

## COVERAGE AND QUALITY PROBLEMS WITH EXISTING DATA RESOURCES FOR FREIGHT TRANSPORTATION

Moderator: T. Q. Hutchinson, U.S. Department of Agriculture

---

**Speaker: David Greene, Oak Ridge National Laboratory**

Freight data are needed to determine the extent to which the highway system should be expanded, how well the system is operating now, what costs are involved, and how these costs should be allocated among users and providers.

The basic data needed are truck numbers, truck miles and commodity trips. Within these data elements several breakdowns are needed, such as:

- 1) Truck numbers: configuration (tractor, trailer, single unit), ownership by type of carrier, operating and registered weights, all by type of commodity hauled.
- 2) Truck miles: number of miles driven for each of the above categories.
- 3) Commodity trips: Number and distance driven carrying each commodity type within each of the above cells.

Without considering specific origins and destinations--and some origin/destination (OD) data is needed--160 million cells are required. Considering OD data, 6.4 trillion are needed.

Obviously the data must come from more than one source, and interrelate various data sources. In fact, that is what we do.

Since several sources, several analysts and several models are involved, the national policy analyst must attempt to achieve coordination among surveys and analysts to achieve comparability of data among the several sources.

Oak Ridge National Laboratory has completed a study of the quality and coverage of six sources of truck vehicle miles and numbers. We conclude that the basic source of VMT data, by truck type, state, highway class and weight should be the Federal Highway Administration's Highway Performance Monitoring System and the associated Truck Weight Study Data because the best way to collect such data is to directly sample vehicles operating on the road network. What is needed is for all the states to cooperatively agree to a set of uniform collection procedures and standard definitions of truck type and weight classes.

Truck Inventory and Use Survey (TIUS) is the best source of truck and owner characteristics. TIUS tells us about truck characteristics, such as body type, engine

type, etc., and affords uniform definitions of truck weight classes.

Nationwide Truck Activity and Commodity Survey (NTACS) is not going to give us the comprehensive commodity flow data that some want, but it is going to give us commodity trips on the highway network. One advantage is that it is integrated with TIUS. Current problems are that NTACS does not follow TIUS closely enough in time and, since it samples trucks rather than commodity movements, it cannot tell us all we need to know about intermodal movements.

Using these data sources together--and no single source will do--requires cooperation, coordinating surveys, and data collections to yield comparable data.

**Speaker: Paul Roberts, Trans-Mode Consultants**

There are three kinds of data needed to answer both public-policy issues and the issues faced by management: 1) data on demand for freight movement; 2) data on the supply of facilities and equipment used (which we will examine shortly); and finally, 3) data on the operating entity--financial statements and the like.

Commodity flow data: There are really only five major aspects of all freight movements: what moved, when did it move, from where to where, who moved it and how much did it cost to move it. More detail is needed to answer each of these questions.

*Under what moved:* what was the commodity, how large was the shipment, and what packaging was required?

*When did it move:* Date and times of pickup and delivery.

*From where to where:* Geocoding becomes important.

*Who moved it:* The carrier, the mode, the type of service. The carriers may well not own any transportation equipment and information tends to be proprietary.

*What did it cost:* I am always a little suspect of costing systems, so I am going to say the cost to the shipper is equal to the revenue to the carrier. Subcontractors and/or handling parties can complicate things.

To understand the ability to move freight we need data covering both facilities and equipment. Usually we use a jointly owned network with individual links. We need data concerning these networks, how are they

organized and this requires definitions of the system and its components. We need to know physical attributes, capacity, utilization, age of equipment, condition, etc.

*Operating entities:* These are the organizations that provide the actual freight movements for a fee. You need financial information on their operating characteristics, something about the number of vehicles they own, the number of owner-operators, the number of trips, and miles, the amount of utilization.

Existing data sources are uneven and of differential quality. The Commodity Transportation Survey of 1977 is an antique more than anything else. The truck data area seems to be where the least data is available. The biggest gaps are in truck data, but especially those with no relationship to the ICC. For railroads we lack intermodal data. In the maritime area, operating statistics are very sporadic. A lot of the players are offshore.

The data supply problem can be solved, and at a reasonable price. The answer lies in using lots of small, disaggregate samples designed to answer specific questions.

**Speaker: Rolf Schmitt, U.S. Department of Transportation**

Tremendous unevenness exists in data resources from mode to mode.

In aviation, information is maintained by the federal government for each plane. On the other hand, we do not really know how many trucks there are in the country. We have a good handle on those registered in the states through the Truck Inventory and Use Survey, but government-owned vehicles are excluded (leaving out, for example 99,000 Jeeps owned by the Postal Service).

The quality of data varies all over the map, and it has been affected by some sources of decline over the past few years, and it has been even more affected by some sources of change that have been alluded to earlier today. There are three major sources of decline.

The first big source of change is money. As costs of traditional data-collection programs go up, we run into problems stemming from the federal deficit.

Second, deregulation has taken its toll. Using the aviation example again, aircraft manufacturers who design aircraft based on information that is collected from the regulatory regime and government agencies who use that data for a variety of nonregulatory

purposes managed to salvage their data during airline deregulation.

In trucking, we have had some signs of recent change. The Census Bureau is now collecting through surveys some data once collected through reporting requirements. So, for strategic planning and national scale analysis, we are starting to fill some of the data gaps, but deregulation has taken its toll.

Third, paperwork reductions. Respondent burden is a very real problem. People are tired of filling out long questionnaires, thus, long questionnaires lower response rates. Respondent burden is also a political issue. Under the Paperwork Reduction Act, the Office of Management and Budget (OMB) is responsible for clearance of questionnaires and has insisted on a lengthy review process.

Survey designers tend to ask for everything because the cost of getting the survey through the process is so high. Thus, you end up with a fairly large and complex questionnaire to avoid coming back time after time with little surveys.

A lot has changed in the transportation industry since deregulation, and our data collection programs must not be targeted to the old way in which transportation business was done. For example, intermodalism affects the rail waybill data. A container is moving under a container rate, but what is inside the box is unknown. Freight-of-all-kinds rates also conceal what is moving and contract rates conceal the real freight bill.

The role of third parties also muddies traditional categories. At what point does a customs broker that provides local pickup and delivery stop being a broker and become a trucker? When you have a motor carrier subsidiary of a railroad that is using independent owner-operators, the chances for double counting are horrendous.

The technology we are using to measure things is changing the nature of the game. Once we stopped trucks to weigh them and to ask the driver what was onboard. New automatic vehicle identification and weigh-in-motion equipment can monitor a truck going 55 miles an hour, but not what is inside the truck. So there is better information about the weight distribution of trucks, but not about commodity movements.

Our methods of analysis are also affecting our data needs. Microcomputers run geographically specific models which create a demand for detailed data because local planners can process it. So technology is creating both problems and opportunities for us.

As previously mentioned, uniform definitions are a must. In relying on state furnished data, we learn that all 50 states, Puerto Rico and the District of Columbia have different definitions. To link separate data sources or to be able to overcome differences in definitions of things like, "what is a truck," we need some overlap in the kinds of data we collect.

**Speaker: Frank Smith, Eno Foundation for Transportation**

In the 1950s and 60s there was a lot more data because there was a lot more regulation and under regulation it was collected from everybody whether they liked it or not. Congress did not police the process. Most of the deregulation of data has been administrative, it has not been by Congress.

Who used this data? The regulators, obviously. Investors, obviously, to analyze the financial stability of carriers. Fortunately for them, if they are going to put their money in, they still get a lot of data. Consultants--they, obviously, want as much as they can get, especially if it doesn't cost them anything, if the government provides it. Finally, policy makers. They judge regulatory needs for legislation.

For aviation, adequacy of data is excellent. Massive financial and operating data is available for each carrier. As far as individual commodities, Census data tend to be outdated but are useful as guidelines.

Oil pipeline data: The Federal Energy Regulatory Commission still collects fairly detailed data from all of the carriers. Unfortunately they do not collate or publish them. the Oil and Gas Journal, Oil Pipeline Research Institute, Association of Oil Pipelines, American Petroleum Institute and the U.S. Department of Energy all publish some information.

Rail freight data sources: the ICC collects only from Class I carriers, and the Association of American Railroads(AAR) publishes it. Data on Class I carriers is quite comprehensive.

In the past, regional, shortlines, terminal and switching railroads have been neglected, but the AAR is trying a survey of their own. The non-Class I carriers account for about 10 percent of carloads and about 13 percent of ton-miles.

Truck data is covered pretty well, but most of it is published by the American Trucking Associations. They offer financial data for Class I and II carriers, but omit a lot of the others.

Highway statistics are published by the Federal Highway Administration. Maybe the information could be a lot better, but what is available--registrations, user taxes, and vehicle miles--is very useful.

Finally, water carriers. The best source of date are the Corps of Engineers' reports. These are quite comprehensive. Unfortunately they tend to have a built-in delay, and nobody collects revenue and expense data.