TRANSPORTATION RESEARCH

# CIRCULAR

Number E-C205 March 2016

# Commodity Flow Survey Workshop

October 29, 2015
Washington D.C.

TRANSPORTATION RESEARCH BOARD

## TRANSPORTATION RESEARCH BOARD
## 2016 EXECUTIVE COMMITTEE OFFICERS

**Chair: James M. Crites**, Executive Vice President of Operations, Dallas–Fort Worth International Airport, Texas
**Vice Chair: Paul Trombino III**, Director, Iowa Department of Transportation, Ames
**Division Chair for NRC Oversight: Susan Hanson**, Distinguished University Professor Emerita, School of Geography, Clark University, Worcester, Massachusetts
**Executive Director: Neil J. Pedersen**, Transportation Research Board

## TRANSPORTATION RESEARCH BOARD
## 2015–2016 TECHNICAL ACTIVITIES COUNCIL

**Chair: Daniel S. Turner**, Emeritus Professor of Civil Engineering, University of Alabama, Tuscaloosa
**Technical Activities Director: Ann M. Brach**, Transportation Research Board

**Peter M. Briglia, Jr.**, Consultant, Seattle, Washington, *Operations and Preservation Group Chair*
**Alison Jane Conway**, Assistant Professor, Department of Civil Engineering, City College of New York, New York, *Young Members Council Chair*
**Mary Ellen Eagan**, President and CEO, Harris Miller Miller and Hanson, Inc., Burlington, Massachusetts, *Aviation Group Chair*
**Barbara A. Ivanov**, Director, Freight Systems, Washington State Department of Transportation, Olympia, *Freight Systems Group Chair*
**Paul P. Jovanis**, Professor, Pennsylvania State University, University Park, *Safety and Systems Users Group Chair*
**D. Stephen Lane**, Associate Principal Research Scientist, Virginia Center for Transportation Innovation and Research, *Design and Construction Group Chair*
**Hyun-A C. Park**, President, Spy Pond Partners, LLC, Arlington, Massachusetts, *Policy and Organization Group Chair*
**Harold R. (Skip) Paul**, Director, Louisiana Transportation Research Center, Louisiana Department of Transportation and Development, Baton Rouge, *State DOT Representative*
**Ram M. Pendyala**, Frederick R. Dickerson Chair and Professor of Transportation, Georgia Institute of Technology, *Planning and Environment Group Chair*
**Stephen M. Popkin**, Director, Safety Management and Human Factors, Office of the Assistant Secretary of Transportation for Research and Technology, Volpe National Transportation Systems Center, Cambridge, Massachusetts, *Rail Group Chair*
**Robert Shea**, Senior Deputy Chief Counsel, Pennsylvania Department of Transportation, *Legal Resources Group Chair*
**Eric Shen**, Director of Transportation Planning, Port of Long Beach, *Marine Group Chair*
**David C. Wilcock,** Vice President and National Practice Leader for Rail and Transit, Michael Baker, Jr., Inc., Norwood, Massachusetts, *Public Transportation Group Chair*

TRANSPORTATION RESEARCH CIRCULAR E-C205

# Commodity Flow Survey Workshop

October 29, 2015
National Academies of Sciences, Engineering, and Medicine
Washington D.C.

*Supported by*
Office of the Assistant Secretary for Research and Technology
Bureau of Transportation Statistics


Kathleen Hancock
*Virginia Polytechnic Institute and State University*
*Editor*


February 2016


Transportation Research Board
500 Fifth Street, NW
Washington, DC 20001
www.TRB.org

TRANSPORTATION RESEARCH CIRCULAR E-C205
ISSN 0097-8515



**Conference Planning Team**

Alison Conway, *Chair*, *City College of New York*
Scott Drumm, *Port of Portland*
Ronald Duych, *Bureau of Transportation Statistics*
Kathleen Hancock, *Virginia Polytechnic Institute and State University*
James Hinckley, *United States Census Bureau*
Bruce Lambert, *Institute for Trade and Transportation*
Catherine Lawson, *University of Albany, SUNY.*
Douglas MacIvor, *California Department of Transportation*
Michael Meyer, *WSP–Parsons Brinckerhoff*
Mathew Roorda, *University of Toronto*
Michael Sprung, *Bureau of Transportation Statistics*

**TRB Staff**

Scott Babcock, *Transportation Research Board*
Thomas Palmerlee, *Transportation Research Board*
Michael Miller, *Transportation Research Board*

Transportation Research Board
500 Fifth Street, NW
Washington, DC 20001
www.TRB.org

Lea Camarda, Production Editor; Jennifer Correro, Proofreading and Layout

# Preface

The Commodity Flow Survey (CFS) is the core data source for many freight planning activities. The TRB Freight Transportation Data Committee collaborated with the U.S Department of Transportation (DOT) Bureau of Transportation Statistics (BTS) to initiate a workshop to provide input to the next application of the survey. Convened after the release of the most recent detailed CFS data and coinciding with plans for the 2017 CFS, the workshop offered an interactive format for a diverse set of users to engage in productive dialogue.

An ad hoc committee, chaired by Alison Conway of the City College of New York and selected by the sponsoring committee, carried out the detailed planning for the conference. This e-circular consists of individually attributed summaries. No language should be construed as consensus findings or recommendations on the part of the conference; the planning committee; the rapporteur; TRB; the National Academies of Science, Engineering, and Medicine; or the Freight Transportation Data Committee.

The planning committee represented CFS producers, analysts, and modelers. A list of these attendees can be found in Appendix B. The 81 attendees represented the following organizations:

| | |
|---|---|
| U.S. Department of Transportation | 20 |
| U.S. Census Bureau | 19 |
| State governments | 3 |
| Local and regional ports | 2 |
| Consultant–private sector | 13 |
| Universities | 12 |
| Other | 12 |

The Office of the Assistant Secretary for Research and Technology provided funding to support travel and onsite expenses for planning committee members and selected speakers. The planning committee thanks Kathleen Hancock for her work preparing this e-circular, as well as TRB staff and the presenters for their contributions to the success of the workshop. The committee also extends a special thanks to BTS for providing the support that made this workshop possible.

# Contents

## CONCLUDING REMARKS



## POSTER PRESENTATIONS: APPLICATIONS OF THE COMMODITY FLOW SURVEY



## APPENDIXES

# Introduction

**ALISON CONWAY**
*City College of New York*

Since its inception in 1993, the Commodity Flow Survey (CFS) has provided an essential data source for understanding freight flows generated by economic activity in the United States. The survey, now conducted every 5 years as part of the Economic Census, is a joint effort of the U.S. Census Bureau and the U.S. Bureau of Transportation Statistics (BTS) and remains the primary information source for national- and state-level domestic shipments in many sectors and the only publicly available source of commodity data for the highway mode.

CFS data are critical for enabling federal, state, and local authorities to make planning and policy decisions that enhance the performance of the national and regional freight transportation systems that support U.S. industry. In previous iterations of the survey, data have been provided to users, including public agencies, industry practitioners, and academic researchers, in the form of aggregated national, state, and metropolitan level flows by mode, commodity group, and trip distance. For the first time, 2012 CFS data products also included shipment-level data in the form of a Public Use Microdata Sample (PUMS).

To remain an effective tool to measure nationwide goods movements, it is important that the CFS be updated in each subsequent implementation to recognize rapidly changing economic conditions and to take advantage of continuously improving data collection technologies and analytical methods. Following in the tradition of previous successful workshops held in 2007 and 2011, the aim of this workshop was twofold. First, the workshop was convened to provide an opportunity for direct feedback from the community of CFS data users to discuss potential near-term improvements and methodological changes for the 2017 CFS. Second, the event also served as a forum for discussing broader emerging technologies and logistics trends that may have near- and long-term implications for the design, implementation, and use of the CFS.

To serve these aims, the workshop included a variety of activities. First, in the opening session, agency and academic leaders laid out the important role that the CFS plays in understanding U.S. economic activity and related freight transportation demand. They also introduced the rapidly changing landscape for freight transportation data. Next, a panel of agency authorities focused on the content of the 2012 CFS and on experience from both the United States and Canada in developing and implementing commodity surveys. This session presented updates and improvements implemented for the 2012 CFS; the content, potential uses and limitations of the PUMS; and recent experiences of Statistics Canada and Transport Canada in development of a new Canadian CFS. During a midday poster session, researchers from government, industry, and academia presented their completed and ongoing work using CFS data. An afternoon panel consisting of experts from government, academia, and industry focused on trends that could influence the future of the CFS. Presenters discussed recent advances in data collection technology that have generated new opportunities to measure and observe freight transportation activity, advanced modeling applications employing CFS data, and rapidly emerging logistics trends made possible by communications technologies.

Following these panels, attendees participated in three breakout groups where they were tasked with discussing the following questions:

- What revisions or updates may be needed to the CFS structure (e.g., classification systems)?
- What revisions or updates may be needed to the CFS scope (e.g., commodities, geographies)?
- What revisions, updates, or additions to CFS sampling methods are needed?
- How can CFS tools and functionality be improved for current and future uses?

In the closing session, leaders from each group reported on their discussions. Opportunities were then provided for attendee discussion of the points presented. The workshop closed with comments from the Census Bureau and BTS leadership.

The remainder of this document details the activities of the workshop. This E-Circular serves as an information source for the broader community of existing and potential CFS data users as well as an invitation for those interested to provide input to the Census Bureau and BTS in support of their continuous efforts to improve the CFS in 2017 and beyond.

# The Big Picture

## *What Is the Role of the Commodity Flow Survey?*

**PATRICIA S. HU**
*Bureau of Transportation Statistics*

Transportation and the economy are interlinked and their relationship is multidimensional. First, transportation is, itself, an economic activity. Second, and perhaps more important, transportation is an enabler of economic activity through moving goods between places and linking workers with jobs. Finally, transportation is a reflection of economic activity. The relationship between freight transportation and the economy expands on and broadens the latter dimension and is represented by the freight transportation service index (TSI) developed by BTS. Freight TSI measures month-to-month changes in freight shipments. The analysis of TSI for the past 25 years has suggested that freight activity leads the economy by 4 to 6 months. Figure 1 shows the relationship between freight activity and the growth cycle in the general economy. The grey bars represent times when the economy is below the long-range growth trend. The labeled peaks show that TSI generally peaked before a growth slowdown. The TSI then bottoms out (troughs) before the growth slowdown ends. This reveals a close and consistent relationship between freight transportation and economy.

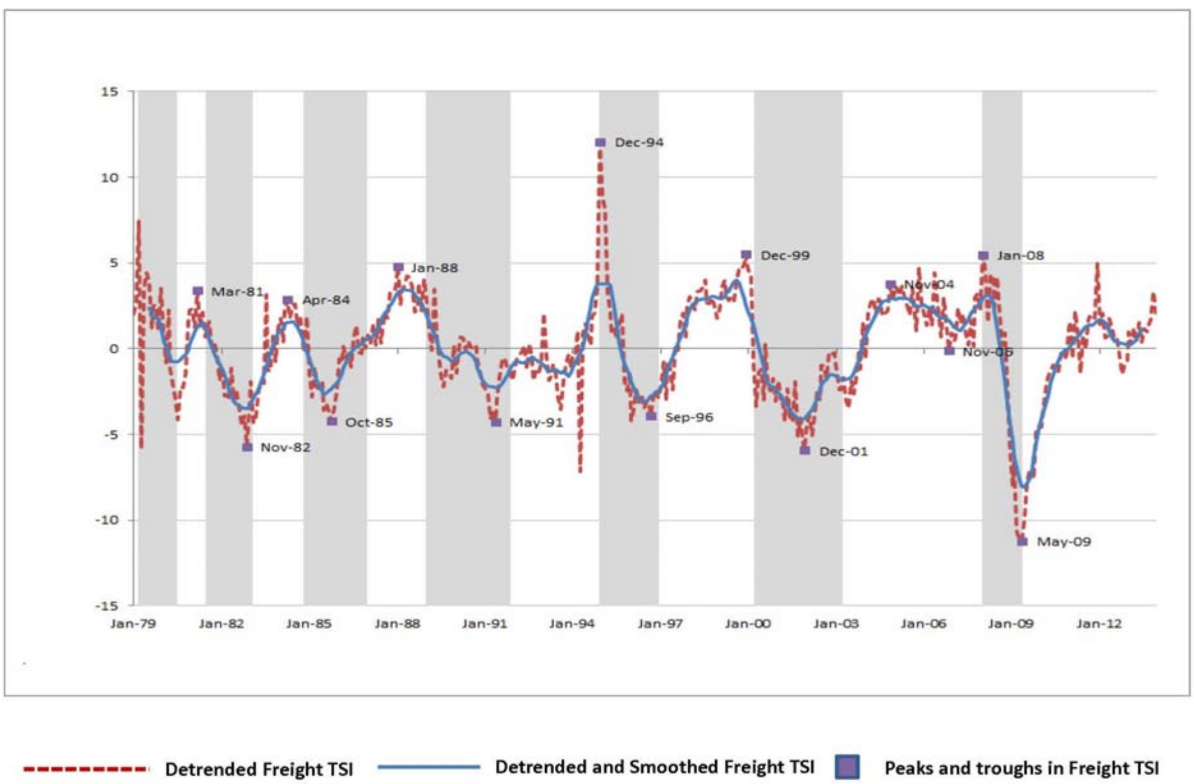


**FIGURE 1 Freight TSI as a reflection of economic activity.**

The freight TSI is generated from the Freight Analysis Framework (FAF) which relies on the CFS in combination with other data sources. The CFS remains the most significant data contributor to FAF and accounts for approximately 60% of the tonnage data and 70% of the value data in FAF. Without the CFS, the FAF would not exist and without the FAF, the freight TSI would not be possible.

The role of the CFS is to support the private and public sectors in preparing for the future of freight transportation both in the near term for the next survey and in the longer term. The challenge is understanding the factors that will drive the future of freight transportation such as shifts in global trade and the global economic gravity; rapidly expanding technologies; changing policies; expansion and evolution of the shared economy and personal fabrication; sharing of infrastructure; shifts in manufacturing; and demographic changes. How should the CFS evolve into a responsive freight data program?

One thing is certain, the demand for freight transportation will increase significantly. The U.S. Department of Transportation (DOT) projects that the volume of freight moving on the nation's transportation network will increase by 45% to 29 billion tons over the next 15 years. Figure 2 shows the Organization for Economic Co-operation and Development's (OECD) international transport projections, indicating that the north Atlantic will remain an important freight corridor with freight volume increasing 270% by 2050 and the north Pacific surpassing the north Atlantic to become the primary trade corridor, reflecting that Asia is becoming the global economic center of gravity.

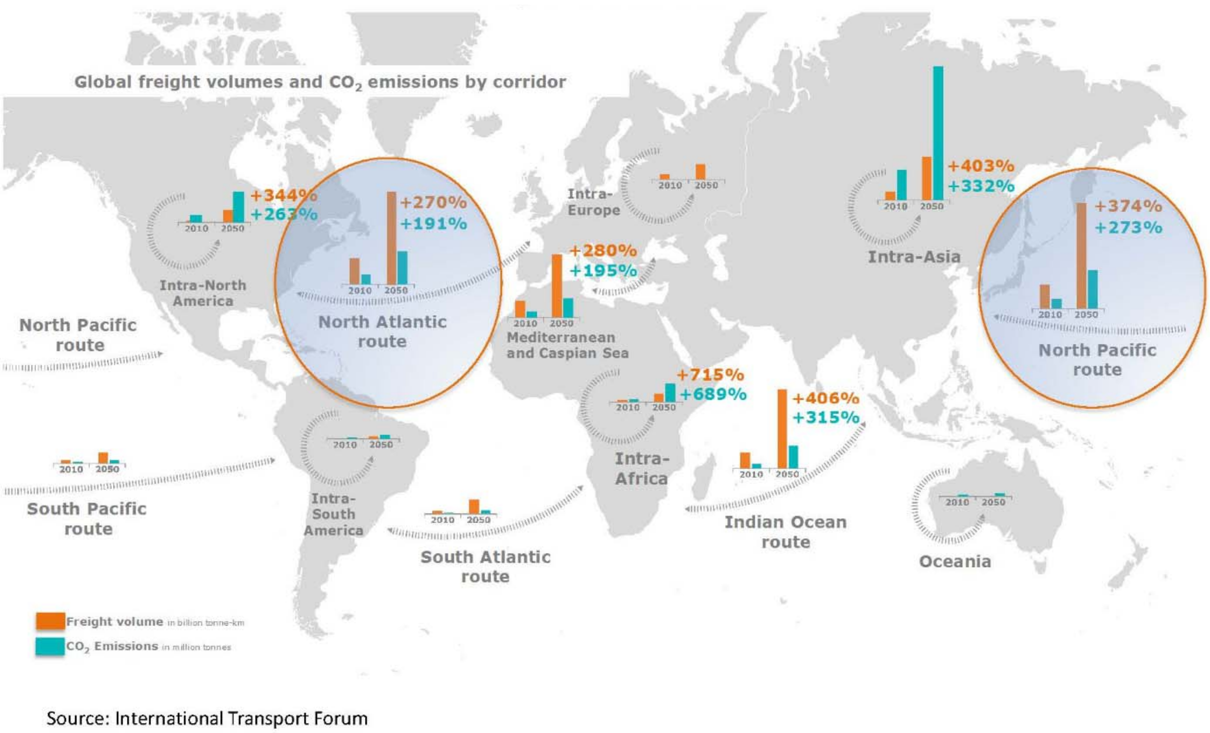


**FIGURE 2 Global freight volume and $CO_2$ emission by corridor: projects to 2050.**

So what does this mean? By 2050 there could be five or six more CFS surveys. How should the program evolve to address the potential changes that are being predicted? In response, consider the following:

- What information is needed to predict the future of freight transportation?
- What information is needed to prepare for the future of freight transportation?
- How is this information incorporated into the development and design of the next generation of the CFS?

Because of importance of freight transportation to the economy, consideration of both short-term enhancements (the next CFS in 2017) and long-term evolution (beyond 2017) of the CFS freight data program is critical to preparing the United States for the future. Past CFS workshops have been very effective in helping BTS and the Census Bureau improve the usefulness of CFS and its products. Many of the recommendations and suggestions from the last workshop were implemented in the 2012 CFS, most notably the creation of the CFS PUMS.

This workshop provides an opportunity to discuss ideas for enhancing the 2017 CFS, making it even more relevant and useful. It also provides the opportunity to begin the discussion for a future freight transportation data program that takes advantage of rapidly changing technological capabilities and addresses evolving logistical trends.

# The Role of the Commodity Flow Survey
## *Ever Evolving*

**John Thompson**
*U.S. Census Bureau*

The Census Bureau conducts many surveys of U.S. businesses, fulfilling its mission to serve as the leading source of quality data about the nation's people and economy. The largest of these surveys is the Economic Census, the cornerstone of the United States' system of economic statistics. Its data products provide the foundation for key measures of economic performance for many sectors of the economy, including mining, manufacturing, wholesale, retail, warehousing, and transportation services, among others. The CFS, a component of the Economic Census, provides a snapshot of a specific activity—the transportation of goods—across these economic sectors.

For more than 20 years, the Census Bureau and the BTS have joined to conduct the CFS, providing data users important statistics about the movement of goods. Estimates are produced for the value, tons, ton-miles, and average miles per shipment by commodity; mode of transportation; and origin and destination geography. Additionally, data are available for hazardous materials and exports. All of these data are used by a wide variety of people and organizations. From all levels of government (federal, state, and local) to students, consulting firms, and transportation companies, CFS statistics are in high demand.

As transportation evolves to meet the needs of the global economy, the CFS estimates improve understanding of this evolution and help data users make informed decisions. As transportation continues to change, the greater the need for the CFS. As a result, the surveys and data products need to evolve as well. Demand for data continues to rise. Everyone wants more data, more detail, more often. Meeting these demands requires resources in an uncertain budget environment. It is in this environment that innovative ways to improve the CFS and other economic surveys are important.

The CFS has already made a step to provide more data with the June 2015 release of the PUMS that so many have already spent hours analyzing. BTS and the Census Bureau listened to data user feedback at the previous workshop and provided data on approximately 4.5 million shipments.

The program must evolve in more ways than just data products. The Census Bureau strives to reduce respondent burden, automate more operations, and reduce costs. In the 2012 CFS, an Internet-based reporting instrument was introduced to respondents. This will be expanded in 2017 with the goal of moving to a paperless survey. For several surveys, including the Economic Census and the CFS, Census Bureau also wants to work with businesses to collect data electronically in their own format rather than through a typical survey questionnaire. This should reduce respondent burden and, in the case of CFS, provide more shipment data than is possible through the current survey methodology. More data, more often at a reasonable cost, expands the opportunity to provide even more meaningful, timely data to users.

Obviously, the flow of goods is not limited to the U.S. borders. Canada is well on its way to conducting its own CFS. Learning from the Canadian experiences and seeing their results will expand knowledge about the transportation of goods outside the United States and across its borders.

This workshop was an important step in planning for the 2017 CFS. The discussions provided information for meeting the evolving needs of CFS data users. Building on their successful partnership, the Census Bureau and BTS will continue to work together to improve the CFS.

# The Future of Freight Data

**JOSEPH SCHOFER**
*Northwestern University*

The link between the CFS and the future of freight data is that the CFS has the opportunity to evolve to meet some of the future data needs of the freight transportation system. From the first presentation, we were reminded that freight transportation is an essential contributor to the state of the economy. As a result, the context for discussing the future of freight data and the CFS is the evolving state of the economy and society. When we consider the business case for freight data, we must recognize the themes that affect data users and uses, changing data attributes, data sources and channels, and then bring these together to paint a picture of the future. The following discussion presents my ideas about that picture, along with some counterpoints and implications.

When we consider the business context, three evolving themes are important: the sharing economy, changing location patterns of production and consumption, and vehicle automation. The sharing economy is now part of society. Today's young adults have been participating for a long time. This includes sharing rides, cars, bikes, accommodations, any asset that can be shared with others. Uber is coming to freight which could have dramatic effects on the last mile and urban delivery.

Without data and information, without the technologies for moving that information around, this sharing could not take place. This leads to modes that are organized differently, increased reliance on information technology, more data, and perhaps more access to flow data that can be important in system planning and management.

And the counterpoint: "Is this trend real, is it substantive, is it going to stay? Will people share these data?" The answers are yes; it is already happening, it is a done deal. How does this relate to freight? Much of the resulting information describes behavior on and of the transportation network, which is important for managing the movement of freight. It means there's a lot more information available and the challenge is how to access and use that data.

The dynamics of location patterns are changing. Location decisions are influenced by characteristics such as freight mobility needs, labor performance, and labor costs of the transportation system, including capacity and reliability, and by changing technology. There is evidence that freight costs and reliability are influencing reshoring and nearshoring of manufacturing. 3-D printing has the potential to move some manufacturing into or closer to homes, which then impacts the flow of raw materials and finished products. The trend toward higher density living and working is shifting consumption locations. Multichannel retailing is diversifying flow patterns.

The counterpoint is that there are always location dynamics but the perturbations are small and will continue to happen. However, because of the rapid technological and value changes, and forces from the global economy, these location patterns are changing more rapidly and substantially. The implication is that as we collect data about freight and freight movements, we need to understand the trends in origins and destinations, rather than relying on historical patterns. Where are the centers of production and consumption and how are they changing? What data collection techniques are available that follow the freight rather than sticking with the old Os and Ds, particularly in a survey that collects data at fairly large time intervals.

Automation is coming and, in particular, it is coming to trucks. A wide variety of driving assist devices are already in trucks and more are coming rapidly to market. With current driver shortages and increasing pressures to improve productivity and safety, the incentive to expand truck automation is enormous. These technologies rely on multiple sensors either as part of the new technologies or already embedded in vehicles, and this implies the potential for a huge flow of data that could be tapped.

The counterpoint is that we cannot access that data because of proprietary concerns. However, we have already seen examples where these concerns have been addressed, usually through a trusted third party. As discussed below, the real potential comes from industry's ability to monetize this resource. The implication is that we should expect to pay for new data streams.

From the big picture, I want to shift to the future of freight data uses and attributes.

One of the concerns is that there are a lot of people who are very good at building models but who do not necessarily understand what decision makers do with those models and what their ultimate value is. Considering the uses of freight data, it is important to frame the approach that we take for the future in terms of decisions about freight system development and management and how data are mapped to those decisions. This is illustrated in Figure 3, with some decisions on the left and data needs on the right.

Freight is about economic development, and thus freight data are also about economic development. The value of transportation links and nodes is reflected in what is flowing across them. We've learned from transportation leaders that average daily traffic is not what's important; instead, we need to know what traffic (commodities) is moving on the link and its value. That drives whether or not investments should be made.

The counterpoint is that data are not available about what products are on a particular link. But that is the kind of information that we need. Economic and social development outcomes are the raison d'etre for transportation investments, and good data helps make smart choices. The current CFS is an entry point for estimating flows on the links, but more detailed and timely data are needed for strategic decision making.

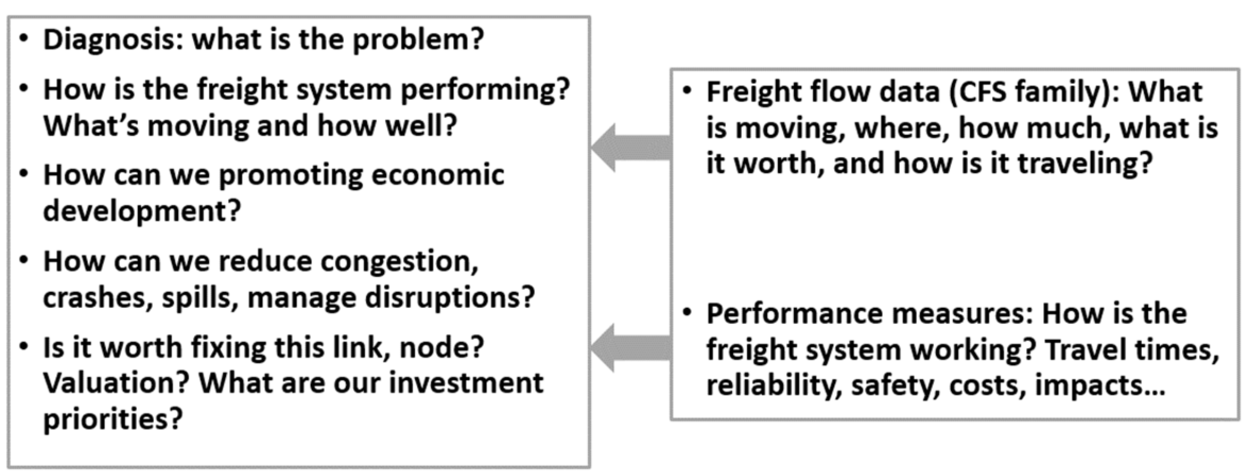


**FIGURE 3 Mapping freight data to decisions.**

Fresh data are the best data. Making 2016 policy using 2007 data is problematic. It may be the best we can do today but it is not good enough. Stale data can not only undermine credibility of a decision, it can result in addressing yesterday's instead of tomorrow's problems. It also makes it difficult to argue for resources for data programs if we are still distributing data that are multiple years old. The counterpoint is that it takes a long time to collect, process, and deliver data. However, we should consider other programs like the American Community Survey and evaluate whether there are ways to engage in a continuous approach to data collection.

There has been a push for acquiring and disseminating real-time freight data, particularly given the widespread availability of close-to-real-time traffic data and the public interest in real-time tracking of packages, but real-time data are not what we need for strategic decision making. Some have described it as mere eye candy. We need to show real-time data on a website to be considered relevant. But real-time data are about what is happening now—well, not really—approximately but not exactly now. The red links on the map on your smartphone are congested now but may be clear when you get there 30 min from now. We cannot use these data for strategic system development and management decisions. What is important is fresh data, data that is recent enough to show current and emerging trends that may guide major decisions. So the effective age of data is an attribute that is relative to the decision. For example, Figure 4 shows the trends of crude oil tank car loadings in the United States from 2009 to 2014, with the numbers in white giving the percentage increase year by year. The chart shows the large growth during the earlier years of exploitation of the Bakken crude oil, followed by a slowdown over the last couple of years. The large year-over-year changes suggest the need to collect data at fairly frequent intervals if we want to track and respond to such phenomena.

The equation in Figure 4 suggests that the value of real-time data lies in extracting performance trend measures from real-time data to support policy and investment decisions. We suggest that the interest in real-time freight data is really a demand for timely performance

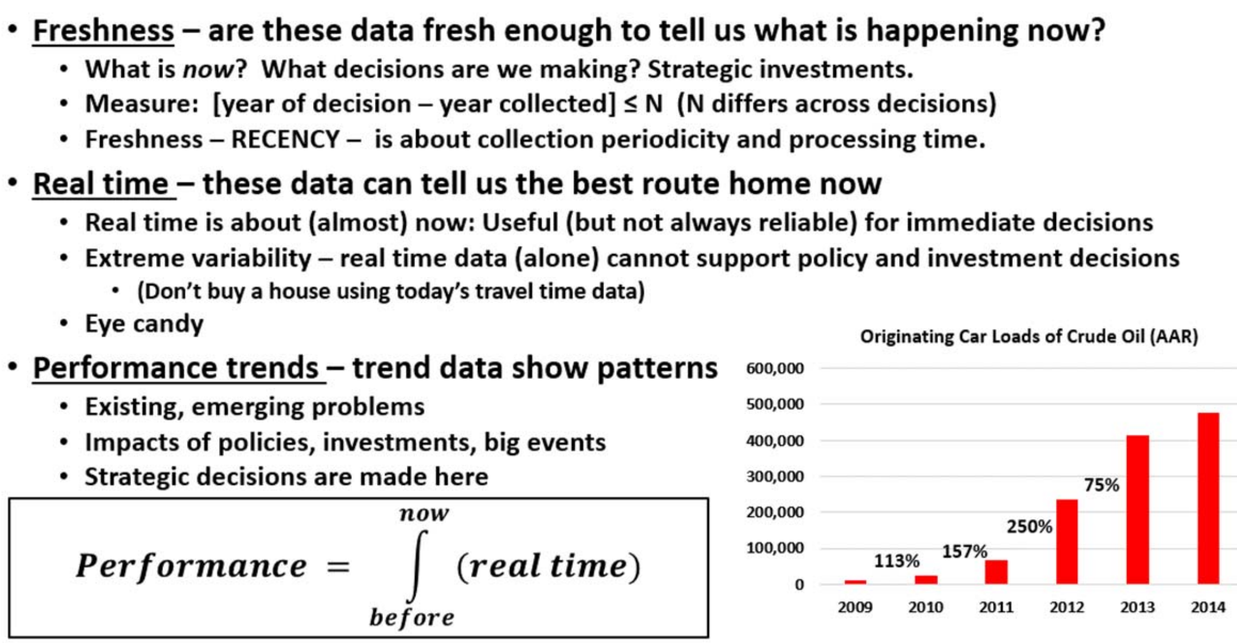


**FIGURE 4 Data freshness and decision making.**

measures that may be derived from real-time data. Table 1 provides examples of decisions with the corresponding ages of data to support those decisions.

Some transportation agencies have felt pressure to host real-time data. At least at the federal level, hosting performance data is more important. The private sector is very good at providing access to real-time traffic data and we should leave that task to it, instead focusing on meeting the interest in and need for performance data. The counterpoint to this discussion is that the CFS focuses on large-scale, long-term commodity flows, not operational performance. The implication is that we should use performance trends to link future CFS to key decision issues such as bottlenecks and system vulnerabilities.

Freight data are particularly important for understanding disruptions. Over the last 15 years, we have seen major disruptions on our Interstate system, which is the highest-performing highway system in the United States. I-5, I-10, I-15, I-65, I-95, and I-580 all suffered major disruptions, some lasting many weeks. Severe weather disruptions and infrastructure failures will continue and potentially accelerate in the future; understanding the impact of these disruptions is important for short-term management but also for strategic planning and for increasing the resilience of our transportation system. To do this, we need fresh performance data. FHWA has been able to evaluate the impact of recent disruptions such as I-65 in West Lafayette, Indiana, and I-10 in the California desert, observing detour patterns and congestion using GPS tracking data.

The counterpoint is that capturing disruption impacts requires quick response data collection and interpretation. However, fresh freight flow and performance data will help manage and potentially avoid disruptions. It is about freshness, it is about performance data, it is about integrating these data sources to address decisions that need to be made right now.

Big data is here. Tracking data, administrative records, sensor data, imagery, and social media are currently available. Opportunities exist because of the availability of enhanced detail, extensive coverage, electronic storage and transmission, accessibility, and powerful computational capabilities. We need to take advantage of those opportunities to understand freight transportation system performance, disruptions, and obstacles.

**TABLE 1 Examples of Decisions and Data Age**

| Decision–Question Answer | Data Age |
|---|---|
| Bridge washout truck detour route and route performance | Recent and near-real time truck volumes, origin–destination patterns, and commodities moved |
| Assessment of time of day route choice impacts of daytime truck toll increase | Current monthly trends, 3 years before, after initiation |
| Risk management–Bakken crude by rail<br>Regulation, routing | Fresh commodity flow trend (3to 12 months) |
| Columbia River crossing (bridge replacement) | Long-term vehicle and commodity flow trends (3 to 10 years) |
| I-75 pavement deterioration due to heavy trucks<br>Pavement design, truck size, and weight | Long-term weigh-in-motion and pavement condition data (15-plus years) |

The counterpoint is that it may be difficult to get access to such data, and it is not always clear how useful it will be. Ultimately, we will have access to more than we want. The challenge is to extract the value from the data, to find the information that is important to the decisions that need to be made and not to be distracted into exploring data just to see what is there.

Freight moves according to supply chains and that is what industry and consumers care about. As a result, we need to measure how freight really moves from that perspective. The freight fluidity initiative underway at FHWA provides a unique opportunity to make the connection between freight data and supply chains.

The counterpoint is that it is difficult to get supply chain data. Who knows how freight moves within the supply chain and physically on the transportation network? But knowing this is important from a public policy and economic development standpoint. We have the best domestic freight system in the world. The challenge is to make sure that we maintain that going into the future, and that we protect both the domestic and the export economies. Gathering supply chain data is a must.

Local freight will become even more important as manufacturing locations shift, retail delivery options diversify, and people select more dense living settings. Retail trade is changing rapidly with multiple ways to purchase and deliver items. The last mile options are changing as well. Items are delivered through the sharing economy, by parcel delivery, by courier, and soon by drone. The increasing diversity of moves will make tracking even more challenging. We will need to work with local agencies, carriers, and trade associations to measure local freight flows and performance.

The technology to track and report goods flows is becoming ubiquitous and cost-effective. As data are monetized, the availability of data will increase dramatically. The counterpoint is that it is proprietary data and people will not reveal what is in the box. However, once a system of exchange is in place that ensures that privacy is protected and that money is attached to that exchange, the data will become available. The old barriers are going to erode and new barriers will arise. We will need to address these as they become apparent.

Access to private data will increase. New players, new paradigms and enhanced technology for sharing data will result in more access to more detailed information. The public sector needs to grow its partnerships with the private sector and step aside when the private sector can do a better job in data collection and dissemination, purchasing data as needed.

The headline message is to “expect to pay for data.” Data providers, vendors, and data owners see the value of their data and it will not be free, and it has never been truly free since somebody had to pay for it. An implication of this is that decisions makers should consider the value of data to their decision making and include it as a necessary expense.

The biggest requirement for the future is to “tell the freight story.” Data are the foundation of that story. Freight is both invisible and essential. We need to use data to emphasize the link between freight system performance and economic competitiveness, and to show the economic outcomes of an effective system. This includes both a compelling narrative and the supporting visualizations that clearly depict the linkages, their causes and effects. These stories need to be crafted for and told to the general public as well as decision makers. Consumers now expect to receive products within a day or even an hour of ordering without understanding the requirements that make this expectation possible and the costs that are associated with ensuring that those expectations are met. This is an example of one story that needs telling.

Ideas for moving forward include the following:

- The government role in collecting, analyzing, and disseminating freight data remains strong, but it is changing.
- CFS will continue to be important, but its characteristics should evolve in response to technologies and user needs.
  - Include supply chain and performance data.
- Freight data are not only about CFS—there are or will be multiple freight data sources.
  - Public–private boundaries are blurred and shifting.
  - The challenge of data fusion must be addressed.
- Prepare to buy data.
  - Can it be win–win–win–win for users, agencies, shippers–carriers, and vendors?
- Big data is here, it just needs to be tapped.
  - There are needs for standardize access and sharing agreements and security tools.
  - Technical, institutional, and proprietary barriers need to be addressed.
  - The roles for commercial vendors and trusted third parties should be resolved.
- Performance data and trends are an important component of freight data.
  - Make room for real-time data, integrated over time to support performance measures.
  - Leave disseminating real-time data to others?
- Keep data fresh.
  - Continuous sampling
  - Fresh data + important uses = value.

We have taken a peek into the future of freight data. One thing to remember about such predictions is that they are likely to be wrong in some ways. We can be certain that the future will be different than what we see today, and that it will come much sooner than any of us expect. And finally, the future is changing rapidly and for better or worse, we will be making freight transportation decisions regardless of what data we have in our files. It would be good to have the right data to support those decisions.

# Content and Scope of the 2012 Commodity Flow Survey

**RYAN GRUBE**
*Bureau of Transportation Statistics*, *recorder*

A panel consisting of Ron Duych, BTS; Scot Dahl, Census Bureau; and Larry McKeown, Statistics Canada, spoke during the session about their unique involvement and experience with the CFS. Main points discussed during the session included 2012 CFS content and changes from the 2007 CFS, the release of a first-generation CFS PUMS file, and Statistics Canada and Transport Canada's involvement with a Canadian CFS.

## PANELIST COMMENTS

**Ronald Duych**
*Bureau of Transportation Statistics*

Ronald Duych presented an overview of the 2012 CFS highlighting changes that occurred since the 2007 CFS. As a project manager of the 2007 and 2012 CFS, he has the unique opportunity and perspective to discuss differences between both surveys. While providing an overview of the 2012 CFS, Duych mentioned the following items:

- CFS collects and publishes information on commodities shipped, shipment value and weight, ton-miles generated, origin and destination information, and mode of transportation.
- CFS is the only source of highway freight data for hazardous materials.
- Establishments are requested to provide a sample of their outbound shipments for one week (reporting week) during each quarter of the survey year.
- For the 2012 CFS, approximately 37% of all establishments sampled submitted a questionnaire for all four quarters, while approximately 57% of all establishments sampled submitted a questionnaire for at least one quarter.

Based on feedback from data users at the 2010 CFS workshop and from known data gaps in the 2007 CFS, multiple updates were implemented for the 2012 CFS. Changes to the survey were divided into three categories:

1. Internal changes were developed and executed in-house by BTS and Census staff. These changes, or data edits, supported an effort to improve data quality for the 2012 CFS. A sample of these includes:
    - Improvement to the mileage routing software.
    - Revision of the parcel definition to align with industry standards. Parcel shipments are now defined at 150 lb or less.

- Development of edit checks that flag questionable commodity–mode pairs, incompatible commodity–North American Industry Classification (NAICS) pairs, truck shipments traveling excessive distances for certain commodities, etc.

2. Questionnaire and *Standard Classification of Transported Goods* (SCTG) *Manual* changes were largely based on feedback received from data users and other CFS stakeholders. Measures to improve the questionnaire and SCTG Manual included:

- The addition of a temperature control question and rush-delivery question to the 2012 questionnaire.
- Removal of the third-party logistics (3PL) and intermodal questions from the questionnaire. Results had shown that those completing the questionnaire typically lacked the knowledge to answer these types of questions.
- Updates to the SCTG manual included new coding schemes for certain commodities, inclusion of ethanol and biodiesel, and more descriptive text and examples for easier use of the manual.

3. The main external change to the survey was related to geography. The number of counties that comprise a CFS metropolitan area is dependent upon the census definition (size) for that metropolitan area which results in the potential for the size of census metropolitan areas to fluctuate based on the decennial census.

The discussion closed with participants identifying challenges to making changes to the CFS and the traditional mail-out, mail-back survey. Going forward, other methods to collect freight data, during and between survey years, will likely need to be examined.

**Scot Dahl**
*U.S. Census Bureau*

Scot Dahl of the U.S. Census Bureau spoke about the 2012 CFS PUMS focusing on three areas of the PUMS file: contents of the PUMS, accessing PUMS and supporting documents, and limitations with the microdata.

1. Contents of the 2012 CFS PUMS file include:

- More than 4.5 million shipments, reported by approximately 60,000 establishments;
- Establishment data such as location of shipping establishment and NAICS code;
- Shipment data such as shipment ID, quarter in which commodity was shipped, shipment value and weight, SCTG code, temperature control status, hazardous material indicator, shipment destination, transportation mode, distance shipped in miles, export indicator, and export country indicator; and
- A tabulation weighting factor used to scale up the PUMS data to produce estimates for the total population.

2. The 2012 CFS PUMS file is accessible from the Census Bureau's CFS website. Data users have the option to download the file as a statistical analysis software (SAS) dataset or a comma-separated values (CSV) file. Accompanying the data file on the website is a PUMS User's Guide and Data Dictionary. The supporting documents help data users with using the microdata and in distinguishing differences between the PUMS file tabulations and estimates from the published CFS tables.

3. The primary limitation associated with the PUMS is confidentiality. Because of the unique nature of microdata files, actions were taken to protect the confidentiality of data in the PUMS file including:

- Extremely large shipment values and shipment weights were top-coded based on commodity;
- Extremely large tabulation weighting factors were reset to 975,000;
- Multiplicative noise was applied to shipment value and shipment weight;
- Shipment value, shipment weight, and distance quantities were rounded to the nearest integer; and
- For rare combinations of origin geography, commodity, and mode, the level of detail shown was reduced or suppressed. For example, under a reduced origin geography detail scenario, a shipment record with an origin in Chicago (CFS metropolitan area) would be reduced to an origin in Illinois (state level) instead.

Observations made by individual workshop participants included an interest in more geographic detail below the CFS metropolitan area such as at the county or zip code level and for the addition of more variables and descriptors, such as employment figures.

**Larry McKeown**
*Statistics Canada*

Larry McKeown of Statistics Canada spoke about Canada's efforts in developing their own CFS and FAF. Like the existing CFS, the Canadian CFS would be a joint venture between Statistics Canada and Transport Canada. Transportation policies in Canada have driven efforts to track freight movement. Such transportation initiatives and concerns included tracking gateway and corridor freight flows, having an ability to assess network capacity to move freight, environmental concerns, and more security and safety awareness.

To gather information to support these decisions, Statistics Canada and Transport Canada intend to create a Canadian FAF. The framework data dimensions would be comprised of geography data: 76 economic regions (ER) or 45 urban areas (census metropolitan area) and rest-of-province, commodity data, and modal and routing data.

Statistics Canada and Transport Canada are currently determining the best method to collect the data, evaluating whether to use a carrier-based approach or shipper-based approach. Transport Canada currently collects carrier data which gives insight into trucking commodity origin–destination (O-D) pairs, monthly rail waybill data, and marine data. A shipper-based approach, like the existing U.S. CFS, can provide comprehensive and integrated freight flows from origin to destination while providing specific shipment information.

During 2015, Statistics Canada tested concepts of the survey on potential CFS establishments. Their questions targeted the following issues:

- Willingness of establishments to provide data,
- Ability of shippers to provide the requested data,
- Overall comprehension about the purpose of the survey,
- Ability to obtain shipment records and enter information correctly on the shipment section of the questionnaire, and
- Understanding of the commodity coding scheme (SCTG).

From the testing, Statistics Canada identified the following:

- Export port is unknown for shipments by parcel delivery or freight forwarding company.
- Lack of ability to target 3PLs and distribution centers in the NAICS.
- Lack of clarity about what data to use within shipment records. Could empty shipment container movements be captured and is the information worth capturing?
- Establishments did not want to sample their shipments. They would rather provide all shipment information or nothing at all (i.e., they prefer to provide all their shipments for a reporting week versus providing every 129th shipment for that reporting week).

From ongoing work with developing a Canadian CFS, examples of outstanding issues include:

- Whether to track extractive industry movements (forestry, agriculture, crude oil, etc.) currently not sampled in the U.S. CFS.
- Whether to include additional methods to track routing of shipments. Use of vehicle device data versus modeling expected routes.
- How to address commodity classification: SCTG or North American Product Classification System (NAPCS)?
- Electronic data collection capability. Can varying formats for electronic shipment data from respondents be uploaded and processed correctly?

The primary challenge that Canada faces is one of time versus modernization. To have the Canadian CFS in the field for 2017, Statistics Canada and Transport Canada will potentially need to replicate the current U.S. CFS. In doing so, they will be limited to the existing methods and data collection practices without the opportunity to modernize the survey to meet current and expected future demands of CFS stakeholders.

## POST-PRESENTATION DISCUSSION

Following the presentations, workshop participants had the opportunity to ask questions and provide suggestions for possible improvements to the survey. Some of the questions and suggestions raised included:

- In considering traditional CFS and Big Data CFS, are there options for new data collection methods? Can current tracking systems be used to help with shipment routing and as a source of additional CFS data?
- Can CFS tap into vehicle devices being used for safety purposes? Can the information collected from these devices provide more than safety information? If so, can the CFS make use of the non-safety data collected?
- Would data users prefer that the SCTG coding scheme continue as is or could it be modified to match with NAICS?
- Could a 2007 CFS PUMS file be created? What would data users like to see in a 2017 CFS PUMS file?

- Could application programming interface (API) tools be included for CFS data manipulation?
- Where could CFS developers focus efforts to resolve problems with the survey? As a data problem or knowledge problem? If a data problem, it was suggested to collect more data. If framed as a knowledge problem, it was suggested to make use of models to infer data gaps between survey years. Can Census and BTS focus on more than data problems to lessen the knowledge problem or is the mission of Census and BTS to focus solely on data collection?

*Joy Sharp, Bureau of Transportation Statistics, presided over this workshop.*

# Emerging Trends and the Future of the Commodity Flow Survey

**JANINE BONNER**
*Bureau Transportation Statistics, recorder*

A panel consisting of Cavan Capps from the U.S. Census Bureau, Jose Holguin-Veras from Rensselaer Polytechnic Institute, and Kenneth Allen from H-E-B (retired) each spoke on the topic of Emerging Trends and the Future of the CFS.

## REPORT ON MIT WORKSHOP: USING NEW FORMS OF INFORMATION FOR OFFICIAL ECONOMIC STATISTICS

**Cavan Capps**
*U.S. Census Bureau*

Cavan Capps provided information from the perspective of the data collector in a world with big data. He highlighted discussion from a recent big data workshop at Massachusetts Institute of Technology (MIT) that included representatives from the Census Bureau, BTS, and MIT. Goals of the workshop were to discuss how the uses of big data can improve data timeliness, geographic detail, and product details, as well as identify other data for possible linkage to increase the informational power of the data. With increasing pressure to get the most current data released quickly and the ability to take advantage of private industry electronic transaction standards, he urged that stakeholders think outside the box, explore new possibilities, and consider a new perspective on how data can be collected beyond the traditional paper survey.

Capps described some challenges faced by the CFS in order to highlight the need for new methods of data collection, including the time it takes to process the data, the lack of data links to foreign trade data, and the lack of data sources that detail what products are inside of known shipments.

The opportunity for harnessing big data is available now that companies electronically store operational shipping data, including product shipment information for orders, sales, and GPS tracking. Capps stated that those analyzing the CFS data would profit from a better understanding of business logistics and individual shipping operations to better understand how to request company data and to better understand a company's value proposition; i.e. what incentives are the most appealing to companies to encourage participation.

Several participants mentioned the potential benefits of having representatives from the business community and shippers join the discussion to hear their perspectives. That workshop will also be used to investigate issues of data bias, such as self-selection bias, passive economic data bias and other systematic biases.

A CFS workshop attendee commented that, based on their experience with this type of activity, ATRI would be a good participant for the next workshop to connect with businesses to better understand individual shipping perspectives and communicate the government confidentiality laws that are in place to protect businesses.

## ESTIMATION OF NATIONWIDE FREIGHT DEMAND MODELS WITH THE COMMODITY FLOW SURVEY MICRODATA: EXPERIENCES, CHALLENGES, AND POTENTIAL

**José Holguín-Veras**
*Rensselaer Polytechnic Institute*

José Holguín-Veras presented ideas on how to strengthen the CFS through the use of estimation models and knowledge development. He highlighted what he perceives as weaknesses in the CFS; primarily the lack of route, mode, and logistics data, as well as the lack of geographic detail due to confidentiality requirements. He argues that the CFS issues are not merely data issues, but more knowledge or information problems, and that the combination of knowledge of the data combined with appropriate models yields the desired data. He stated that collecting more data will not solve the current problems and that, additionally, that approach is too expensive, time consuming and will not increase knowledge. As a counterpoint, he proposed ways in which the data that is currently captured can be modeled to develop the desired data.

He proposed that use of time-dependent effects are important in models since they can capture systematic freight demand changes. He believes that a redesign of the CFS using a cross-sectional design would result in more data by using all historical CFS data in combination with current year data to increase record counts. This method, which is currently available, makes use of big data techniques and should be used. In addition, he encouraged use of shipper attributes in the model, broadening geographic coverage to assess spatial differences, and use of all obtainable data.

Comments from the audience included:

- From the current CFS, there is not enough information for urban areas and estimates could be obtained from the ideas presented during Holguín-Veras' presentation.
- More modeling could fill in data gaps, and universities are good resources for supporting this.
- An important message from the presentation is to identify ways to connect data sources to be more useful, and that enhanced modeling using secondary data is useful for filling data gaps.

## MAJOR SUPPLY CHAIN TRENDS WITH IMPLICATIONS FOR THE COMMODITY FLOW SURVEY

**Kenneth Allen**
*H-E-B Stores (retired)*

Kenneth Allen gave a unique perspective of the logistical operations of a retail business and shared his thoughts on the evolution of the supply chain system. As a retired senior vice president of a major regional grocery chain, he was able to detail the supply chain and logistics of H-E-B and their competitors. H-E-B specializes in carrying the maximum amount of product variety within and between stores, as they offer different in-store products dependent on the community served by each store. Their philosophy is to ensure the highest quality at the lowest prices, and to do this, they purchase full truckloads from their vendors. To maximize quality and minimize store-based storage costs, they warehouse products and provide multiple daily deliveries to each store.

In contrast, Costco works on a very different model, where they offer a very wide variety of types of products but very limited selection within each type; toothpaste, televisions, tires, and food are all sold in the store but only one or two brands of each product are available. They serve a much narrower demographic of people and their warehouses offer most products in bulk. Amazon, as another model, offers over 200 million different products, of which 20 million are available for delivery within 2 days, and 1 million items within 2 h. Amazon has sales of $110 billion annually, which is remarkably different than a traditional supply chain. Since they have no stores, all items are shipped directly from warehouses, often one at a time, which affects transportation network usage in ways that no traditional stores do. H-E-B has made a substantial investment to develop an Amazon-like experience for their customers and expects the future of shopping to be radically different from what it is now.

In response to questions from workshop participants, Allen also shared the following information:

- H-E-B stores 80% to 82% of their items in a warehouse and 18% to 20% directly in the store. As a result, it estimates 95 million truck delivery-miles per year.
- It is possible for H-E-B to track the quantity that they ship from zip code to zip code since all mixed shipments are coded. However, the data are hard to extract through their current systems.
- H-E-B shipment data could be available at the household level if the household provides an address.
- Amazon would be considered a shipper within a CFS survey.
- Large e-commerce companies like Amazon are becoming a challenge for the CFS because only 40 shipments (observations) weekly are being requested for reporting, which will never measure the geography or diversity of shipments from a company like Amazon.

*Michael Meyer, WSP–Parsons Brinckerhoff, presided over this workshop.*

# User Breakout Discussions

Following the formal sessions, workshop participants were divided into three breakout sessions and were tasked with providing ideas for improvements for the future conduct of the CFS. Each group was asked to respond to four questions based on their own experiences and information from the earlier sessions.

- What revisions or updates may be needed to the CFS structure (e.g., classification systems)?
- What revisions or updates may be needed to the CFS scope (e.g., commodities, geographies)?
- What revisions, updates or additions to CFS sampling methods are needed?
- How can CFS tools and functionality be improved for current and future uses?

Summaries of each breakout session are presented in the following sections.

## BREAKOUT DISCUSSION 1

**Monique Stinson**
*Massachusetts Institute of Technology, recorder*

### Introduction

This breakout session included federal agency officials, lead CFS managers and developers, individuals from other agencies, users of the data, and those from academia. The discussion covered four topics: classification systems, the scope of the CFS, survey methods, and tools related to the functionality of using the data. The discussion of classification systems received the most discussion in terms of time, but this was partly due to information gaps regarding new classification systems and their relation to the CFS SCTG system. Key themes emerging from the session included:

- A need to better understand the NAPCS and its relationship to both the CFS and the SCTG codes.
- Identifying the best ways to modify the CFS scope, with some specific ideas regarding inclusion of import-related questions and future consideration of the sampling frame.
- Exploring the feasibility of "big data" as an alternate way for larger establishments to provide shipment records.
- Specific ideas regarding potential tools (such as web-based interactive tools) to enhance the user experience.

These and other discussion elements are discussed further below.

**Commodity Flow Survey Structure**

The NAPCS is gaining more widespread use. For example, it will be used in the 2017 Economic Census. Participants discussed whether NAPCS could be used in the upcoming CFS. A number of participants expressed concern with this approach, pointing out that NAPCS is not necessarily as suitable for transportation as the SCTG codes currently in use. For example, waste is a category in SCTG but not in NAPCS. Ultimately, if NAPCS is used in the future, a variant of it will likely be needed for it to be suitable for transportation analysis.

Correspondence between NAPCS and SCTG was flagged as an issue that warrants more detailed attention in the long run. There is currently no one-to-one correspondence between the two systems. This may create issues when users of the data conduct analyses based on data from these two different systems. Some felt this was a minor issue given historical CFS processes. However, the counterpoint was made that since commodity flows are a form of economic and trade data, consideration could be given in the future to decide whether to somehow combine Census economic data with SCTG-generated figures.

Similarly, one participant suggested that if the use of SCTG continues, it could be reviewed to account for changes in the economy since the last review. For example, SCTG Class 36 may be too broad. Revisions that will accord well with NAPCS could be considered when this review is conducted. However, another participant noted that classification issues between the CFS SCTG codes and NAPCS could not be addressed in detail for the upcoming CFS, as the survey documents would be finalized in the near future.

Many participants indicated that much more information is needed about the status of the upcoming NAPCS–Economic Census. A possible idea was to conduct a more formal study on NAPCS versus SCTG to develop a better understanding and to have the opportunity to provide input to other teams and agencies that are deciding what to use for their own systems.

**Commodity Flow Survey Scope**

The discussion on CFS scope focused on two areas. First, the participants discussed whether a more comprehensive range of industries could be covered in the sampling frame. Second, participants discussed potential ways to use unused questionnaire space that had been devoted to questions that are no longer in use.

The CFS focuses primarily on manufacturers and wholesalers. Due to the sampling frame, important components of freight movement are not captured in the survey. These include, for example, movements from farms to grain elevators, from oil fields to refineries, and household and business moves. It was pointed out that these flows comprise a significant amount of total flows and are currently estimated through post-processing when the FAF, a CFS-based data product, is developed. Some participants also noted that including these originating businesses would require modifications to the sampling process. Further, other sources are available such as U.S. Department of Agriculture and U.S. Department of Energy reports.

Participants discussed several options for using available space on the questionnaire. Ideas included capturing:

- Domestic portion of waterborne shipments.
- U.S. segments of import movements (e.g., from the Port of Los Angeles to final destination in United States). This may be obtained from port owners and operators. Port Import–

Export Reporting Service (PIERS) data may be used for at least part of this. However, it is not currently clear whether missing segments between ports, terminal operators, and PIERS can be identified.

- General information regarding import activities for individual firms. An example question that was posed was "What is the purchase price of imports versus domestic goods by your firm in the last 3 years?"

Including ports in the sampling frame was suggested as a means to collect import data.

Participants also discussed possibly using the cognitive field testing to explore the feasibility of collecting imports or other information.

**Commodity Flow Survey Sampling Methods**

Participants primarily discussed the potential to collect big data from certain firms. For example, large firms may provide all records for 1 week rather than a sample. This could help address issues related to using the same maximum number of sampled shipments for all businesses regardless of size which is the current process for the third stage of the sampling process. Gathering more data from bigger shippers would also help capture more representative moves by large businesses.

**Commodity Flow Survey Tools and Functionality**

Some participants mentioned several tools and methods that may have the potential to enhance functionality:

- Adding API capability;
- Providing an ability to select the suppression level, i.e., suppressing geography or commodity;
- Incorporating an NTS-style add-on program with a note that this would only include inbound information; and
- Engaging in multistate data collection, similar to NCHRP, to obtain outbound data from other states' inbound data.

Although current users seem comfortable working with the raw data, other individual participants noted an interest in developing interactive regional freight data applications similar to those sponsored by FHWA in the SHRP 2-C20 program.

Several aspects of CFS data dissemination were discussed. These included:

- Expanding engagement with the private sector, state DOTs, and consultants to communicate the value of the CFS for providing responses to the questionnaire and use of the data.
- Expanding engagement with the private sector, state DOTs, and consultants to identify improvements to make the data more useful and valuable.
- Using an app or API that accesses data in the cloud. A suggestion was made to adopt the process used by some transit agencies of make their data available and to allow interested users to create their own APIs.

- Using off-the-shelf visualization tools such as Tableau and FlowMapper.

Two critical factors related to the CFS were reiterated: (1) confidentiality of companies must remain protected and (2) mechanisms are required to make sure that users are aware that the sampling process can impact results and that results may be less accurate due to that sampling process as the data are parsed. Regarding the latter, a suggestion was made to offer ways to collect an expanded sample for municipalities that are interested. For example, guidance could be provided to other agencies on how to conduct this type of survey or an add-on sample process similar to the National Household Travel Survey.

During the discussion, a one participant made the suggestion of better engaging private-sector firms, state governments, and other entities to learn more about what is going on with freight movements. Another suggestion was to include a marketing component in the November 2015 outreach in hopes of engaging the private sector more in the CFS planning effort.

*Scott Drumm, Port of Portland, presided over this discussion.*

## BREAKOUT DISCUSSION 2

**Julie Parker**
*Bureau of Transportation Statistics, recorder*

This group consisted mostly of participants from the public sector. Topics of greatest interest to the group were disaggregated data, big data, and more detailed state-level data. The FAF was discussed somewhat interchangeably with the CFS in regard to improvements and which of the two is more appropriate depending on the research question. Items for discussion in this session included

- Training materials for users;
- Enhanced promotion;
- Geography–commodity tradeoffs;
- Data gaps in rural states; and
- Import and inbound shipment information.

These and other discussions are summarized in the following sections.

### Commodity Flow Survey Structure

Little input was provided on this topic and attendees expressed overall satisfaction with the current CFS structure. Some participants thought that having a crosswalk to NAICS classification could be helpful.

### Commodity Flow Survey Scope

Many participants expressed interest in more disaggregated geography at various levels including establishment, zip code, county, and metropolitan statistical area (MSA). With more disaggregated geography, travel demand models created by metropolitan planning organizations (MPOs) could be improved. Participants discussed having the ability to generate a state-level PUMS file particularly for rural states.

Participants expressed a desire for obtaining modal splits as a method to help understand the delivery chain.

Various participants expressed contentment with the ethanol issue from the last CFS and that the temperature control question should be maintained.

In comparing the FAF to the CFS, some participants indicated that the datasets do not agree in regard to outbound shipments. This led to a discussion on receivers and how that information could be gathered by adding a few simple questions. Having information about outbound shipments could help calibrate dataset models and other samples. Further, knowing if the shipment was imported could also be helpful in modeling.

Some discussion was directed at the volatility of certain commodities and the difficulty in capturing this over time with a 5-year CFS. Timelier data would be beneficial although potentially not through a continuous survey. A possible option was proposed for a special study that looks at commodity volatility and considers establishments that could easily provide data. This would likely be a separate effort.

**Commodity Flow Survey Sampling Methods**

Big data was discussed in the context of obtaining more information. Some participants thought establishments may be more willing and able to upload a block of data than to extract samples per the current the survey. This could potentially improve the results given that an analyst would process the data as opposed to the respondent attempting to meet the specifications. A few participants talked about various data cleaning and ways to retrieve required and desired CFS data items. Based on the earlier presentation by Kenneth Allen, participants discussed the possibility of a pilot program with one or two grocery chains to plug into their system and follow individual transactions. The goal would be to understand current business practices. Although the data volume and velocity are unknown, participants posited that the result could be a value-added product and that participation could be compensated in some way.

Participants discussed some of the challenges with information associated with geography. A few ideas from individual participants included:

- An indication about whether a shipment experienced an intermediate stop;
- An indication of whether a shipment was an import;
- An indication of whether establishments report drop shipments; and
- Merging or calibrating the CFS with foreign trade data to understand point-of-entry to domestic destination movements.

Interest was also expressed in evaluating the CFS related to other data sets as a data quality enhancement activity.

Participants were interested in obtaining the questions and comments from those filling out the survey about the survey itself. It was mentioned that these kinds of metadata about the survey would be helpful to the research community.

### Commodity Flow Survey Tools and Functionality

Participants discussed the possibility of having disaggregated summary data. Preference was towards county-level data but MSA-level was also discussed. Participants also explored the idea of tradeoffs between geography and other characteristics.

As noted earlier, participants were interested in a state-level release of the PUMS. Another idea was to provide support for states and MPOs to implement their own CFS using a local survey document that was appropriately vetted.

A few participants discussed getting seasonality coefficients to provide insight into what goods were moving when.

Promotion of and training for the CFS was discussed. Participants were interested in the possibility of having a webinar or video on how to use the data. Another idea was to provide an opt-in list-serve that would notify users about new items and changes to CFS tools and products.

*Bruce Lambert, Institute for Trade and Transportation, presided over this discussion.*

## BREAKOUT DISCUSSION 3

**Matthew Roorda**
*University of Toronto, recorder*

### Introduction

This summary outlines the comments and suggestions raised by 27 participants who came from all levels of government, an association, the private sector, and academe.

### Commodity Flow Survey Structure

The discussion about the CFS structure recognized that several classification systems are used for a variety of different purposes. For example, classifications include NAICS for industry classification and SCTG for commodity classification. Classification systems can be represented at different levels of detail (e.g., NAICS two-digit to six-digit levels of disaggregation, or the SCTG two digit to five digit) and are occasionally updated to reflect changes in the nature of commodities or industries. To integrate information from the CFS with other data products, some participants suggested that "data crosswalks" are a necessary supplementary product for appropriate and consistent use of the CFS data and that crosswalks could to be developed and updated to reflect any adjustments in the classification systems used in the CFS. Crosswalks between classification systems at different levels of disaggregation (e.g., two-digit, three-digit, or six-digit) could be investigated and consistency between the different levels of disaggregation was considered to be important.

It was also suggested that a classification system could be designed to be consistent with economic data. While it was noted that the NAPCS is designed for such consistency, the SCTG classification has been used in the CFS because it breaks down the attributes of commodities in terms of their physical characteristics (e.g., dry versus liquid goods), which are important for understanding the movement of freight.

It was acknowledged that most shippers are also receivers, yet the CFS only collects information about outbound shipments. Two possibilities were raised to potentially enhance understanding of supply chains, and to provide better information for the estimation of mode choice models. The first was to ask the shipper questions about inbound shipments. The second was to ask the shipper for information about the receiver of the goods they ship to identify the industry of the recipient.

**Commodity Flow Survey Scope**

The lack of representation of oil in the CFS was noted as an important omission. Oil transport has not been included in previous CFS implementations because (*a*) the transport of oil, which is primarily by pipeline, is incompatible with the CFS collection of information for discrete shipments, and (*b*) other data sources are available to provide information about oil transport. However, rail transportation (including mode transfers from trucks) have more recently become a competitive mode of transportation for oil, particularly the movement of shale oil by truck and rail from states such as North Dakota and Pennsylvania.

Geography was also discussed. It was noted that megaregions are increasingly becoming an important policy focus. However, those that include portions of the "rest of state" CFS zones, may be challenging to assess given the current CFS sampling methodology. To address this, it was suggested that CFS tabulations could be improved by distinguishing between establishments that are in urban versus rural areas and that increasing sampling of smaller establishments, which are assumed to be more heavily represented in urban areas, could improve the representation of freight activity within megaregions.

The potential to add questions about the expected delivery or travel time of the shipment was also raised. Such variables could address the question of whether the freight system is meeting shippers' expectation or desire for delivery times. Challenges with such a question were identified, including whether the shipper would be able to answer such a question. It was also noted that a question tested on the most recent survey related to time constraints for deliveries was not considered to be very successful.

Attitudinal questions were also raised for potential inclusion in the CFS. It was not clear whether a need for these existed or what the appropriate attitudinal questions would be, but some participants indicated that modifications to the CFS to enhance understanding of shipper behavior would be desirable.

**Commodity Flow Survey Sampling Methods**

One participant suggested that the CFS could take advantage of the potential for using models (e.g., freight generation models) to represent industries that have stable behavior for goods production. By representing such stable industries with modeled behavior, CFS resources could potentially be more effectively applied to more rapidly changing industries. The point was made that the most efficient allocation of CFS resources would involve a combination of freight generation or trip generation models and data collection

Some individual participants raised concern about underrepresentation of the following types of establishments and shipments:

- Smaller establishments in urban areas.

- Establishments in rural areas. In particular, movements of agricultural products from farms to the first stop are not covered in the CFS. It was noted that CFS depends on the U.S. Department of Agriculture to provide information about these shipments. Also, municipal solid waste is generally destined for rural areas and such movements are not included in the CFS.
- Very large establishments, such as Amazon distribution centers. The CFS sampling approach, which requests 40 shipments per establishment, is not a good representation of large distribution centers, which generate an extraordinarily high number of shipments per day. Instead of asking for a fixed number of shipments per establishment, it was suggested to explore "differential sampling" (increased sample of shipments for companies with a larger number of shipments). The potential to obtain a full set of all shipments (i.e., a data dump) for large companies could be explored (and may in fact be easier for responding establishments to provide). It was noted that differential sampling was already being considered for the 2017 CFS.
- Shipments with higher economic value could be considered for oversampling.
- Shipments associated with reverse logistics. It was noted that reverse logistics (returns of unwanted items) were generally not recorded in the CFS. Given that e-commerce, Internet shopping, and home delivery has the potential to greatly increase the number of returns, it is currently an area that could be explored.
- Other elements out of scope include moving furniture, pallets, etc., within the same firm, without a sale.

Finally, it was suggested that feedback could be incorporated into the sampling process. For instance, if some shipment types are found in one CFS survey to have attributes that are highly variable (e.g., value-to-weight ratios), then the next CFS survey could use this information to design a more efficient sampling strategy such as high variability shipment types being oversampled.

**Commodity Flow Survey Tools and Functionality**

Participants applauded the addition of the PUMS file. A suggestion was made that the value of the microdata file could be improved by linking shipment information with information about the shippers and, potentially, also some information about the receivers. An option would be to include a shipper ID number so that multiple shipments from the same shipper in the PUMS file could be linked.

The methods for accessing the CFS data include American Fact Finder, the published CFS tables, and the PUMS files. The question was raised whether people were using the data in all of these forms. Several people in the breakout session acknowledged that they had used either the American Fact Finder or the predefined CFS tables. One participant suggested that the location of file formats other than pdf for CFS tabular information be more obvious. The handling of suppressed information in American Fact Finder could also be improved. Reducing time between surveys was also suggested by another participant.

*Michael Meyer, WSP–Parsons Brinckerhoff, presided over this discussion.*

# Concluding Remarks

**ALISON CONWAY**
*City College of New York*

**KIMBERLY P. MOORE**
*U.S. Census Bureau*

**ROLF SCHMITT**
*Bureau of Transportation Statistics*

## COMMODITY FLOW SURVEY: OPPORTUNITIES AND CHALLENGES

**Alison Conway**

Even in the emerging world of "big" freight data, the CFS remains a critical source of information for strategic planning of and decision making for the nation's freight transportation system. While new data sources (e.g., near-real time GPS) that are useful for understanding some aspects of freight transportation system performance have rapidly emerged, none address the key contribution of the CFS: to relate industry-specific shipment values to the multimodal flows moving on the U.S. freight infrastructure.

### The 2012 Commodity Flow Survey

A number of key improvements were implemented for the 2012 CFS to improve the survey instrument, the data quality, and the usefulness of the data products for researchers. Problematic questions regarding 3PLs and intermodal movements were removed from the questionnaire, and replaced with questions on temperature control and rush delivery. BTS updated its internal methods for data processing, including improving routing software and developing new checks to identify infeasible shipment combinations. Updates were also made to the SCTG manual, both to improve text and visuals and to introduce new coding schemes for specific commodities.

The most notable change for the 2012 CFS was the introduction of a public use micro-dataset (PUMS), which includes more than 4.5 million shipment records from about 60,000 establishments. To make these data available and useful to the public while maintaining the confidentiality of survey respondents, a number of processing steps were required. Weighting factors were developed to scale up the PUMS to represent the full population. Data were top coded and noise was added to protect shipper confidentiality, and for some combinations of origin geography, commodity, and mode, data details were suppressed. The PUMS is now available for download as a SAS or CSV file.

### Future Improvements to Consider

While these changes for 2012 have been well received by the user community, gaps and further opportunities for improvement remain. Additionally, logistics in many sectors have shifted dramatically in recent years due to changes in manufacturing processes, warehousing and vehicle technologies, retail models, and consumer shopping choices. These changes pose new challenges for conducting the CFS. Many specific concerns and potential improvements were discussed by attendees for the 2017 survey and beyond.

**Survey Instrument**

The CFS remains a traditional mail-out survey that is somewhat burdensome for a shipper to complete. As many shippers are using electronic record-keeping systems, there may be opportunities moving forward to work with the private sector to access the large volume of data these systems produce. Evidence from Canada suggests that shippers would be more willing to share full datasets than to extract a sample as required for the current CFS. However, improving public access to private data will potentially require development of an adequate value proposition, especially as shippers recognize and monetize the value of this information.

**Shippers and Commodities**

Some attendees pointed out that there were a number of specific establishment and commodity types that received limited or no coverage in the 2012 CFS. Shipment types and commodities identified for consideration included household and business moves, municipal solid waste, reverse logistics, rural movements including farms to grain elevators and oil fields to refineries, and high value goods. Other attendees also noted that while oil has traditionally not been included in the CFS due to difficulty in defining pipeline shipments, recent growth in the use of rail, particularly for shale oil movements, may warrant revisiting its inclusion. Specific establishments included small businesses in urban areas as well as very large, diverse shippers like Amazon, for whom the current sample size of 40 shipments is not representative. Ideas for improving sampling methods included oversampling, differential sampling for businesses of different sizes, development of guidance for states and local agencies to conduct their own additional sampling, and the application of models to estimate shipments in stable industries so that scarce resources can be used to measure more volatile sectors.

**New Data Elements**

A number of attendees noted that the survey currently does not include supply chain or performance data. While the current survey is limited to outbound shipments, many shippers are also receivers. Some attendees suggested two potential approaches to collect information about upstream and downstream flows: asking the shipper questions about inbound shipments and asking the shipper for information about the receiver of the goods they ship. In addition, a few attendees additionally noted the lack of links to foreign trade data, general information regarding import activities, domestic portions of waterborne shipments, and U.S. segments of import movements. Including ports in the sampling frame could potentially address some of these gaps. To assess system performance, attendees offered suggestions to include additional questions related to expected delivery or travel time and attitudinal questions to understand shipper behavior.

### Classification Systems

Several attendees expressed concern about the relationship of SCTG codes currently used in the CFS with both the NAICS and NAPCS classification systems. One possible approach that was mentioned was to develop "data crosswalks" as a supplementary data product to enable a more seamless linkage between CFS data and records using other systems.

### Data Products

While many attendees recognized the PUMS as a major enhancement to CFS data products, a number also noted that the suppression of submetropolitan geographies limits the value of the data for local analysis. Another participant noted that the PUMS does not include any shipper characteristics, which are useful for modeling demand. One suggestion was to include the ability to select the suppression variable (e.g., geography versus commodity). Availability of county or zip code level data could also address two additional geography-related challenges identified: difficulty observing changes over time in metro areas whose boundaries change with each decennial census and difficulty assessing megaregions that might include portions of areas currently classified as "rest of state." A few attendees also suggested improving accessibility to non-pdf tables and employing APIs and off-the-shelf visualization tools.

### Survey Frequency

One more consideration that was identified by multiple speakers and attendees, particularly in the big data era, is the 5-year time between surveys. As noted above, the pace of change in logistics and supporting technologies is rapid, meaning that data may be outdated by the time it is processed and made available for public consumption. Some potential approaches offered to address this challenge included more frequent cross sectional surveying and development of models recognizing time dependent effects that could be employed in near real-time to measure changes.

## LOOKING AHEAD: THE 2017 COMMODITY FLOW SURVEY

### Kimberly P. Moore

From the perspective of the Census Bureau, this workshop has provided important information both to those who work on the survey and to those who use the resulting data and products. This workshop provided us with insight into how it is used, what is important to users, and what changes could make it more useful. As noted earlier, the CFS needs to continue to evolve and this workshop plays an important role in setting that direction.

The 2017 CFS is already underway. Planning started well before the PUMS file was released in the summer of 2015. The pre-canvas to identify shippers of in-scope NAICS codes will be 100% Internet based and begins in January 2016. The results will be used to identify and determine who is eligible to be included in the 2017 CFS, with that data collection starting in January 2017.

To meet those dates and to keep the CFS on schedule, the questionnaire will be finalized in early 2016 which is just a few months away. Staff will be in the field doing cognitive testing including testing questions in the questionnaire, so evaluating ideas for improvements related to data collection will be performed quickly.

The CFS must evolve to meet the needs of data users. Census and BTS continue to work together to accomplish that goal. Input from this workshop included ideas for changes in potential scope, classification, sampling methods and data dissemination as well as consideration of alternative means of data collection given the continuing challenge of maintaining an acceptable response rate. This also includes identifying ways to reduce the burden on responders.

Working with BTS, the Census Bureau will continue to improve the CFS not only in 2017 but beyond.

## FOUNDATION OF FREIGHT MOVEMENT KNOWLEDGE

**Rolf Schmitt**

Our understanding of freight movement has been founded primarily on the CFS since the 1990s, and on the predecessor of the CFS (the Commodity Transportation Survey) in the 1970s. This workshop is held every 5 years to get user input on how the CFS should continue to evolve and improve that understanding. Hopefully, the workshop 5 years from now will cover the CFS and the Canadian version of the CFS.

The CFS is the foundation of our knowledge about freight movement over the transportation system, about the economic geography of the nation (links among places and not just among industries), and provides the context for transportation performance measures. The CFS covers a broad spectrum of freight transportation, including the very different worlds of bulk and high velocity high value commodities. The world looks different from shipper and carrier perspectives, such as whether or not to include container weight in shipment tonnage. The CFS measures both bulk and high velocity from the shipper side to go with various measures that we have from the carrier side, all of which ends up in the FAF.

Much has been said in this workshop about the importance of modeling, which is underscored by our dependence on models to assign O-D flows to specific parts of the transportation network and to disaggregate FAF data to smaller geographic units. BTS is increasingly involved in FAF because it meets our mandate to create an Intermodal Transportation Database. We expand the CFS into current-year estimates, forecasts, and complete coverage through the FAF, and are looking to improve the linkages between FAF and CFS so that users can drill down from FAF tables to the CFS for greater detail on pieces of the FAF.

Our basic freight monitoring strategy has been to create benchmarks and update with supplemental data. This is the same approach that the Bureau of Economic Analysis uses to measure the structure of the economy. Should we be looking to other strategies to measure freight, such as continuous surveys or continuous indicators or something else? Or should we concentrate on using big data as the supplement to the CFS benchmarks for improved freshness and geographic and temporal detail. No one has said the day of the CFS is over, though clearly CFS is only part of a freight monitoring strategy.

I have been involved with this survey since my first paper on freight transportation that was based on the 1972 CFS. As a consequence, I am very happy to hear the general satisfaction with how far we have come. I look forward to the continued involvement of the freight data veterans and especially the new members of the freight data community in improving the CFS and our related programs. Thank you all for helping shed light on the fascinating world of freight.

# Applications of the Commodity Flow Survey

## USING COMMODITY FLOW DATA TO ASSESS THE ECONOMIC IMPACTS OF HIGHWAY INVESTMENTS ON FREIGHT-DEPENDENT INDUSTRIES

**Stephen Fitzroy, Paul Bingham, and Derek Cutler,** *EDR Group*

The ability to analyze and prioritize investments in projects designed to address freight transportation can be significantly improved by including information available in the CFS, the FAF, and other sources of data on specific commodities moving on the highway system. The key is to assess how improvements in network performance affect costs of freight transportation and then translate ways that these performance improvements affect costs faced by shippers and beneficial cargo owners. Since the ability to absorb, pass through, or adjust costs through supply chain and logistics adjustments varies by industry sector, analytic methods that account for the effects of freight system performance and industry-sector logistics practice must be addressed in any analysis that assesses the economic impacts of highway network investments.

This poster presentation demonstrated, practical ways that CFS-based data sources can be used with economic models to evaluate regional economic dependence on freight, and to further refine that linkage to show how transportation investments that enhance the performance of the freight network can benefit industries on a sector-by-sector basis. Using multiregion analysis and their respective differences in industry mix, we show how the improvements in resolving major chokepoints or bottlenecks can affect different parts of the same region or state. This can have important implications as to how the impacts accrue to a region and can affect decisions on how investments designed to reduce the costs of freight transportation are prioritized.

The sequence of steps involved in testing this assertion involves first identifying freight flows moving in relation to a region or corridor. Next the associated freight was profiled to identify a mix of goods being shipped. Then improvements in network performance that generate cost savings to businesses were assessed. These cost savings were then translated into economic impacts for specific industry sectors that affect their supply chain operations in a variety of ways, including measures of reliability, capital costs, perishability measures, and inventory management costs. By accounting for these factors, we show the process with which the vehicle-related cost savings affect industry sectors and regional economies. This approach also shows how important it is to understand the effects of commodity movements on certain industries, and how different corridors affect different industry sectors. This has important implications for decisions that link transportation system investment with economic development objectives (e.g., sustaining existing businesses, development of new or emerging industry clusters, or other competitiveness objectives of transportation investment policy.)

The poster session demonstrated a method for taking future expectations, trends, and forecasts of economic industry performance and building a case for quantifying their effects on active freight transportation. One side effect of this research was the realization of how important it is to stress the distinction between economic and freight flows, and how differences in both could be identified, characterized and potentially used to improve future models.

## OIL ON RAIL FLOW: A SEAMLESS ADDENDUM TO THE COMMODITY FLOW SURVEY

**Shih-Miao Chin and Ho-Ling Hwang,** *Oak Ridge National Laboratory*
**Jiaoli Chen,** *University of Tennessee*

Freight transportation is an integral part of our national economy and its growth is anticipated to be significant as the nation's economy expands in the future. Policy makers and transportation professionals are facing the challenges to maintain, manage, and plan for the existing and future freight transportation system. The CFS, cocompiled by U.S. Census Bureau and BTS, is a comprehensive freight data series that can be used to perform policy studies. However, the CFS does not cover the emerging crude oil transportation practice that has evolved from the shale oil exploration industry.

Advances in horizontal drilling and hydraulic fracturing have made shale oil and gas extraction profitable. High gasoline prices and shifting energy security policies have also contributed to the shale oil and gas exploration boom. The boom has induced the growth of shipping oil by rail train. Crude on rail traffic has increased dramatically since 2010. While shipping shale crude on rail to refineries, unintended and undesirable consequences do happen. Specifically, crude on rail operation information is a critical part in formulating remedial measures to counter the negative impacts induced by shale oil exploration and subsequent transportation logistics. Information on energy-generating commodity production origin, consumption destination and associated mode of transportation is also needed to formulate integrated holistic resolutions. There is a need to compile an oil–train flow database that can be used seamlessly with related energy commodity O-D and logistic information provided by the CFS.

This poster session illustrated the data requirement, information sources, and imputation methodologies to compile the oil-train operation information. Most of the data used in this research effort was obtained from official information published by Energy Information Administration (EIA), Census Bureau, and Surface Transportation Board. Due to strong resistance from railroad companies on sharing their oil train routing information to the public, the collection of oil train routing information is rather challenging. Although railroad companies were mandated by the Federal Railroad Administration to furnish routing information to state emergency management authorities for public safety purposes, two major railroad operators have gone to court in efforts to stop states from releasing routing information to the public. To combat this data gap, a significant portion of the findings presented in this poster was based on unofficial oil-train flow information found on the internet.

## USE OF THE COMMODITY FLOW SURVEY DATA TO ESTIMATE FREIGHT MODE CHOICE MODELS

**José Holguín-Veras and Shama Campbell,** *Rensselaer Polytechnic Institute*

This poster session provided a brief description, key findings, and a detailed summary of the process to be used in analyzing freight mode choice as a part of NCFRP Project 44. The objective of this project was to study current modal patterns, analyze the factors influencing mode choice of different freight agents, and to develop analytical methods to assess various policy implications on modal patterns. It is important for public policies to use various modes of transportation effectively to minimize energy consumption and externalities while fostering economic growth. To achieve this, it is crucial to have a better understanding of the variables and behavior of the freight agents that influence mode choice decisions. However, there is woeful lack of research in the field of freight mode choice. This study begins to fill this void by providing insights into freight mode choice to enable more accurate demand forecasts and better quantification of the impacts of freight activity.

The authors used the CFS 2007 microdata at the Census Research Data Center in New York to estimate freight mode choice models. CFS data, together with the shipment size, value, and other characteristics of the establishment, offer a unique opportunity to conduct freight mode choice analyses to the establishment level, and to the zip code level geographically. However, analyzing the modal pattern between all zip codes is overwhelming as there are more than 40,000 zip codes in the United States. To overcome this, the team obtained the potential zip–zip O-D (ZOD) pairs for the analysis using a random sampling procedure that gives an overall picture of freight flows in the United States. To select ZODs, firstly the state–state O-Ds are selected using the CFS data. Then, the team obtains ZOD pairs using the employment data of zip codes in each state since previous research found that employment is a good indicator of freight generation. This poster summarized the methodology followed and the results obtained in selecting the potential ZODs from the CFS 2007 data.

Once the ZODs were finalized, the team processed the CFS data for the modal share of each ZOD to get an overall picture of modal patterns across the United States. Comparing modal share with the establishment characteristics from the Census Bureau Business Register at each zip code, the mode choice models are estimated as a function of establishment-level variables. For example, the probability of choosing truck or rail for given shipment size, commodity, or industry segment is obtained. In addition, aggregate models at the state or country level are estimated that provide market shares of each mode as a function of decision variables such as cost, transit time, and reliability. These models will serve as a potential tool for the public sector to analyze policies and their influence on freight mode share.

## USE OF THE COMMODITY FLOW SURVEY DATA TO ESTIMATE FREIGHT GENERATION AND FREIGHT TRIP GENERATION MODELS

**José Holguín-Veras and Carlos Gonzalez-Calderon,** *Rensselaer Polytechnic Institute*

This poster session showed the research findings and advances in the estimation of freight generation (FG) (amount of cargo) and freight trip generation (FTG) (number of freight vehicle trips) models using the CFS data. This research is part of NCFRP Project 25. The objective of this project is to develop a handbook that provides improved freight trip-generation rates or equivalent metrics for different land use characteristics related to freight facilities and commercial operations to better inform state and local decision making.

FG and FTG have received increased attention by researchers and practitioners in the past few years. However, a lack of research and freight data still affects all facets of transportation modeling. In particular, there is a great need for research into the quantitative aspects of FG. A better understanding of the variables driving the generation of freight demand would enable more accurate demand forecasts, and better quantification of the traffic impacts of freight activity.

One of the unique characteristics about the work presented in the poster session is that this is the first time that FG models are developed using the CFS microdata. The microdata are the most important source of freight demand data in the country as it provides detailed data pertaining to movement of goods in the United States. This includes commodities shipped, their value and weight, mode of transportation, and shipment O-Ds, making the CFS microdata ideal for FG. The Longitudinal Business Database (LBD) provides measures of economic activity at the establishment level with attributes such as employment, revenue, NAICS code, and other business characteristics. The combination of CFS microdata and LBD provided a strong foundation to produce solid FG models which include both dependent and explanatory variables.

Various processes using econometric techniques were analyzed, including ordinary least squares (regression analyses). The team estimated FG models (constant, linear, nonlinear) for New York, California, Texas, Wyoming, Ohio, and the United States.

For the creation of FTG models, the authors used disaggregated data at the establishment level collected from receivers and carriers in New York City and Albany Capital Region in the state of New York. With the data, the researchers were able to produce models for both freight trip attraction and production. Additional FG (both attraction and production) and service trip attraction (STA) models were also estimated using regression analyses. After consideration of the explanatory variables, the number of employees per establishment was the main independent variable used. Conceptual validity, statistical significance, and the root mean square error were the criteria used to assess the suitability of the functional form to estimate FG, FTG, and STA.

## MODIFICATIONS TO THE METHODOLOGY AND CONTENT OF THE 2012 COMMODITY FLOW SURVEY

**Hossain Eftekhari-Sanjani and Ronald Duych,** *Bureau of Transportation Statistics*

This poster session highlighted major revisions included in the methodology of the 2012 CFS, and provided highlights of the data on temperature-controlled and expedited shipments captured in the latest survey.

The CFS is a joint effort by the BTS and the U.S. Census Bureau. This survey, which is conducted every 5 years as part of the Economic Census, is the primary source of national- and state-level data on the movement of goods, and is the only publicly available source of data for shipments by the highway mode. The CFS provides data on the type of commodity being shipped, along with the value, weight, mode(s) of transportation, origin and destination, hazardous material status, and the distance and ton-miles of each shipment. The CFS data covers approximately 100,000 establishments in mining, manufacturing, wholesale, auxiliaries, and selected retail and services trade industries located in the 50 states and the District of Columbia.

Compared to the prior rounds of the CFS, the 2012 CFS methodology included updates and modifications that have impacted the questionnaire content and survey methodology. Data users should be aware of these changes and exercise caution in analyzing results, especially when comparing to previous years' data. The CFS updates for the 2012 include, but are not limited to: providing an electronic option for respondents to report their data; expanding the number of geographic areas from 123 in 2007 to 132 in 2012; updating the mileage calculation program by adding new routing network and modal categories; developing new edit procedures to identify incompatible or unreasonable data; updating and creating new detailed codes in the SCTG to better capture fuels and the commodities used for biofuels (e.g., ethanol, and biodiesel) or to simplify identification of the commodities; and updating the boundaries of MSAs.

In addition to methodology changes, the 2012 CFS collected data and provided estimates on shipments requiring temperature-controlled transport. The survey also included an additional item, at the establishment level, inquiring about the percentage of the reported shipments that were sent using expedited shipping methods. Finally, for the first time in the CFS series, the published data included a 2012 CFS PUMS file.

## A NON-LIGHT-DUTY ENERGY AND GREENHOUSE GAS EMISSION IMPACTS ASSESSMENT TOOL DEVELOPED BASED ON THE COMMODITY FLOW SURVEY AND FREIGHT ANALYSIS FRAMEWORK

**Yan Zhou and Vyas Anant,** *Argonne National Laboratory*

Argonne National Laboratory developed a new energy and greenhouse gas (GHG) emission assessment tool, NEAT, which provides estimates of the potential end-use energy consumption, upstream energy consumption, and GHG emission impacts through 2050 for a base case and user-defined alternative case(s) related to five domestic freight modes and their use of alternative fuels. The five modes are: (1) intercity freight-carrying trucks, (2) freight rail, (3) domestic freight marine, (4) domestic freight aviation, and (5) pipeline. This analytical tool was developed to evaluate freight scenarios accounting for long-time full fuel-cycle energy and GHG emissions impacts resulting from changes in freight ton-miles, mode shares by commodity, modal energy intensity (EI) due to technological, regulatory, and operational reasons, and changes in fuel mix, such as more usage of natural gas or biofuels. It also assesses changes in petroleum consumption for given years in a defined scenario.

The tool is a spreadsheet model which is transparent, flexible and user-friendly. First, NEAT incorporates data from the FHWA's FAF projections and EIA's Annual Energy Outlook (AEO) projections up to 2040. Then projections are made to 2050 using growth rates derived from two projections where possible. The tool includes 36 commodity types which consist of 30 commodity combinations from the 43 FAF commodities and six additional energy-related commodities. Second, mode shares were developed for each commodity for five modes. The methodology allocates ton-miles by multiple modes and unknown or other modes to five known modes. Third, commodity-level energy intensities were developed which were not available in FAF and later adjusted so that total modal energy use matched with known data in the Transportation Energy Data Book published by Oak Ridge National Laboratory. Rail EI by commodity is from AAR's Railroad Facts and the CFS. Truck EI estimates were obtained from the 2002 Vehicle Inventory and Use Survey and 2002 CFS, update to 2010. Future EIs reflect improvements projected in the AEO. Fourth, full fuel cycle GHG emissions or upstream energy consumption for freight modes are from Argonne National Laboratory's GREET model. Users can specify their own scenario by providing additional input to one or more of five data items: (1) ton-mile change factors over 2010 values by commodity; (2) ton-mile shares by mode within commodity; (3) modal EI (Btu/ton-mile) by commodity; (4) fuel shares by mode (petroleum fuels, biofuels, electricity); and (5) electricity generation primary fuel shares (% kWh/fuel).

2015 NEAT base case projects that total ton-miles grow 96% from 2010 to 2050 while energy use during the same time grows only 90% mainly due to mode shift, use of alternative fuels and efficiency improvements. In the same time, such changes cause GHG emissions to increase 20% from 2010 to 2050. However, upstream energy use would grow by 110% during the same time due to the use of alternative fuels.

APPENDIX A

# List of Abbreviations and Acronyms

| | |
|---|---|
| 3PL | third-party logistics |
| AAR | Association of American Railroads |
| AEO | Annual Energy Outlook |
| API | application programming interface |
| ATRI | American Transportation Research Institute |
| BTS | Bureau of Transportation Statistics |
| CFS | Commodity Flow Survey |
| CSV | comma-separated values |
| EI | energy intensity |
| EIA | Energy Information Administration |
| ER | economic regions |
| FAF | Freight Analysis Framework |
| FG | freight generation |
| FTG | freight trip generation |
| GHG | greenhouse gas |
| GPS | Global Positioning System |
| LBD | longitudinal business database |
| MIT | Massachusetts Institute of Technology |
| MSA | metropolitan statistical area |
| NAICS | North American Industry Classification System |
| NAPCS | North American Product Classification System |
| NCFRP | National Cooperative Freight Research Program |
| O-D | origin–destination |
| OECD | Organization for Economic Co-operation and Development |
| PIERS | Port Import–Export Reporting Service |
| PUMS | Public Use Microdata Sample |
| SAS | Statistical Analysis Software |
| SCTG | Standard Classification of Transported Goods |
| STA | service trip attraction |
| TSI | transportation service index |
| ZOD | zip code-to-zip code origin–destinations |

APPENDIX B

# List of Attendees

Ken Allen
H-E-B Stores (retired)
San Antonio, Texas
kallen3315@gmail.com

Kayleigh Axtell
Texas Department of Transportation
Austin, Texas
Kayleigh.Axtell@txdot.gov

Scott Babcock
Transportation Research Board
Washington, D.C.
Sbabcock@nas.edu

Dan Biggio
U.S. Census Bureau
Suitland, Maryland
daniel.biggio@census.gov

Joseph Boccardo
U.S. Census Bureau
Suitland, Maryland
joseph.david.boccardo@census.gov

Thomas Bolle
U.S. Department of Transportation/OST-R
Washington, D.C.
thomas.bolle@dot.gov

Bonner, Janine
Bureau of Transportation Statistics
U.S. Department of Transportation
Washington, D.C.
janine.bonner@dot.gov

William Bostic
U.S. Census Bureau
Suitland, Maryland
william.g.bostic.jr@census.gov

Kenneth Cannon
Econometrica, Inc.
Bethesda, Maryland
kcannon@econometricainc.com

Cavan Capps
U.S. Census Bureau
Washington, D.C.
cavan.paul.capps@census.gov

Paul Ciannavei
IHS Global Insight
West Hartford, Connecticut
paul.ciannavei@ihsglobalinsight.com

Alison Conway
City College of the City University of New York
New York, New York
aconway@ccny.cuny.edu

Scot Dahl
U.S. Census Bureau
Washington, D.C.
scot.alan.dahl@census.gov

William Davie Jr.
U.S. Census Bureau
Washington, D.C.
william.c.davie.jr@census.gov

Patrick Donovan
Planning Center of Expertise for Inland Navigation and Risk-Informed Economics Division
Huntington, West Virginia
patrick.j.donovan@usace.army.mil

Scott Drumm
Port of Portland
Portland, Oregon
scott.drumm@portofportland.com

Ronald Duych
Bureau of Transportation Statistics
U.S. Department of Transportation
Washington, D.C.
ronald.duych@dot.gov

Chester Ford
U.S. Department of Transportation/OST-R
Washington, D.C.
chester.ford@dot.gov

Makarand Gawade
University of South Florida
Tallahassee, Florida
makarand@mail.usf.edu

Carlos Gonzalez-Calderon
Rensselaer Polytechnic Institute
Troy, New York
gonzac8@rpi.edu

Ryan Grube
Bureau of Transportation Statistics
U.S. Department of Transportation
Washington, D.C.
ryan.grube@dot.gov

Charlie Han
MacroSys, LLC
Arlington, Virginia
charlie.han@macrosysrt.com

Kathleen Hancock,
Virginia Tech
Falls Church, Virginia
hancockk@vt.edu

James Hinckley,
U.S. Census Bureau
Washington, D.C.
james.hinckley@census.gov

José Holguín-Veras
Rensselaer Polytechnic Institute
Troy, New York
jhv@rpi.edu

Patricia Hu
Bureau of Transportation Statistics
U.S. Department of Transportation
Washington, D.C.
Patricia.Hu@dot.gov

Ho-Ling Hwang
Oak Ridge National Laboratory
Knoxville, Tennessee
hwanghl@ornl.gov

Laura James
U.S. Census Bureau
Suitland, Maryland
laura.james@census.gov

Ron Jarmin
U.S. Census Bureau
Washington, D.C.
ron.s.jarmin@census.gov

Nicole Katsikides
Federal Highway Administration
Washington, D.C.
Nicole.katsikides@dot.gov

Bruce Lambert
Institute for Trade and Transportation Studies
New Orleans, Louisiana
ittsdirector@excite.com

Virgil Langdon
Planning Center of Expertise for Inland Navigation and Risk-Informed Economics Division
Huntington, West Virginia
virgil.l.langdon@usace.army.mil

Catherine Lawson
University at Albany
Albany, New York
lawsonc@albany.edu

Yin Jin Lee
Massachusetts Institute of Technology Center for Transportation and Logistics
Cambridge, Massachusetts
yinjin@mit.edu

Mark Lepofsky
FACTOR
Arlington, Virginia
mlepofsky@essentialfactor.com

Stephen Lewis
Research and Innovative
Technology Administration
Washington, D.C.
steve.lewis@dot.gov

Berin Linfors
U.S. Census Bureau
Suitland, Maryland
berin.linfors@census.gov

Andrew Ludasi
New Jersey Department of Transportation
Trenton, New Jersey
andrew.ludasi@dot.nj.gov

Donald Ludlow
CPCS Transcom Limited
Washington, D.C.
dludlow@cpcstrans.com

Brian McElroy
U.S. Census Bureau–Commodity Flow Survey
Suitland, Maryland
brian.j.mcelroy@census.gov

Lawrence McKeown
Statistics Canada
Ottawa, Ontario, Canada
Larry.McKeown@statcan.gc.ca

Dominic Menegus
Bureau of Transportation Statistics,
U.S. Department of Transportation
Washington, D.C.
dominic.menegus@dot.gov

Michael Meyer
WSP–Parsons Brinckerhoff
Atlanta, Georgia
mm39prof@gmail.com

Michael Miller
Transportation Research Board
Washington, D.C.
mmiller@nas.edu

Brett Moore
U.S. Census Bureau
Hyattsville, Maryland
brett.c.moore@census.gov

Kimberly Moore
U.S. Census Bureau
Washington, D.C.
kimberly.p.moore@census.gov

Bengt Muten
IHS, Inc.
Lexington, Massachusetts
bengt.muten@ihs.com

Vidya Mysore
Federal Highway Administration, Office of
Technical Services Resource Center
Atlanta, Georgia
vidya.mysore@dot.gov

Sreevatsa Nippani
Maricopa Association of Governments
Phoenix, Arizona
SNippani@azmag.gov

Nick Orsini
U.S. Census Bureau
Washington, D.C.
nick.orsini@census.gov

Tom Palmerlee
Transportation Research Board
Washington, D.C.
tpalmerlee@nas.edu

Birat Pandey
Federal Highway Administration, Office of
Technical Services Resource Center
Baltimore, Maryland
birat.pandey@dot.gov

Julie Parker
Bureau of Transportation Statistics,
U.S. Department of Transportation
Boston, Massachusetts
julie.parker@dot.gov

Alan Pisarski
Consultant
Falls Church, Virginia
alanpisarski@alanpisarski.com

Steven Riesz
U.S. Census Bureau
Washington, D.C.
Steven.Riesz@census.gov

Matthew Roorda
University of Toronto
Toronto, Ontario, Canada
roordam@ecf.utoronto.ca

Douglas Scheffler
U.S. Coast Guard
Washington, D.C.
Douglas.W.Scheffler@uscg.mil

Rolf Schmitt
Bureau of Transportation Statistics,
U.S. Department of Transportation
Washington, D.C.
rolf.schmitt@dot.gov

Joseph Schofer
Northwestern University
Evanston, Illinois
j-schofer@northwestern.edu

Joy Sharp
Bureau of Transportation Statistics,
U.S. Department of Transportation
Washington, D.C.
Joy.Sharp@dot.gov

Christine Sherman
RSG
Arlington, Virginia
christinelsherman@gmail.com

Jeffrey Short
American Transportation Research Institute
Marietta, Georgia
jshort@trucking.org

Michael Sprung
Bureau of Transportation Statistics,
U.S. Department of Transportation
Washington, D.C.
michael.sprung@dot.gov

Steve Stalikas
U.S. Army Corps of Engineers
Buffalo, New York
stephen.m.stalikas@usace.army.mil

Monique Stinson
Massachusetts Institute of Technology
Cambridge, Massachusetts
mstinson@mit.edu

Ed Strocko
Bureau of Transportation Statistics,
U.S. Department of Transportation
Washington, D.C.
ed.strocko@dot.gov

Tianjia Tang
Federal Highway Administration
Washington, D.C.
Tianjia.tang@dot.gov

John Thompson
U.S. Census Bureau
Washington, D.C.
john.h.thompson@census.gov

Kimberly Vachal
North Dakota State University
Fargo, North Dakota
kimberly.vachal@ndsu.edu

Al Vallens
U.S. Census Bureau
Suitland, Maryland
alexander.s.vallens@census.gov

Jennifer Whitaker
U.S. Census Bureau
Suitland, Maryland
jennifer.n.whitaker@census.gov

Rosalyn Wilson
Parsons
Vienna, Virginia
Rosalyn.wilson@parsons.com

Joel Worrell
Florida Department of Transportation
Tallahassee, Florida
Joel.Worrell@dot.state.fl.us

Supin Yoder
Federal Highway Administration
Matteson, Illinois
supin.yoder@dot.gov

Jiashen You
Bureau of Transportation Statistics,
U.S. Department of Transportation
Washington, D.C.
jiashen.you@dot.gov

Jessica Young
U.S. Census Bureau
Suitland, Maryland
jessica.m.young@census.gov

Feifei Yu
IHS, Inc.
Lexington, Massachusetts
Feifei.Yu@ihs.com

Yan Zhou
Argonne National Laboratory
Argonne, Illinois
yzhou@anl.gov

Johanna Zmud
Texas A&M Transportation Institute
Washington, D.C.
j-zmud@tti.tamu.edu

*The National Academies of*
SCIENCES · ENGINEERING · MEDICINE

The **National Academy of Sciences** was established in 1863 by an Act of Congress, signed by President Lincoln, as a private, nongovernmental institution to advise the nation on issues related to science and technology. Members are elected by their peers for outstanding contributions to research. Dr. Ralph J. Cicerone is president.

The **National Academy of Engineering** was established in 1964 under the charter of the National Academy of Sciences to bring the practices of engineering to advising the nation. Members are elected by their peers for extraordinary contributions to engineering. Dr. C. D. Mote, Jr., is president.

The **National Academy of Medicine** (formerly the Institute of Medicine) was established in 1970 under the charter of the National Academy of Sciences to advise the nation on medical and health issues. Members are elected by their peers for distinguished contributions to medicine and health. Dr. Victor J. Dzau is president.

The three Academies work together as the National Academies of Sciences, Engineering, and Medicine to provide independent, objective analysis and advice to the nation and conduct other activities to solve complex problems and inform public policy decisions. The Academies also encourage education and research, recognize outstanding contributions to knowledge, and increase public understanding in matters of science, engineering, and medicine.

Learn more about the National Academies of Sciences, Engineering, and Medicine at **www.national-academies.org**.

The **Transportation Research Board** is one of seven major programs of the National Academies of Sciences, Engineering, and Medicine. The mission of the Transportation Research Board is to increase the benefits that transportation contributes to society by providing leadership in transportation innovation and progress through research and information exchange, conducted within a setting that is objective, interdisciplinary, and multimodal. The Board's varied committees, task forces, and panels annually engage about 7,000 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.

Learn more about the Transportation Research Board at **www.TRB.org**.

TRB
TRANSPORTATION RESEARCH BOARD
500 Fifth Street, NW
Washington, DC 20001