ALTERNATIVES TO PUBLIC DATA SOURCES

Moderator: K. Eric Wolfe, Association of American Railroads

Speaker: William H. Oderwald, ALK Associates, Inc.

I am going to talk about how to better use the data we have through transportation data enhancement and network analysis. ALK enhances the ICC's Carload Waybill Sample by adding mileages and other geographic codings, and also uses data with networks. Networks are a principal representation of transportation problems and opportunities. In essence, data sources can be thought of as the supply and the network as the demand. The network shows you the structures, the sources, the origins, destinations, and whatever, in a coherent way.

Networks allow researchers to not only verify the data through visualization but also analyze it. Without networks, irregularities or errors could slip through in the reams of printouts associated with a large project. To be useful networks must have several attributes. These include (1) transportation attributes associated with each link such as distance or service quality, (2) analysis attributes which allow assignment of traffic to the network--the software, (3) some sort of geographical reference such as BEAs, and (4) timely maintenance of the network.

For example, the ALK highway model contains over 40,000 links at the present time and is anticipated to have over 60,000 by the end of 1989. These links will cover over one half million miles of highway. Each link contains various attributes related to the type of analysis performed, such as recent studies for clients involved routing of hazardous materials. With population densities associated with the links in the network, risk assessments could be made. Studies answering a variety of "what if" questions could also be performed.

In conclusion, though data may be less than satisfactory, the application of some network tools can create some analyses that enhance the information.

Speaker: Joseph B. Riker, Reebie Associates

Reebie Associates has been providing freight transportation consulting services for over 20 years. In the early 1970s, to support our consulting practice, a highly detailed point-to-point data base was established using a combination of then-public sources and proprietary carrier data. From this work, our TRANSEARCH project began. TRANSEARCH in its most basic form is a data base of U.S. domestic freight movements. It is a comprehensive source which provides detailed information on movements among 285 specific market areas. Commodities are defined at the four-digit level of Standard Transportation Commodity Code (STCC) detail and separate volumes are shown for seven modes of transportation. Our current data base contains over 1 million records, with an aggregate volume of over 5 billion tons. Market areas can be defined as BEA's states or other specialized areas. This data base is built on a combination of public and private sources, each of which is individually enhanced and normalized into a common format so that it may be combined into a single data base.

TRANSEARCH has had a very wide range of applications by its users, who represent all segments of the transportation sector of the economy. An important feature of our service is the ability to customize TRANSEARCH--in terms of geographic area, units of measure and shipment characteristics--to meet specific user requirements.

One of our biggest challenges has been to develop and improve TRANSEARCH over time as fundamental changes have occurred in the availability and quality of some of the underlying data sources. The recent trend toward the shrinking of available public sector data has not been--and is not likely to be--offset by additional data provided by carrier organizations. While carriers want data that will help their operations, they do not want to provide a resource that will be helpful to competitors--either within their own mode or in other modes.

Thus, we see the need for increased primary data collection by firms in the information business. These private efforts, however, cannot hope to achieve the efficiency or scope of a well conceived federally-based effort. Instead, a combined effort--with a significant Federal role in primary data collection and private sector firms focusing on customizations/enhancements --would generate the highest quality information at the lowest cost to the user. This in turn, would maximize the use of freight movement data in the analysis of transportation issues.

Speaker: Forrest Baker, Transportation Research and Marketing

Transportation Research and Marketing (TRAM) operates the National Motor Transportation Data Base

(NMTDB). Unlike those presented earlier, it is a homemade, do-it-yourself project.

We interview, in depth, approximately 25,000 long-haul truck drivers yearly. Each questionnaire contains over 40 questions and takes about five minutes to complete. Interviews are taken at 19 locations across the county. The locations were predetermined by inventorying every truck stop in the U.S. in the early 1970s, determining gallonages sold by truck stop and route, and then locating interview sites to cut the heavily-travelled, long-haul, competitive truck routes. We do not generate data on short-haul traffic.

The data base generates approximately 50,000 sets of data annually, which contain movement origin-destination pairs, commodity moved, and trailer type. All data is mileages by point of interview. We include mileage of all of each movement's component segments: the deadhead miles to load the outhaul; the deadhaul between the two loads; the mileages of headhauls and backhauls.

The data is not necessarily commodity-specific, market-penetration data; in fact, in many instances, it is poor data for that use. We generate data on the equipment: the width, length, horsepower, make and model of the tractor, year of manufacture; domicile of the driver; ownership of the equipment; and if the driver is not an owner-operator, was he ever one? Driver profiles are also drawn based on reported age, years of driving, length of time employed with his present company, driving characteristics, hours driven, and miles driven per month and year.

Purchasing habits of the drivers are also defined--where he fueled; how much fuel he bought; how much he spent for food; how he is eating. We find the meal consumption pattern of drivers is the best barometer of the economic welfare of the driver. In the 1970s he was eating four meals a day. Now he is down to less than 2.5.

We have gathered this data since 1977. We have a data set that goes through the period of deregulation. You can watch the rate patterns, you can watch driver compensation, you can watch driver turnover; you can see the age of the equipment being stretched out. The TRAM data base profiles the long-haul trucking industry quite accurately.

Speaker: Alan E. Pisarski, Falls Church, Virginia

A central concern about data collection activities is technical skills that are a requisite for designing and carrying out a program. Large amounts of money and large amounts of logistical skills--the ability to move people, paper, and things around as needed and to make things happen--are clearly also required. These requirements place severe pressures on the types of institutions that conduct and sponsor large scale data collection efforts. It has always been the institutional problems that were the foremost concerns in establishing viable, continuing data series.

As we approach perhaps the most important legislative decisions our country has seen in transportation--including highways, transit, and aviation --the needs are great to be effective and to form viable policies. It would be wonderful to be assured that the best available data and analytical resources are being brought to bear on these multi-trillion dollar public decisions that will leverage even larger private decisions over the coming years. Yet we have no national travel survey to know about passenger flows, no bus survey to know about intercity bus travel, no commodity flow surveys to know about freight flows. Our ability to forecast future flows is even worse. Not a great score card! We approach these major national decision issues in ignorance.

Clearly our institutions have failed us. There has been lack of interest, lack of support, and lack of recognition of need from our public institutions. The decline of regulation permitted a much-needed attic cleaning in our available data programs. While we should not mourn the loss of some of these sources, we should regret the failure to take the opportunity to replace them with better public or private programs. The push to privatize impelled a policy of getting government out of the data business as well. Thus no public institutions have moved forward to fill the gap. We may have to invent those institutions.

Can the private sector fill this void? How do they fit in for the long term? After ten years the answers are still very tentative. Most of our vendors are value-added purveyors of public sources, and they would be the first to support better public sources. If the private sector is going to be a serious provider, then the government, often the main buyer, is going to have to use serious resources as a buyer, or there will be no marketplace for private vendors to sell in.

For better institutions and more resources to come about, the government will have to at least recognize its needs for data to illuminate its own decision, much less recognize any responsibility to others to produce a data rich environment for pubic and private decision-making. Government will have to recognize that better data are a serious part of the national transportation infrastructure for which it has responsibility.