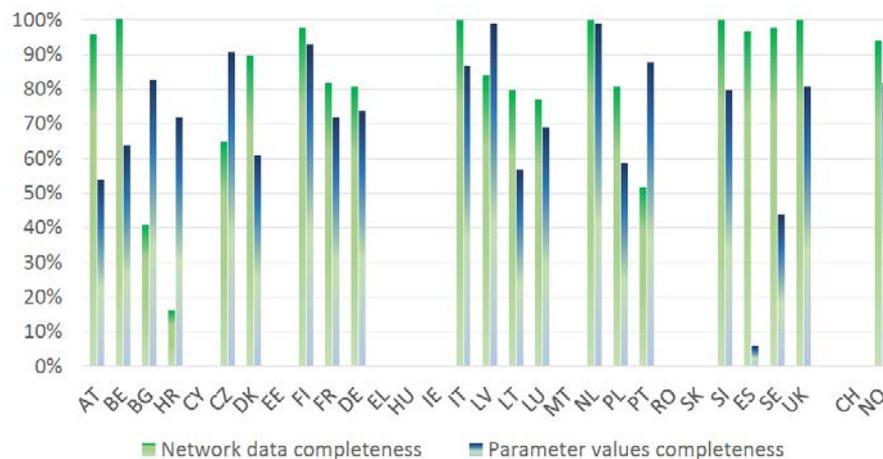
Source: European Rail Agency

5.   Often these data gaps reflect difficult problems that statisticians face in collecting data. The Netherlands Central Bureau of Statistics, in particular highlighted the challenges associated with market liberalization of the rail sector. The number of rail companies is increasing in many countries due to privatization (see Figure 2 for the example of the Netherlands). While many view this as a step toward increasing transport efficiency, private operators typically have an increased desire for confidentiality and often operate across borders (particularly in the European Union). This adds complexity for statisticians both from the perspective of both data collection and data processing where data must be disaggregated by country in the data validation stage. For its part, the United Kingdom of Great Britain and Northern Ireland noted the increased complexity in collecting freight rail data due to the difficulty in establishing memorandums of understanding (MoUs) with so many different private enterprises, although through goodwill they typically maintain a high response rate.

Figure 2
Number of rail companies active in the Netherlands
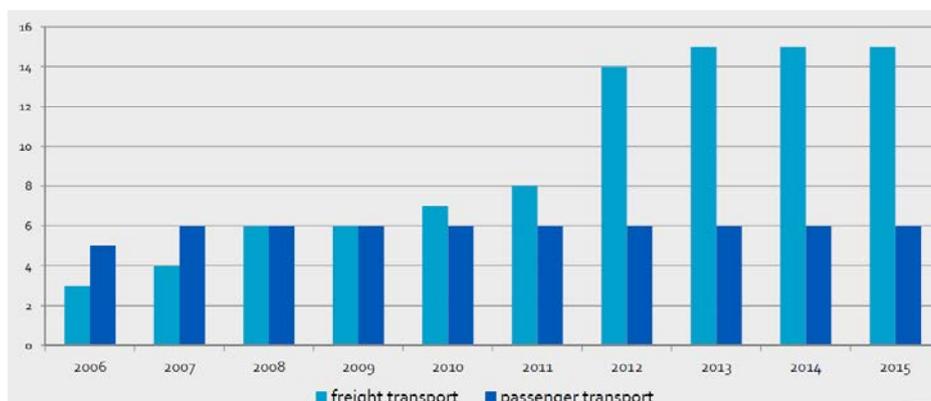
Source: Netherlands Central Bureau of Statistics

6. Despite these obstacles, many statistical offices are able to provide valuable data to policymakers and other data users. JSC Russian Railways demonstrated the sometimes dramatic changes over time in rail loading, notably around the 2009 financial crisis where there was a large decrease followed by a slower recovery in subsequent years. As such, the data can be used as a leading economic indicator. Another example is the data reported by International Union of Railways (UIC) on the change in the coverage of railway lines since 1970 in Europe. Figure 3 shows the decrease in the total length of lines in Europe between 1970 and the early 2000s followed by its stabilization in the last decade.

7. Participants highlighted innovative method*s* throughout the workshop, including the use of data from "train spotters" in the Netherlands. These data, collected by amateur train enthusiasts and often published for free on the internet, are sometimes used in the country for validation of train composition and commodities types. A more typical modernization effort is also ongoing in the United Kingdom where national statisticians at the Office of Road and Rail (ORR) are developing a web form to facilitate the transmission of data by rail operators. Once implemented, this method will streamline data collection efforts as well as reduce the burden on smaller operators by reducing the data required in previous questionnaires.

8. Developing collaborative relationships with rail operators is another step many statistical offices are taking to improve their statistics on rail. In the Netherlands, agreements are in place with some private operators to provide data directly to the statistics office. In the United Kingdom, the ORR has a detailed MoU in place to require the provision of passenger rail data through the country's passenger rail operators (Network Rail). The agreement includes the delivery medium, type of data to be received and the format for data supply. By reviewing this agreement each year, the relationship remains up-to-date and ensures that business needs are met.

9. Relationships such as these also help to harmonize data collection and can greatly improve the timeliness of data. With policymakers placing increased importance on making informed decisions quickly, this is another factor that data providers must consider when developing and updating their data collection and dissemination methods. Statistics Canada noted their own efforts in this area in publishing data on monthly railway car loadings less than two months after the reference period. The United Kingdom has time scales defined in their memorandums of understanding in order to feed directly into their publication schedule as efficiently as possible.

Figure 3
Total length of rail lines in Europe (differences relative to 1990)
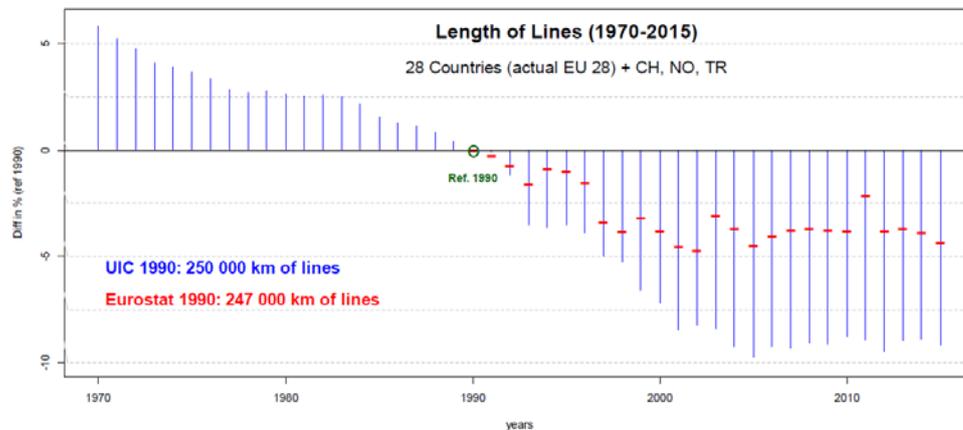
Source: International Union of Railways

10**.** Several participants and presenters noted a lack of data consistency between international or regional organizations such as UIC, ERA, Eurostat and UNECE among others. UIC showed that, at a region-wide level, small to moderate differences persist between data published by international organizations (see Figure 3 for a UIC and Eurostat comparison of total lengths of rail lines in Europe). Different organizations may use divergent methods for estimation, have slightly different definitions or collect on the different timelines. This can create confusion for data users. UNECE, Eurostat and the International Transportation Forum (ITF) collaborate continuously on harmonizing definitions and use a common web questionnaire to both increase consistency between agencies and decrease the response burden on their members. While these efforts have improved data consistency over recent years, much work remains to be done, with a major step to be undertaken in the coming years with the next update of the joint ITF/Eurostat/UNECE publication of the Glossary for Transport Statistics. This update will be a global effort and incorporate some of the newer innovations in transport across all transport modes, including rail.

III. Conclusions

11. The June workshop was a useful step in increasing the sharing of information between different rail data providers and users and it should be viewed as part of ongoing conversations that UNECE will continue to foster on this topic. UNECE has plans with the aforementioned update of the Glossary for Transport Statistics and with the activities of other relevant Working Parties to continue to promote communication and collaboration between countries on the data collection of rail statistics.

———————