18

WORKSHOP B6 REPORT: THE ACQUISITION OF LONG-DISTANCE FREIGHT DATA

Kara Kockelman, The University of Texas at Austin, USA
Michael Browne, University of Westminster, London, UK
Jacques Leonardi, University of Westminster, London, UK

The following paper is a pre-print and the final publication can be found as Chapter 18 in Bonnel, P., M. Lee-Gosselin, J. Zmud and J.-L. Madre (eds.) *Transport Survey Methods: Keeping Up with a Changing World*, 2009, Emerald Group Publishing, U.K.

Keywords: long distance freight; freight surveys; survey methods; vehicle utilization; supply chains; freight data

INTRODUCTION

The acquisition of long-distance freight data workshop discussions focused on the strengths, limitations and uses of existing data sets, both long established and more recent, innovative endeavors. Twelve participants engaged in an active dialogue, quickly filling the workshop's eight hours. Starting from the review of available data collection methods, different type of survey examples were assessed, and an analysis of possible solutions to overcome current data gaps was provided. In the discussion, the workshop was mainly devoted to organisational and practical questions on how to better fulfil the policy and company needs for long-distance freight data. The key issues were the main purpose for data collections, the differences and similarities in survey methods (mechanisms), the sampling, and last but not least the motivation for companies to participate in freight surveys. In the following workshop report, the structure of this discussion is maintained. It is starting from the presentation of the actual status, continuing with the identification of major gaps and challenges, further progressing with findings on some existing or potential opportunities, and concluding with the suggestions of innovative solutions.

KEY TOPICS

Freight transport data collection was clearly a priority of the ISCTSC conference, as this workshop complemented an earlier workshop which focused on intra-urban truck movements. Key topics of discussion in this latter, long-distance freight workshop included survey needs and uses, optimal methods of data acquisition, status of current freight surveys around the world, and sampling issues. Major public- and private-sector ***motivations for freight transport data collection*** include an appreciation (and forecasting) of the following:

- Traffic loads for operations management and infrastructure investment (including pavement maintenance, capacity expansion, etc.),
- Trade linkages and economic interactions across firms, industries, regions and nations (to anticipate opportunities for enhancing such linkages and facilitating efficient operations),
- Vehicle utilization (including driver and fuel use, and capacity utilization by weight and volume),
- Logistical operations (including scheduling, trans-shipment activities, intermodality, and other supply-chain linkages),
- Safety in freight movements, and
- Other areas for policymaking and planning, including the introduction of new technologies (e.g., new vehicle designs), regulation (e.g., border control, fuel economy standards and driver rest policies), and pricing (via gas taxes, road tolls and registration fees, both fixed and variable).

A greater appreciation of these many variables across shippers, carriers and commodities enables an evaluation of firm- and system-based performance, competitiveness, costs, benefits and environmental impacts. This includes opportunities for new, more sophisticated and hopefully more robust modelling efforts, thereby facilitating forecasting traffic, revenues and other impacts, for more comprehensive evaluation of transport and trade policies, including service provision and energy standards. Policymakers, transportation planners, and regulators stand to benefit greatly from access to detailed data sets on freight transport, assuming sampled movements fairly represent the travellers of interest.

In recent years, multiple ***mechanisms for data collection*** have been added to the set of established freight transport surveys, and this variety was well presented and discussed during the workshop. Unfortunately, none of the existing survey methods is presently without limitations. Past and present approaches for data acquisition include the following:

- Vehicle-based surveys, such as EUROSTAT's extensive, continuous survey for EU countries, and its intra-national complements, including:
    - the UK's Continuing Survey of Road Goods Transport (CSRGT) (DfT, 2009),
    - Spain's Encuesta Permanente Transporte Mercancías Carretera (EPTMC) (Pérez-Martinez, 2008),

  - France's TRM and SiTRAM surveys (SESP, 2008),
  - U.K.'s Key Performance Indicators (KPI) survey (McKinnon 2007),
  - and the U.S.'s recently terminated Vehicle Inventory and Use Survey (VIUS)(U.S. Census Bureau, 2003),
- Shipper-based surveys, such as the U.S.'s Commodity Flow Survey (CFS) (U.S. Census Bureau, 2009),
- Carrier-based surveys, such as the DLR freight company survey presented by Julius Menge at the COST355 final conference in Annecy (Menge and Lenz, 2008),
- Supply-chain interviews, data collection and case studies, as presented at COST355 final conference by Michael Browne (Browne et al., 2008),
- Roadside interview surveys (rather common throughout the world, including a four-nation collaboration for long-distance freight crossing the Alps and Pyrenees [Houée, 2008]),
- Shipment-based surveys (where an item is tracked using electronic tags or other information from point to point, start to finish, like the French ECHO survey[1], and
- Other, often existing, complementary opportunities, such as transponder tag reads on an instrumented (typically tolled) highway system, tax data required on transactions, and customs data (on imports and exports).

Discussions noted the rise in third-party logistic firms (3PLs), to manage shippers' transport needs. Such firms can be key in obtaining proper data, and may store a great deal of the needed details. Of course, ***sampling and other issues*** emerge, with the rise in 3PLs, concerns about privacy protections, sample coverage, burden and non-response. For example, to avoid any opportunity for unique identification of a firm's trading, the CFS shipments are bundled by state or super-region, resulting in dramatic geographic-information losses. Commodity flow data at the county or small-region level (with industry-level aggregation, to preserve anonymity of responses) would allow the analyses that modelers envision while providing the information base decision-makers truly need for infrastructure planning and the application of new policy. In addition, survey scope definition is critical in determining sample frame and questionnaire design. For example, are service trips in company vehicles considered as freight or as passenger transport data? Are public sector vehicles included? How about pipeline transport?

Another key issue is the ***motivation of company and decision makers for participation*** in data collection. For example, there is a need to maintain interest in the U.K.'s Key Performance Indicators (KPI) surveys (McKinnon 2007), by providing indicator values directly back to responding firms, thereby supporting the within-firm competitiveness evaluation process. In time-consuming, repeat government-required surveys, it was noted that some firms are seeking to pay the fine for non-response. In this volume, McKinnon and Leonardi's resource paper notes that support of a major trade association, the backing of respected industry champions, and a continuing relationship with sampled firms are key ingredients of a relatively high response rate. One key conclusion of the workshop is that the data providers need to

[1] ECHO stands for Enquête CHargeurs Opérateurs. Undertaken in 2003/2004, this French Shipper Survey sampled 3,000 companies and monitored roughly 10,000 shipments – from sender to final receiver, across vehicle types, illuminating entire supply chains. Non-road mode shipments were over-sampled, to ensure higher confidence on related estimates.

understand the value of aggregate data sets, ideally in a personalized way (e.g., firm-based KPIs tabulated relative to competitors). Reporting is burdensome, and, unless firms are willing to allow electronic identification of their vehicles and shipments or share their entire delivery data base with researchers or a third party (who can "scrub" identifying information and tabulate needed values), it is likely to remain burdensome.

PRESENTED PAPERS

The workshop was informed by a resource paper, complemented by a discussion paper, two contributed papers and stimulating dialogue. Highlights of these papers are summarised below.

Titled "The collection of long-distance road freight data in Europe", the resource paper (presented by Jacques Leonardi and included in this volume) emphasized data requirements for long-distance road-based freight transport. The authors described various data collection systems used for statistics, models, market and scientific studies on long-distance road-based freight transport. They presented first a list of what would be required in an ideal data set, to enable high-quality research while providing reliable guidance for freight operators and policy makers. They then introduced data collection systems emerging in Europe, and discussed some essential shortcomings that need to be addressed. A key criticism relates to the limited number of freight indicators collected in national surveys across Europe. Such feedback is not consistent with the freight market's complexities and does not address various important policy questions. Finally, the third part of their resource paper evaluates the strengths and limitations of alternative approaches to collecting and analysing road freight data. (McKinnon and Leonardi 2009)

Gaps include a lack of detail on vehicle fuel consumption and cargo-space utilization, and on intermodal movements. The paper ended by proposing that there is considerable value to be gained from adopting a KPI-style data collection approach, enhanced by lessons learned in the KPI process. These lessons include involving major professional organisations (to help promote participation in the data collection process), showing sampled firms past survey outcomes and derived analyses, and demonstrating how such information can complement company decision-making. One particularly valuable option is the benchmarking of company performance values, for comparisons within sectors along with potential explanations for differences, focused on improving the freight transport system and fuel efficiency.

Michel Houée's paper (titled "The progressive elaboration of a multinational harmonised database for freight transit through Alps and Pyrenees") highlighted the benefits of adopting a common data approach across France, Spain, Switzerland and Italy. The paper identified limitations in such surveys (including a lack of roadspace for stopping fully one-percent of all trucks at certain times of the day and year, and the need to abandon some questions such as

initial origin and final destination of goods[2]) and scope for improvements. These include the combination of such results with data on maritime transport (which serves as a substitute mode to the truck and rail data obtained) and inclusion of better data on passenger transport in the region (an emerging issue, as congestion in these alpine corridors grows). The current survey takes place every five years – in future by combining data from other sources more frequent updates may be possible.

Pedro Pérez-Martínez's contributed paper (titled "The vehicle approach for freight road transport energy and environmental analysis in Spain") used 1997-2003 EUROSTAT data[3] for analysis of freight transport's carbon footprint, by vehicle type, across Spain. The work reviewed several key indicators of vehicle efficiency and performance. The EUROSTAT data are based on a stratified random sampling, and over the 7 year period studied, represents an audit of 41,600 vehicles per year (each over the course of a week). Although the study data exhibit wide variation in absolute values between 1997 and 2003, energy requirements per ton-kilometer fell just 0.2% over the seven-year period. The analyses suggest new approaches for assessing vehicle utilization and fuel efficiency, along with the result that larger commercial trucks are delivering far more ton-kilometers per gram of carbon dioxide than smaller trucks.[4]

STATUS OF LONG-DISTANCE FREIGHT SURVEYS

Dialogue regarding the status of long-distance freight surveys took as its starting point Christophe Rizet's discussion notes. These focused primarily on the comparison of the EU's Heavy Goods Vehicle (HGV) vehicle-based survey and the U.S. CFS (a shipper-based approach).

The CFS started in 1954 and targets shippers. Europe's various national HGV surveys began with France in 1952, and were harmonized over several decades. The CFS is undertaken every five years, in league with the U.S. Economic Census of businesses, while the EU HGV is continuous in nature. Both rely on mandatory reporting requirements of goods shipments, but only the EU HGV requests information on actual price paid (in for-hire transport). Over the years, the CFS sample size has varied between 50,000 and 200,000 establishments (each assigned four one-week reporting periods), while the HGV obtains a sample of 85,000 vehicles (for an entire week of use).

Each survey exhibits various strengths and limitations. The HGV survey offers better information on vehicle activity and utilization (including flows and load factors), but nothing

[2] True origins and final destinations are often not known by drivers, and will be reported with error. Related to this, detailed route information also is difficult to obtain, due to roadside-interview time constraints and driver memory limitations.

[3] Spain satisfies EUROSTAT data requirements via its Encuesta Permanente Transporte Mercancías Carretera (EPTMC) survey.

[4] Estimates of tkm per kg of CO_2 are 14 for trucks over 26.1 metric tonnes versus 3.0 for those in the lightest weight category (3.6 to 7.1 tonnes). Of course, energy efficiency is influenced by several factors, including engine technology and transport optimization strategies.

on transport chains or shipper activity levels. There is no way to identify the true origin and destination of HGV cargos, and there is nothing on non-truck modes (such as rail, pipeline, air and maritime transport). Advantages of the CFS include mode choice and some view of transport chains (though these are often inferred by third-party analysts, at Oak Ridge National Laboratory). However, there is no route information or supply chain details across shippers. Completing the CFS survey may be viewed as a burden by respondents (possibly influencing firms' selection of reported shipments) and there are some trips omitted (for example, links to farms as a source of raw materials/foods).

Possibilities exist to complement both types of surveys. In the case of the CFS it would be feasible to perform carrier surveys with the goal of enhancing inference of route, intermodality, trip chaining and other trade-system attributes. In addition, smaller update surveys could be completed to enhance data timeliness. In the case of the HGV survey (and specifically in France) it is possible to include data on other modes (by means of the SitraM database[5]).
In addition, shipper surveys were carried out to improve the knowledge of transport chains and the links with the economy.

Of course, both types of survey provide complementary information, and there is no winner here. Workshop participants support a policy of developing and delivering both types.

KEY CHALLENGES

As suggested, no single survey can do it all. Workshop participants commented several times on the importance of identifying the key questions of interest before selecting a sample frame and sample size. Major gaps and challenges were identified, as follows:

1. *Tying survey purpose to instrument & frame:* It was agreed that there was a need for greater clarity and precision in defining the purpose of the surveys discussed. The various surveys discussed clearly serve different purposes, so approaches and frames will/should differ. Some surveys (such as the EU's HGV survey and the U.S.'s CFS) may be able to serve several purposes, and researchers can use them for a range of analysis. Nevertheless, the group felt there is a need for caution when using some of the results for forecasting and other initiatives. For example, there is little evidence that improvements of load factors in one country or in one sector will be transferable to another. The details of business operations and product designs often hinder the simple transfer of specific solutions.

2. *Variety:* There is great variety in the types of questions one may ask and uses to which a data set may be put. There is great variety in the types of commodities transported,

[5] The SitraM database (Système d Information sur les TRAnsports de Marchandises) was created by the French Ministry of Transport in 1975. It contains data on transport flows between every pair of districts in France, in tons, by road, rail and inland waterway, and at NST 3 level. Coastal shipping, pipelines and air transport are not included (SESP, 2008).

and the drivers, vehicles and modes that move them, along with involved industries, locations and routes. Freight surveys are profoundly affected, in terms of the range of companies that need to be surveyed and the complexity of the interactions of decisions (e.g., mode choices, trans-shipment points, route choices and trip chaining). The different and evolving responsibilities within a supply chain all have implications for the survey design, robustness of the results and subsequent uses of the data sets obtained.

3. *Missing information:* Participants agreed that current gaps in freight transport information are substantial, despite the scale of some of the existing surveys. Among the most important missing elements are values and volumes moved, shipper costs and prices by mode, international and intermodal ties, use of public vehicles, vehicle fuel consumption, use of small vans for long distance trips, the impact of new technologies, investments, policies and organizational structures on performance, and the costs and benefits of different policy options.

4. *Speed of change:* Technology and supply chains, both local and global, are evolving rapidly. Within a region or nation there may be stability in terms of the total freight flows (volumes and distances); however, as a result of changes in company sourcing and/or company logistics and supply chain strategies, there may be dramatic differences in the freight flow patterns over the network. The speed of such change has increased with rising globalisation, and there are many implications for data acquisition, assembly and distribution, including frequency of surveys, and timeliness in data delivery. There also is an evolution in the technologies available for surveys (such as global positioning systems [GPS] and global system for mobile communications [GSM]), which should change the design of surveys.

5. *Survey continuation:* The desirability of maintaining data collection over the long run was discussed by workshop participants, with significant benefits to be gained, such as robust and consistent trend comparisons. Nevertheless, many nations are experiencing an increased reluctance by government departments to meet the costs of surveys and to burden industry with mandatory (statutory) surveys to complete. For example, the U.S.'s Vehicle Inventory and Use Survey (VIUS) was recently scrapped. One solution may be great diffusion and use of automatic data processing devices, on-board and in companies, for continuous, low-marginal-cost data collection.

6. *Data linkages:* The need for and importance of linking survey results effectively across nations, levels of spatial and industry detail, carriers and shippers, modes and acquisition methods seems clear. Workshop participants agreed that survey methods should seek to facilitate such linkages, in order to exploit the complementary nature of freight surveys at different levels, using different approaches. In particular, many data sets can be enriched by acquisition of data through new technologies, such as roadside toll-tag readers, GPS records, and other types of on-board and off-board devices. However, privacy concerns, cultures and customs tend to slow such adoption.

OPPORTUNITIES AND RECOMMENDATIONS

Many avenues exist for enhancing long-distance freight data set acquisition, and participants chose to highlight the following:

Complementary data sets: Most EU countries appear to be lacking establishment-based data sets and records of non-truck shipments, while the U.S. is now lacking carrier- or vehicle-based surveys. Whatever new data are obtained, there generally is scope to use complementary approaches, thereby enriching existing data sets. However, in order to do this, pilot initiatives are needed, to show what can be done and what is cost-effective (and robust, in research terms), along with a multinational research exercise comparing the establishment- and vehicle-based survey designs and results.

A more global view, up & down the supply chains: By not recognizing complete supply chain linkages, from source (of raw materials) to consumers, we run the risk of collecting much data but having too limited an understanding of critical variables (such as shipment scheduling, trans-shipment decisions, and ultimate destinations). Increasingly, transport decisions are taken in a logistics and/or supply chain context (down to the retail store or end user via home deliveries), where the decision is influenced by organisations upstream and downstream of the intermediate decision-makers. There needs to be a way to acquire more information on these upstream and downstream influences and cross-actor linkages. Moreover, final consumers ideally should be made aware of their purchasing decisions' implications (e.g., carbon impacts) by providing supply chain and freight transport information through to their home site.

Sharing instruments & harmonizing data sets: Workshop participants agreed that major opportunities exist in harmonizing future survey instruments and sharing existing instruments. This includes formal agreement on and consistency in terminology (such as a glossary of terms, for researchers and respondents) and greater collaboration across agencies and countries.

Mechanisms for protecting confidentiality while releasing useful results & products: Given the wealth of data that may become available from new technologies being widely implemented in the transport and logistics arenas (e.g., electronic toll collection, high-resolution satellite images, and GPS on-board units) there is an urgent need to find appropriate and robust ways in which commercial confidentiality can be assured. At present data that could be valuable in a policy context is not being released by companies because of privacy concerns. Use of a third-party intermediary to scrub identifying information and prepare data in accordance with analyst and agency needs may be highly desirable and merits thoughtful examination.

Demonstrate value & use of data collected, educating stakeholders and students: Funding, coverage, data quality and response rate issues could all be addressed, to some extent, by demonstrated value in data collected. Researchers and other users need to become better at

marketing the benefits of their work with long-distance freight survey data. At present many companies view data provision as a burden and do not appreciate the benefits of improved transport policymaking and infrastructure management decisions. The information and other benefits that result from the analysis of freight surveys need to be communicated to the many stakeholders and to students in a more proactive and timely way. Researchers need to ensure that students on transportation programmes (who represent the next generation of freight data collectors and users) are also familiar with the scope for improvement and the value of high quality and insightful data collection activities. By doing this the research community will strengthen the willingness of companies to participate in the surveys exercises.

SUMMARY AND CONCLUSION

To summarize, the workshop participants and transportation community at large look forward to future research activity and data acquisition innovations in the freight transport arena, both long- and short-distance, across modes and nations, to facilitate economic analysis, environmental policy, transportation system management and the like. There was great benefit in comparing and sharing information and ideas within the workshop's international group, and the ideas and suggestions emerging from our session will inform future freight-related surveys. Heavy-duty truck transport accounts for upwards of 15 percent of a nation's vehicle-miles travelled and most pavement damage. Both trucks and trains congest tunnels and track, and many seaports and airports regularly reach capacity. Freight transport consumes roughly 10 percent of a developed nation's petroleum, resulting in a significant carbon footprint along with other, regulated emissions. Good data are needed now, for modelling, planning, and policy-making.

REFERENCES

Browne, M., C. Rizet, J. Leonardi, J. Allen (2008). Analysing Energy Use in Supply Chains: The Case of Fruits and Vegetables and Furniture. In: *Supply Chain Innovations: People, Practice and Performance Proceedings of the Logistics Research Network Annual Conference 2008 University of Liverpool, UK, 10th-12th September 2008* (A.C. Lyons, ed), pp. 395-401

DfT - Department for Transport (2009). *Road Freight Statistics 2007*. London. Available at http://www.dft.gov.uk/pgr/statistics/datatablespublications/freight/goodsbyroad/roadfreightstatistics2007

Houée, M. (2008). The progressive elaboration of a multinational harmonised database for freight transit through Alps & Pyrenees. Presented at the 8th International Conference on Transport Survey Methods, Annecy, France, May 2008

McKinnon, A.C., Leonardi, J (2009). The Collection of Long Distance Road Freight Data in Europe. In this volume.

McKinnon, A. C. (2007). Synchronised Auditing of Truck Utilisation and Energy Efficiency - A Review of the British Government's Transport KPI Programme. In: Presented at the World Conference on Transportation Research, Berkeley, California. Available at http://www.greenlogistics.org/PageView.aspx?id=150

Menge, J., Lenz, B. (2008). Company survey approach. Paper presented at the COST355 final conference, in Annecy, France.

Pérez-Martinez, P. (2008). The vehicle approach for freight road transport energy and environmental analysis in Spain. Contributed Paper presented at the 8th International Conference on Transport Survey Methods, Annecy, France, May 2008

SESP - Ministère de l'Écologie, de l'Energie, du Développement durable et de l'Aménagement du territoire (2008). *Enquête transport routier de marchandises (TRM) and SiTRAM (all modes).* Paris. http://www.statistiques.equipement.gouv.fr/rubrique.php3?id_rubrique=33.
U.S. Census Bureau (2003). *Vehicle Inventory and Use Survey*. Available at http://www.census.gov/svsd/www/vius/products.html

U.S. Census Bureau (2009). *2007 Commodity Flow Survey Advance Data Files* Service Sector Statistics Division, Commodity Flow Branch. Available at http://www.census.gov/svsd/www/cfsmain.html

Participants

Mike Browne (U.K.)
Bastian Chlond (Germany)
Michel Houée (France)
Kara Kockelman (USA)
Jacques Léonardi (U.K.)
Mariame Mbacké (France)
Alan McKinnon (U.K.)
Julius Menge (Germany)
Arnim Meyburg (USA)
Peter Ottmann (Germany)
Danièle Patier (France)
Pedro José Pérez-Martínez (Spain)
Christophe Rizet (France)