SPECIAL REPORT 276

A Concept for a

NATIONAL FREIGHT DATA PROGRAM

TRANSPORTATION RESEARCH BOARD
OF THE NATIONAL ACADEMIES

SPECIAL REPORT 276

A Concept for a

NATIONAL FREIGHT DATA PROGRAM

Committee on Freight Transportation Data:
A Framework for Development

TRANSPORTATION RESEARCH BOARD
OF THE NATIONAL ACADEMIES

Transportation Research Board
Washington, D.C.
2003
www.TRB.org

Transportation Research Board Special Report 276

Subscriber Category
VIII freight transportation (multimodal)

Transportation Research Board publications are available by ordering individual publications directly from the TRB Business Office, through the Internet at www.TRB.org or national-academies.org/trb, or by annual subscription through organizational or individual affiliation with TRB. Affiliates and library subscribers are eligible for substantial discounts. For further information, contact the Transportation Research Board Business Office, 500 Fifth Street, NW, Washington, DC 20001 (telephone 202-334-3213; fax 202-334-2519; or e-mail TRBsales@nas.edu).

NOTICE: The project that is the subject of this report was approved by the Governing Board of the National Research Council, whose members are drawn from the councils of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine. The members of the committee responsible for the report were chosen for their special competencies and with regard for appropriate balance.

This report has been reviewed by a group other than the authors according to the procedures approved by a Report Review Committee consisting of members of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine.

The study was sponsored by the Bureau of Transportation Statistics of the U.S. Department of Transportation.

Library of Congress Cataloging-in-Publication Data

National Research Council (U.S.). Committee on Freight Transportation Data: A Framework for Development.
A concept for a national freight data program / Committee on Freight Transportation Data: A Framework for Development, Transportation Research Board of the National Academies.
p.cm.—(Special report ; 276)
ISBN 0-309-08570-5
1. Freight and freightage—United States—Data processing. 2. Information storage and retrieval systems—Freight and freightage. 3. Transportation—United States—Planning. I. Title II. Special report (National Research Council (U.S.). Transportation Research Board) ; 276.

HE199.U5N37 2003
388′.044′028557—dc22

2003060678

THE NATIONAL ACADEMIES

Advisers to the Nation on Science, Engineering, and Medicine

The **National Academy of Sciences** is a private, nonprofit, self-perpetuating society of distinguished scholars engaged in scientific and engineering research, dedicated to the furtherance of science and technology and to their use for the general welfare. On the authority of the charter granted to it by the Congress in 1863, the Academy has a mandate that requires it to advise the federal government on scientific and technical matters. Dr. Bruce M. Alberts is president of the National Academy of Sciences.

The **National Academy of Engineering** was established in 1964, under the charter of the National Academy of Sciences, as a parallel organization of outstanding engineers. It is autonomous in its administration and in the selection of its members, sharing with the National Academy of Sciences the responsibility for advising the federal government. The National Academy of Engineering also sponsors engineering programs aimed at meeting national needs, encourages education and research, and recognizes the superior achievements of engineers. Dr. William A. Wulf is president of the National Academy of Engineering.

The **Institute of Medicine** was established in 1970 by the National Academy of Sciences to secure the services of eminent members of appropriate professions in the examination of policy matters pertaining to the health of the public. The Institute acts under the responsibility given to the National Academy of Sciences by its congressional charter to be an adviser to the federal government and, on its own initiative, to identify issues of medical care, research, and education. Dr. Harvey V. Fineberg is president of the Institute of Medicine.

The **National Research Council** was organized by the National Academy of Sciences in 1916 to associate the broad community of science and technology with the Academy's purposes of furthering knowledge and advising the federal government. Functioning in accordance with general policies determined by the Academy, the Council has become the principal operating agency of both the National Academy of Sciences and the National Academy of Engineering in providing services to the government, the public, and the scientific and engineering communities. The Council is administered jointly by both the Academies and the Institute of Medicine. Dr. Bruce M. Alberts and Dr. William A. Wulf are chair and vice chair, respectively, of the National Research Council.

The **Transportation Research Board** is a division of the National Research Council, which serves the National Academy of Sciences and the National Academy of Engineering. The Board's mission is to promote innovation and progress in transportation through research. In an objective and interdisciplinary setting, the Board facilitates the sharing of information on transportation practice and policy by researchers and practitioners; stimulates research and offers research management services that promote technical excellence; provides expert advice on transportation policy and programs; and disseminates research results broadly and encourages their implementation. The Board's varied activities annually engage more than 4,000 engineers, scientists, and other transportation researchers and practitioners from the public and private sectors and academia, all of whom contribute their expertise in the public interest. The program is supported by state transportation departments, federal agencies including the component administrations of the U.S. Department of Transportation, and other organizations and individuals interested in the development of transportation. **www.TRB.org**

www.national-academies.org

Committee on Freight Transportation Data: A Framework for Development

ARNIM H. MEYBURG, Cornell University, Ithaca, New York, *Chair*
PAUL H. BINGHAM, Global Insight, Inc., Washington, D.C.
KENNETH D. BOYER,* Michigan State University, East Lansing
ROBERT COSTELLO, American Trucking Associations, Inc., Alexandria, Virginia
DAVID L. GANOVSKI, Maryland Department of Transportation, Baltimore
J. SUSIE LAHSENE, Port of Portland, Oregon
CATHERINE T. LAWSON, State University of New York at Albany
ROBERT E. MARTÍNEZ, Norfolk Southern Corporation, Norfolk, Virginia
ROBERT TARDIF, Ontario Ministry of Transportation, Downsview, Canada
C. MICHAEL WALTON, NAE, The University of Texas, Austin

Transportation Research Board Staff
JILL WILSON, Study Director

Consultant
RICK DONNELLY, PBConsult, Inc., Albuquerque, New Mexico

*Kenneth D. Boyer chaired the committee from May through October 2002.

Preface

On November 14–15, 2001, the New York State Department of Transportation (NYSDOT) hosted a conference in Saratoga Springs, New York, entitled "Data Needs in the Changing World of Logistics and Freight Transportation."[1] This conference, organized by NYSDOT and the Transportation Research Board (TRB), was sponsored by the American Association of State Highway and Transportation Officials, the Bureau of Transportation Statistics (BTS), the Federal Highway Administration, and the Northeast Association of State Transportation Officials. The main objective was to provide transportation officials concerned about the economic competitiveness of their regions with a broader understanding of data issues associated with global economic competition and its implications for the existing transportation infrastructure, trade corridors, and market areas.

Conference participants—who were drawn from both the public and the private sectors and represented most aspects of freight transportation and associated data needs[2]—determined that currently available regional and national data are inadequate to support the requirements of analysts and policy makers and that market area data are not readily available. However, participants also agreed that freight data collection, storage, and distribution are expensive activities, so any effort to collect new freight data should be preceded by an understanding of why such data are needed.

---

[1] The conference synthesis and additional material on conference presentations can be found on the NYSDOT website (www.dot.state.ny.us).

[2] A list of participants is included in the conference synthesis.

Development of a strategic freight data business plan to guide future data collection efforts and an associated freight data framework emerged from the Saratoga Springs meeting as one of the highest-priority action items. Conference participants recommended that this framework be structured so that

- Underlying reasons for freight movements are considered,
- Data sets are compatible across geographical and functional aggregations,
- A time frame for data updates is included to keep the data current,
- Joint efforts (partnerships) between the public and private sectors are represented, and
- The latest developments in information technology are used to track shipments and vehicle movements.

After the Saratoga Springs meeting, BTS asked TRB to conduct a study to recommend a framework for the development of national freight data. This framework is not intended to be a detailed data collection plan. Instead, it is to articulate the types of freight data needed by the variety of users in transportation and the roles of different data providers.

In response to BTS's request, TRB convened a 10-member committee with expertise in freight transportation planning and logistics, transportation policy and infrastructure, freight transportation data and modeling, and survey methodology and data collection. To expedite the study process, Rick Donnelly of PBConsult, Inc., was appointed as a consultant to the committee and charged with preparing a freight data business plan under the committee's guidance. Dr. Donnelly's commissioned paper, presented in Appendix A, formed a foundation for committee discussion of a new approach to freight data collection. The committee's commentary on Dr. Donnelly's plan for a national freight data program is provided in Chapter 3. At the committee's request, Dr. Donnelly also prepared the review of freight survey collection techniques presented in Appendix B.

The committee met three times between June and November 2002. The first two meetings included information-gathering activities, details of which are given in Appendix C. The final meeting was devoted to deliberative discussions and preparation of the committee's report.

The committee members reached consensus on all the findings and recommendations, which are presented at the outset of the report, immediately after the highlights. The committee also reached consensus on Chapters 1 and 2. However, one member, Kenneth D. Boyer, dissented from Chapter 3, where a concept for a national freight data program is discussed. Dr. Boyer's dissenting statement is presented in Appendix D. In accordance with National Research Council policies, this appendix provides the opportunity for the expression of views not shared by the majority of the committee.

The report commences with a highlights section, which summarizes very briefly the reasons freight data are needed, the deficiencies of current data, and a new approach to data collection aimed at remedying these deficiencies. The highlights section is followed immediately by the committee's findings and recommendations. Subsequent chapters address in some detail the need for freight transportation data (Chapter 1), current data limitations (Chapter 2), and a concept for a national freight data program (Chapter 3).

This report has been reviewed in draft form by individuals chosen for their diverse perspectives and technical expertise, in accordance with procedures approved by the Report Review Committee of the National Academies. The purpose of this independent review is to provide candid and critical comments that will assist the authors and the National Academies in making the published report as sound as possible and to ensure that the report meets institutional standards for objectivity, evidence, and responsiveness to the study charge. The review comments and draft manuscript remain confidential to protect the integrity of the deliberative process. The committee thanks the following individuals for their participation in the review of this report: Lillian C. Borrone, Avon-by-the-Sea, New Jersey; Theodore K. Dahlburg, Delaware Valley Regional Planning Commission, Philadelphia, Pennsylvania; Robert E. Gallamore, Northwestern University, Evanston, Illinois; Lance R. Grenzeback, Cambridge Systematics, Inc., Cambridge, Massachusetts; John J. Lawson, Transport Canada, Ottawa, Ontario; Mary Lynn Tischer, Virginia Department of Transportation, Richmond; and Charles A. Waite, CBW Consulting, Falls Church, Virginia.

Although the reviewers listed above have provided many constructive comments and suggestions, they were not asked to endorse the committee's conclusions or recommendations, nor did they see the final draft of the report before its release. The review of this report was overseen by Lester A. Hoel, University of Virginia, Charlottesville, who was responsible for making certain that an independent examination of the report was carried out in accordance with institutional procedures and that all review comments were carefully considered. Responsibility for the final content of this report rests entirely with the authoring committee and the institution.

The committee thanks the many individuals who contributed to this study through presentations at meetings, correspondence, and telephone calls. The assistance of Thomas Bolle and Jack Wells of BTS in responding to committee requests for information is gratefully acknowledged. The committee would particularly like to thank Rick Donnelly of PBConsult, Inc., for developing the initial concept for a freight data framework, and for his enthusiasm in working with the committee to expand and refine this initial concept.

Jill Wilson managed the study under the supervision of Stephen R. Godwin, Director of Studies and Information Services. Frances E. Holland assisted in logistics and communications with the committee. Suzanne Schneider, Associate Executive Director of TRB, managed the report review process. The report was edited and prepared for publication by Norman Solomon under the supervision of Nancy Ackerman, Director of Publications.

Contents

Highlights

Demands on the U.S. transportation system continue to evolve in response to changing patterns of goods movement and passenger travel and heightened concerns about transportation security. In the case of freight, the growth of international trade, the shift from a manufacturing to a service economy, deregulation, and the advent of freight logistics have all resulted in changes in the nature and volumes of goods shipped and the origins and destinations of shipments. At the same time, growing congestion on the nation's roads and at transportation hubs, such as ports and airports, not only inconveniences travelers but also threatens to undermine the reliable and timely movement of goods so critical to the national economy and quality of life.

The effectiveness and efficiency of the freight transportation system are heavily dependent on reliable data to inform a range of decisions at all levels of government and in the private sector about economic and infrastructure investments and policy issues. Data on goods movements are needed to identify and evaluate options for mitigating congestion, improve regional and global economic competitiveness, enable effective land use planning, inform investment and policy decisions about modal optimization, enhance transportation safety and security, identify transportation marketing opportunities, and reduce fuel consumption and improve air quality. While data alone cannot guarantee good decisions, informed choices are not possible without good data.

Data on goods movements are collected by federal agencies and other public- and private-sector entities that monitor or analyze transportation and trade activities on a regional, state, national, or international level.

Because these data collection efforts are not coordinated, the resulting data sets are of varying quality and reliability and provide an incomplete picture of the universe of freight movements. Furthermore, difficulties in combining data from the diverse sources limit the usefulness of current data sets for the purposes of freight transportation analyses. To remedy these deficiencies, a national freight data framework is needed.

A concept for a framework to guide the development of a national freight database and related data collection and synthesis activities is proposed in this report. This conceptual framework focuses on increasing the linkages between different sources of data and filling data gaps to develop a comprehensive source of timely and reliable data on freight flows. The national freight database aims to fulfill the major needs of a wide variety of users by capturing the important characteristics of freight movements—namely, shipment origin and destination; commodity characteristics, weight, and value; modes of shipment; routing and time of day; and vehicle or vessel type and configuration. The database also forms a foundation on which users can build their own more specialized data sets.

In its eight recommendations, the study committee offers the U.S. Department of Transportation and the Bureau of Transportation Statistics specific guidance on developing a multiyear program to implement the framework concept. In particular, the committee strongly recommends that a freight data advisory committee composed of stakeholders and experts drawn from both the public and the private sectors play a key role in guiding program development and implementation. The proposed initiative will require a sustained effort over many years and involve many technical and organizational challenges. The amount of data required is large, and some of the information needed by decision makers has not previously been collected in the United States. The committee highlights the development of innovative, low-cost methods for data collection and of procedures to protect the confidentiality of data providers as critical to a successful final outcome.

Findings and Recommendations

In 1998 the nation's transportation system moved more than 15 billion tons of goods valued at more than $9 trillion. The Federal Highway Administration (FHWA) of the U.S. Department of Transportation (USDOT) estimates that by 2020, this volume will have grown by nearly 70 percent and the value of the goods moved will be almost $30 trillion (FHWA 2002). While these burgeoning commodity flows reflect growing economic activity, some of their impacts—particularly at the local level—are raising concerns about the challenges facing the nation's transportation system. For example, congestion at ports, airports, and railheads and on streets, roadways, and highways slows not only the movement of freight, but also that of passenger traffic using the same transportation facilities.

Although concern about freight movements is not new, recent changes in the U.S. economy have highlighted a number of issues relating to freight transportation. Because of the growing importance of global markets and international trade, goods are being transported over longer distances than was the case 20 years ago. Furthermore, the patterns of goods movement have changed as a result of the growth of trade with Pacific Rim nations and among the North American Free Trade Agreement partners. Other changes in the business environment have also altered production, distribution, and logistics requirements. The shift from a manufacturing to a service economy, deregulation, and the advent of freight logistics have all resulted in changes in the nature and volumes of goods shipped and the origins and destinations of shipments (FHWA 2002). Much of the nation's transportation infrastructure was built to accommodate patterns of goods movement and passenger travel very different from those of today. As a result, the transportation system is increasingly challenged to provide the levels of efficiency and reliability in goods movement needed to ensure the international competitiveness of U.S. products and services and to sustain regional and local economies and quality of life.

The effectiveness and efficiency of the transportation system are heavily dependent on reliable information. In the case of freight movements, a range of decisions must be made at all levels of government and in the private sector about issues such as capacity utilization and potential shortfalls, congestion, safety, security, and environmental impacts. As noted in a report from the Bureau of Transportation Statistics (BTS), "good decisions require having the right information in the right form at the right time" (BTS 1998).

Recognizing the pressing need for reliable freight transportation data within the context of the changing U.S. economy and business environment, BTS was one of the sponsors of the conference "Data Needs in the Changing World of Logistics and Freight Transportation" held in Saratoga Springs, New York, in November 2001 (Meyburg and Mbwana 2002).[1] Conference participants representing a wide range of freight data interests concluded that freight transportation data are needed by a multitude of data users for varied reasons, ranging from broad policy issues to specific business logistics analyses. They also observed that existing freight data consist of a patchwork of data sets resulting from uncoordinated data collection efforts by diverse entities. Most of these efforts were not designed for the purposes of freight transportation analyses, and the resulting data are less than ideal for such applications. The most important action item identified by the Saratoga Springs conference sponsors and conferees was to develop a strategic freight data business plan to guide future data collection efforts. This plan "would identify all freight data users and their needs," and on the basis of these needs "a national or international freight data architecture or framework would be developed" (Meyburg and Mbwana 2002, 23).

[1] Other supporting sponsors of the conference were FHWA, the American Association of State Highway and Transportation Officials, and the Northeast Association of State Transportation Officials. The conference was organized by the New York State Department of Transportation and the Transportation Research Board (TRB).

After the Saratoga Springs conference, BTS asked TRB to conduct a study to "recommend a framework for the development of national freight data." The findings and recommendations of the committee formed to conduct this study are presented below. In accordance with its statement of task, the committee focused on developing a conceptual framework rather than a detailed data collection plan. The committee recognizes that the implementation of a national freight data framework will require sustained effort and funding over a number of years but believes that important benefits would accrue from streamlining data collection efforts, filling major data gaps, and harmonizing data sources for greater compatibility.

FINDINGS

Data for Decision Making

Finding 1. Reliable, consistent, comprehensive, and timely data on the movement of freight in the United States are essential for informing decisions at all levels of government and in the private sector about (a) economic and infrastructure investments and (b) policy. Such data are also needed to support informed decisions about the operation and optimization of the transportation system as a whole—decisions aimed at ensuring the safe and efficient movement of both passengers and freight.

Transportation-related strategic planning in both the public and the private sectors relies on data on the movement of goods and people. Analysis of these data provides the information needed to inform policy and investment decisions. A 1992 report on data requirements for national transportation policy making noted that "without good data, decisions will be arbitrary, options overlooked, and solutions reactive" (TRB 1992, 5).

Policy issues affecting freight transportation include the identification of options for alleviating congestion in suburban and inner-city areas, the assessment of regulations governing the shipment of hazardous materials with a view to developing safer and more cost-effective approaches, and the identification of opportunities to enhance the security of goods movements without impeding the timely flow of commodities on which the nation's economy depends. Decisions on all these issues require data on items such as the origin and destination of shipments, the commodities moved, the modes of shipment, vehicle/vessel type, routing, and time of day. These data need to be provided in a timely fashion so that decisions can be taken on the basis of up-to-date information.

The private sector needs freight transportation data to identify underserved and emerging markets as well as potential efficiency improvements meriting investment. Data that help match loads to empty capacity are particularly important in allowing shippers to increase capacity utilization at very low marginal rates, thereby reducing shipping costs. Since much of the nation's freight transportation infrastructure is privately owned and almost all freight is carried by private firms, private-sector investments have an important influence on the transportation system as a whole.

Finding 2. The current disjointed patchwork of freight data sources is costly to generate and maintain but does not provide decision makers with the data they require. To remedy this deficiency, a national freight data framework is needed to guide the development of a national freight database and related data collection and synthesis activities with the potential to meet users' data requirements.

Many transportation analyses require freight data that are not available from any single source. Thus, it is frequently necessary to combine data from different sources. For example, carrier data from the Port Import Export Reporting Service database have been combined with shipper data from the U.S. Bureau of Customs and Border Protection to inform maritime infrastructure planning and analysis.

Even data sources covering all modes fall short in meeting many of the data needs of analysts and researchers. The Commodity Flow Survey (CFS), conducted every 5 years by BTS and the Census Bureau, aims to provide reasonably comprehensive information on the flow of goods by mode of transport within the United States. However, gaps in coverage limit the usefulness of the CFS data. For example, because the survey samples only domestic shipper establishments, it does not provide reliable data on the flow of goods into the country—data that are increasingly needed because of the growing importance of international trade to the U.S. economy.[2] Furthermore, a survey of shipper establishments alone cannot pro-

[2] Imports are included in the CFS at the point at which they leave the importer's domestic location—which may not be the port of entry—for shipment to another location in the United States.

vide a complete picture of freight movements. For example, a shipper may specify a date and time by which certain goods need to arrive at their destination but may depend on the carrier to select an appropriate means of transportation. Consequently, the modal information from shippers captured by the CFS is incomplete in some cases. A further deficiency of the CFS is that the survey does not provide the level of geographic detail required by many state and metropolitan planners and engineers, who need to assign commodity and vehicle flows at least to corridors, if not to major highways and rail lines. The CFS also provides little information about unused, available existing capacity in the freight system. As a result of these limitations, the CFS data are often supplemented by data from other sources for the purposes of analysis and modeling.

The combination of data from different sources, often known as "data fusion," is frequently problematic. Most existing sources of data on freight transportation were developed independently of each other to meet the needs of specific users. Thus, these sources vary considerably in terms of their modal coverage, data collection strategies, and data definitions. The possibility of using these disparate sources to populate a comprehensive national freight database raises concerns about the quality and comparability of the resulting combined data.

A further deficiency of existing sources of freight transportation data is that some of the information required by decision makers is simply not available. For example, informed efforts to alleviate highway congestion require data on routes traveled, time of day, and the types of trucks and commodities caught in congestion—data that are rarely collected, at least in the United States.

Both the committee's discussions with users and the personal experience of individual members revealed a sense of frustration with existing freight data. The disjointed array of data sources is cumbersome and difficult to use, lacking in geographic detail, and notably deficient in covering increasingly important motor carrier flows. Several users also expressed concern about the unnecessary burden on data providers, who may be asked to provide similar data to different organizations—sometimes in different formats. This heavy respondent burden is likely to hinder efforts to gather quality data.

The committee concluded that the present patchwork of uncoordinated and incomplete freight data sources needs to be reengineered in the context of a national freight data framework that provides for a more integrated approach to freight data collection and synthesis. In the committee's view, it would be impractical—and prohibitively expensive—to subsume all existing data sources into a new monolithic freight data framework. Furthermore, existing data requirements and the need for continuity of data to inform trend analyses limit the possibilities for "wiping the slate clean" and initiating an entirely new set of data collection programs. Therefore, the committee envisages a framework with a flexible, modular structure that leverages existing data sources as far as possible and avoids costly new data collection initiatives unless they are needed to fill data gaps.

Potential Benefits of a National Freight Data Framework

Finding 3. A national freight data framework offers potential advantages for enhancing transportation security and sustaining and growing the economy.

A national freight data framework—such as that shown schematically in Figure F&R-1—that facilitates data fusion and fills data gaps would aid in developing a comprehensive picture of freight flows. Such a picture offers opportunities to enhance the security of goods movement by identifying vulnerabilities. For example, data on routing and time of day for bulk shipments of hazardous materials could be used in identifying high-risk scenarios and targeting appropriate security measures. In the longer term, an improved understanding of normal freight flows would also provide a baseline against which to identify anomalies linked to possible acts of terrorism. A recent report on countering terrorism identified an understanding of normal patterns of transportation activity and behavior as a key research need for transportation security (NRC 2002).

The improved freight data resulting from the framework approach could also increase the international competitiveness of U.S. products and services through more effective use of and improvements in the nation's transportation system. In addition, these data could help sustain the current strengths and stimulate the development of regional and local economies through informed decisions that take appropriate account of freight transportation needs and opportunities.

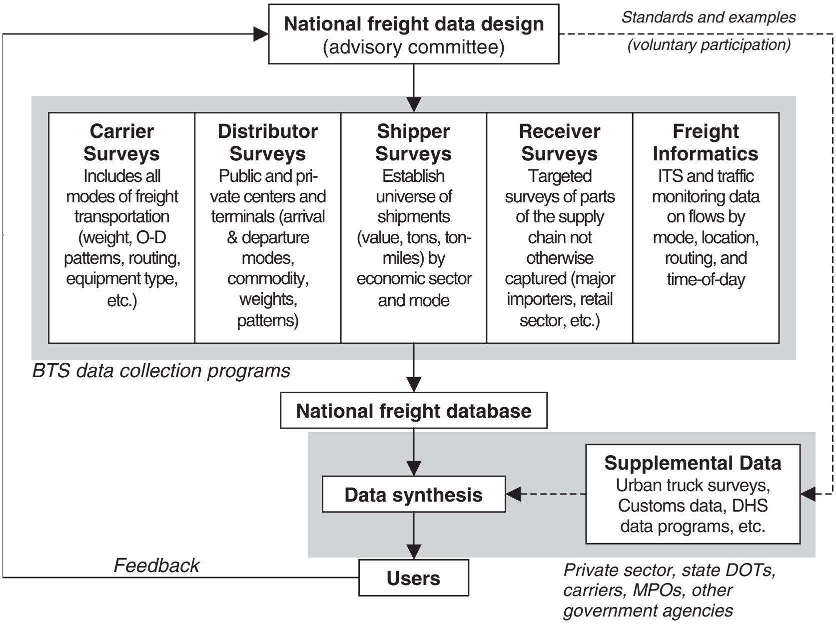

Figure F&R-1 Proposed framework of a national freight data program. [BTS = Bureau of Transportation Statistics; O-D = origin–destination; DHS = Department of Homeland Security; MPO = metropolitan planning organization; state DOT = state department of transportation. Source: Adapted from a paper prepared for the committee by R. Donnelly (Appendix A).]

Although many of the issues in freight transportation are well known, comprehensive high-quality data may be useful in pinpointing underlying causes and prioritizing policy and investment decisions. Such data may also enable research aimed at solving freight transportation problems—research that cannot currently be undertaken because the necessary data are not available.

FHWA recently completed a 3-year project to develop the Freight Analysis Framework (FAF)—a national-level policy analysis tool (FHWA 2002). The FAF, which is based on a composite of many databases, has played, and continues to play, an important role in raising awareness of freight issues. It is also a "point of departure for further examination of policies, programs, and initiatives that might be undertaken by decision makers at all levels of government" (FHWA 2002, 2). In many instances, such examination will require the kinds of data that a national freight data framework would provide.

What Data Are Needed?

Finding 4. The following data would capture important characteristics of freight movements and meet the major needs of a wide variety of data users in the public and private sectors:

- *Origin and destination;*
- *Commodity characteristics, weight, and value;*
- *Modes of shipment;*
- *Routing and time of day; and*
- *Vehicle/vessel type and configuration.*

The committee's discussions with data users indicated that providing all the data needed for all applications would be beyond the scope of any national initiative. In the committee's judgment, a national freight data framework program needs to focus on populating a national freight database that provides robust data to meet basic user requirements. Providing options to enhance the data set for more specialized purposes would be extremely beneficial in meeting a wider range of user needs. On the basis of its discussions with users, review of the literature,[3] and the experience of individual members, the committee identified the data items listed above as particularly important for a wide variety of applications. Not all the items would be used for every analysis; users would "pick and choose" the data they require.

No source, public or private, currently provides reliable data on all the items listed in Finding 4. Only the CFS and the related Transearch database from Reebie Associates even recognize the need to provide a complete description of freight flows in the United States, and both have important deficiencies. As discussed under Finding 2, the CFS has inherent limitations because it samples only domestic shipper establish-

[3] Useful tables of information on freight data needs for different applications are provided in a report on a possible ITS Archived Data User Service (Margiotta 1998) and in a report from BTS on data gaps (1998).

ments. The survey also suffers from gaps in industry coverage because it excludes transportation and service establishments and most retail establishments, as well as agricultural shipments from the farm to the first point of assembly. While the Transearch database attempts to fill some of the gaps in the CFS, its coverage is also incomplete because of limitations in the data sources available for input. In addition, the Transearch database is generated through proprietary methods, and information about data reliability is not reported.

Many of the data items listed above are already available from different sources, but, as noted earlier, combining data from different sources is challenging. Furthermore, there are gaps in existing data that are extremely difficult to fill other than by data synthesis. In terms of modal coverage, the most significant gap is in motor carrier flows—an important growth area that is not well covered by any existing data collection efforts. Of the data items listed in Finding 4, routing and time of day are sparsely covered by current data sets, even though they are important for assessing congestion mitigation strategies, evaluating system capacity, and ensuring the security of shipments.

Need for Leadership

Finding 5. The federal government is uniquely positioned to provide the proactive leadership needed for the success of a program to develop and implement a national freight data framework.

No single organization by itself has the resources and expertise necessary to develop and implement a national freight data framework. The interest and cooperation of a range of public- and private-sector organizations will be essential to the overall success of the proposed framework initiative. The participation of industry will be particularly important given that almost all freight in the United States is carried by private firms.

In addition to a team effort involving public- and private-sector participants, the framework initiative will require strong leadership to coordinate the data collection activities of diverse entities within the context of an overall strategy. Given the national significance of the proposed framework, the diffuse nature of some of its "public good" benefits, and the need for a systemwide approach involving all levels of government and the private sector, the committee concluded that an organization within the federal government would be best positioned to assume the leadership role.

RECOMMENDATIONS FOR DEVELOPING A NATIONAL FREIGHT DATA FRAMEWORK

The following recommendations address the initial programmatic and technical steps required to move forward with the implementation of a national freight data program.

Public- and Private-Sector Roles

Recommendation 1. USDOT should assume a leadership role in developing and implementing a national freight data framework similar to that described in this report.

In the committee's view, USDOT has the knowledge and expertise needed to lead the development and implementation of a national freight data framework. Furthermore, the ultimate goal of the framework—namely, providing a comprehensive picture of freight flows in the United States—is consistent with the department's mission of ensuring a transportation system that meets vital national interests and enhances the quality of life of the American people.[4]

The modal administrations within USDOT are already involved in mode-specific data collection programs for operating and administrative purposes. Their expertise is complemented by that of BTS in multimodal freight surveys (the CFS) and in statistics and survey methodology in general. The experience of FHWA in developing the FAF could also be helpful in implementing the proposed national freight data framework initiative.

Recommendation 2. USDOT should establish a freight data advisory committee to guide the development and subsequent implementation of a national freight data framework.

[4] The USDOT mission statement is given in full on the department's website (www.dot.gov).

Although USDOT has considerable knowledge and expertise in the area of freight data, the national framework initiative is too broad in scope to succeed without the support and involvement of the wider freight data community. Therefore, the committee envisages that the development and implementation of the framework will be guided by a freight data advisory committee of stakeholders and experts. The membership of this advisory committee should reflect the broad spectrum of freight data users and providers and should include representatives of federal, state, and local jurisdictions as well as a range of private-sector stakeholders. The latter group should include consulting companies, representatives of different modes of transportation (air, maritime, pipeline, railroad, trucking), shippers and receivers, third-party logistics companies, and academic researchers. Since national defense activities, such as those of the Army's Military Traffic Management Command, could benefit from improved freight flow data, the committee suggests that the advisory committee membership include an expert in defense logistics.

Recommendation 3. BTS should play an important role in developing and implementing the framework, with guidance from the freight data advisory committee.

The legislation authorizing the establishment of BTS requires the agency's statistics to support transportation decision making by all levels of government, transportation-related associations, private businesses, and consumers [49 U.S.C. 111(c)(7)]. Furthermore, as a federal statistical agency, BTS is charged with the continual development of more useful data "to provide information that is accurate, timely, and relevant for changing public policy needs" (NRC 2001). In light of these requirements and the agency's experience and expertise in survey methodology and statistics, the committee considers it appropriate for BTS to play a major role in the freight data framework initiative.

Details of the BTS role will require further investigation as plans are developed for implementing the framework. For example, it could be argued that BTS should assume full responsibility for developing and maintaining a national freight database because of the need for transparency (no proprietary data or methods)[5] and the likely need to work with confidential microdata. On the other hand, extensive expertise in data fusion and manipulation exists in the private sector. For this reason, some would argue that BTS should focus on data collection and leave the development of a national freight database to others. More detailed definition of the tasks involved in database development—and other aspects of framework implementation—will be needed before appropriate roles for various public- and private-sector groups can be assigned on an informed basis.

Guiding Principles

Recommendation 4. USDOT/BTS, under the guidance of the freight data advisory committee, should explore the potential for meeting the needs of freight data users through the implementation of a cost-effective national freight data framework. The tasks undertaken are likely to include

- ***Examination of the relative merits of different kinds of surveys (carrier, distributor, shipper, receiver)—and combinations thereof—for gathering the types of data users require;***
- ***Identification and preliminary assessment of opportunities to use data sources other than surveys—for example, electronic data interchange (EDI) and intelligent transportation system (ITS) data—to collect data more cost-effectively, fill data gaps, and reduce respondent burden;***
- ***Identification of opportunities for facilitating the integration of data from different sources; and***
- ***Investigation of techniques to assist a range of users in combining data from supplemental sources (urban truck surveys, Customs data, security data, etc.) with national freight data to meet their individual requirements.***

The committee envisages that implementation of the framework will involve the development of a freight data road map that builds on the con-

[5] A recent National Research Council report states that a federal statistical agency "should be open about its data and their strengths and limitations" (NRC 2001, 8).

cepts presented in this report. Such a road map will need to assess the state of various data collection, synthesis, and fusion techniques; identify opportunities for coordinating data collection efforts and filling data gaps; and identify promising areas for research and innovation to support implementation of the framework.

Assessment of the strengths and weaknesses of current data sources and of different types of surveys will be a prerequisite to setting priorities for further investigation and development. The committee does not envisage that all possible approaches will be pursued concurrently or that equal effort will be devoted to each. The freight data advisory committee will need to weigh the costs of obtaining data against the potential benefits of making the most effective use of limited resources.

Because of the large amounts of data required and the relatively high costs of surveys, implementation of the framework will need to take advantage of nonsurvey data streams. Opportunities to use low-cost passive data collection appear particularly promising. For example, ITS roadway surveillance data, which are generated continuously and at a very detailed level, could be used for congestion monitoring and intermodal facilities planning, over and above their use in real-time control strategies. Preliminary requirements have already been identified for an ITS Archived Data User Service to support a wide range of stakeholder activities (Margiotta 1998).

Modifications of current data collection efforts that facilitate the integration of data from more than one survey or of survey data with ITS or EDI data could be a particularly cost-effective way of developing more useful freight data. As noted in a recent National Research Council report, "when separate datasets are collected and analyzed in such a manner that they may be used together, the value of the resulting information and the efficiency of obtaining it can be greatly enhanced" (NRC 2001, 7). The committee anticipates that a major strength of the framework will be the provision of opportunities for states, metropolitan planning organizations (MPOs), railroads, transportation-related associations, and the like to build on a national freight database in developing their own data sets. To facilitate the combination of specialized and national data, the framework will need to provide standard survey methodologies and examples of their use, together with recommended best practices for activities such as data collection and processing. The committee envisages a leadership role for BTS, as a statistical agency within USDOT, in developing standard methodologies and encouraging the use of best practices in freight transportation surveys.

Recommendation 5. The development and implementation of a national freight data framework should be guided by the following principles:

- ***Focus on providing real data—as opposed to imputed or synthesized data—when possible.***
- ***Ensure that data are sufficiently timely to meet the needs of users.***
- ***Ensure that data collection and synthesis methods are appropriately documented so that data users can assess the resulting data quality and reliability.***
- ***Encourage the use of compatible data elements to facilitate the combination of data from different sources.***
- ***Incorporate mechanisms that encourage continuous feedback from users and subsequent refinement of the framework.***

On the basis of its discussions with users, review of the outcomes of the Saratoga Springs meeting (Meyburg and Mbwana 2002), and the experience of individual members, the committee understands that the provision of real (as opposed to synthesized) data in a timely fashion, together with information on data quality and reliability, is important to users. These user expectations will need to be taken into account in developing and implementing the framework and in populating the national freight database. In particular, a decision by the freight data advisory committee concerning the level of geographic detail to be provided by the national freight database will involve balancing the needs of different users (federal government agencies, state departments of transportation, MPOs, consulting companies, academic researchers, etc.) for data on international, national, state-to-state, and intrastate goods movements and the costs of collecting, processing, and disseminating such data in accordance with requirements for timeliness, reliability, and the like.

While a national freight database cannot provide all data for all users, it can form a basis on which to build by facilitating linkages to other data sets. One of the major deficiencies of current freight data is the lack of harmonization among different databases. In the committee's view, it is imperative to move as rapidly as possible toward a more integrated approach that eliminates unlinked data "silos." The use of compatible data elements, standard survey methodologies, and other techniques for facilitating data fusion will be essential to the successful implementation of the national freight data framework.

The committee anticipates that the content and detailed structure of the framework will evolve over time to reflect research findings, new opportunities for data collection, and practical experience with different data sources. Feedback from users will be essential in guiding this evolutionary process, and the freight data advisory committee will need to establish mechanisms for encouraging interactive development and implementation of the framework.

Stakeholder Involvement

Recommendation 6. USDOT/BTS should actively encourage the participation of data providers as partners in the development and implementation of the freight data framework by

- ***Explaining clearly why data are being collected and for what purposes they will be used,***
- ***Avoiding overly burdensome reporting requirements, and***
- ***Respecting the imperative to maintain data confidentiality.***

New data providers[6] will need to see some payoff for supplying data, while current providers will need encouragement to expand the scope of data they supply or adapt to new data collection methods. In many instances, diffuse "public good" benefits are unlikely to constitute a particularly attractive incentive for participation in data collection programs. The promise of specific benefits, such as the availability of data on empty movements to help carriers increase load factors, are likely to be more persuasive.

The need to maintain the confidentiality of individual firms is a potentially fatal flaw that must be addressed early in the framework development

[6] Throughout this report, the term "data provider" has been used to designate survey respondents and other data subjects. Unless explicitly stated, the term does not include third parties who undertake activities such as data fusion and data synthesis.

process. Data providers will not participate in framework activities if they see any risk of their competitors gaining access to commercially confidential information. Thus, in seeking to encourage private-sector participation, USDOT/BTS will need to recognize that much of the proprietary data collected for legitimate business planning and investment purposes cannot readily be converted into public use data.

A recent National Research Council report notes that federal statistical agencies need to treat data providers fairly. In addition to adopting policies and procedures for maintaining data confidentiality, agencies need to seek the advice of respondents in designing data collection procedures and determining data products (NRC 2001, 10). Such practices appear particularly pertinent in the context of efforts to develop and implement a national freight data framework.

The International Trade Data System (ITDS) project, which seeks to streamline the collection and dissemination of international trade data, may provide helpful insights for implementing the national freight data framework. In contrast to current practice, which often involves multiple reporting requirements, traders will submit standard electronic data only once. ITDS will then distribute these data to federal agencies on a "need-to-know" basis; each agency will receive only information relevant to its mission.[7] Such standardized data collection could help reduce the burden on providers of freight data, and a system of strict controls on data access could allay the concerns of these providers about inadvertent release of commercially sensitive information.

> ***Recommendation 7. USDOT/BTS, with guidance from the freight data advisory committee, should investigate opportunities to stimulate investment in data collection and synthesis by a range of private and public-sector organizations. Such activities would be invaluable in supplementing federal government efforts to implement a national freight data framework.***

The willingness of various groups in both the public and the private sectors to participate in framework implementation will largely depend on their perceptions of how useful the resulting data are likely to be. To encourage broad participation, USDOT/BTS will need to understand the

[7] ITDS Background (www.itds.treas.gov/itdsovr.html).

perspectives of different stakeholders and present the anticipated benefits of the framework accordingly. FHWA's FAF program involved extensive outreach to freight stakeholders for the purposes of improving understanding of the nature of freight movements, identifying challenges to improving freight productivity and security, and developing strategies to increase freight productivity (FHWA 2002). Implementation of the national freight data framework is likely to involve comparable outreach activities to explain the value of better data in addressing various freight issues, solicit suggestions for framework development and implementation, and encourage participation in the framework initiative.

Program Continuity

> ***Recommendation 8. Since the implementation of a national freight data framework is likely to require a sustained effort over a period of 7 to 10 years, USDOT/BTS should establish the necessary planning, development, and management capabilities to provide program continuity.***

Some form of institutional structure, such as a program office, will be needed to coordinate activities within a national freight data framework program, support the freight data advisory committee, and provide a focal point for the framework initiative. A framework program office would also be responsible for facilitating continuous feedback and refinement of the framework, identifying emerging data collection opportunities and encouraging related research investigations, and ensuring the sustainability of data collection activities so that the national freight database can be expanded and updated. Over time, the focus of program activities is likely to shift from feasibility studies and concept development through implementation to updating and maintenance to ensure long-term viability. The institutional program structure will require sufficient flexibility to accommodate this evolutionary process.

REFERENCES

Abbreviations

BTS	Bureau of Transportation Statistics
FHWA	Federal Highway Administration
NRC	National Research Council
TRB	Transportation Research Board

BTS. 1998. *Transportation Statistics Beyond ISTEA: Critical Gaps and Strategic Responses.* BTS98-A-01. U.S. Department of Transportation, Washington, D.C.

FHWA. 2002. *The Freight Story: A National Perspective on Enhancing Freight Transportation.* FHWA-OP-03-004. U.S. Department of Transportation, Washington, D.C. www.ops.fhwa.dot.gov/freight/publications/freight%20story/freight.pdf.

Margiotta, R. 1998. *ITS as a Data Resource: Preliminary Requirements for a User Service.* Office of Highway Policy Information, Federal Highway Administration. www.fhwa.dot.gov/ohim/its/itspage.htm.

Meyburg, A. H., and J. R. Mbwana (eds.). 2002. *Conference Synthesis: Data Needs in the Changing World of Logistics and Freight Transportation.* New York State Department of Transportation, Albany. www.dot.state.ny.us/ttss/conference/synthesis.pdf.

NRC. 2001. *Principles and Practices for a Federal Statistical Agency,* 2nd ed. National Academy Press, Washington, D.C.

NRC. 2002. *Making the Nation Safer: The Role of Science and Technology in Countering Terrorism.* National Academy Press, Washington, D.C.

TRB. 1992. *Special Report 234: Data for Decisions: Requirements for National Transportation Policy Making.* National Research Council, Washington, D.C.

1

The Need for Freight Transportation Data

More than 6 million business establishments in the United States rely on the nation's transportation system to engage in local and interstate commerce and international trade (BTS 1998). In 1999, the nation's freight transportation bill was $562 billion—approximately 6 percent of gross domestic product (Wilson 2001). The effective and efficient movement of freight is critical to the nation's economy and must be assured in the face of changing circumstances, such as an increased emphasis on global markets, shifts in domestic economic activity to the suburbs and suburban malls, new transportation patterns for improved logistics efficiency, growing congestion on the nation's roads, heightened concerns about transportation security and capacity, and increased maintenance requirements associated with the aging U.S. transportation network.[1]

[1] A more detailed discussion of trends affecting freight movements is provided in a series of working papers on the changing nature of freight demand, carrier operations and infrastructure, and public policy, prepared for the Federal Highway Administration's Office of Freight Management and Operations (www.ops.fhwa.dot.gov/freight/theme_papers/theme_paper_index.htm).

While the focus of this report is on freight transportation and associated data needs, it is important to remember that freight activity and passenger travel both affect the demand for transportation facilities and services. Since most major freight nodes (ports, airports, and railheads) and the origins and destinations of most shipments are in cities, freight must compete with passenger traffic for the use of transportation facilities (TRB 2003). Thus, limitations imposed by the transportation infrastructure affect the movement of both goods and people. For example, road congestion around New York's John F. Kennedy Airport is making access difficult for cars carrying passengers and for trucks carrying freight.

Changes in the transportation infrastructure and related operations cannot be made without influencing the complex transportation network as a whole and all its users. For example, efforts to alleviate highway congestion are likely to benefit passengers as well as freight. Conversely, attempts to accelerate the flow of trucks carrying freight on the highway without affecting passenger traffic would appear problematic. Despite this interrelationship between freight and passenger movements, transportation-related problems affecting passengers (e.g., delays due to congestion) may attract more attention than problems affecting freight because "parcels cannot talk."

One consequence of the focus on passenger needs—and complaints—is a lack of widely available data to inform decisions about freight transportation issues. For example, analysis of the effectiveness and costs of alternative options for mitigating congestion in urban areas requires better data on patterns of freight movements. Metropolitan planning organizations typically have detailed origin–destination data on commuting patterns by industry and occupation from the Census Bureau's journey-to-work data. However, data on truck movements in metropolitan areas are often scant and lack details of trip origins and destinations and the industry/commodity breakdown for products being carried. Furthermore, many of the existing metropolitan data on truck movements miss delivery of business products and services by car, van, and light-delivery vehicle (Weisbrod et al. 2001). Another important gap is the lack of shipment weight data for U.S. exports by land modes of transportation (truck, rail, pipeline, and mail). This deficiency hampers efforts to estimate and forecast the impact of international trade on domestic transportation infrastructure—efforts that are critically dependent on accurate shipment weight data (BTS 2002).

CHARGE TO THE COMMITTEE

In November 2001, the New York State Department of Transportation and the Transportation Research Board (TRB) convened a conference in Saratoga Springs, New York, to address data needs in the changing world of logistics and freight transportation. The main conference objective was to provide transportation officials concerned about the economic competitiveness of their regions with a broader understanding of data issues associated with the changing focus of the global competitive market and accompanying implications for the existing transportation infrastructure, trade corridors, and markets (Meyburg and Mbwana 2002). At the conclusion of the conference, participants agreed on a major action item—to develop a strategic freight data business plan to guide future data collection efforts. They proposed that the plan identify all freight data users and their needs. On the basis of these needs, a national or international freight data framework would be developed.

After the Saratoga Springs conference, the Bureau of Transportation Statistics asked TRB to conduct a study to recommend a framework for the development of national freight data. The framework was to identify

- The data requirements for the various users of freight data (public, including various levels of government, and private); and
- Appropriate federal, state, and private-sector roles in the development and dissemination of freight data.

The framework was to be conceptual in nature and *not* a detailed data collection plan. Instead, it would articulate the types of freight data needed by the variety of users in transportation and the roles of different data providers.

The findings and recommendations of the study committee are presented at the outset of this report. In the remainder of Chapter 1, needs for freight transportation data are identified, and examples of policy and investment decisions and associated data requirements are provided. In

Chapter 2 an overview of existing freight transportation data is provided, and some of their limitations for the purposes of transportation analysis are identified. The conceptual model for a national freight data program, which was developed through extensive interactions between the committee and its consultant, Rick Donnelly, is discussed in Chapter 3. Dr. Donnelly's commissioned paper, *A Freight Data Business Plan,* is presented in Appendix A.

WHY DATA ARE NEEDED

Although the public and private sectors have a different focus, their requirements for reliable, high-quality data are similar. Such data allow investment decisions—many of which may be long-lived—to be taken in the knowledge that alternative strategies have been developed and compared on the basis of reliable information.

Reliable data on the movement of freight are needed to inform public policy decisions on issues such as congestion mitigation, transportation security, air quality improvement, economic development, and land use. Many of these decisions are more difficult today than they would have been 20 years ago. There are fewer opportunities to add capacity to the transportation system, and the need to meet environmental and other social goals often results in delays and increased project costs. Consequently, optimizing modal and intermodal choices requires much better information about the freight transportation system to evaluate the advantages and disadvantages of a limited number of expensive options. A recent report notes that infrastructure projects costing from $100 million to several billion dollars are becoming more common, especially in urban areas (TRB 2003).

Reliable, high-quality freight transportation data are also needed by the private sector to inform a range of strategic investment decisions relating to topics such as equipment utilization, new market opportunities, and business relocation. For example, information on the movements of empty shipping containers and railroad cars could assist freight carriers in identifying opportunities to reduce unprofitable movements of empty equipment. The associated improvements in operational efficiency could allow lower pricing of freight transportation services and increase the potential for economic growth.

Box 1-1

Major Reasons for Needing Freight Data

- Promote transportation efficiency and mitigate congestion.
- Improve regional and global economic competitiveness.
- Enable effective land use planning.
- Inform investment and policy decisions about modal optimization.
- Enhance transportation safety and security.
- Identify transportation marketing opportunities.
- Reduce fuel consumption and improve air quality.
- Understand economic geography.
- Reduce incremental operating costs for all users.
- Provide information for national accounts.*
- Enhance national defense activities.

*For further information on transportation satellite accounts, see BTS 1999.

Within the broad context of policy making and investment decisions, the committee identified 11 major reasons for needing freight data (Box 1-1). Some of these reasons are related; for example, reduced fuel consumption is likely to result in reduced incremental operating costs. While these reasons are intended to be illustrative rather than exhaustive, the related examples and case histories presented in the remainder of this chapter reflect the diverse and wide-ranging social and economic impacts of freight transportation.

EXAMPLES AND CASE HISTORIES

Promote Transportation Efficiency and Mitigate Congestion

Efforts to improve transportation efficiency require data to identify problem areas and to evaluate proposed improvements. For example, Metroplan Orlando recently developed its Freight, Goods and Services

Mobility Strategy Plan to identify and address concerns about the movement of freight, goods, and services in the Orlando Urbanized Area in central Florida (Metroplan Orlando 2002). One of the key underpinnings of the plan was a data set describing freight traffic movements to, from, and through the area. The development of this data set constituted an important part of the entire effort and required considerable specialized technical expertise.

A review of the National Highway System (NHS) freight connectors also required considerable effort to obtain the necessary data (U.S. Department of Transportation 2000). These connectors are public roads linking the NHS to seaports, airports, and major intermodal terminals, and are therefore vital to the national intermodal freight system. However, because the connectors are frequently local, county, or city streets that cannot accommodate high volumes of heavy truck traffic, they can be major bottlenecks. Identification of major connectors required data on the level of activity of a terminal and its importance to a particular state. The participation of national associations, state departments of transportation, metropolitan planning associations, terminal operators, and the staff of the U.S. Department of Transportation modal administrations was needed to obtain the necessary information.

Data issues encountered in implementing a regional transportation plan for Portland, Oregon, are discussed in Box 1-2.

Local problems, such as bottlenecks at a hub airport or rail center during peak periods, can have a severe impact on airline or railroad transportation networks, can quickly spread hundreds or thousands of miles from their source, and can be costly for operators (TRB 2003). For example, if a cargo plane misses its landing slot and the associated distribution bank of flights, the operator may need to charter another aircraft in an effort to deliver goods on time.[2] Aggregate data from national freight data sets may not be sufficiently detailed to reveal such local capacity constraints. Instead, they may simply indicate that an average link is operating below capacity over an average time period. The lack of geographic detail in many current data sets is particularly damaging to decisions made at the metropolitan and county levels, where the

[2] This example was provided to the committee by Mort Plumb, Director, Ted Stevens Anchorage International Airport, Alaska.

Box 1-2

▼

Improved Freight Mobility Around Portland, Oregon

Portland, Oregon, has developed a regional transportation plan, which includes a freight network. As part of the plan implementation, projects to improve freight mobility into, around, and through the region are being solicited. However, the lack of relevant data will impede assessment of the potential merits of different projects and associated funding decisions. Although the volume of trucks on any particular route is known, available data do not provide satisfactory information on where these trucks are going, what goods they are carrying, whether the shipments are part of a series of intermodal moves, or the extent to which timely delivery of the goods affects just-in-time production schedules.

bottlenecks and greatest congestion delay costs occur. Because most of the available data concerning freight flow are at the state level, decision makers do not have the information they require to identify specific freight demands on metropolitan and county infrastructure or evaluate the relative merits of proposed improvements. Difficulties are also encountered in obtaining data for metropolitan areas that extend across state boundaries, rather than being located within a single state.

The need for better data to inform decisions about options for mitigating congestion on the Upper Mississippi River–Illinois Waterway is discussed in Box 1-3.

Improve Regional and Global Economic Competitiveness

Accurate freight data are needed to support investment decisions aimed at improving regional and global economic competitiveness. The example of the Northeast North American Free Trade Agreement (NAFTA) corridor is discussed in Box 1-4. Such data could also provide insights into opportunities for supply chain efficiencies and possible business relocation.

Box 1-3

▼

Upper Mississippi River Locks

The U.S. Army Corps of Engineers recently conducted a study evaluating proposals for improvements to locks and other navigation facilities on the Upper Mississippi River–Illinois Waterway. The purpose of these improvements is to relieve increasing waterway congestion, particularly for grain being transported to New Orleans for export. During a 300-hour trip from the Upper Midwest to New Orleans, a barge may typically experience 30 hours of congestion-induced delay.

The study proved controversial, and a National Academies committee was asked to review the methods used. This committee identified one of the problem areas as a lack of adequate data for forecasting future levels of barge traffic across the entire navigation system. Without such data, the potential benefits of proposed improvements to locks and other navigation facilities cannot be assessed satisfactorily. The committee recommended that the corps obtain a database of grain and other relevant freight shipments by barge and alternative modes that includes quantity, origin and destination, and price. The committee suggested that such data take account of seasonal effects and that grain flow surveys seek information from both barge terminal operators and grain shippers.

Sources: NRC 2001; TRB 2003.

For example, better freight data could reveal the effect of improved information technology on the location of production in the United States. Measuring which commodities have been affected most strongly by changes in information technology and observing how commodity flows have changed for these commodities would yield an understanding of the effects. This understanding could be used to help identify opportunities for further efficiencies.

Box 1-4

▼

NAFTA Corridors

The New York State Department of Transportation is seeking to identify freight transportation improvements that could result in better linking of NAFTA shipments, thereby making the Northeast transportation network a more effective contributor to the NAFTA corridors. The Northeast states require accurate strategic freight shipping data (origin and destination, commodity, mode of transport, route, etc.) to understand the role of the Montreal–Boston–New York–Washington corridor vis-à-vis competing routings. At the national level, such data would help inform national transportation policies on NAFTA corridors and allow an objective assessment of related funding requirements.

Source: Boardman 2002.

Enable Effective Land Use Planning

In some urban areas, there may be opportunities to improve the transportation system through effective land use planning. Possible strategies include the relocation or consolidation of rail terminals and track, the use of brownfields and other vacant lands as freight centers, and the improvement of access to intermodal terminals. For example, the Baltimore Regional Rail Corridor Study is considering the possibility of building a new rail route to free up old railroad tunnels and other rights-of-way in congested areas. These existing corridors could then be used for commuter rail or light rail service. However, planners need good data on freight movements into, out of, and through the area of interest, including route structures and use, to identify promising options and likely outcomes. In another example, the New Jersey Transportation Planning Authority is planning to reclaim abandoned industrial brownfield sites near northern New Jersey ports, airports, and rail terminals for the development of a "freight village" (Meyer 2001). Again, effective planning of these freight- and trade-related distribution facilities will require comprehensive and reliable data on freight movements.

Inform Decisions About Modal Optimization

Improved freight data would help in identifying opportunities to substitute rail for truck transportation and vice versa. Currently, much of the information on freight flows does not provide enough detail on commodities or on specific routes to allow planners to understand when one mode can substitute for the other and when the two modes complement each other. Thus, state secretaries of transportation do not have the information they need to make good decisions about investing limited capital in more highway lanes or more rail capacity. More detailed geographic and commodity data would help transportation planners and marketing experts conduct better analyses to identify promising opportunities for modal diversion. The example of the I-81 corridor in Virginia is presented in Box 1-5.[3]

Enhance Transportation Safety and Security

Information on vehicle weight is important for many transportation planning applications and is even more useful if it can be linked to route data. Many traffic data collection systems count numbers of vehicles passing fixed points on the highway system but do not distinguish between passenger vehicles and trucks. Highway investments are often based on vehicle counts because data on vehicle type and weight needed to provide a more informative picture of road use are not available. Infrastructure projects may be adversely affected if engineers have to design pavements without the benefit of reliable estimates of loads over the expected life of the structure. Since pavement damage increases rapidly with increasing axle weight,[4] the vehicle mix has an im-

[3] An ongoing National Cooperative Highway Research Program project (Project 8-42) is addressing the subject of rail–freight solutions to roadway congestion. Two of the project tasks are particularly data intensive: the evaluation of the likely impacts of diverting various levels of freight traffic from truck to rail and the identification of short- and long-term trends in freight movements and land use affecting congestion. Recognizing likely data deficiencies, the project also calls for an assessment of the adequacy of currently available data for conducting these two tasks.
[4] The effects of vehicle weight on pavement performance are usually estimated on the assumption that the effective load on the pavement varies as axle weight to the fourth power.

Box 1-5

Expanding Capacity on the I-81 Corridor in Virginia

I-81, an Interstate highway running through Virginia, is a major truck route. The state has a plan for several billion dollars in highway improvements over the next 20 years to accommodate expected traffic growth on I-81, notably in truck traffic. Norfolk Southern Corporation proposed to the state of Virginia that it consider public investment in improvements to the company's Shenandoah rail corridor as an alternative to widening I-81. As a result, the Virginia state legislature instructed the Virginia Department of Transportation to study the feasibility of shifting traffic in the I-81 corridor from highway to rail. Extensive data on goods shipped, shipment origin and destination, mode and cost of transport, routes, and the like are needed to compare the benefits and costs of the two alternative approaches to expanding freight capacity in the I-81 corridor. Data on the nature and amount of freight moving by truck are particularly important for identifying opportunities to shift freight from truck to rail. Many of these data are not available or are not sufficiently reliable to constitute a robust foundation for decision making. Consequently, investment decisions may need to be taken without the benefit of well-informed assessments of road and rail options.

Source: TRB 2003.

portant influence on the service life of a highway, and reliable data on vehicle weight are critical to the design process. Without such data, structures may be over- or underdesigned for their intended use and life. Such inefficient designs not only result in an inappropriate allocation of scarce funds but also pose a risk to the public if structures such as pavements deteriorate rapidly because of unanticipated patterns of use by heavy vehicles.

Comprehensive and reliable freight transportation data offer the potential to enhance the security of goods movements within and into the United States. For example, to support their mission of protecting critical infrastructure, policy makers at the state and local levels need accurate data on the movement of goods to make better resource allocation decisions about the types and levels of protection to put in place. Even before September 11, 2001, terrorist attacks on the United States using the global transportation logistics network were identified as a possible risk to homeland security—for example, by smuggling weapons of mass destruction into the country in oceangoing cargo containers (Flynn 2001). Measures have now been taken to reduce the likelihood of such a scenario. These include amended Customs regulations requiring advance presentation of certain manifest information prior to lading at the foreign port and the implementation of container screening procedures (*Federal Register* 2002). However, screening of all the nearly 6 million cargo containers off-loaded at U.S. seaports every year is clearly impractical. Therefore, cost-effective methods for pinpointing high-risk shipments that require special screening are important in ensuring that security procedures are effective but do not impede the timely movement of goods on which the nation's economy depends. By providing a picture of normal shipment patterns, good freight transportation data (and data analyses) would establish a baseline against which to identify unusual or suspicious patterns meriting special scrutiny.

Identify Transportation Marketing Opportunities

Freight transportation providers seeking to develop and market their services to meet evolving customer requirements need freight transportation data for logistics modeling. Data on origin and destination, commodities moved, and the stage of completion of manufactured products help identify opportunities for expediting the efficient flow of raw materials, work-in-process inventory, and finished goods from supplier to customer. Freight carriers and third-party logistics providers can then market their services to potential clients on the basis of anticipated time and cost savings. For example, air cargo carriers providing service into and out of Anchorage, Alaska, would like better data on international freight movements. Data on the true origins and destinations of shipments—not just the segments into, out of, or through the international air cargo hub in Anchorage—would help firms identify business development opportunities.[5]

Trucking firms are increasing partnering with railroads to move more of their domestic long-haul shipments from the highway to rail. These partnerships marry the long-haul cost advantages of rail with the delivery flexibility of truck, resulting in win-win solutions for these commercial enterprises, as well as for government entities facing major investment in highway infrastructure. Despite the benefits, the resulting service and capacity impacts on the railroad can prove significant. Intermodal growth requires long-term investment in yard and track infrastructure. These investments must be financially sound and justified by accurate capacity and revenue estimates. Though trucking partners may share some information on a case-by-case basis, capacity and service planning requires accurate freight data on an aggregated basis, by commodity and by market. Primary data at this level are not available today.

The proposed expansion of Norfolk Southern's rail line paralleling the I-81 corridor presents one example of this deficiency (see Box 1-5). For the railroad, planning for this project requires estimating the total amount of freight traffic that could be handled on the expanded rail line. These calculations require the volume of traffic currently moving along the I-81 corridor, along with the actual origin and destinations. The lack of accurate freight movement information along the corridor has made planning difficult, limiting the amount of potential investment. Better capacity planning also enables the railroads to provide their trucking partners with better rail service. Service requirements for intermodal transportation are stringent, and scheduling must be planned in concert with other traffic requirements. More complete freight flow information will allow planners to determine the amount of demand in key markets and the capacity needed to meet this demand. Adequate capacity will ensure that intermodal transportation maintains a service standard that allows it to attract greater motor carrier support.

[5] This example was provided to the committee by Mort Plumb, Director, Ted Stevens Anchorage International Airport, Alaska.

CONCLUDING REMARKS

The objective of the 1991 Intermodal Surface Transportation Efficiency Act was "to develop a National Intermodal Transportation System." As a result, the last decade has seen an increased emphasis on multimodal approaches to transportation issues, including those affecting freight. At the same time, the U.S. economy experienced the longest uninterrupted expansion in its history over the period from 1991 to 2001 (TRB 2003). The resulting demands on the nation's transportation infrastructure, coupled with greatly heightened concerns about transportation security, have resulted in an urgent need for better data on freight movements to inform public-sector policy and investment decisions. Such data are also needed by the private sector to identify underserved and emerging markets and potential efficiency improvements. In the next chapter, the limitations of existing freight data are summarized, and the need for a new approach to provide the data required to inform public- and private-sector decision making is identified.

REFERENCES

Abbreviations

BTS Bureau of Transportation Statistics
NRC National Research Council
TRB Transportation Research Board

Boardman, J. 2002. Foreword. In *Conference Synthesis: Data Needs in the Changing World of Logistics and Freight Transportation* (A. H. Meyburg and J. R. Mbwana, eds.), New York State Department of Transportation, Albany. www.dot.state.ny.us/ttss/conference/synthesis.pdf.

BTS. 1998. *Transportation Statistics Beyond ISTEA: Critical Gaps and Strategic Responses.* BTS98-A-01. U.S. Department of Transportation, Washington, D.C.

BTS. 1999. *Transportation Satellite Accounts: A New Way of Measuring Transportation Services in America.* BTS99-R-01. U.S. Department of Transportation, Washington, D.C.

BTS. 2002. *International Trade Traffic Study (Section 5115): Measurement of Ton-Miles and Value-Miles of International Trade Traffic Carried by Highways for Each State* (draft). U.S. Department of Transportation, Washington, D.C.

Federal Register. 2002. Presentation of Vessel Cargo Declaration to Customs Before Cargo Is Laden Aboard Vessel at Foreign Port for Transport to the United States. Final Rule. Vol. 67, No. 211, Oct. 31, pp. 66,318–66,333.

Flynn, S. 2001. Cargo Clearance, Security, and Safety: National Security. In *Conference Proceedings 25: Global Intermodal Freight: State of Readiness for the 21st Century,* TRB, National Research Council, Washington, D.C., pp. 75–78.

Metroplan Orlando. 2002. Freight, Goods and Services Mobility Strategy Plan. www.metroplanorlando.com/msplan.

Meyburg, A. H., and J. R. Mbwana (eds.). 2002. *Conference Synthesis: Data Needs in the Changing World of Logistics and Freight Transportation.* New York State Department of Transportation, Albany. www.dot.state.ny.us/ttss/conference/synthesis.pdf.

Meyer, D. M. 2001. Delivering the Future: e-Freight. Presented at the Roundtable Discussion of e-Freight and Metropolitan Implications, Foundation for Intermodal Research and Education, March 7–9. www.intermodal.org/FIRE/meyerpaper.html.

NRC. 2001. *Inland Navigation System Planning: The Upper Mississippi River–Illinois Waterway.* National Academy Press, Washington, D.C.

TRB. 2003. *Special Report 271: Freight Capacity for the 21st Century.* National Research Council, Washington, D.C.

U.S. Department of Transportation. 2000. *NHS Intermodal Freight Connectors: A Report to Congress.* www.ops.fhwa.dot.gov/freight/pp/nhs%20final%20report.DOC.

Weisbrod, G., D. Vary, and G. Treyz. 2001. *NCHRP Report 463: Economic Implications of Congestion.* TRB, National Research Council, Washington, D.C.

Wilson, R. A. (ed.). 2001. *Transportation in America,* 18th ed. Eno Transportation Foundation, Washington, D.C.

2

Freight Transportation Data: Current Limitations and Need for a New Approach

In the United States, information on freight movements is collected by federal agencies and other public and private entities that monitor or analyze transportation and trade activities on a regional, state, national, or international level. Widely used federal databases providing freight transportation data at the national level include the Commodity Flow Survey (CFS) from the Bureau of Transportation Statistics and the Census Bureau, the Rail Waybill Sample from the Surface Transportation Board (STB), the Vehicle Inventory and Use Survey (VIUS)[1] from the Census Bureau, and the Waterborne Commerce of the United States database from the U.S. Army Corps of Engineers. Data from these sources are available free of charge but may be subject to confidentiality constraints. For example, limited data from the Rail Waybill Sample are made available in a public use file, but more detailed commercially sensitive data can be provided to certain parties upon approval by STB. Freight data avail-

[1] The VIUS was formerly known as the Truck Inventory and Use Survey.

able at a charge from private organizations include trucking industry directories, economic forecasts, and customized databases, such as the Transearch database from Reebie Associates and the Port Import Export Reporting Service (PIERS) database from the Journal of Commerce.

A 1997 report (Cambridge Systematics, Appendix C) identifies approximately 35 data sources containing particularly useful information on freight transport activity and demand. A somewhat shorter but more recent list of currently available freight data sources is provided by Meyburg and Mbwana (2002, Appendix 1), and Southworth (1999) identifies principal sources of data on national and international goods movement. The reader is referred to these reports for a more detailed discussion of the diverse sources of freight transportation data.

The limitations of current data sources for the purposes of freight transportation analyses are discussed in this chapter, and a need for a new, coordinated approach to data collection to fulfill the data requirements discussed in Chapter 1 is identified.

LIMITATIONS OF CURRENT DATA

At first sight, the large number of freight transportation data sources appears to indicate a plethora of information. However, transportation analysts seeking to use data from these diverse sources often encounter problems that detract from the usefulness of the available information. The following are among the problems:

- Variations in reporting that complicate the interpretation, comparison, and combination of data from different sources;
- Incomplete coverage of freight movements;
- Lack of geographic detail, particularly at metropolitan and local levels;
- Lack of information on data reliability; and
- Difficulties in using databases designed for purposes other than transportation analysis.

These issues are addressed in the following sections. The growing need for more timely data to accommodate shorter technology cycles and decision times is discussed in Chapter 3.

Variations in Reporting

Because existing databases have been developed independently over time to meet the specific demands of different users, they vary considerably in terms of the way that information is reported on items relevant to transportation analyses, such as origin and destination, shipment characteristics, and transport characteristics. While these variations may not present any problem for the intended users of a database, they frequently cause difficulties for analysts seeking to interpret, compare, and combine data from different sources for the purposes of characterizing freight flows.

Origin and Destination

True origin–destination (O/D) flows are defined as movements of goods between locations where they are produced and locations where they are consumed. Some databases include true O/D information, but others frequently do not. For example, databases covering a single mode may report the origin and destination of the portion of a multimodal freight movement made using that mode.

The PIERS database on waterborne freight includes O/D data taken from vessel manifests listing addresses of the shipper and consignee, although it does not include information on shipment routing or landside modes used in transporting shipments to and from inland origins and destinations. Historically, the biggest problem with the PIERS O/D data has been confusion between the location of the owner or bill-to party and the physical origin or destination of the shipment.

Trade databases in general tend to be problematic in providing true O/D data. In addition to the PIERS example discussed above, confusion can occur between administrative and physical ports of entry or exit. For example, in the case of goods entering or leaving at land ports, data providers are asked to indicate the port of entry or exit where the cargo actually entered or exited the U.S. land border. In practice, however, the reported port is often the administrative port where information about the transaction was filed—not where the cargo physically moved. Hence, some data show imports coming by land from Canada entering the United States in Dallas, Texas (BTS 2002). Clearly, such anomalies confound efforts to characterize freight flows on the U.S. transportation system.

Shipment Characteristics

Different industries and transport modes use different measures of shipment size—differences that are reflected in freight and trade databases. Although shipment weight is widely used, commodity-based sources often specify shipment volume and may use specialized volumetric units (e.g., bushels of grain, barrels of petroleum), thereby complicating efforts to generate comprehensive flow data (see Box 2-1). Trade databases may also specify shipments in terms of their total value, particularly for import/export purposes.

Box 2-1

▼

Barrels and Tons

The U.S. Army Corps of Engineers' Waterborne Commerce Statistics Center (WCSC) decided to change its primary source of foreign waterborne import, export, and in-transit data to Customs' manifest data. Following publication of the statistics for 2000, it was discovered that an incorrect factor had been used to convert barrels of inbound crude petroleum to tons. The size of such shipments reported in vessel manifests may be in barrels (volume), tons (weight), or both. No correction was needed for shipments where weight was reported, but data for shipment sizes reported only in barrels required correction. The correction process resulted in a nationwide decrease of 7.1 percent in WCSC's published weight of foreign inbound crude oil; the corrections for individual ports ranged from 0 to –13 percent.

Source: U.S. Army Corps of Engineers 2003.

The diversity of measures of shipment size is illustrated by the example of cargo passing through Port Canaveral in Florida (Metroplan Orlando 2002). This cargo falls into one of four categories, each of which uses different measures. Thus,

- Dry bulk and breakbulk cargoes are measured by weight,
- Liquid bulk is measured by volume,
- Roll on/roll off cargoes are counted in units and converted to a weight measure, and
- Containerized cargoes are measured in 20-foot equivalent units.

Information about commodities shipped also varies considerably in both form and level of detail. In some cases, industry-specific descriptive categories may be used to describe commodities, or goods may simply be described in terms of their handling characteristics (bulk, container, breakbulk). More often, commodities are described by using classification codes. Trade databases use product-based classification systems such as the Harmonized Schedule of Foreign Trade (HS) and the Standard International Trade Classification, whereas transport-oriented databases use classifications such as the Standard Transportation Commodity Classification and the Standard Classification of Transported Goods (SCTG). Each of these classification systems uses a hierarchical series of identifiers to provide different levels of commodity detail. In the HS system, for example, preparations of vegetables, fruit, nuts, or other parts of plants are assigned a two-digit identifier (20); fruit and vegetable juices are assigned a four-digit identifier (2009); and frozen orange juice is designated by a six-digit identifier (200911). In the SCTG system, fruit and vegetable juices are designated by a different four-digit identifier (0724) and the two-digit code "20" designates basic chemicals. Conversions between the various commodity coding systems have been developed, although there is always some loss of accuracy in the translation (Southworth 1999).

Transport Characteristics

Current data sources provide a variety of information on transport characteristics such as modes of shipment and equipment used. Some sources, such as the Rail Waybill Sample and the VIUS, cover a single mode; others, notably the CFS and the Transearch database, cover all major modes. Information on equipment type may be fairly specific. For example, the Rail Waybill Sample identifies railcar type, and the PIERS database identifies the carrier and vessel name—information that can be correlated with data from other sources, such as Lloyd's Register of Shipping, to learn more about equipment detail. In contrast, there is a notable lack of data on truck type and weight.

Incomplete Coverage

Although transportation and trade databases provide considerable amounts of data on freight activity, some aspects of freight movements are poorly characterized—at least using publicly available data. In particular,

- Coverage of different economic sectors and transportation modes is uneven;
- Coverage of international shipments is incomplete; and
- Some specific items of information, such as travel time from origin to destination, are reported only infrequently.

Disparities exist between the freight data available to business and the data available to government (Meyburg and Mbwana 2002, 7). For example, few data are available to government on the movement of freight by express urban delivery vehicles, despite the contribution of such vehicles to congestion in areas of high-density commercial activity. However, companies such as FedEx and United Parcel Service have detailed information on the nature and movements of their vehicle fleets (number and type of vehicles, number and time of day of stops, deliveries made per mile, value of commodities carried, etc.). Since such commercially sensitive data are likely to remain confidential, the following discussion of deficiencies in freight data coverage focuses on limitations in publicly available data.

Many discussions of freight data gaps take the CFS as their starting point (see, for example, BTS 1998). While this survey does not aim to provide complete coverage of all economic sectors or supply all the data items users may need, it is the major source of publicly available nationwide data on the movement of goods by all modes (air, motor carrier, rail, water, and pipeline) and intermodal combinations. The Transearch database is a private-sector derivative that attempts to fill some of the gaps in the CFS. The CFS and Transearch are the only two national freight databases that recognize the need for a complete picture of freight flows.

Economic Sectors and Modes of Transport

In terms of economic sector coverage, the CFS captures data on shipments originating from manufacturing, mining, wholesale, and selected retail establishments. It does not cover establishments involved in farming, forestry, fishing, construction, and crude petroleum production; households; governments; foreign establishments; and most retail and service businesses (BTS 1998). Some of these gaps can be filled by using data from sources covering single modes, such as the Rail Waybill Sample and the Corps of Engineers and PIERS databases on waterborne freight, all of which provide comprehensive coverage across economic sectors.

While the CFS provides some coverage of all modes of transport, there is widespread agreement among data users that the coverage of air and truck shipments is limited. Air freight is particularly important for establishments not included in the CFS. Although air freight data can be obtained from other sources, there is a general lack of detailed information on commodity movements by air for both domestic and international freight (BTS 1998). The CFS collects data on freight movements by for-hire and private trucks. However, the activities of many firms involved in the trucking industry have proven difficult to track (Southworth 1999), and truck freight data are widely acknowledged to be both complex and deficient. One of the short-term action items suggested by participants in the 2001 Saratoga Springs meeting was to concentrate on truck data "since this is where the bulk of the problems reside" (Meyburg and Mbwana 2002, 25). The deficiencies in truck data are a major concern because of the importance of motor carriers for moving freight within the United States. In 1997, single-mode movements by for-hire or private truck accounted for more than 60 percent of the tonnage moved in U.S. commercial freight shipments and almost 30 percent of the ton-miles (BTS 2001). Trucks also play an important role in movements involving multiple modes, such as truck-and-rail and truck-and-water.

A schematic representation of the coverage of freight movements provided by important data sources for different economic sectors and modes of transport is given in Figure 2-1.

International Movements

Information about the domestic portion of international goods movements has become increasingly necessary as a result of the growing importance of international trade to the U.S. economy and expanding competition among transportation providers in North America (BTS 1998). The CFS captures some information on these movements but is notably deficient in its coverage of imports. Because the survey samples domestic shipper establishments, it cannot capture information on shipments from foreign establishments. Imported products are included in the CFS at the point where they leave the importer's domestic location

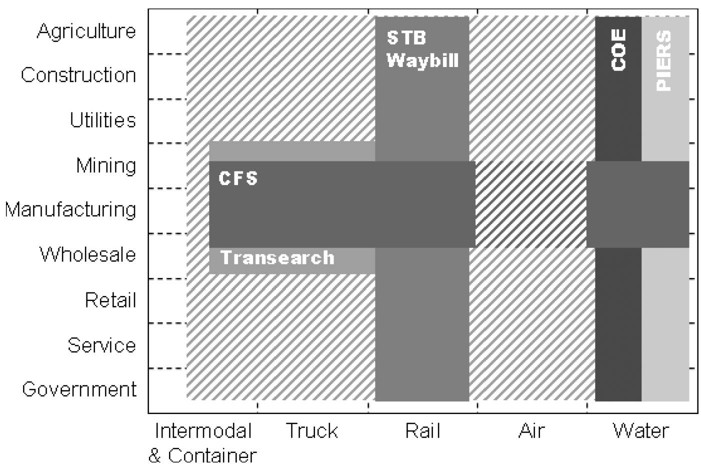

Figure 2-1 Coverage of freight movements for different economic sectors and modes of transport. (CFS = Commodity Flow Survey; STB Waybill = Surface Transportation Board's Rail Waybill Sample; COE = Army Corps of Engineers' Waterborne Commerce of the United States database; PIERS = Port Import Export Reporting Service database. Source: Rick Donnelly, PBConsult, Inc.)

(which is not necessarily the port of entry) for shipment to another location in the United States. Thus, the CFS does not cover the first leg of import shipments (TRB 2003a).

Foreign trade data may be helpful in filling gaps in CFS import data and in establishing comparisons among different data sources. However, such data have traditionally focused on economic transactions rather than physical commodity flows and often provide only limited information about the transportation aspects of the transaction. In the case of the Canadian trade statistics, the data are likely to show the dollar flows between one large multinational firm and another or between divisions or factories within a single firm. The actual flow of goods may take place between subsidiary locations far from those physically handling the import or export transactions.

Specific Data Items

Data on shipment frequency, travel time from origin to destination, and costs of transportation would be useful in assessing the performance of the freight transportation system in terms of cost and timeliness. Data on routing and time of day would also be valuable in assessing strategies for alleviating congestion and identifying opportunities for modal diversion. The CFS does not collect such data because of concerns about overburdening survey respondents, and publicly available data on costs, travel time, and system reliability from other sources are too limited to compile a complete national picture (BTS 1998). A recent report on freight capacity notes that the available aggregate data are generally inadequate for identifying local constraints on transportation systems (TRB 2003b).

Lack of Geographic Detail

Most national data sets, such as the CFS, do not provide the level of geographic detail required by analysts seeking to understand freight transportation at metropolitan area and county levels. There is an outstanding requirement for more focused data on freight movements at the metropolitan level to provide insights into transportation demand, the relationships between freight movement and business patterns, and freight flows through key corridors (BTS 1998). These metropolitan-level data need to be linked to data on global, national, and regional freight movements, since the latter have an important influence on what happens at the metropolitan level.

Much of the data already collected by the CFS could be useful in providing the levels of geographic detail required by analysts. However, Title 13 restrictions on the disclosure of company-specific information prevent the release of such data. The CFS data products present aggregate data to prevent the identification of individual establishments. The public release of much of the detailed trade data is similarly restricted to protect the confidentiality of data providers, notably carriers.

Sample size constraints also contribute to the lack of geographic detail in the published CFS data. As the sample size decreases, the statistical variability of the data increases. If the sample size is too small, the data may not be sufficiently reliable to be useful for analysis at the required level of geographic detail (TRB 2003a). The CFS provides information on state-to-state flows, but is not intended to provide detailed coverage of local freight movements. The information needed by local decision makers far exceeds the capacity of any single national survey, and local data collection efforts that supplement national surveys appear to be a more promising approach to remedying the current lack of detailed geographic data (BTS 1998).

Lack of Information on Data Reliability

Because many transportation-related investment decisions involve large expenditures and have long-lived implications, it is important to assess the reliability of data used to inform these decisions. Some sources, such as the CFS, are well documented in terms of the origin, quality, and limitations of the data they provide, thereby allowing users to assess data reliability. Others provide only limited insights into how data were obtained and processed, with the result that users risk providing information and advice on the basis of questionable data.

One of the most important distinctions affecting data reliability is that between "real" and synthesized data. The synthesis of missing data may sometimes be the only feasible method of providing timely information; this approach also avoids the expense of data collection and processing. However, synthesized data depend on the assumptions built into the analytical models used to generate them and how well these models replicate reality. Furthermore, while such models may be helpful, data are required to validate them. In principle, real data obtained from survey questionnaires, bills of lading and other trade-related documentation, and continuous electronic data streams generated by traffic monitoring and control systems provide a more solid basis for decision making, although care is still needed in assessing data reliability.

The design of any data collection effort affects the reliability of the resulting data. For the many existing freight data sources based on surveys, sample size is a key determinant of data quality and reliability. The methods of data collection are also important. For example, data taken from vessel manifests and bills of lading are not subject to the same concerns about nonresponse bias as data gathered by using survey questionnaires.

Most federal data sets present real (as opposed to synthesized) data and provide much of the information needed by analysts to assess the reliability of these data. Commercial data sets tend to be less transparent. For example, they may incorporate proprietary data whose origin and reliability are not reported. They may also use proprietary methods to combine different data sets and synthesize data and may interpret reported data to generate information requested by clients. In many cases, little information is provided about the assumptions used in generating derived data products, with the result that their reliability is unknown.

Difficulties in Using Databases Designed for Other Purposes

Commodity flow–type freight databases typically include data on commodity type, shipment origin and destination, shipment weight, and shipment value for one or more modes of transport. Supplementary freight databases do not contain any commodity flow data but provide information that can be used in the development of freight models. For example, the VIUS microdata provide information on percentages of loaded miles by commodity for vehicles on an annual basis. (The records are masked to avoid possible disclosure of individual vehicles or owners.) However, the survey does not provide data on when during the year these commodities are carried or on the origin and destination of shipments. Trade databases contain information about freight movements to and from, and sometimes within, the United States.[2]

The supplementary freight and trade databases were not designed to track freight flows. For example, VIUS data are used to perform cost allocation analyses, study safety issues, determine user fees, and estimate per-mile vehicle emissions.[3] Many trade databases provide information on goods entering and leaving the country to inform decisions about the legal admissibility of and duty applicable to imported merchandise, the safety of food products for consumption, and the like. Consequently, the data reported may be incomplete for the purposes of transportation analysis or may include inconsistencies that are difficult to reconcile when developing information on freight flows. Two examples illustrate these difficulties.

The source documents (ship manifests, bills of lading) used to compile the PIERS database on waterborne freight are intended to provide evidence of liability and are not developed primarily as a source of freight transportation data.[4] For example, a bill of lading functions as a receipt, as evidence of a contract of carriage, and as a document of title to the goods being shipped. Because the ocean carrier may be responsible for transporting the goods only from Port A to Port B, the bill of lading may not show the true O/D of the shipment, and the PIERS database will be similarly deficient. A transportation analyst seeking to establish a complete picture of the freight movement will, therefore, need to find other data describing the initial movements from the true origin to Port A and the final movements to the true destination from Port B.

The province of Ontario in Canada relies heavily on trucking knowledge gained from roadside surveys of carriers. While such surveys provide reliable vehicle data and relatively accurate commodity data, they cannot provide data on commodity value—data that are critical to informing investment decisions and building linkages with economic input/output models. In light of the economic importance of trade with the United

[2] For further discussion of trade databases and the two major categories of freight database (commodity flow and supplementary), the reader is referred to *Freight and Trade Data Information*, Center for Transportation Research and Education, Iowa State University, Ames, Iowa (www.ctre.edu/research/bts_wb/cd-rom/freight_intro.htm).

[3] Further information on the use of VIUS data is provided on the Census Bureau's website (www.census.gov/econ/www/viusmain.html).

[4] As reported to the committee by Bill Ralph, PIERS, September 19, 2002.

States and emerging issues relating to trade corridors, attempts have been made to obtain commodity values from other sources. However, data from shippers, receivers, the census, taxation officials, and Customs include different interpretations of commodity value, depending on the intended use of the data. Thus, the reported commodity value may reflect the insured value (possibly inflated), value added at different stages of manufacture, deliberate underreporting of value to Customs, nonreporting of in-bond or tariff-exempt flows, self-assessed value with the intent to avoid interference and subsequent secondary customs inspection, wholesale value, retail value, manufacturer's estimate of costs to manufacture, transportation costs, and so on. Thus, there is no single reliable source of unambiguous data on commodity value to complement the information from roadside surveys.

NEED FOR A NEW APPROACH

Participants in the 2001 Saratoga Springs meeting on freight data needs observed that data collection, storage, and distribution are expensive activities, and they urged data users to make full use of available data (Meyburg and Mbwana 2002). They also noted that "any effort to collect new freight data should be preceded by an understanding as to why the new data are needed" (p. 23). Having established the reasons why freight data are required (see Chapter 1), the study committee initially considered the possibility that a comprehensive national picture of freight flows might be developed by combining existing data sources without the need for modifications to current data-gathering procedures or additional data collection initiatives. The committee quickly concluded that creating a comprehensive national freight database by patching together existing data sources is not feasible. The lack of coordination among different data collection efforts creates fundamental problems not readily overcome. In particular, the niche data products can be linked only with considerable difficulty and accompanying loss of accuracy, and even then important data gaps remain.

The committee concluded that a new approach is required to provide the freight data needed to inform important policy and investment decisions such as those described in Chapter 1. In the committee's view, the coordination of freight data collection efforts through the implementation of a national freight data framework offers improvements over the current approach. Such a framework needs to recognize and build on the strengths of current data sources, but also establish linkages among different data sources in an effort to eliminate unconnected data "silos." Similarly, any new data collection initiatives for filling data gaps need to be designed so that the resulting data can be readily integrated with data from other sources.

A recent National Research Council report notes that "when separate datasets are collected and analyzed in such a manner that they may be used together, the value of the resulting information and the efficiency of obtaining it can be greatly enhanced" (NRC 2001, 7). Thus, the committee anticipates that a coordinated approach to collecting freight data should be more efficient than current efforts. The uncoordinated collection of trade data by different federal agencies provides an interesting analogy to the situation with freight data. Currently, traders importing or exporting goods provide information to each individual trade agency by using a variety of different automated systems, numerous paper forms, or a combination of systems and forms. The United Nations Conference on Trade and Development has estimated that submission of redundant information and preparation of documentation amount to 4 to 6 percent of the cost of the merchandise. The International Trade Data System (ITDS), a federal government information technology initiative, aims to develop a system that will allow traders to submit one set of standard electronic data for imports or exports. ITDS will then distribute the data to the relevant federal agencies on a need-to-know basis. Requirements for traders to submit redundant information to multiple federal agencies will be eliminated, with associated reductions in cost and respondent burden.[5]

A new, integrated approach to gathering freight data will need to offer advantages similar to those of ITDS in terms of more cost-effective data collection and reduced respondent burden. The approach will also need to be more effective than the current fragmented system in providing data that meet the needs of a range of users. A concept for a national freight data program that aims to meet these ambitious goals is presented in the next chapter.

[5] ITDS Background (www.itds.treas.gov/itdsovr.html).

REFERENCES

Abbreviations

BTS	Bureau of Transportation Statistics
NRC	National Research Council
TRB	Transportation Research Board

BTS. 1998. *Transportation Statistics Beyond ISTEA: Critical Gaps and Strategic Responses.* BTS98-A-01. U.S. Department of Transportation, Washington, D.C.

BTS. 2001. *Transportation Statistics 2000.* BTS01-02. U.S. Department of Transportation, Washington, D.C.

BTS. 2002. *International Trade Traffic Study (Section 5115): Measurement of Ton-Miles and Value-Miles of International Trade Traffic Carried by Highways for Each State* (draft). U.S. Department of Transportation, Washington, D.C.

Cambridge Systematics, Inc. 1997. *NCHRP Report 388: A Guidebook for Forecasting Freight Transportation Demand.* TRB, National Research Council, Washington, D.C.

Metroplan Orlando. 2002. Freight, Goods and Services Mobility Strategy Plan. www.metroplanorlando.com/msplan.

Meyburg, A. H., and J. R. Mbwana (eds.). 2002. *Conference Synthesis: Data Needs in the Changing World of Logistics and Freight Transportation.* New York State Department of Transportation, Albany. www.dot.state.ny.us/ttss/conference/synthesis.pdf.

NRC. 2001. *Principles and Practices for a Federal Statistical Agency,* 2nd ed. National Academy Press, Washington, D.C.

Southworth, F. 1999. *The National Intermodal Transportation Data Base: Personal and Goods Movement Components* (draft). Oak Ridge National Laboratory, Oak Ridge, Tenn.

TRB. 2003a. Letter Report on the Commodity Flow Survey. National Research Council, Washington, D.C. gulliver.trb.org/publications/reports/bts_cfs.pdf.

TRB. 2003b. *Special Report 271: Freight Capacity for the 21st Century.* National Research Council, Washington, D.C.

U.S. Army Corps of Engineers. 2003. Corrections to CY2000 Published Crude Oil Import Statistics. www.iwr.usace.army.mil/ndc/usforeign/index.htm.

3

Concept for a National Freight Data Program

A conceptual plan for a national freight data program is presented in Appendix A under the section entitled "Product Definition." The plan is based on an initial concept proposed by the committee's consultant, Rick Donnelly, and further developed by Dr. Donnelly after extensive discussions with the committee at its second and third meetings. This chapter provides the committee's commentary on Dr. Donnelly's proposed plan.

The framework for a national freight data program illustrated in Figure 3-1[1] and described in Appendix A proposes establishing an advisory committee to oversee the design and implementation of a multifaceted data collection program. An integrated program of freight surveys and a freight informatics initiative that gathers data from electronic data streams, such as those associated with intelligent transportation systems (ITS) and electronic data interchange (EDI), would provide the data needed to populate a national freight database. These data would be supplemented by

[1] Figure 3-1 is adapted from Figure A-1 of Appendix A.

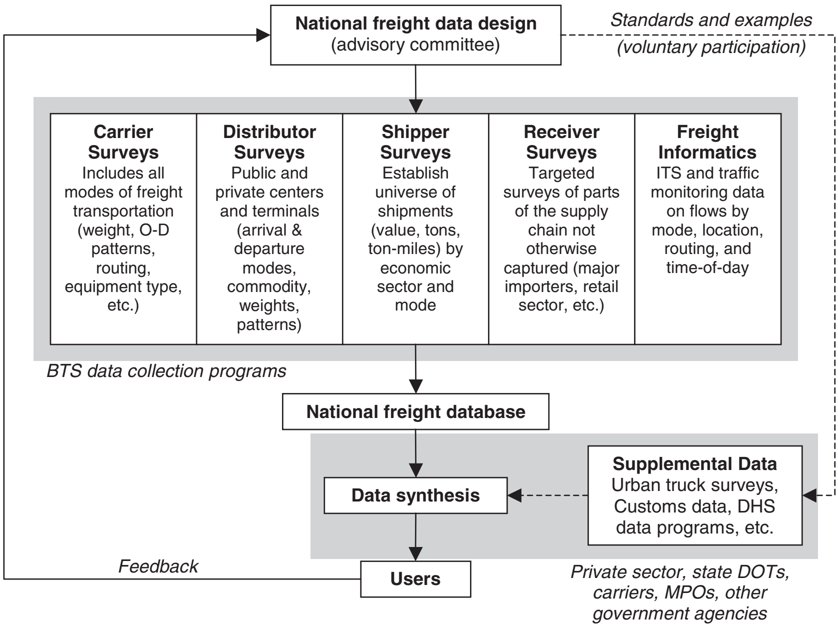

Figure 3-1 Proposed framework of a national freight data program. [BTS = Bureau of Transportation Statistics; O-D = origin–destination; DHS = Department of Homeland Security; MPO = metropolitan planning organization; state DOT = state department of transportation. Source: Adapted from a paper prepared for the committee by R. Donnelly (Appendix A).]

data from other sources, such as urban truck surveys, and by synthesized data. The resulting databases would be made widely available to the user community, whose comments and feedback would help inform further development of the framework.

In the committee's opinion, the framework shown in Figure 3-1 forms a guide for improving on the current patchwork of uncoordinated freight data collection efforts by a more systematic approach that eliminates unlinked data "silos." The proposed framework focuses on increasing the linkages between different sources of data and filling data gaps to develop a comprehensive source of timely and reliable data on freight flows.

The committee recognizes that the implementation of a national freight data framework such as the one proposed will require a sustained effort—and funding—over many years and will involve many technical and organizational challenges. From a technical perspective, the amount of information required is large, and some of the information needed by decision makers—such as comprehensive data on route, time of day, and commodity for highway freight movements—has not previously been collected in the United States. Research will be needed in areas such as survey methodology and data processing, and the effort will not succeed without innovative, low-cost data collection strategies.

From an organizational perspective, the success of a national freight data program to implement the framework will depend on the participation of diverse public- and private-sector organizations at various levels. The assignment of responsibilities for tasks such as the conduct of surveys, database development, and data synthesis will require further discussion and elucidation as part of program development and implementation activities. For example, the committee does not envisage the Bureau of Transportation Statistics (BTS) assuming responsibility for all the data collection activities (survey programs and the freight informatics initiative) grouped together in Figure 3-1 under the designation "BTS data collection programs."

Instead, the committee anticipates that much of the data will continue to be collected by the same organizations as today (U.S. Army Corps of Engineers, Surface Transportation Board, Census Bureau, etc.). A coordinating body such as BTS, working under the guidance of an advisory committee of stakeholders and data experts,[2] will take measures to encourage harmonization of these data collection efforts and will coordinate access to the data as appropriate. Public- and private-sector roles are discussed in the section of this chapter addressing challenges in implementing a national freight data framework.

The committee has deliberately proposed a flexible data framework that can evolve as research results indicate which data collection strategies are likely to be the most fruitful. For example, several types of survey are proposed, but it is unlikely that all will be pursued concurrently or that equal effort will be devoted to each. Since resources are limited, it will be necessary to identify the most promising avenues for development and implementation and prioritize funding allocations accordingly. The committee recognizes that, in some instances, further investigation

[2] The composition and role of the proposed advisory committee are discussed in the Findings and Recommendations chapter.

may reveal certain features of the conceptual plan to be impractical or not cost-effective.

Despite the many challenges to implementing a national freight data framework, the committee observed widespread agreement among data users—including high-level decision makers—that current freight data sources are not meeting the need for reliable data to inform investment, planning, and policy decisions. Thus, many stakeholders may be willing to support and participate in a national freight data program that offers benefits for diverse data users—and ultimately for the national transportation system as a whole. Nonetheless, it will be important to encourage the participation of data providers by clearly defining anticipated benefits. If the imposition of focused private costs on survey respondents appears to offer only diffuse benefits, the necessary broad participation is unlikely to be achieved.

The rationale for the conceptual plan presented in Appendix A is described in the remainder of this chapter, and additional insights into some of the challenges to be addressed in implementing the plan are provided.

RATIONALE FOR CONCEPTUAL PLAN

The conceptual plan for a national freight data program comprises five major components:

- A national freight data framework,
- An integrated program of freight surveys,
- A freight informatics initiative,
- Freight data synthesis, and
- Standard survey methodologies.

The supporting rationale for each of these components is discussed below.

National Freight Data Framework

Because it would be impractical for the federal government to meet all the freight data needs of all users, the proposed national freight data framework facilitates opportunities for combining data from a variety of sources and identifies possible roles for a range of stakeholders, including the federal government, the private sector, state departments of transportation, metropolitan planning organizations (MPOs), and research organizations. The framework comprises a series of modules, as follows:

- A national freight database will be populated through an integrated program of freight surveys and a freight informatics initiative. The survey program itself comprises another series of modules, namely, carrier surveys, distributor surveys, shipper surveys, and receiver surveys.
- A freight data synthesis program will fill data gaps, particularly in the short term before all of the proposed data sources are fully established.
- Supplemental data collection activities will provide additional specialized data to complement the data in the national freight database. Standard survey methodologies and examples of their use will be provided to guide supplemental data collection and help ensure compatibility with national data.

The committee anticipates that different users will use the various modules in different ways. For example, users interested in local transportation issues will gather supplemental data from their own jurisdictions to obtain the level of geographic detail they require. Following the practices and guidelines in the standard survey methodologies will help these users ensure that their data can then be combined with data from the national freight database. Other users will combine data from the national freight database with synthesized data to fill gaps that are too difficult—and expensive—to fill by using surveys or other data-gathering techniques. Yet others will leverage federal economic or trade data programs, such as the International Trade Data System (ITDS) of the Bureau of Customs and Border Protection, to obtain supplementary data that they will combine with data from the national freight database.

The provision of training and education in the use of freight data resources and methods will be an important component of efforts to implement the framework. As discussed in a report from BTS (1998), the availability of sophisticated models, complex analytical methods, and large data sets to relatively inexperienced users necessitates new approaches to training. Data customers are no longer limited to analysts in a few federal agencies and large consulting companies, and training programs will need to recognize user diversity. While the proposed framework offers different users the option to employ the various modules in different ways, many will require training so they understand how to use the modules correctly and obtain meaningful results.

Since funding for new freight data initiatives is limited, simultaneous exploration of all possible data collection and synthesis options is not realistic. The advisory committee responsible for guiding the design and implementation of the national freight data framework will need to weigh the costs of obtaining data against the potential benefits when advising on program priorities and a timetable for development. The proposed modular approach lends itself to an incremental development process whereby various options can be investigated as resources permit, as new technologies become available, and as opportunities arise to leverage ongoing activities. Under this approach, implementation of the framework could focus initially on developing the national freight database rather than devoting equal effort to all aspects of the initiative. A similar phased approach could be adopted in populating the database. For example, ITS offers the possibility of collecting routing, time, carrier, and origin–destination data for trucks but cannot at present routinely determine the commodity carried or certain truck characteristics. Therefore, a carrier survey could be used initially to obtain the commodity and vehicle data needed to complement shipper survey data. The carrier survey could then be phased out and replaced by more sophisticated ITS data collection methods as these become available.

Table A-1 (Appendix A) proposes a schedule for a national freight data program to implement the proposed framework. The committee views the task breakdown and sequence outlined in Table A-1 as a helpful and appropriate basis for discussion and development by the freight data advisory committee. Further definition and detailed sequencing of the tasks will be needed, probably in the form of a road map. Clearly, the availability of funding and the pace of research will determine how many years the entire process will take.

Integrated Program of Freight Surveys

Comprehensive data on freight flows cannot be collected by using a single type of survey. As discussed in Appendix B, the complexity of supply chains and the number of agents involved in moving goods from their origin to final destination make it impossible to design a single survey process that would provide data on all aspects of freight flows. The shipper, receiver, customer, carrier, regulator, and distributor all make or influence decisions about freight movements, and most have only limited knowledge of the factors influencing decisions made by others. For example, a shipper may specify a date and time by which certain goods need to arrive at their destination but may not know (or need to know) the route traveled or the modes of transport used. Thus, a shipper survey is unlikely to provide good route data and may provide only incomplete modal information. Additional data from carriers are needed to provide some of the information required for a national freight database. Some analysts already fuse carrier and shipper data, albeit with some difficulty, to inform their investigations. For example, carrier manifest data from the Port Import Export Reporting Service database have been combined with customs data supplied by shippers to inform maritime infrastructure planning and analyses.

Some understanding of the supply chain is necessary to identify the best ways of gathering the various items of data needed by different users.[3] For example, in deciding how to gather data on goods movements, it is important to understand that a carrier who transports goods from a warehouse to a distribution center may not know the true origin of the goods or their final destination. Thus, although a carrier survey (for example, a roadside truck survey) can potentially provide good data on mode, routing, time of day, and vehicle size and configuration, such a survey may well be of limited use in providing reliable data on the true origin and final destination of goods movements.

A dramatic increase in the importance of warehousing and distribution centers over the last two decades suggests that distributor surveys could provide useful data on freight movements (see Appendix B). When goods move through an intermediate point, such as a marine port or truck terminal, the visible linkage between the original shipper and ultimate receiver is broken. However, surveying a transportation terminal or distribution center could provide information on movements into and out of this intermediate location, as well as on the transition between the two movements. Despite their potential advantages, distributor surveys are conducted only

[3] The purpose of the national freight data program is not to develop a quantitative description of the entire supply chain.

infrequently. Possible impediments to such surveys include the need to secure the cooperation of private-sector owners of distribution hubs and concerns over protecting the confidentiality of clients.

A judicious balance among the different types of survey (and other data collection methods) will likely be needed for cost-effective development of a national freight database. This balance is expected to change over time as experience is gained and new data collection opportunities arise.

Because different types of surveys have different data gaps, they can, in principle, be used to complement each other. Southworth (1999) notes that two or more data sources can be combined by a process known as "data fusion" to create more complete movement information without the additional costs of further data collection. However, as discussed in Chapter 2, different surveys have been developed independently to provide data to meet specific, and diverse, needs. As a result, the technical challenges in integrating data from different surveys are often formidable, in large part because differing data collection strategies and data definitions raise concerns about the quality and comparability of the resulting combined data. In addition, errors may occur because of confusion arising from the different ways of describing and quantifying shipments. Therefore, an important feature of the integrated program of freight surveys is the development of survey designs providing data that can readily be fused to provide users with the information they require.

Data fusion is a complex and challenging process, as the examples discussed by Southworth (1999) illustrate. Thus, specialized technical expertise will be needed to identify and develop approaches for facilitating the combination of data from different sources. Common data elements across surveys—such as commodity classifications, geographic information, and mode definitions—would not overcome all the difficulties but could aid in the fusion process. For example, geographic information systems could be used to connect different surveys that collect precise information on vehicle location (latitude and longitude) at specified times.

Freight Informatics Initiative

The budget for the 2002 Commodity Flow Survey (CFS), which collected data on 2.7 million shipments from 50,000 domestic shipper establishments, is $13 million. Order of magnitude cost estimates indicate that developing a comprehensive national freight database by using only data collection strategies similar to those currently used for the CFS is unrealistic. Affordable collection of the large amounts of data needed to provide users with high-quality data at useful levels of geographic and commodity detail will require innovative and less costly strategies.

Although the use of technology—for example, electronic reporting using Web-based survey questionnaires—offers opportunities to improve data quality and reduce both cost and respondent burden in survey programs, it is unlikely to result in the major cost reductions needed to populate a national freight database of the type envisaged. In contrast, passive data collection methods that take advantage of continuous electronic data streams from traffic-monitoring systems or mine transaction data from EDI systems promise large quantities of data at low cost. The purpose of the freight informatics initiative is to investigate the technical and institutional aspects of such passive data-gathering approaches. While current practical experience with passive data collection is limited, the committee believes that without such alternatives to conventional surveys the costs of developing a comprehensive national freight database will be prohibitive. Furthermore, passive data collection may be the most promising approach for gathering the reliable data on shipment routing and time of day required by many users. Given the importance of low-cost data collection methods for the overall success of the framework initiative, the committee would like to see an early emphasis on freight informatics pilot studies as part of the framework implementation process.

Freight Data Synthesis[4]

Ideally, the national freight database will contain "real" data, gathered by using a range of surveys and passive data collection methods, rather than synthesized data generated by simulation techniques. However, participants in the November 2001 Saratoga Springs meeting noted that data collection, storage, and distribution are expensive activities and stressed that data users should make full use of available data and where possible

[4] The term "data synthesis" is used here to describe the generation of missing data by the use of simulation techniques.

"use analytical models to fill in data gaps" (Meyburg and Mbwana 2002, 23). Gathering the data needed to fill many of the current gaps will take time and resources, and, at least in the short to medium term, the national freight database is likely to contain important amounts of synthesized data. In the longer term, the replacement of much of the synthesized data by real data should allow users to have more confidence in the data they use to inform decisions.

Standard Survey Methodologies

The freight data needed for all proposed applications exceed current and expected future capabilities of national freight data sources, as illustrated by the examples given in Chapter 1. It is not clear that a national freight database can, or should, provide the large amounts of data required to capture the diversity of establishment sizes and inbound/outbound flows for all MPOs, counties, and local jurisdictions nationwide. Furthermore, even within a category such as MPOs, there is considerable variability in data requirements. For example, some MPOs require detailed data to address serious congestion problems, whereas others have little need for such information. In the committee's view, the national freight database should focus on providing a large number of users with frequently requested data items—such as origin and destination, commodity information (characteristics, weight, value), modes of shipment, routing and time of day, and vehicle or vessel type and configuration.

The national freight database will form a foundation on which users can build their own data sets. Thus, users concerned with regional markets and metropolitan, county, and local issues will be able to supplement the national database with localized data to obtain the degree of geographic resolution they require. The inclusion of standard survey methodologies in the national freight data framework is intended to assist users in generating supplemental data compatible with the national freight database.

The committee envisages that standard survey methodologies will address survey design issues, including determination of the sample size needed to provide the required level of geographic detail with an acceptable degree of reliability. Different methods of data collection (mail, Web-based survey, telephone, etc.) and anticipated response rates will also need to be considered. To help ensure consistency, key items such as origin and destination will need to be clearly defined, as will standard categories of truck type and size. Recommended best practices will address a range of survey topics, including the use of commodity categories with the potential to "roll up or down" to broader or narrower categories. Advice will also be required on assumptions about the first point of rest for freight shipments and how to avoid double counting. Examples of well-designed survey instruments that have yielded quality data will be provided for guidance.

CHALLENGES IN IMPLEMENTATION

The purpose of this report is to provide a blueprint to guide the reengineering of today's disjointed patchwork of freight data into a more integrated and useful national freight data program, rather than to describe the detailed implementation of the framework illustrated in Figure 3-1. However, the committee considers it important to highlight some of the challenges likely to be faced in the implementation process. Some general principles pertinent to this implementation are discussed in the following sections, and issues relating to data quality and timeliness, new data collection opportunities, confidentiality, and the roles of the public and private sectors are identified. Some specific issues that may well arise during implementation of the framework are discussed in Appendix D.

General Principles

Although many current sources of freight data are far from ideal for the purposes of freight transportation analyses, the national freight data program will need to be developed in the context of these sources. The content and detailed structure of the data framework will evolve to reflect research findings and practical experience, but there will be a continuing need to provide consistent data for trend analysis. "Wiping the slate clean" by initiating a set of totally new data collection programs risks jeopardizing the ability to study trends in freight movements over time—a subject of considerable importance for many investment and policy decisions. Thus, implementation of the national freight data framework will need to be an evolutionary, rather than a revolutionary, process that builds on experience with a range of surveys, takes account of data classifications and standards, and establishes links to previous data sets for the purposes of time series analyses. The data in the national freight database will also need to be updated on an ongoing basis. These updates will need to take account of the evolving nature of the data framework in such a way as to permit trend analyses.

Initiatives to modify or replace current surveys will need to recognize the strengths and weaknesses of current data sources, as illustrated by two examples. First, the CFS has been criticized because of its incomplete coverage and lack of geographic detail. Nonetheless, because the CFS is coupled to the Economic Census, response is mandatory and response rates are correspondingly high (approximately 75 percent for the 1997 CFS). Decoupling any successor to the CFS from the Economic Census could well result in a reduction in response rate—and an associated deterioration in data quality—unless an alternative mechanism could be established to require response to the new survey. Furthermore, diverting resources away from the CFS to other freight data initiatives could result in a reduced sample size and a resulting reduction in the usefulness of CFS data. Nevertheless, new data collection initiatives could remedy some of the deficiencies of CFS data, resulting in an overall improvement in the coverage and quality of multimodal freight data.

Second, data sets from the Surface Transportation Board (Rail Waybill Sample) and the Corps of Engineers already provide useful carrier data on rail and waterway movements, respectively. Rather than duplicate these efforts, it would seem appropriate for the national freight data program to focus on filling gaps and providing better data where current coverage is inadequate—for example, in the areas of motor carrier and air freight data.

Data Quality and Timeliness

Users need to have confidence in the data they use to inform decisions. In particular, most users would like to have information on how data were obtained, together with an estimate of their reliability. Although the widespread use of various commercial databases demonstrates the existence of a role in the marketplace for data that are not entirely transparent, many decision makers indicated to the committee that they would like to see a clear distinction made between real and synthesized data. Data also need to be provided in a timely fashion if they are to be useful for decision making. Participants in the Saratoga Springs meeting observed that government long-term planning now has to provide for infrastructure capacity needs and efficient operations with shorter lead times because of the shorter life cycles of technology products (Meyburg and Mbwana 2002). Short times for decision making require more complete and detailed information sooner than is provided by data sources such as the 5-yearly CFS. The development of survey designs for the integrated program of freight surveys will need to take account of these user requirements for timely and reliable data.

The freight data business plan presented in Appendix A proposes moving away from periodic (5-yearly) surveys toward continuous surveys to obtain more timely data for decision making. The motivation for this proposed move is very similar to that which stimulated the Census Bureau to develop the American Community Survey (ACS). Rapid demographic and technological change meant that a static, once-in-a-decade snapshot of the nation's population was no longer a satisfactory basis for informing a wide range of policy decisions and the allocation of billions of dollars of federal funds. Therefore, Census Bureau managers concluded that the design of the decennial census needed to be simplified and long form data collection made more timely in response to stakeholders' requests. The ACS development program, which collects demographic and socioeconomic data in 3-month cycles, divides a huge nationwide workload into manageable pieces over a longer time frame, with the result that data products can be released much sooner than under the conventional episodic approach (Census Bureau 2001). From a development perspective, continuous surveys offer the potential to obtain timely feedback on new data collection strategies and other survey design features, rather than having to wait several years to learn how a new approach is working in practice. Thus, if continuous data collection is adopted for the integrated program of freight surveys, the committee hopes that relatively rapid progress can be made in developing and refining survey designs. The cost implications of a move to continuous data collection are dependent on the details of the survey design (sample size, frequency, coverage, and the like) and will require further study.

Research will be needed to optimize the designs of the proposed rolling series of cross-sectional freight surveys. For example, a better understanding of the changed (and changing) nature of commodity distribution patterns would inform the development of effective sampling strategies. New business practices, such as freight logistics and just-in-time manufacturing, have changed the nature and volumes of goods shipped and the frequency, origins, and destinations of shipments. Thus, while the movement of bulk goods (e.g., grains, coal, and ores) still makes up a large share of the tonnage moved on the U.S. freight network, lighter and more valuable goods (e.g., computers and office equipment) now make up an increasing proportion of what is moved (FHWA 2002). A better understanding of the patterns of movement of these nontraditional commodities, possibly obtained through analysis of ITS data, would help inform decisions about survey design, notably sampling.

Wherever possible, survey research in support of the national freight data program should draw on the experience of public- and private-sector organizations in the United States and overseas. For example, there are valuable lessons about carrier surveys to be learned from Canada's National Roadside Study (NRS), a joint effort by federal and provincial transportation officials to conduct roadside carrier surveys capturing information on vehicle, cargo, trip, driver, and carrier for both international and domestic intercity trips. The 1999 NRS collected information from approximately 65,000 truck intercepts at 238 data collection sites spread across the Canadian road network. The surveys attempt to provide much of the detailed geographic information needed by analysts and planners. For example, data are gathered on highways used, stops along the trip, international border crossings, provincial boundary crossing points, and both trip and commodity origin and destination in an effort to identify both carrier and shipper logistics.[5]

The use of technology to collect survey data is expected to result in improved data quality, with the possible added benefits of reduced cost and reduced respondent burden. For example, Canadian roadside truck surveys currently involve photocopying waybills for reference purposes. In the future, it may be possible to take a digital picture of the truck with a personal digital assistant to confirm the observations of those conducting the survey. Web-based surveys also show promise for a range of applications. For example, the implementation of a Web-based questionnaire

[5] Further information on the NRS is available on the website of the Eastern Border Transportation Coalition (www.ebtc.info).

for the 2002 Economic Census will allow businesses to extract data directly from their own spreadsheets rather than having to transcribe information onto a questionnaire form, thereby reducing the likelihood of transcription errors. This electronic reporting is expected to significantly lower the respondent burden and associated reporting costs incurred by some large businesses, as well as save the Census Bureau both time and money.[6]

During the early stages of implementing the national freight data framework, the emphasis is likely to be on the quality and timeliness of data in the national freight database, and particularly the data gathered as part of the integrated program of freight surveys. The committee identified a clear role for BTS, as a federal statistical agency within the U.S. Department of Transportation, to initiate and stimulate activities aimed at improving data quality and timeliness. In this context, consideration could also be given to including quality control procedures in the data framework, perhaps by incorporating predefined performance measures against which to assess survey designs and evaluate the statistical reliability of data sources.

In the longer term, the lessons learned from developing the freight survey program are expected to inform the development of the standard survey methodologies. Experience with the Canadian NRS suggests that special efforts may be needed to encourage organizations to follow survey guidelines designed to generate quality data. While there is consensus among the provinces that driver interviews should be conducted by local staff familiar with regional travel and vehicle characteristics, important variations still occur. Different groups with different objectives (enforcement, planning, policy making, etc.) gather the data, reflecting each province's reasons for participating in the NRS. Some focus on collecting data on vehicle weight and dimensions for enforcement purposes, and others focus on collecting data on trip details (origin and destination, highway used, border crossings) for planning purposes.[7] While such differences are understandable, it is important to capture local detail in a well-planned and consistent manner when national data for a wide range of uses are collected.

---

[6] 2002 Economic Census: Electronic Reporting (www.census.gov/epcd/ec02/ec02electronic.htm).

[7] As reported by committee member Robert Tardif, Ontario Ministry of Transportation, during committee discussions.

Also in the longer term, the harmonization and streamlining of data collection efforts should result in data quality improvements as data from different sources are used to perform consistency checks and cumbersome and error-prone data-gathering methods are superseded.

New Data Collection Opportunities

In the short term, conventional transportation surveys (shipper surveys, carrier surveys, and the like) are expected to be the main source of data in the national freight database. However, because such surveys are frequently expensive and a burden on respondents, it will be important to identify and exploit additional data sources as the national freight data program develops. New data sources may emerge in the future, but the committee identified two current opportunities as potentially promising—security data and passive data collection technologies.

Security Data

Concerns over national security may drive more timely and efficient collection of freight transportation data, including real-time data on goods movements and more detailed information on the nature and value of shipments. Much of this information could be useful for modeling and planning applications and for identifying opportunities to improve capacity utilization. However, the extent to which such security data will be made available for nonsecurity applications is unclear. There is even concern in some quarters that certain data currently available to the public may no longer be generally accessible because of their potential to undermine national security.

Regardless of the unanswered questions about access to security data, the committee believes that the national freight data program could provide the Department of Homeland Security with important information on freight movements. In particular, establishing a picture of normal freight flows would be valuable as a baseline against which to identify anomalies associated with possible security threats. In the committee's view, there are mutual benefits for BTS and the Department of Homeland Security in working together to ensure that (*a*) security data feed into the national freight data framework as far as possible within disclosure constraints and (*b*) the framework is designed to support security-related data needs.

Passive Data Collection

As discussed in the earlier section on the freight informatics initiative, the national freight data program will need to take advantage of non-survey data streams. Research will be required to investigate

- Opportunities offered by new technology for low-cost passive data collection; and
- Methods for sampling and processing the large quantities of data generated by monitoring and control systems that function 24 hours per day, 7 days per week.

For example, the tracking of electronic transmitters on shipping containers may result in low-cost, high-quality data, but new data sampling and processing strategies will be required to make the most effective use of such large quantities of information.

Opportunities have already been identified for using ITS data for a much wider range of applications than originally anticipated (see, for example, BTS 1998 and Margiotta 1998). Most ITS systems are designed to manage day-to-day or minute-to-minute conditions, and many of the data collected for these monitoring and control purposes are not saved. If ITS data are to be used in populating a national freight database, procedures for data integration and archiving will be required. Legal issues, privacy concerns, and limitations on the use of proprietary data will also need to be addressed. In the longer term, the possibility of developing ITS further to meet specific data collection needs may merit investigation.

The involvement of data providers in developing and implementing new data collection methods is likely to be important to the ultimate success of such efforts. For example, the trucking industry will need to be involved in discussions about the possible collection of data from Global Positioning System–based truck-tracking systems to ensure that proposed approaches are not only technically feasible but also compatible with normal operational practices.

The committee recognizes that transitioning to passive data collection is "not a panacea" (BTS 1998, 18). Setup costs may be high, and the availability of continuous data streams will necessitate innovation in data management and processing. Nevertheless, passive data collection technologies are developing rapidly in coverage and sophistication and have the potential to generate large amounts of useful, high-quality data at low cost. Therefore, the committee anticipates that passive data collection is likely to increasingly replace active data collection (i.e., surveys) as the preferred method for populating the national freight database. For example, as ITS and EDI data become more widespread, traditional surveys are likely to become less important as sources of freight transportation data. The changing role of active and passive data collection approaches is illustrated schematically in Figure 3-2.[8]

The relative importance of active and passive data collection methods is likely to vary across modes of transportation as different modes embrace new technologies in different ways and on different schedules. For example, rail and marine carriers are already fairly advanced compared with other modes in implementing EDI. Thus, EDI may provide useful data for these modes in the relatively short term but may not capture a representative sample of freight movements by all modes of transportation for some years to come.

Figure 3-2 is also likely to look somewhat different for different data items. For example, because current truck detection systems can provide vehicle counts and information on vehicle type and speed, passive data collection could provide a high percentage of such data in the near term. However, these detection systems cannot determine the commodity being carried, so such information will need to be collected by active means, such as roadside surveys, pending the development of more sophisticated detector technology.

In the committee's view, a phased approach to exploring the potential benefits of EDI and other electronic data streams offers important advantages. Useful data may be obtained in the relatively short term, and the lessons learned can be applied in subsequent development—for example, as other modes adopt EDI or as sensors used in passive data collection devices become more sophisticated.

[8] Figure 3-2 is identical to Figure A-2 of Appendix A.

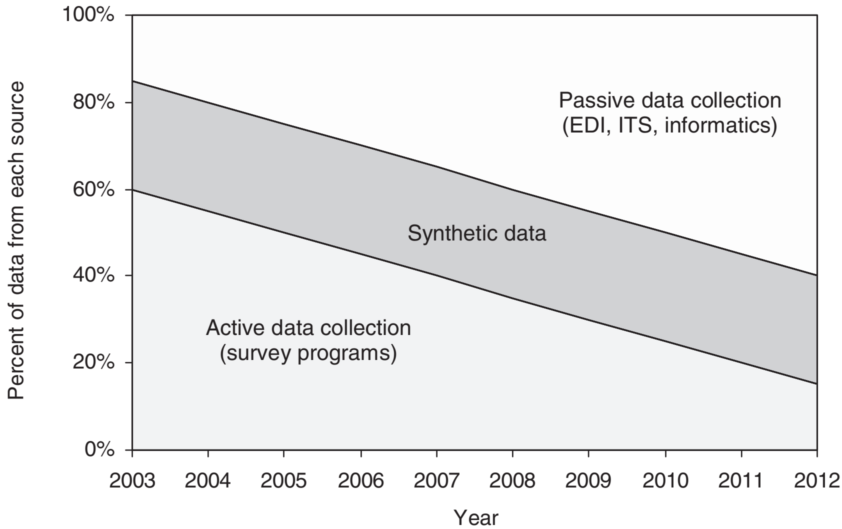

Figure 3-2 Schematic representation of the roles of various data sources over time. Source: A paper prepared for the committee by R. Donnelly (Appendix A).

Confidentiality

The need to safeguard the confidentiality of data providers—and in particular to protect commercially sensitive information—will be critical in determining how data in the national freight database are made available to the public. Federal agencies such as the Bureau of Customs and Border Protection and the Census Bureau already collect data that could be useful for transportation analysts, modelers, and planners. However, legal limitations protecting the confidentiality of individual establishments prevent the release of raw data on freight movements and shape the presentation of data released to the public. These limitations are essential to obtaining the cooperation of data providers, without whose input survey programs such as the CFS would not be possible.

To meet the needs of users, particularly at state and local levels, the national freight data framework is intended to provide opportunities for improving the geographic resolution and level of commodity detail of freight data. However, a lack of accompanying measures to safeguard the competitive position of individual firms and protect the confidentiality of data providers would be a fatal flaw in any efforts to implement the framework. Therefore, the committee considers it essential that a research effort to examine options for ensuring the necessary levels of confidentiality be initiated as one of the first steps toward implementation of the framework. Data providers will need to be involved in discussions of confidentiality requirements to ensure that proposed approaches address their concerns.

Research into confidentiality protection is likely to focus on a variety of models already in use or proposed. For example, limited data from the Rail Waybill Sample are available in a public use file, but more detailed information can be provided to certain parties upon approval by the Surface Transportation Board.[9] For example, data may be provided to a contractor who has been asked to prepare investment advice for private-sector clients, but the clients themselves do not have access to the raw data. A particularly interesting model from the perspective of the national freight database is the hierarchical access system being developed for the ITDS. The ITDS will distribute standard data to federal agencies, but each agency will receive only information relevant to its mission.[10] In the case of the national freight database, such varying levels of access to the data could eventually be provided through an interactive database system.

Other possible research areas include the application of disclosure limitation methods (see, for example, NRC 1993) and ways of protecting the confidentiality of items in electronic data streams being mined for passive data collection.[11]

Roles of Public and Private Sectors

Implementation of the proposed national freight data program will require the participation of a variety of public- and private-sector organizations at various levels. The former group will include federal agencies,

[9] Access to nonpublic use rail waybill data is automatically granted to state governments. Access is granted to other parties on a need-to-know basis.

[10] ITDS Background (www.itds.treas.gov/itdsovr.html).

[11] In February 2003, BTS inaugurated a series of seminars to discuss confidentiality and access issues. Further details are available on the agency's website (www.bts.gov/confidentiality_seminar_series/index.html).

state departments of transportation (DOTs), MPOs, and local jurisdictions. The latter will include consulting companies, representatives of different modes of transportation, shippers, receivers, third-party logistics companies, and academic researchers. Since much of the nation's freight transportation infrastructure is privately owned and almost all freight is carried by private firms, industry involvement will be critical to the success of the national freight data program.

In view of the diverse participation, broad scope, and complexity of the proposed program, the committee believes that federal government leadership will be needed to provide a key link among participants and to coordinate their activities. Federal government leadership does not imply that the federal government should bear the full financial burden of the national freight data program. For example, the Canadian NRS is coordinated at the federal level by Transport Canada but involves cost sharing with DOTs in the Canadian provinces and U.S. border states. The provincial DOTs are responsible for all data collection and related quality assurance. For both the 1995 and the 1999 NRS, a memorandum of understanding between Transport Canada and the provincial DOTs addressed areas such as study objectives, the federal formula for cost sharing, survey design, standardization of data collection processes, and data processing and dissemination. Similar approaches aimed at sharing costs and responsibilities among participants in the proposed national freight data program may merit investigation.

The freight data business plan presented in Appendix A suggests that BTS should be responsible for coordinating data collection activities to populate the national freight database; other organizations (private sector, state DOTs, MPOs, carriers, etc.) should be responsible for the collection and synthesis of supplemental data. The committee agrees in principle with this broad division of responsibilities between the federal government and other parties, although it recognizes that the detailed assignment of responsibilities will need to be worked out as part of the process of implementing a national freight data framework.

Experience with national surveys such as the CFS indicates the value of federal government involvement in providing transparent "core" data that are widely used by many different groups for a variety of purposes. For example, CFS data are used by the private sector to develop value-added data products, such as the Transearch database from Reebie

Associates and customized state and local databases constructed by consultants for specialized studies. The committee believes that the collection of specialized data is generally better left to organizations outside the federal government but considers it important for the national freight data framework to facilitate the linkage of such supplemental data to the national freight database.

Implementation of the national freight data framework will provide opportunities for various organizations to build on their strengths and experience in areas such as survey design, data collection, and data analysis. For example, a federal statistical agency such as BTS could play a key role in researching new sampling strategies for passive data collection systems and in developing standard survey methodologies to guide supplemental data collection efforts. Similarly, the Federal Highway Administration (FHWA) and a number of private-sector organizations have developed expertise in data fusion through their work on the FHWA Freight Analysis Framework policy analysis tool. This expertise could be valuable both in developing the national freight database and in combining national and supplemental data for specialized applications.

The proposed framework includes a feedback loop from data users to the freight data advisory committee. This feature is intended to encourage ongoing dialogue between the advisory committee and data users to inform development and refinement of the national freight data program. A variety of feedback mechanisms will likely be needed to accommodate the diverse data users. Options include a feedback button on a website that allows users to send comments by e-mail, as in the case of the Vehicle Inventory and Use Survey, and meetings to facilitate the interchange of ideas among users and those responsible for survey development and design. The committee envisages the private sector playing a particularly valuable role in providing feedback, given its importance to the transportation enterprise as a whole and the underlying need to ensure confidentiality within the national freight data program.

A number of private-sector organizations currently provide freight data to meet specific client needs. The committee envisages that activities of this type will continue within the broad context of the national freight data program, which will need to recognize the private-sector role in selling value-added data to meet client requirements. In some instances, these data may be less transparent than those in the national freight database; for example, different firms may fuse data sets differently using proprietary techniques for specific applications.

Implementation of the national freight data framework will require careful consideration of the roles of public- and private-sector parties in data fusion and synthesis to ensure that the desired levels of transparency are maintained and that users have access to unambiguous information on data reliability. Figure 3-1 indicates that it will be necessary to fuse data from the freight informatics initiative with survey data to populate the national freight database. This fusion process may be complicated by two factors. First, some of the freight informatics data—ITS data, for example—may not represent a random sample. Therefore, it will be necessary to develop statistical methods for fusing survey data collected from a random sample of respondents with nonrandom "informatics" data. Second, some of the data may not be collected by federal government agencies. When fusing such data with data from federal government surveys, it will be important to ensure that the sources and reliability of the resulting data are clearly reported. Under certain circumstances, private-sector groups could decide to follow the federal government lead in disclosing more information about the origin and reliability of their proprietary data. However, the extent of any such disclosure is difficult to anticipate. Thus, conferring the imprimatur of the federal government on fused data from a combination of federal and nonfederal sources could be potentially misleading to users whose expectations for federal government data are based on experience with the CFS and similar surveys. In the committee's view, the public- and private-sector roles in a variety of data fusion activities merit further discussion.

Figure 3-1 also suggests that data synthesis should be the responsibility of parties outside of the federal government. Such data synthesis is a very different exercise from the data imputation performed during the analysis of raw survey data.[12] In the committee's opinion, there may be benefits in different groups making different assumptions and using different methods to obtain results aimed at meeting diverse user needs. As

[12] A process of imputation may be used to fill gaps in survey data resulting from item nonresponse. For example, when a survey question is unanswered or the response provided fails computer or analyst edits, the missing or erroneous value in a respondent's survey questionnaire may be replaced by a predicted value obtained from an appropriate model.

noted earlier, the private sector has traditionally played an important role in developing data sets to meet the specialized needs of a range of users.

In the committee's view, detailed definition of the public- and private-sector roles in implementing the data framework would be premature at this conceptual stage of the national freight data program. For example, decisions about strategies for protecting the confidentiality of data providers will be important in determining who has access to what data. Levels of data access will, in turn, largely determine who is in a position to undertake data-processing activities requiring access to survey microdata. Some data providers would like to see data stripped of any identifiers before being given to a regulatory agency such as the U.S. Department of Transportation. If implemented, this approach would have important implications for the development and maintenance of the national freight database.

NEXT STEPS

In the committee's view, the conceptual plan for a national freight data program presented in Appendix A is a goal toward which BTS and others should aspire in seeking to respond to the recommendations of the 2001 Saratoga Springs conference. The committee recognizes that achieving the desired objective—a comprehensive picture of goods movement in North America—will take time and require considerable effort and resources. It is apparent from the preceding discussion that implementation of a national freight data framework will require careful analyses of various options to map out an appropriate strategy. Many technical and institutional issues will require investigation, and program planning, development, and management capabilities will be needed to help ensure the necessary continuity of effort. The committee's recommendations address the initial programmatic and technical steps required to move forward with the implementation of a national freight data program.

REFERENCES

Abbreviations

BTS	Bureau of Transportation Statistics
FHWA	Federal Highway Administration
NRC	National Research Council

BTS. 1998. *Transportation Statistics Beyond ISTEA: Critical Gaps and Strategic Responses.* BTS98-A-01. U.S. Department of Transportation, Washington, D.C.

Census Bureau. 2001. *Meeting 21st Century Demographic Data Needs—Implementing the American Community Survey. Report 1: Demonstrating Operational Feasibility.* www.census.gov/acs/www/Downloads/Report01.pdf.

FHWA. 2002. *The Freight Story: A National Perspective on Enhancing Freight Transportation.* FHWA-OP-03-004. U.S. Department of Transportation, Washington, D.C.

Margiotta, R. 1998. *ITS as a Data Resource: Preliminary Requirements for a User Service.* Office of Highway Policy Information, Federal Highway Administration. www.fhwa.dot.gov/ohim/its/itspage.htm.

Meyburg, A. H., and J. R. Mbwana (eds.). 2002. *Conference Synthesis: Data Needs in the Changing World of Logistics and Freight Transportation.* New York State Department of Transportation, Albany. www.dot.state.ny.us/ttss/conference/synthesis.pdf.

NRC. 1993. *Private Lives and Public Policies: Confidentiality and Accessibility of Government Statistics.* National Academy Press, Washington, D.C.

Southworth, F. 1999. *The National Intermodal Transportation Data Base: Personal and Goods Movement Components* (draft). Oak Ridge National Laboratory, Oak Ridge, Tenn.

Appendix A

A Freight Data Business Plan

Rick Donnelly, PBConsult, Inc.[1]

The goal of a national freight data program is to make the freight transportation system as efficient, safe, and secure as possible through informed planning and investment decisions. These goals cannot be reached without information about the magnitude and economic importance of freight flows at the national, regional, and urban levels. However, this information is not readily available. In this appendix, a national freight data program that will meet the needs of transportation planners and decision makers at the local, state, and federal levels is outlined.

PROBLEM

The safe and efficient movement of freight has become a significant transportation policy issue. The contribution of freight to a regional

[1] This appendix was prepared by Rick Donnelly with extensive input from the committee.

economy and the role it plays in the nation's economic competitiveness have become important topics at all levels of government. Productivity gains realized by industries over the last 15 years are in part a result of improvements in transportation infrastructure, deregulation, technology, and business process improvements such as supply chain logistics. Improving the efficiency of freight movements will be necessary to permit further increases in productivity.

Congestion in urban areas is now seen as an impediment to economic competitiveness. For this reason, analysts are increasingly called upon to document the economic contribution of freight and the costs and benefits associated with public–private partnerships and infrastructure investment. However, the factors, agents, and dynamics associated with freight movements are complex and often affected by labor, pricing, and regulatory mechanisms. For example, although most of the nation's freight transportation assets are privately held, freight often moves upon or through publicly provided infrastructure.

Most communities lack information on freight flows or an appreciation of the impact of freight movement on their local economy and transportation system. The regional nature of freight flows and the connections of these flows between metropolitan areas are often not understood. Data from the national level cannot inform local policy making, both because of the data's focus on the state level and lack of integration across various modal and industry sources. Thus, transportation planners currently lack the information necessary to inform decision making, with the result that the utilization of funds and resources is often inefficient and uncoordinated.

POTENTIAL CLIENT BASE

For the purposes of this business plan for a national freight data program, the primary clients are public-sector transportation planners and policy makers at the local, state, and federal levels. These clients include planners seeking to understand and forecast freight flows and their economic significance, as well as officials concerned with state and federal taxation, regulation, and monitoring. The data must also be useful to policy researchers in universities. Other clients include the private sector, which has many of the same information needs as government and would be well served by data programs that enable companies to identify underserved or emerging markets and clientele. The needs of this sector should be taken into account in the design of a freight data program. Information that helps private enterprise make more profitable investments would contribute to making the entire transportation system more efficient and competitive. In the longer term, such a national freight data program would foster better joint decision making and public–private financing.

ASSESSMENT OF THE COMPETITION

There are several sources of freight transportation data in the public and private arenas. Most of them focus on intercity and international movements by rail and marine modes of transport, although they vary widely in purpose, scope, extent, and level of detail. Public-sector sources include the Commodity Flow Survey (CFS), the Vehicle Inventory and Use Survey (VIUS), rail waybill statistics, foreign trade data, and traffic monitoring data. Private-sector sources include the Port Import Export Reporting Service (PIERS) from the Journal of Commerce and the Transearch data from Reebie Associates. These sources are adequate to meet the needs for which their data are collected, but users face difficulties in combining and comparing data from the different sources because of incompatible metrics and definitions.

Only two sources of data on domestic freight transportation—the public CFS and the private Transearch database—attempt to depict a picture of freight movements by all modes of transportation at the national and state levels. Both have the potential to illuminate trucking, intermodal, and container flows. Therefore, to be considered viable, alternatives to these sources should add significant value to what CFS and Transearch already provide.

The CFS is conducted as part of the Economic Census. It is jointly funded by the Census Bureau of the U.S. Department of Commerce and the Bureau of Transportation Statistics (BTS) of the U.S. Department of Transportation (USDOT). In its current form, the CFS was conducted in 1993 and 1997. The 2002 survey is in the field, and results are expected to become available in 2004. The CFS surveys shippers from selected mining, manufacturing, and wholesale firms. As such, it can capture only a part of total freight movements. Moreover, available evidence suggests that it significantly underreports international trade movements.

The high cost of the CFS, coupled with a sense of dissatisfaction among users with several of its aspects, has prompted BTS to examine alternatives. In November 2000, the bureau sponsored a Freight Data Round Table in Washington, D.C., to solicit input on the program's future. The CFS was widely viewed as very important to the participants, who represented a broad spectrum of users. They reported that the data were essential for a number of transportation planning, regulatory, and research activities at the local, state, and federal levels. However, dissatisfaction with several aspects of the CFS emerged. The timeliness of the data's availability, the geographic scale of reporting, and the imperfect coverage of the shipper industry were cited as key limitations of the data. Many of these same issues were raised a year later in Saratoga Springs, New York, at the conference "Data Needs in the Changing World of Logistics and Freight Transportation."[2] Participants in both meetings underscored the importance of consistent and comprehensive freight data. These information needs and the growing dissatisfaction with the content and coverage of CFS data have prompted BTS to reconsider how freight data are collected, summarized, and disseminated to its customers.

The Transearch database is maintained and distributed by Reebie Associates. Data are available for purchase at many different levels of modal, geographic, and commodity detail. Historically, these data have been available only at the state level, probably owing to their basis in the CFS and other public sources. Data at the county level have become available more recently. The limitation of these data has more to do with what is not known about them. The vendor fuses public data sources and augments them with information it collects from carriers and shippers. However, little more is known about how the database is constructed. The Transearch data appear to offer little insight into container traffic and only limited information on intermodal movements. Despite these limitations, the product is widely used by transportation planners at all levels

[2] Meyburg, A. H., and J. R. Mbwana (eds.). 2002. *Conference Synthesis: Data Needs in the Changing World of Logistics and Freight Transportation.* New York State Department of Transportation, Albany. www.dot.state.ny.us/ttss/conference/synthesis/pdf.

of government. Reebie Associates is the market leader in providing such data to the public sector and presumably to the private sector as well. The Transearch data are the yardstick by which any alternative initiative should be measured.

THE INFORMATION GAP

At present, holistic data on freight flows at the local level are virtually nonexistent. State and metropolitan agencies rarely collect data on freight movements, and fewer still use forecasting or operational models to assess their area's needs. Thus, there is a major unfilled need for data on freight movements by all modes of transportation in greater geographic detail than is provided by most national data sets. Such data would illuminate flows in urban areas as well as between them.

A further problem is the lack of comprehensive freight data covering all modes. In particular, the largest and fastest-growing segments of the transportation market by almost any measure—motor carrier flows, truck–rail combinations, and containerization—are not well covered by any existing efforts. Furthermore, all competing data sources are beset with incompatibilities, ranging from how commodities are classified to modal definitions to how the data are reported. The lack of integration of data frustrates users who are trying to combine data to summarize total freight demand by all modes. An obvious solution would be to work with the vendors to standardize definitions. However, preliminary contact with many vendors revealed that most seem content with the current structure of their products and are, at best, indifferent to the idea of harmonized modal, commodity, and flow metrics. A product capable of incorporating information from these third-party sources would be a significant advance that none of the products alone could match.

All available freight data are affected by a fundamental conflict between confidentiality and utility. Almost all users of the data are frustrated by the data's coarse spatial resolution, which makes the data incapable of usefully informing analyses at the scale of most appraisals. However, the data's level of abstraction is not arbitrary. Most data are provided by private firms with a legitimate business need to protect the confidentiality of their customers. Access is closed to the raw data that would otherwise provide public-sector planners with data at the scale they seek. A product capable of resolving this conflict would have tremendous utility.

As a result of the deficiencies outlined above, decision makers in the public and private sectors lack important and relevant information on freight flows. Meeting the information needs of the entire spectrum of users is well beyond the capability of any single data provider. However, there are basic elements of information that are useful to almost all users and investors. These include the origin, destination, intermediate stops, weight, frequency, commodity classification, modes of transport and interconnectivity, and value of the shipment. Additional data that reveal differences between long- and short-haul shipments are important, as are data about routing. These data will usefully inform public-sector planning and policy making, while providing limited and potentially less useful information to the private sector.

PRODUCT DEFINITION

This plan for a national freight data program builds upon innovative ideas from both the BTS-sponsored Freight Data Round Table and the Saratoga Springs conference (see above). It outlines the program's essential elements and adds several components that have not previously been considered. The magnitude of the freight data problem hinders attempts to create a single data source or program capable of addressing all needs. Indeed, there is a readily apparent need to create an architecture[3] or high-level context within which to place a candidate freight data program. A high-level view of the proposed architecture for a national freight data program is shown in Figure A-1. An advisory committee will oversee the detailed design of a multifaceted survey program. The survey programs would be supplemented with data from other sources and data derived by synthetic techniques. The resulting databases would be made widely available to the user community, whose comments and feedback would help inform the evolution of all parts of the architecture.

The goal of such a freight data program will be to provide a timely, comprehensive, and consistent database of multimodal freight flows in

[3] The word "framework," which was used throughout the committee's report, has the same meaning as "architecture" as used in this appendix.

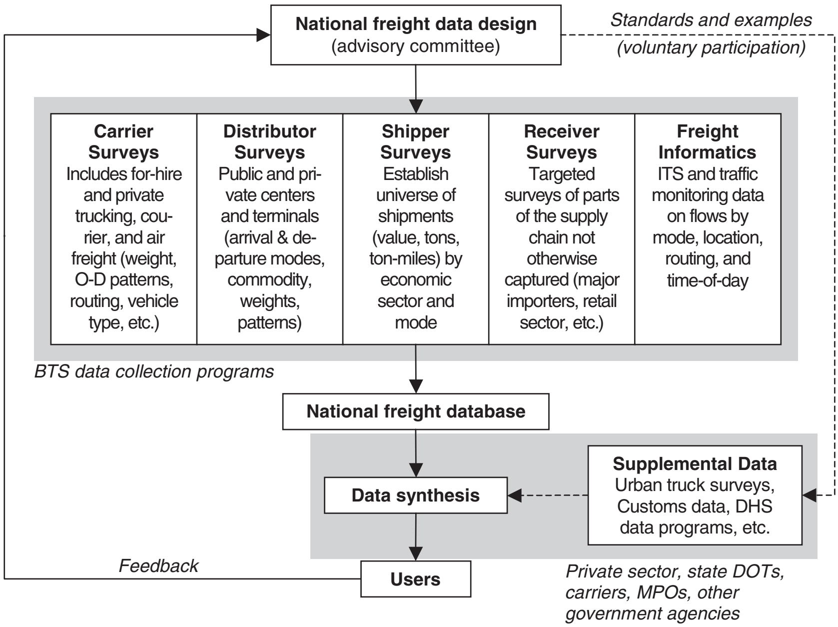

Figure A-1 Proposed architecture of a national freight data program. (BTS = Bureau of Transportation Statistics; O-D = origin–destination; DHS = Department of Homeland Security; MPO = metropolitan planning organization; state DOT = state department of transportation.)

North America. The data will include domestic as well as international flows. It must be capable of providing data at varying levels of geography. The program will be developed over time; the program design and quick implementation of carrier surveys will be undertaken first. The design of more complex components, such as establishment surveys, will be carried out in later years.

Several components of the system will be necessary for meeting these goals.

National Freight Data Architecture

An ideal data program will encompass several surveys and establish linkages to other databases. Careful design of an overall architecture that will facilitate the access to and fusion, reporting, and maintenance of these data will be essential. The architecture will also specify future directions for the program. It is anticipated that the eventual program goal will be to depict the movement of goods through the entire supply chain, although the initial work may only capture many of the same segments of freight flow as do current surveys. An important part of the data architecture will be the description of supply chain linkages and the data required to understand them.

An Integrated Program of Freight Surveys

This program will replace the CFS and the survey's singular emphasis on shipments from selected industries. It will attempt to capture a wider range of freight movements by expanding its sampling frame to include receivers as well as shippers and samples from all sectors of the economy. Moreover, it is recommended that this program be conducted continually. Despite the intuitive appeal of panel surveys, it is suggested that the freight surveys be structured as a rolling series of cross-sectional surveys rather than a time series of selected establishments. The resulting data will be summarized at the state and metropolitan area levels, as is currently done for the CFS. Other survey programs, such as the American Community Survey, have successfully transitioned to rolling sampling, and this experience could be useful in implementation of continuous surveying by a national freight data program.

An important departure from the current CFS will be the availability of microdata for planners and researchers with sensible but sufficient safeguards to protect the privacy of the parties involved in the surveyed transactions.

It is envisioned that the program will initially be capable of collecting a sample as large as that of the current CFS. This will probably be adequate for reporting the universe of freight movements in North America, but only at the state or regional level. The sampling frame necessary to capture comparable data at the county or subcounty level, even for a few states, would far exceed the resources likely to be made available for such a program. This major limitation will be overcome in two important ways, which are discussed next.

Freight Data Synthesis

It is impractical to design a survey program capable of meeting the needs of all transportation planners across the country. A synthetic process of generating these data—especially in the short term before all of the different but complementary surveys become fully operational—offers a tractable compromise. A variety of simulation techniques can be employed that provide flexibility to accommodate data needs on an individual case basis, permitting the use of all available local data as well as relevant third-party sources. A port city, for example, might place particular emphasis on intermodal accessibility and the clustering of transportation firms near it, as the cities of Portland and Los Angeles have recently done.

Data synthesis would provide the same type of information available from the different surveys but at the local or subregional level. Constraints can be imposed to match urban cordon counts, Highway Performance Monitoring System data, or just about any other source of information. A likely scenario would involve the use of economic input–output tables to help allocate flows by commodities to the appropriate producing and consuming industries in an urban area. Other information sources, such as the VIUS, might provide marginal distributions to help condition each simulation.

Synthetic data will also be capable of illuminating gaps in the coverage and content of the different freight surveys. This will allow adjustments in the survey instrument, methodology, and sampling frame. The ongoing nature of the national freight data program will allow these changes to be made quickly. For example, the data collection process in urban areas might be augmented with data on the local economy. Changes in the local economy would trigger adjustments in both survey methodology and synthesis.

Freight Informatics Initiative

Electronic data interchange and automated data collection methods have the potential to greatly reduce the cost and respondent burden associated with data collection. However, currently operational systems are tailored to individual firms and lack linkages to the type of information typically used by planners, such as origin, destination, mode and intermodal connections, and commodity classification. Moreover, these systems are currently deployed more often by carriers, who may have only incomplete information on commodity classification or no knowledge of the shipment's place in the overall supply chain.

The freight informatics initiative would examine the institutional and technical issues surrounding electronic data collection and its potential for use in freight data collection. It is envisioned that a prototype system will be designed and placed into limited field testing. Linkages with intelligent transportation system (ITS) efforts at USDOT and other agencies will permit the program to benefit from the considerable progress that has already been made in commercial vehicle operations and electronic data interchange initiatives.

Standard Survey Methodologies

The freight data programs suggested here and other federal initiatives will improve the quality of data and information available. However, they are unlikely to meet all the data needs of local users, particularly those states and urban areas that want to construct freight travel models or conduct detailed corridor or facility studies. Many of these agencies have little or no experience in designing and conducting freight surveys.

A standard survey practice would enable states and urban areas to expand the freight survey samples in their jurisdiction so they can gain additional local information to complement the national survey. A similar approach has been successfully adopted for the National Household Travel Survey. This task will involve the development of standard prototype survey instruments and methods for urban areas. The consistent data that can be harvested as a result will be useful at both the local and the national levels and will facilitate improvement of both the national freight database and data synthesis procedures.

Summary

The interactions among these parts of the national freight data program are shown in Figure A-1. The national freight data design defines the

Table A-1 Timing of Business Plan Implementation

Immediate Actions (this fiscal year)	Short-Term Actions (1–4 years)	Long-Term Actions (5 years and beyond)
Development of the national freight data design Begin mining and publishing existing ITS data	Refinement of the national freight data design Design and begin carrier survey program Design and begin distributor survey program Design and pilot test freight informatics Begin data synthesis	Continued refinement of the national data design Continued refinement and conduct of carrier and establishment surveys Full implementation of freight informatics initiative Design and implement shipper and receiver surveys Refine and continue data synthesis

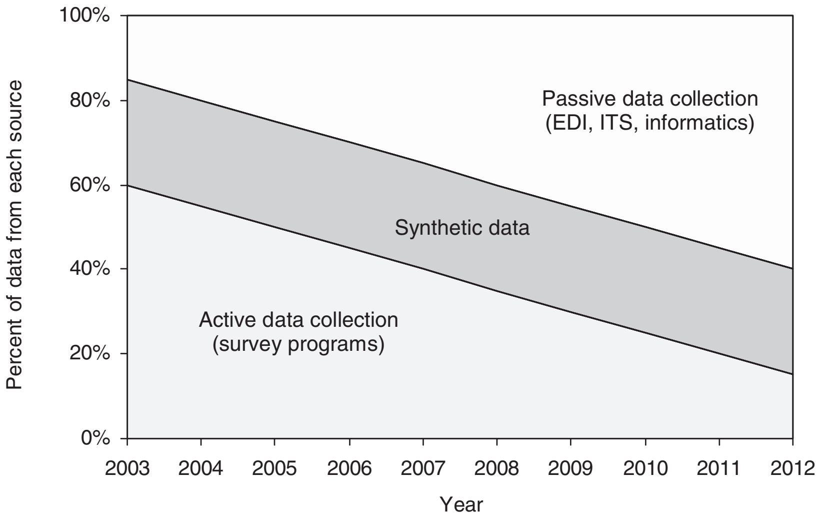

Figure A-2 Schematic representation of the roles of various data sources over time.

various components of the survey program, some of which may be developed sooner than others. A notional view of the timing of each part of the architecture is shown in Table A-1. The resulting database will evolve over time. For reasons of constraints on cost and knowledge, there will always be a role for data synthesis in the process. This will be especially true in the program's early years, before enough survey data can be acquired to portray a statistically valid picture of freight movements, especially at the metropolitan level or between urban areas. This process is illustrated in Figure A-2. As electronic data collection and the mining of real-time transaction data become more commonplace, the importance of traditional surveys and data synthesis can be expected to diminish. The former will still be the most efficient means of gathering data in some instances, while the latter will always have a place in generating data that cannot be measured in society.

BTS can play an important role in a national freight data program. With its mission and resources, it has the capability of defining an agenda for freight data collection and dissemination and the means of implementing it.

BENEFITS

A national freight data program will provide better-quality and more complete information than is currently available through any existing data source. It will address needs that cannot be met through any combination of currently available freight transportation data, whether provided publicly or privately. The data synthesis element of the program will focus on providing the timely data needed to inform planning and investment decisions at urban and local levels. When analytical requirements dictate higher levels of accuracy and precision, these synthetic data will offer insights that can then be supplemented with locally collected data.

Appendix B

Review of Freight Survey Collection Techniques

Rick Donnelly, PBConsult, Inc.

The movement of freight is the economy in motion. Firms and households trade with and compete against one another in the marketplace. The marketplace in which many firms operate today is truly global. Almost every person and business in the world obtains goods and services that are provided through a supply chain whose inputs come from around the world. Firms use information technology to make real-time adjustments to production and inventories, the latter of which may be stored in "rolling warehouses" of containers, trucks, and railcars.

Transportation planners in both the public and the private sectors need to understand the dynamics of freight if they are to wisely allocate resources designed to improve the efficiency and safety of freight movements. These movements are considerably more complex than person travel in urban areas and change far more quickly than do commuting patterns. The picture is further complicated by the recent emergence of supply chains, which portray the often lengthy process of production in a spatial context. A few chains are very short, as in the case of fresh fish taken from the dock to a nearby restaurant for consumption that evening. The production of automobiles and aircraft is at the opposite end of the spectrum. They are assembled from components produced in many different places, which are in turn made up of smaller components from other suppliers. The production of a single automobile is contingent on parts supplied by hundreds of vendors, who move their products to the final assembly point.

The importance of understanding supply chains can be illustrated further by considering foods on the shelves of supermarkets. They are almost always shipped from the initial producer to a regional distribution center. There they are assembled for delivery to the local supermarket. The movement from initial producer to distributor is often invisible to decision makers in local communities, who encounter trucks making the rounds from the few distribution centers to many local supermarkets. Upset about congestion, they wonder why more shipments cannot be pushed back onto rail. The movement from initial producer to distributor might be a candidate for intermodal shipment. But the comparatively short end of the chain from distributor to retailer, which might amount to 20 miles—the contributor to congestion in the urban area—most certainly is not.

It is only with a complete picture of how freight moves through the transportation system that policies and investments can be formulated that will enhance the economic competitiveness and reduce the costs associated with freight transportation. The need for information on how the freight transportation system operates is readily apparent. Most transportation agencies, especially those in the public sector, play no role in the production or consumption of the freight. Thus, what they know about it must come from direct observation, usually in the form of user and carrier surveys.

The sheer complexity and number of agents involved in even a smaller supply chain make a single survey process that would illuminate all parts of it impossible to design. Different parts of the chain operate at different spatial and temporal scales and involve different business processes. A multitude of agents influence freight movements, including the shipper, receiver, customer, carrier, regulator, and distributor. Each makes travel choices or influences those made by others at one or more points in the supply chain. Moreover, many of these agents have limited or incomplete knowledge of the factors influencing the choices made by the other agents. Thus, in order to understand the system as a whole, one must examine and synthesize information from each major step in the supply chain. The challenge in collecting such information is to acquire it as efficiently as possible, without disrupting the target of interest.

Freight transportation surveys are typically carried out at the places where goods are produced or distributed, since they tend to be relatively small in number and easy to locate. At least three classes of survey are used to collect data at these points of concentration. Establishment surveys typically target the producers and consumers of goods. Historically, there has been a much greater emphasis on understanding producers than consumers, since the former are considered the linchpin of the economy, while the latter are often considered as a large and widely spread marketplace. Distributor surveys can capture information on goods movement through intermediaries, a fast-growing segment of the transportation industry. Both can be supplemented with information from carrier surveys. These surveys are focused on the physical agents of distribution, the business patterns of which are felt in terms of flows on the transportation network. The key benefits of each of these approaches are discussed in the following sections.

ESTABLISHMENT SURVEYS

Freight flows reveal the patterns of trade between firms and households in the marketplace. Collecting data from one or both ends of this transaction makes sense for a number of reasons. One of the most important is that it allows generalizations from the survey to be applied to the economy or local marketplace as a whole. If something is known about the shipment of widgets from surveys of a few factories that produce them, the total traffic of widgets can be inferred by simply counting the number of widget factories and ascribing the observed behavior to all of them. Many commodities are produced by only a small number of firms but consumed by a larger population. Collecting data at a few locations (the producer) is far more efficient and less costly than attempting to capture the same data at the much more diffuse level of the consumer.

Most establishments both produce and consume goods. Thus, they are both *shippers* and *receivers.* If the economy is assumed to be a closed system, a survey at either location can be assumed to illuminate the same patterns of movement between them. Aside from the obvious differences in scale at each end, there are some subtle differences that distinguish the two types of establishments.

Shipper surveys are often favored because shippers tend to be fewer in number, easier to classify and characterize, and easy to identify and locate. In addition, they are often considered the true drivers of the economy. Shippers usually either own the means of transportation to the customer or have a distribution manager who is familiar with the details of how the goods are shipped. The goods they produce are most often identified with the sector of the economy they belong to. Food companies produce food, for example, and paper is produced by paper companies. Thus, linking the output of firms segmented either by commodity produced or their industrial classification is typically a simple exercise. Finally, the production cycles of firms often result in fairly regular intervals of shipments.

Shipper surveys have typically been oriented toward firms in the mining and manufacturing sectors of the economy, under the assumption that such firms produced the majority of goods transported by freight carriers. This was probably a valid assumption 25 years ago, before industrial consolidation became common. Many industries were vertically integrated and produced most of the components of their products. The automobile industry is an example of such a business model. Until recently the industry not only produced the final product but also all of the intermediate products within it, such as a upholstery, glass fixtures, stereos, tubing and hoses, batteries, and the like.

The shipment from the shipper to the receiver was thought to portray most of the time and distance traveled between production and consumption of most goods. Even when a manufactured product was sold through intermediaries, such as wholesalers and retailers, they typically were located close to the ultimate consumer of the product. This may still hold true for certain types of commodities, but it is no longer universally true.

The advent of supply chain logistics has all but eliminated the ability to understand freight movements through shipper surveys. A recent study of trucks crossing the U.S.–Canada border revealed that half were bound to or coming from a warehouse, distribution center, or transportation terminal. Indeed, almost a quarter of the trips were between such transshipment points on both ends of their journey. The shipper is rarely visible in these transactions. Moreover, a shipper survey can capture information about exports to other countries, but these transactions are often handled by a domestic intermediary who is characterized as the receiver of the goods. Finally, in a supply chain environment many shippers produce only parts of a larger product, one that may undergo several incremental stages of assembly before it is ready for delivery to the final consumer.

Many planners wishing to better understand freight have advocated surveying receivers in addition to or instead of shippers. While the universe of consumers is far larger than the universe of producers, it has been argued that their purchasing decisions directly determine the level of goods being moved. Moreover, it is thought that the presence of thousands of receivers in urban areas, each consuming goods produced by several sectors and in relatively small increments, gives rise to the large number of urban truck movements. Such urban truck movements are seen as both major contributors to and victims of congestion. Many analysts have pointed out that almost all of the inefficiencies of the transportation network are encountered in congested urban areas, which receivers are more sensitive to than shippers.

Receiver surveys can reveal the mix of goods consumed by firms and households, including goods purchased from abroad. Given the large deficit of trade in America, where imports amount to three to four times exports of merchandise, this constitutes a significant portion of goods movement that cannot be captured in shipper surveys. The trend toward globalization and the transition of the U.S. economy from industrial production to a service economy underscore the importance of foreign imports in the domestic supply chain.

The sheer number of consumers in the economy is a significant limitation to the use of receiver surveys. A single truck leaving a factory may deliver goods to a handful of intermediate destinations, each delivery in turn being divided into many individual shipments to the final consumer. Thus, hundreds of surveys of receivers would need to be carried out to capture the information about the output of a single factory. Many final consumers are households. Individual households may only purchase durable manufacturing goods on an infrequent basis, but collectively in an urban area they may consume the same products over regular intervals.

The large sampling frame required for receiver surveys has kept them from being widely adopted. However, such surveys are the most effective means of collecting information about certain types of commodities, such as consumer electronics and textiles. Many of these goods are produced abroad, and their largest concentration in the domestic supply chain occurs in distribution centers close to the point of final consumption.

It should finally be noted that the distinction between shipper and receiver is perhaps unhelpful and leads to disjointed data collection efforts. As previously noted, almost all sectors of the economy are both producers and consumers of goods. A survey program that would focus on both patterns simultaneously would offer more value and insight than each completed separately.

DISTRIBUTOR SURVEYS

The importance of warehousing and distribution centers has increased dramatically in the last two decades. When goods move through one of these intermediate points the visible linkage between the original shipper and the ultimate receiver is broken. The shipper most often no longer knows who the end purchaser will be, nor does the final consumer deal with the original shipper. The intermediary, who then controls the distribution of the goods from shipper to receiver, is the only agent in the supply chain who understands these linkages. Among such intermediaries are transportation terminals (including marine ports, intermodal centers, and truck terminals), warehouses, and distribution centers. Customs brokers and some freight forwarders also function as distributors, although they do not take physical control of the shipments.

It is apparent that surveys of such distributors would be superior to either shipper or receiver surveys. Such surveys would capture two legs of the supply chain, as well as the characteristics of the transition between them. Like shippers, distributors tend to be relatively concentrated in number and location, which reduces the size of the required sampling frame. A shipment moving through a distribution center also implies the possibility of a change in mode of transportation.

Despite their apparent advantages, such surveys are only infrequently carried out and are not known to form the basis of any public-sector freight data collection program. The private ownership of some distribution hubs may make them harder to detect, and securing the cooperation of the owners may be more difficult. Other centers may have high flow rates concentrated during relatively few hours, making data capture impractical or disruptive. Finally, many distributors have a legitimate business requirement to protect the confidentiality of their clients or may lack permission to share such information with third parties. These limitations must be overcome through a combination of nonintrusive survey methods and careful recruiting, because distribution centers are rapidly becoming the most efficient place to gather information about freight movements and their metamorphosis during the movement from initial production to final consumption.

CARRIER SURVEYS

Capturing data about shipments while in transit is perhaps the easiest and least expensive way to gather information about them. This is especially true for surveys completed in intercity corridors, where carriers can be stopped without immediately generating bottlenecks and congestion. These surveys are most often used to gather information from truckers, about whom little information can be gleaned from other surveys. Carrier surveys allow not only the collection of origin–destination and commodity information, but also detailed statistics concerning characteristics of the vehicle and the driver. Large numbers of surveys can be collected in a short period of time and can easily be expanded to represent the population of trucks passing the survey point. As such, carrier surveys are more suitable for local studies or characterizing areas that can be enclosed by a survey cordon.

Most carrier surveys are able to gather information about both the weight of the shipment and the vehicle, which are key data required in calculating the design requirements of roadway infrastructure. This information is almost never available from other surveys, where the agents typically focus more on the metrics of shipment size and value. Information about routing, travel times and temporal patterns, and carrier responses to delays and congestion is only available through carrier surveys. Moreover, by randomly sampling from the traffic stream, the full universe of shipments can be illuminated, perhaps revealing important segments of the market that are not covered by traditional surveys or that are growing faster than others.

Carrier surveys also suffer from important limitations. Of all the links in the supply chain, carriers usually possess the least information about the commodities carried or the true origin and destination of the shipment. Carriers often lack any information at all about the contents of containers. Even when they do know the origin and destination of the goods, they often have a difficult time characterizing their industrial classification. This frustrates efforts to understand how the shipments fit into the supply chain or opportunities for alternative means of shipment. The mix of carriers and commodities also varies considerably by region of the country, making it difficult if not impossible to generate a regional or national picture of freight movements.

COMMENTARY

No single survey type is capable of capturing information about the entire supply chain. This is even more so today than in the past, owing to the rapid increase in globalization, industrial consolidation, and the use of intermediate distributors. All of the survey methods discussed are capable of illuminating only a single link in the supply chain, with the exception of distributor surveys, which can capture two. From an efficiency and cost standpoint, distributor surveys would be superior in all cases except for single shipments from original producer to final consumer. While common in the agriculture and fossil-fuel segments of the economy, such simple shipments appear to be rare in the larger realm.

The value of each of these surveys becomes readily apparent when they are used together to capture information about the entire supply chain. A carrier survey, for example, can be used to estimate the universe of commodities flowing on a multimodal freight network. This in turn would reveal which commodities are typically moved through distribution centers. Surveys conducted there would greatly increase knowledge about such movements, as well as about which commodities are likely to benefit from such consolidation. Some commodities will be better represented in shipper or receiver surveys than in distributor surveys. The choice between survey types will depend on the structure and size of the industry. Receiver surveys are likely to be a far more efficient way to gather information about food shipments than any other survey type, for example. In contrast, the distribution of many durable manufacturing products will be better understood through traditional shipper surveys.

Using each survey program to focus on the commodities that are concentrated in its domain during shipment is the only way to paint a lucid and robust picture of freight flows. Using a combination of survey programs will complicate the overall freight data collection and summarization process. Yet it offers the only means of fully understanding the freight transportation system and the factors influencing choices made by the users and managers of the system.

Appendix C

Committee Meetings and Other Activities

FIRST COMMITTEE MEETING, JUNE 4–5, 2002, WASHINGTON, D.C.

The following presentations were made to the committee by invited speakers:

Sponsor Expectations for the Study
Ashish Sen, Bureau of Transportation Statistics

Overview of Freight Transportation Data Sources
Lance Grenzeback, Cambridge Systematics, Inc., and Harry Cohen, Consultant, Ellicott City, Maryland

The Commodity Flow Survey
Michael Cohen, Bureau of Transportation Statistics

The Freight Analysis Framework
Gary Maring and Bruce Lambert, Federal Highway Administration

SECOND COMMITTEE MEETING, SEPTEMBER 19–20, 2002, WASHINGTON, D.C.

The following presentations were made to the committee by invited speakers:

Freight Data Business Plan—Draft Version
Rick Donnelly, PBConsult, Inc. (committee consultant)

Air Cargo Data—Needs and Uses
Mort Plumb, Ted Stevens Anchorage International Airport, Alaska

Transportation of Hazardous Materials
Kevin Coburn, Research and Special Programs Administration

Hazardous Materials Transportation—Security-Related Activities Since 9/11
Tom Sherman, Office of the Secretary of Transportation, U.S. Department of Transportation

Use of Technology to Collect Truck Data
Jeff Patten, Federal Highway Administration

The International Trade Data System
Craig Fundling, *eCP*

MPO Freight Data Requirements
Gerald Rawling, Chicago Area Transportation Study
Dennis Hooker, Metroplan Orlando, Florida

The committee also received a written submission on MPO Freight Data Requirements from Chuck Purvis of the Metropolitan Transportation Commission, Oakland, California.

A Shipping Industry Perspective on Freight Data Needs and Uses
Bill Ralph, Port Import Export Reporting Service

In addition, the committee discussed security data currently gathered by the Federal Aviation Administration and the Transportation Security Administration with Admiral Paul Busick of Battelle.

THIRD COMMITTEE MEETING, NOVEMBER 21–22, 2002, WASHINGTON, D.C.

PRESENTATION OF FREIGHT DATA BUSINESS PLAN

A draft of the freight data business plan (Appendix A) prepared by the committee's consultant, Rick Donnelly, was presented at the 82nd TRB Annual Meeting in Washington, D.C., on Wednesday, January 15, 2003. The details of the session are as follows.

Session 713: National Freight Transportation Data Program Development
Arnim Hans Meyburg, Cornell University, presiding
Sponsored by Committee on Freight Transportation Data

Importance of Freight for Transportation
Emil Frankel, U.S. Department of Transportation

National Freight Data Business Plan
Rick Donnelly, PBConsult, Inc.

Trucking Industry Perspective
Bob Costello, American Trucking Associations, Inc.

Railroad Industry Perspective
Robert E. Martínez, Norfolk Southern Corporation

Forecaster's Perspective
Paul Bingham, Global Insight, Inc.

Appendix D

A Framework for the Development of National Freight Data

Dissenting Statement of Kenneth D. Boyer

This committee was charged with recommending "a framework for the development of national freight data . . . The framework was to be conceptual in nature and *not* a detailed data collection plan. Instead, it would articulate the types of freight data needed by the variety of users in transportation and the roles of different data providers." The intellectually defensible way to come up with such a recommendation is to include a balanced discussion of the benefits of different ways of collecting freight data, along with a discussion of the limitations, constraints, problems, costs, and characteristics of each way of collecting these data. Once the analysis of the problem is laid out in this way, the recommendation should follow from the analysis. The recommended framework should offer the greatest benefits within the constraints identified by the analysis.

The majority of the committee in their draft of Chapter 3 did not do this, but rather focused on the benefits of improved data. The committee neglected to give equal weight to the discussion of the practical limitations of different data collection methods. By contrast, this appendix, representing a minority dissent from the discussion in Chapter 3, lays out an analysis of the practical realities of freight data collection, focusing on three issues: (*a*) the size of the database called for in Chapters 1 and 2, (*b*) confidentiality issues, and (*c*) the need for judgmental data fusion to create the database. This appendix then uses this analysis of the practical constraints implicit in freight data collection to propose a framework for freight data development. This framework is offered not as the only way to achieve the goals of Chapters 1 and 2; it may be discovered that there are other, better ways. It does, however, offer an intellectually coherent recommendation for a framework for freight data development that is missing in Chapter 3. Using the framework, this appendix shows that Chapter 3 errs in several key areas, among them the following:

- Failure to recommend a procedure for dealing with confidentiality issues,
- Confusion on how data series like Waterborne Commerce of the United States and the 1 percent Railroad Waybill Sample should be owned and managed in relation to the new proposed data collection efforts,
- Failure to clearly define the role of third-party data organizers, and
- An apparent recommendation to shift resources away from the current Commodity Flow Survey (CFS) in favor of surveying other participants in the supply chain.

THREE PROBLEMS INHERENT IN CODMRT DATA

Chapters 1 and 2 of this report give a rationale for pushing freight data collection in the direction of extremely fine descriptions of freight traffic flows. The fundamental reason for moving in this direction is to support infrastructure investments that either mitigate congestion or promote regional economic development. Currently data are not available at such a fine disaggregation, but accurate measurement of the benefits of infrastructure investment requires data describing CODMRT—that is, commodity, origin, destination, mode, route, and time of day. This report recommends that a national freight data collection program have as its goal the provision of these ideal data.

The difficulty of this program should not be underestimated. For at least three reasons, there is no example in the world where such a national database has been created. First, the sheer magnitude of the data is almost beyond comprehension. For example, if we assume a modest 1,000 commodities, 1,000 origins, 1,000 destinations, 1,000 routes, 5 modes, and 4 times of day, a database that described each of the elements would have 20 trillion entries, and even this level of detail is too coarse to support a decision on whether to replace a bridge on a particular highway over a particular river. A national program of data collection to support all possible infrastructure investments would be populated by quadrillions of data points—clearly far more than can be published in any form. Realistically, a CODMRT data collection program will consist of a combination of investigations of data for specific projects along with a publication of summaries at a much coarser level of aggregation.

Almost all of the entries in a CODMRT database will be zero, but such is the nature of transportation data—one should not expect to find coal shipped to Newcastle or wheat shipped from Manhattan by any mode or route, much less by rail to Fargo, North Dakota, via I-20 through Shreveport at 3 in the afternoon. But the thinness of the data creates a second problem—that of confidentiality. Even before routes and time of day are added, and even at origins and destinations defined as states, for commodities that are defined more finely than broad aggregates, most of the C-O-D (commodity-origin-destination) entries in the CFS are suppressed, since fewer than four shippers are represented by the data. Once we add mode, route, and time of day to the data description, we can confidently predict that confidentiality requirements will prevent publication of the result in the large majority of cases where CODMRT data are nonzero.

The third and most difficult problem with CODMRT data is that, as a general rule, there is no agent who could be surveyed who knows what is moving between an origin and a destination by a mode and route at any time of day. A shipper fills a container and passes it to a carrier with a contract to deliver it to a particular destination. The receiver can verify that what was received is what was contracted for but may not know the origin of the shipment. The carrier likely will not know what has been hauled beyond a very general description. Neither the shipper nor the receiver will know the route, and in the case of motor carriage, the route and time of day may be known only to the driver, who may be ignorant not only of the contents of the truck, but also of the origin and ultimate destination.

In the future, national security concerns may require that all freight be shipped on a freight bill specifying everything in the CODMRT sextuple, and it is worthwhile monitoring developments in this area to see whether database development motivated by national security concerns can be tapped as a source of transportation data. Until such time as all freight movement requires freight bills, a framework different from simple sampling will have to be used to get the desired data.

These three issues—the vastness of the implied database, the thinness of the individual flows (leading to confidentiality concerns), and the fact that in the general case there is no individual who could fill out a survey to populate single entries in the database—dictate the proposed framework for freight data collection. In fact, it is perhaps misleading to characterize the gathering of CODMRT data as data collection, since such data will require joint inference from records contained in more than one data set.

DATA FUSION

The name "data fusion" will be given to the inference of flow characteristics from data contained in more than one data set, no one of which contains all of the information of interest. For example, one data set may contain records of 100 shipments of coiled sheet steel from Gary, Indiana, to St. Louis. After applying expansion factors derived from the sampling rate of steel firms in Gary, we might infer from these 20 records that 60,000 tons of coiled sheet steel is shipped from Gary to St. Louis annually. We might also have 10 records of machinery from Gary to St. Louis and, by applying expansion factors, we might infer that these represent 60,000 tons annually. From a separate database, we might get the information that among the shortest routes between the two cities, I-55 through Springfield has twice as much traffic as the I-57–I-70 route through Effingham, Illinois. This does not, however, mean that 40,000 tons of coiled sheet were shipped from

Gary to St. Louis via Springfield and 20,000 tons of coiled sheet were shipped via Effingham, that the proportion of machinery was the same on the two routes, or that the diurnal flow of steel or machinery between Gary and St. Louis will match that of traffic generally.

More accurate inferences will be made if one uses the original records or microdata. For example, if one has records of the individual enclosed vans and steel-hauling trucks between the two cities, it may be possible to assume that machinery will match the time and route pattern of steel coil shipped between the two cities, while machinery will have route and time patterns of enclosed vans. It is important to recognize that data fusion is *not* the same as record matching—it is extremely unlikely that a shipment of steel or machinery from Gary to St. Louis will be sampled as a shipment and again sampled as a movement on the truck.

Data fusion is not simply a matter of getting a consistent definition of commodities, origins, destinations, modes, routes, and times of day so that the "data silos" can be merged. Data fusion instead involves assumptions and judgment about matching records of steel shipments against records of truck movements, perhaps none of which will represent shipments of steel.

As noted by Southworth (1999), data fusion is not a mechanical process. All data fusion makes assumptions about how, for example, flows should be assigned to routes or how total shipments from an origin should be assigned to destinations. The accuracy of data fusion is then dependent on the assumptions made by the data fuser. Two different data fusers could make two different estimates of the CODMRT and both would be credible, depending on the different traffic assignment assumptions made by the two practitioners. The result of CODMRT estimations will then not be data of the kind found in the CFS, for example, with known accuracy dependent on the sampling rate in the sample frame. Instead, fused data should be considered conditional estimates, dependent on the appropriateness of the model and accuracy of the assumptions made by the data fuser. The results of a fusion of a movement database with a shipment database are not data in the sense that individual survey responses are data. They are estimates, interpolations, or forecasts and are only as good as the assumptions and judgment of the data fuser.

SEPARATING DATA COLLECTION AND DATA FUSION

Since CODMRT data require fusion of different data sets and since the method of fusion depends on assumptions made by the fuser, innovation and experimentation should be encouraged. The marketplace should then be offered different products made by different organizations employed to query the data archive. After some time, it is reasonable to assume that the market will find some fusion assumptions better than others, though improvements in methods may take years to develop. The encouragement of multiple ways of fusing data argues in favor of data fusion being primarily a private initiative rather than a government enterprise. However, the result of data fusion will be most accurate if the fuser works with individual records, in violation of confidentiality restrictions. Key to the framework allowing the assembly of CODMRT data will be the ability to solve the twin problems of maintaining confidentiality and encouraging imagination and innovation in data fusion.

The U.S. Department of Transportation's Bureau of Transportation Statistics (BTS) has experience in maintaining a confidential database of transportation records. The flagship product of BTS has been the CFS. The CFS will be at the heart of any movement toward CODMRT data assembly. However, most freight data in the United States are collected outside of BTS. There are several modal data sources, for example the 1 percent waybill sample collected by the Association of American Railroads and the U.S. Army Corps of Engineers' Statistics of Waterborne Commerce. The Reebie Transearch data are known to use proprietary truck data in conjunction with the CFS. Individual cities collect cordon count data, in which truck types and volumes are counted on individual highway segments. Import data are in the process of coming online, and intelligent transportation system (ITS) data are collected as trucks pass weigh stations in multiple states. A data fuser would wish to use the individual record microdata from all of these sources to estimate CODMRT data, something that jurisdictional boundaries and confidentiality problems now preclude.

To facilitate the creation of CODMRT data, BTS should become a depository or real-time archive of all forms of freight microdata. BTS should then allow data fusers to access these data if they follow confidentiality rules on publication. BTS should also be in a position to provide incentives to data gatherers to place their microdata in the BTS archive. One possible way to do this is to forbid data fusers who have access to confidential data from the BTS archive to fuse data from proprietary sources outside of the BTS archive. In this way, for example, an organization that continued to use proprietary data outside of the BTS archive would create products that were less accurate than other data sources since it would not have access to microdata in the BTS archive; organizations would then have an incentive to place their proprietary data in the BTS archive so that data fusers using the data could have access to other confidential data as well. Similarly, localities interested in understanding metropolitan freight flows would have an incentive to contribute local data in order to have access to the microdata in the BTS archives. If proprietary data are more valuable to the collector when they are fused with confidential data in the BTS archives, owners of proprietary data will have an incentive to contribute their data set, thus augmenting the whole.

FREIGHT DATA ADVISORY BOARD

The larger is the BTS data archive, the more valuable it will be. If it reaches a critical size, it is reasonable to expect that data collectors, for example the Association of American Railroads and the U.S. Army Corps of Engineers, will voluntarily add their data so the data fusers can get access to the confidential records in it. Success will require an impartial hand overseeing relationships among the archive holder (presumably BTS); data fusers, who would have access to the confidential data in the archive; and the data fusers' customers, who would not have access to confidential data. BTS is not appropriate as an overseer since it will also participate in the process as data archivist. This oversight should instead be given to the Freight Data Advisory Board, composed of representatives of data generators (modes, shippers, EZ-Pass and similar organizations, etc.), data users (including third-party data fusers and their customers), and government data organizers (BTS and state and metropolitan statistical agencies).

The Freight Data Advisory Board should define the division of tasks between the data archive maintainer (BTS) and the data fusers. BTS should not be precluded from publishing summaries of data in the data archives that meet confidentiality conditions, as it now does. BTS would also be expected to impute values of missing data from single data sets using generally accepted methods. However, imputing information about freight flows from combinations of data sets in the archives would be the primary task of independent third-party data fusers, certified as qualified to view confidential data but forbidden to disclose it to any parties outside of BTS. The data summaries provided by data fusers should also meet the criteria for confidentiality that BTS must abide by. In order to maintain confidentiality, the Freight Data Advisory Board should publish guidelines for systematic aggregation criteria to mask activities of individual shippers. In fact, deciding how to aggregate the very thin CODMRT flows to ensure maximum usefulness of route-level data while maintaining confidentiality will be one of the first and most important tasks for the advisory board.

The output of data fusion will in most cases be estimates of the commodities shipped on a particular transportation link at specific times of day, tagged by origin and destination. There are far too many CODMRT combinations for these data to be published on a national basis, but estimates can be expected to be made for specific local projects. As noted previously, the probability is very high that there will be fewer than four shippers of any specific commodity on a specific route at a specific time of day between two specific origins and destinations, and thus confidentiality rules will prevent publication of the data. One way around this problem is for the BTS advisory committee to develop rules to facilitate discussions between data fusers and shippers to waive confidentiality requirements where appropriate.

BRINGING NEW DATA COLLECTION PROGRAMS INTO THE DATA ARCHIVE

BTS, while not invited to be a data fuser, would be expected to be both a creator of data, as in the CFS, and an archivist of data collected by other organizations. The most promising external data sources to be included in the BTS archive are existing electronically collected passive data streams. Chief among these are ITS data that track trucks with appropriate transponders as they cross the country. Not all motor carriers or private truckers have ITS devices, so the data collected cannot be considered a random sample and cannot be used directly as CODMRT data even if commodity identity were collected by the transponders. However, these data should be a rich source of routing information, which could be fused with traditional CFS data covering commodities, origins, and destinations. The International Trade Data System is also a promising source of import data to fill the one significant coverage gap in the CFS. Passively collected data also have the promise of more timely data availability than has been possible in the past.

The Freight Data Advisory Board should also advise BTS on the desirability of starting new data collection efforts to augment the CFS and the data programs for which it acts as an archivist. One promising source of CODMRT data is roadside surveys like those conducted in Canada in which trucks are stopped randomly and the driver is asked to give information on routing, commodity, origin, and destination. If legal authority were found for collecting data in this way, a pilot study might be justified to determine the feasibility, acceptability, and cost-effectiveness of collecting CODMRT data through roadside surveys.

Less promising, but also worthy of consideration by the Freight Data Advisory Board for new data collection efforts to be undertaken by BTS, would be surveys of other participants in the logistics system—receivers, carriers, distributors, and so forth. Actively collected surveys like these are expensive and, like the CFS, do not provide timely information. Surveys of other supply chain participants should only be undertaken if they have information that is not available from other sources and if a sampling frame can be established that permits a random sample.

The CFS survey of shippers provides a long time line of COD data, thus permitting trend analysis. A shift of resources away from shipper surveys to paper surveys of other parties in the supply chain risks shrinking the CFS sample size, thus jeopardizing the reliability and usefulness of the CFS. If the Freight Data Advisory Board is attracted by the prospect of traditional survey instruments for parties other than shippers, pilot projects should be undertaken to ensure the feasibility, acceptability, and cost-effectiveness of these other survey types.

SUMMARY

Chapters 1 and 2 of this report recommend that freight data collection move in the direction of making credible estimates of freight flows specified by commodity, origin, destination, mode, route, and time of day. A framework that supports this goal must provide a mechanism for maintaining confidentiality of information while providing access to individual records of shipments and movements, the combination of which will be necessary if these finely defined transportation data are to be assembled. This appendix recommends the establishment of a Freight Data Advisory Board advising BTS. BTS would be charged with overseeing a freight data archive composed initially of existing databases augmented with passively collected electronic transportation data. The data archive will then be queried by a separate group of third-party data fusers, whose job will be to combine data sets in the archive by using their own assumptions about the data generation process and to create reports under contract to data users. While the data fusers would have access to confidential data in creating the reports, their output would be subject to the same confidentiality conditions as we now have. The relative roles of the data archive holder and the data fusers would be arbitrated by the Freight Data Advisory Board.

REFERENCE

Southworth, F. 1999. *The National Intermodal Transportation Data Base: Personal and Goods Movement Components* (draft). Oak Ridge National Laboratory, Oak Ridge, Tenn.

Study Committee Biographical Information

Arnim H. Meyburg, *Chair,* is Professor in the School of Civil and Environmental Engineering at Cornell University, where he has been a faculty member for more than 30 years. He is also Director of the Transportation Infrastructure Research Consortium of the New York State Department of Transportation, a position he has held since 1995. Dr. Meyburg's research interests include the development and use of models for planning passenger and freight movements; improved methods for surveying travel behavior intended for the development of travel behavior models; and the economics of transportation regulations, infrastructure, and systems management. He is a member of the National Research Council (NRC) Committee to Review the Bureau of Transportation Statistics' Survey Programs and recently prepared the conference synthesis for the November 2001 meeting, "Data Needs in the Changing World of Logistics and Freight Transportation," held in Saratoga Springs, New York. Dr. Meyburg holds a B.A. (equivalent) from the Free University of Berlin, and M.S. and Ph.D. degrees from Northwestern University.

Paul H. Bingham is a principal with Global Insight, Inc., where he manages consulting work for the company's global trade and transportation group. He has 19 years of experience in economic analyses of domestic and international freight transportation and is a specialist in freight transportation data. His work for public- and private-sector clients has included studies on the evolution of freight markets, patterns of international trade flows, and the economics of transportation infrastructure projects. He has managed freight forecasts for more than a dozen state departments of transportation or port authorities and several large metropolitan planning organizations. Mr. Bingham chairs the Transportation Research Board (TRB) Committee on Freight Transportation Data and was a member of the NRC Committee for the Study of Freight Transportation Capacity for the Next Century. He holds a B.S. in economics.

Kenneth D. Boyer is Professor of Economics at Michigan State University, a position he has held since 1984. His current research focuses on the economics of transportation, the economics of antitrust and regulation, and the measurement of the costs and benefits of activity within geographic networks. His recent publications include papers on trucking deregulation, the effects of the North American Free Trade Agreement, and spatial techniques for analyzing transportation demand, and a textbook, *Principles of Transportation Economics.* Dr. Boyer is a member of the Industrial Organization Society and the Transportation and Public Utilities Group of the American Economic Association. He has served on several NRC committees, including the Committee for the Study of Public Policy for Surface Freight Transportation, which prepared the report *Paying Our Way: Estimating Marginal Social Costs of Freight Transportation* (TRB Special Report 246). He holds a B.A. from Amherst College and M.A. and Ph.D. degrees from the University of Michigan, all in economics.

Robert Costello is the Chief Economist and Vice President for the American Trucking Associations (ATA), the national trade association for the trucking industry. In this position, he manages all ATA's collection, analysis, and dissemination of trucking economic information, including several monthly trucking economic indicators, motor carrier financial and operating data, and an annual freight transportation forecast. He also conducts economic analyses of proposed regulations and legislation affecting the trucking industry. Prior to joining ATA, Mr. Costello was an economist with Joel Popkin & Company in Washington, D.C., an economic consulting firm specializing in the analysis of wages, inflation, and macroeconomic trends. He holds a B.A. degree from Winona State University in Winona, Minnesota, and a master's degree from the University of Northern Iowa.

David L. Ganovski is Director of Rail Freight Services with the Maryland Department of Transportation. He administers statewide railroad issues and is the state's primary consultative resource on matters relating to freight movements, particularly by rail. He served on the Maryland Governor's merger task force that managed the recent breakup of Conrail. Mr. Ganovski has participated in several United States–European Union intermodal summit meetings on rail movement and infrastructure planning; he also cochairs the American Association of State Highway and Transportation Officials' rail subcommittee infrastructure task force. Before joining the Maryland Department of Transportation, Mr. Ganovski held management positions with CSX Transportation in Baltimore, Richmond, Pittsburgh, Houston, and Jacksonville. He was responsible for a range of sales, marketing, operations, and operations planning activities, and for many years he was involved in the CSX "Total Quality Process." Mr. Ganovski majored in transportation law and regulation at the University of Baltimore and holds a Certificate of Transportation Management, also from the University of Baltimore.

J. Susie Lahsene is transportation planning manager for the Port of Portland, Oregon. She manages transportation and logistics planning, transportation capital planning, and resource acquisition for the port's surface transportation system. She chairs the TRB Committee on Urban Freight Transportation and is a founding member of the Portland Chapter of the Women's Transportation Seminar. Ms. Lahsene holds a B.A. in urban studies and a master's degree in urban and regional planning, both from the Virginia Polytechnic Institute and State University, and an M.B.A. from the University of Portland.

Catherine T. Lawson is Assistant Professor in Geography and Planning at the State University of New York at Albany, and Interim Director of the university's Master's Program in Urban and Regional Planning. Her research and teaching interests include freight and archived intelligent transportation systems, geographic information systems, transportation planning, and spatial analysis and statistical applications. She recently completed a project for the Oregon Department of Transportation on the development of an effective methodology to survey the freight community and is currently working with the Washington State Department of Transportation on the formulation of performance measures for freight movements in the I-5 corridor using In-Transit Visibility technologies. Dr. Lawson holds a B.A. in economics and accounting from Western Washington University, master's degrees in applied economics and urban and regional planning from Portland State University, and a Ph.D. in urban studies and regional science, also from Portland State University.

Robert E. Martínez is Vice President, Marketing Services and International with Norfolk Southern Corporation, which owns Norfolk Southern Railway Company, a major freight railroad. From 1994 to 1998, Dr. Martínez was the Secretary of Transportation in the Commonwealth of Virginia with oversight responsibility for the development and implementation of Virginia's transportation program. He also had management and budgetary responsibility for the commonwealth's Department of Transportation, Department of Motor Vehicles, Department of Aviation, Department of Rail and Public Transportation, and Port Authority. His previous positions include Manager of Strategic Planning at Norfolk Southern Corporation and Associate Deputy Secretary of Transportation with the U.S. Department of Transportation. Dr. Martínez holds a B.A. degree from Columbia University, an M.A. degree in international relations from Yale University, and a Ph.D. degree in political science, also from Yale.

Robert Tardif is a senior planner with the Ontario Ministry of Transportation in Downsview, Canada. He has 17 years of experience in transportation planning, with emphasis on demand forecasting and strategic studies. In recent years, his work has focused on the conduct and analysis of commercial vehicle surveys, and he was the project manager for Ontario's 1999 Commercial Vehicle Survey. This activity involved developing partnerships with New York and Michigan through the Eastern Transportation Border Coalition to jointly fund data collection at international gateways. Mr. Tardif has developed models for forecasting future freight flows to assist the Ministry of Transportation in its planning and capital allocation processes. He holds a diploma in civil engineering technology from Humber College and has completed several transportation planning courses at the University of Toronto.

C. Michael Walton is Professor of Civil Engineering and holds the Ernest H. Cockrell Centennial Chair in Engineering at the University of Texas. He also has a joint academic appointment in the Lyndon B. Johnson School of Public Affairs. Dr. Walton is actively involved in developing and defining transportation policy through his research, publications, and national service to government and industry. He has contributed to more than 200 publications in the areas of intelligent transportation systems; freight transport; and transportation engineering, planning, policy, and economics. He is a founding member of the Intelligent Transportation Society of America and currently serves as Chair of its Board of Directors. Dr. Walton is a past chair of the TRB Executive Committee and a member of the National Academy of Engineering. He holds a B.S. from the Virginia Military Institute and master's and Ph.D. degrees from North Carolina State University, all in civil engineering.