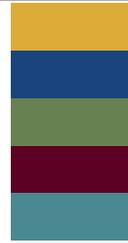


161407-1



# Strategic Research PROGRAM

Opportunities and Challenges Using Passively Collected  
Data in Travel Demand Modeling

October 2014



1. Report No. TTI/SRP/14/161407-1	2. Government Accession No.	3. Recipient's Catalog No.	
4. Title and Subtitle Opportunities and Challenges Using Passively Collected Data in Travel Demand Modeling		5. Report Date October 2014	
		6. Performing Organization Code	
7. Author(s) James A. Mullins III, Ipek N. Sener, and Richard J. Lee		8. Performing Organization Report No.	
9. Performing Organization Name and Address Texas A&M Transportation Institute College Station, Texas 77843-3135		10. Work Unit No. (TRAIS)	
		11. Contract or Grant No. 10727	
12. Sponsoring Agency Name and Address Texas A&M Transportation Institute College Station, Texas 77843-3135		13. Type of Report and Period Covered	
		14. Sponsoring Agency Code	
15. Supplementary Notes Supported by the State of Texas. Project Title: Model Development with Passively Collected Data			
16. Abstract Passive data collection technologies are providing new opportunities for travel demand modeling without the difficulties associated with traveler interaction inherent to traditional survey methods. Given this, the main objective of this project is to assess current use of passive data in travel modeling and identify specific topics of research that would evaluate the potential for increased and wider usage of passive data in travel demand modeling. This research consists of a literature review, a web-based survey, in-person interviews, and a workshop with transportation researchers. The research resulted in identification specific research topics and tasks that would provide answers to questions regarding the ability of passive data to support travel modeling. Although the use of passive data is growing in the travel modeling field, many of the study participants expressed concerns regarding its limitations and capabilities to provide useful data for modeling without further research. Recommendations for future research and potential partnerships are proposed to help practitioners overcome these concerns and to expand the use of passively collected data in travel modeling.			
17. Key Words Passively Collected Data, Travel Demand Modeling, Travel Survey, Cell Phone, GPS, Bluetooth		18. Distribution Statement No restrictions. This document is available to the public through NTIS: National Technical Information Service Alexandria, Virginia <a href="http://www.ntis.gov">http://www.ntis.gov</a>	
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 40	22. Price

# **Opportunities and Challenges Using Passively Collected Data in Travel Demand Modeling**

by

James A. Mullins III  
Research Scientist  
Texas A&M Transportation Institute

Ipek N. Sener  
Assistant Research Scientist  
Texas A&M Transportation Institute

and

Richard J. Lee  
Student Technician II  
Texas A&M Transportation Institute

Strategic Research Program  
Final Report  
Project Number 161407-1

Texas A&M Transportation Institute  
College Station, Texas 77843-3135

October 2014



## **DISCLAIMER**

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

## **ACKNOWLEDGMENTS**

Support for this research was provided by the State of Texas. The authors would like to acknowledge the TTI researchers (listed in the appendix) and Guy Rousseau from the Atlanta Regional Commission who participated in the workshop. The authors would also like to acknowledge Johanna Zmud, Kevin Hall, and Thomas Williams for their review of an earlier version of this document, as well as Michelle Benoit and Chris Pourteau for their help in disseminating the web-based survey. Finally, the authors thank Linda Cherrington and Katie Turnbull for their guidance during the course of this research project.

# TABLE OF CONTENTS

<b>List of Figures</b> .....	<b>vi</b>
<b>List of Tables</b> .....	<b>vi</b>
<b>Executive Summary</b> .....	<b>1</b>
<b>Chapter 1: Introduction</b> .....	<b>3</b>
Background.....	3
Objectives and Approach.....	4
<b>Chapter 2: Modeling Data Sources</b> .....	<b>6</b>
<b>Chapter 3: Survey of Modelers</b> .....	<b>9</b>
Background.....	9
Survey Response Rates and Respondents' General Characteristics .....	10
Summary of Responses.....	11
Follow-Up Questions .....	20
Summary.....	22
<b>Chapter 4: Workshop</b> .....	<b>23</b>
Background.....	23
Workshop Sessions .....	23
Summary.....	26
<b>Chapter 5: Future Research and Testing</b> .....	<b>27</b>
General Recommendations .....	27
Research Areas .....	28
<b>Chapter 6: Conclusions</b> .....	<b>31</b>
<b>Appendix: List of Workshop Participants</b> .....	<b>32</b>

## LIST OF FIGURES

Figure 1. Distribution of Survey Questions.....	10
Figure 2. Organizational Affiliation. ....	11
Figure 3. Has Any Type of Travel Survey Data Been Used to Support Model Estimation/Development/Validation?.....	11
Figure 4. Type of Traditional Travel Survey Data the Region Has Used.....	12
Figure 5. Type of Survey Data the Region Has Used to Support Modeling. ....	13
Figure 6. Type of Passively Collected Travel Survey Data That Has Been Used.....	14
Figure 7. Survey Context in Which Passively Collected Travel Survey Data Have Been Used. ....	14
Figure 8. Attributes of the Passively Collected Data That Have Been Imputed or Otherwise Developed through Linkage/Fusion with Other Data Sources.....	15
Figure 9. Primary Use of Passively Collected Data for Travel Modeling. ....	15
Figure 10. Reasons for Choosing Passively Collected Data. ....	16
Figure 11. Reasons/Concerns That Influenced the Decision to Not Use Passively Collected Data.....	17
Figure 12. Do Passively Collected Data Need to Provide the Same Information in the Same Way as Traditional Data in Order to Be Considered for Use in Model Development? .....	18
Figure 13. Potential Research Areas for Passively Collected Data. ....	19

## LIST OF TABLES

Table 1. Travel Demand Modeling Data Sources.....	7
---	---

## EXECUTIVE SUMMARY

Travel modeling has traditionally relied on survey data acquired through direct interaction with subjects. For years these proven methods have allowed for the collection of detailed trip and traveler information, serving as the basis for travel demand models. More recently, technological advancements and limitations associated with traditional travel survey data—including increasing cost and declining response rates—have resulted in the growing interest in and adoption of “passive” data collection techniques, which do not rely on traveler interaction. These technologies are enticing for their ability to generate massive quantities of travel data at a fraction of the cost of traditional methods.

The main objective of this research was to explore and assess current passive data applications and identify key areas in need of further research. To evaluate the current state of passive data collection and its suitability for travel demand modeling, a web-based survey was developed and distributed to transportation professionals. The survey was disseminated through the Federal Highway Administration’s (FHWA’s) Travel Demand Modeling Improvement listserv and to several other researchers and practitioners who are responsible for travel demand modeling and/or travel surveys in metropolitan planning organizations (MPOs), or who are known to conduct research in the field. The survey, and subsequent follow-up interviews, provided an understanding of the views and concerns of travel modeling practitioners regarding passive data. The data items and information that were sought in this survey were influenced by the results of a stand-alone literature review that identified recent research assessing the reliability and applications of mobile phone positioning, global positioning systems (GPS), Bluetooth®, social networking, and smart card data.

Results of the web survey reveal that cost and sample size was among the primary motivations for using passive data, but many respondents also indicated a reluctance to adopt the technologies. Common concerns included the lack of trip purpose and mode information, data quality, sampling bias, privacy, and a lack of familiarity. Participants wanted further research into these issues, case studies validating these new approaches, easy-to-implement modeling tools, and best practices for processing and analyzing these data. In general, respondents considered passive data suitable only as a supplement to traditional travel data, rather than as a standalone data source. Some in the modeling community expressed the sentiment that the current modeling paradigm will need to adapt to new data collection technologies.

A workshop was held with researchers and practitioners to present preliminary project findings and to discuss future research needs. The resulting identified research topic areas fell into two groups. The first area focuses on evaluation of modeling data items for which passive data are a potential source. Several research areas evaluating the use of passive travel behavior data to replace the use of traditional survey data in the development of model components were identified. Expansion of available data items, either through the use of technology linked to passive data devices or through linkage to third-party commercial data sets was seen as an area in need of research. Additionally, research to establish the equivalency to sampling confidence for passive travel data sets was seen as an important research area for modelers to explore.

A second set of research topic areas resulting from the workshop centered on the evaluation and testing of methods and processes to use passive data to model types of travel for which models

built with traditional data struggle or are non-existent. Modeling of weekend, visitor, and long-distance travel were all topics for which evaluation in the use of passive data was identified.

# CHAPTER 1: INTRODUCTION

## Background

Historically, travel demand analysis and forecasting models/tools have relied on survey data traditionally collected directly from travelers, often capturing trip purpose, length, origin and destination locations, travel mode, and demographic information about the traveler and the traveler's household. The data are then calibrated to an observed base year condition and applied to forecast scenarios to inform intermediate and long-range transportation planning decisions.

Because of costs, these surveys are typically conducted infrequently with small sampling rates and a greater reliance on statistical expansion. Some practitioners, when faced with the cost of data collection, may choose to utilize out-of-date trip rates, borrow trip rates from other locations, minimize sample size, or synthesize trip rates using best judgment. Also, traditional data collection methods have historically and are increasingly limited due to issues such as non-response, sample size, under-reporting, and survey fatigue.

Those collecting and using travel data for travel modeling have become increasingly aware and effected by these cost and data issues and limitations. At the same time, technologies through which passive travel data<sup>1</sup> are collected such as GPS, cell/smartphones, and Bluetooth-equipped devices have been shown to offer many of the trip-oriented data items that are available through traditional data sources. Besides offering several similar travel data items, passive data contain much higher numbers of observations than traditional data and contain trip-oriented data at greater time and geographic detail than data from traditional sources. In addition, passive data do not generally suffer from issues related to under-reporting of travel and survey fatigue since the data are able to be collected without interaction with the survey subject. Passive data can also offer travel information on travel subgroups such as trucks or non-residents of a region that are challenging to capture through traditional survey data collection means.

The lack of interaction with the traveler results in passive travel data lacking demographic and personal information about the traveler and other members of the traveler's household. These are critical items needed for travel modeling and forecasting of future travel. Also, some passive travel data items are simplified or more aggregately categorized than traditional data items. This presents challenges to those seeking to perform newer types of disaggregate travel demand modeling.

Given the now common availability of passive data and the continuing challenges with traditional data collection methods, the travel data collection and travel modeling communities have shown increasing interest in finding ways to expand the use of passive data in modeling. The degree to which this occurs will be influenced by research that evaluates passive data for the ability to provide additional needed travel modeling data and development and testing of procedures and methods to create and use these data for travel modeling.

---

<sup>1</sup> The terms "passive data" and "passively collected data" are used interchangeably throughout this report.

## Objectives and Approach

The potential to respond to the growing interest in passive data and its potential to open new areas of travel modeling research and practice will be heavily influenced by the ability to move beyond current use of this type of data. In this respect, this research has two primary purposes:

- Identify current uses of passive data in modeling so that a baseline from which future research should move can be identified.
- Identify areas of research that are needed to determine if and how passive travel data use can be expanded in travel modeling.

To achieve project objectives, the research was divided into two phases.

### *Modeling Data Needs and Passive Data Use Baseline*

The first phase of the project involved review of the typical travel data needs of travel models, how traditional data are used to satisfy those needs and current and potential use of passive data to meet these needs. This information allowed the establishment of a baseline understanding of the role passive travel data can currently play and may play in the future with research.

Using this information and information gathered through a review of literature and documentation of passive data development and application in modeling,<sup>2</sup> the project sought to gather information on practical applications in the use of passively collected data and then identify future research needs. This report presents the information related to the second phase of the study.

### *Passive Data Applications and Future Research Needs*

The second phase started with a documentation of modeling data sources, which is presented in Chapter 2 of this report. Next, a web-based survey of transportation professionals (primarily modeling practitioners) who have used or have considered using passive data in travel modeling was developed. This survey effort also included follow-up supplementary interviews with a subset of respondents to ascertain additional details on their use of passive data. The survey and interviews provided the opportunity for researchers and practitioners to offer suggestions as to research and/or applications research directions that could expand their use of passive data for travel demand modeling. Section 3 of this report presents the results of this survey-based effort and a discussion on the additional follow-up interviews.

Following the development and implementation of the web-based survey, a panel of modeling and data collection researchers was gathered for a workshop. At the workshop, the participants were presented with the results of the first two aspects of this study, and then asked to discuss and develop ideas for future research topics and/or applications. The presentations consisted of summary results from the literature review and the web-based survey, as well as information

---

<sup>2</sup> Technical Memorandum: “Emerging Data Collection Techniques for Travel Demand Modeling: A Literature Review.” Lee, Sener, and Mullins (September, 2014).

from the proceedings of an FHWA symposium on the use of cell phone data in modeling. Chapter 4 presents the details on this workshop.

Finally Chapter 5 provides the research initiatives directed toward expanding the use of passively collected data in travel demand modeling and identified research areas in the use of passive data in modeling. Chapter 6 presents conclusions from this study.

## CHAPTER 2: MODELING DATA SOURCES

Travel modeling relies on three general types of data:

- Information about travel behavior of individuals.
- Limited personal information about those engaged in the behavior and members of the household in which they live.
- Measurement of collective travel behavior on the transportation system.

Data on travelers and travel behavior are collected through a series of travel surveys in which the subjects are asked to provide information about travel for an entire day and to specific locations to which travel occurred. For those less common forms of travel or travel destinations, targeted surveys are performed. The data items collected from traditional data sources are used in travel demand models to model the amount, frequency, purpose, origin/destination, mode, time-of-day, travel time, and travel distance of individual trips. Traditional data sources are also used to provide data on the linkages of trips made by individuals throughout a day, which have become more important with newer modeling methods.

In order to model the interaction of travel decisions and behavior among household members that is known to affect the characteristics of individual travel such as those mentioned above, traditional data sources include travel information about all members of a household. Traditional data sources also include personal and demographic data about the traveler and other members of the household. The inclusion of travel and limited personal and demographic data from all members of a household is important to the modeling of interactions among individuals in a household that results in travel. Passive data are currently capable of providing a limited amount of the information described above. However, there is still much research needed to evaluate passive data source capability with respect to a great deal of these data items.

Table 1 presents a listing of data items that are used in various aspects of travel modeling, identifies the traditional data source for these items, and categorizes the current availability of these data items from among cell phone, GPS, and Bluetooth passive data sources. The categorization of these items is based on information gathered in the literature review, web survey, and workshop portions of this research project. This categorization also identifies areas where further effort is needed to determine if and how passive data can be a source for the specific data item.

**Table 1. Travel Demand Modeling Data Sources.**

<b>Modeling Data Item</b>	<b>Current Data Source</b>	<b>Passive Data Provides?</b>
Stratified sample of persons/households with known statistical confidence of representativeness	Household, Work Place Surveys	Very large sampling of travelers, but representativeness and confidence uncertain. Potential area of research.
Age, gender, employment/student status of traveler	Household, Work Place, Special Generator, External, Commercial Vehicle Surveys	Requires linkage to other data.
Age, gender, employment/student status of other household members	Household Survey	Currently unavailable. Potential area of research.
Auto availability	Household, Work Place Surveys	Currently unavailable. Potential area of research.
Household income	Household, Work Place, External Surveys	Requires linkage to other data.
Number of tour(s)/trip(s)	Household, Workplace, Commercial Vehicle Surveys	Yes. Superior to traditional data source.
Purpose of tour(s)/trip(s)	Household, Workplace, Special Generator Surveys	Yes. Through imputation. Limited detail. Potential area of research.
Order of tour(s)/trip(s)	Household, Workplace, Special Generator, Commercial Vehicle Surveys	Limited. Potential area of research.
Travel time and distance	Household, Workplace, Special Generator Surveys	Yes. Can be superior to traditional source.
Parking, transit, toll costs for each trip	Household, Workplace, Special Generator, External, Commercial Vehicle Surveys	Requires imputation from/linkage to other data sources. Potential area of research.
Time-of-day of travel	Household, External and Commercial Vehicle Surveys	Yes. Can be superior to traditional reporting.
Activity at all destinations	Household, Workplace, Special Generator, External Surveys	Requires imputation from/linkage to other data sources. Potential area of research.
Travel origin & destination	Household, Workplace, Special Generator, External, Commercial Vehicle Surveys	Yes. Can be superior to traditional source.
Mode of tour(s)/trip(s)	Household, Workplace, Transit, Non-motorized, Toll Facility User Surveys	Limited. Needs research/testing.
Vehicle type	Household, Workplace, Special Generator, External, Commercial Vehicle Surveys	Currently unavailable. Potential area of research.
Vehicle Occupancy	Household, Work Place, Special Generator, External Surveys	Currently unavailable. Potential area of research.
Type of cargo (if commercial vehicle)	Commercial Vehicle, External Travel Surveys	With linkage to secondary data source.
Route	Available, but not of useful sample size	Yes. Superior to traditional sources. Expansion methods in need of research.

Passive data are currently being used in practice to model the amount, origin/destination, time-of-day, travel time, travel distance, and, with limitations, the purpose of individual travel. However, the data needs of many travel models being developed require a greater amount of detail regarding the linkages among trips made by an individual throughout the day. The ability for modelers to make use of passive data will be heavily influenced by the degree to which methods can be developed to obtain personal and demographic data about the traveler and members of the traveler's household. The level of interest among the modeling community in the ability to obtain these data directly from passive data or to develop linkages to other data sources will be a driver in future research.

Travel models are developed for use beyond the analysis of current travel demand. Forecasts and analysis of future travel trends are a primary function of travel demand modeling. Given the challenges in the collection of travel data for travel forecasting, the modeling community needs to determine the degree to which passive data can support forecasting of travel demand. An important aspect of this evaluation should include research to determine the ability to establish the representativeness of passive data. Many of the methods used to collect traditional data are based on a statistically representative sample of a region's population and households. In this way, there is a known statistical confidence in the data's representation of the traveling public. Passive data sets are not gathered via sampling and the representativeness of the data is unknown. Research to evaluate passive data sources for representativeness and the ability to determine statistical confidence could prove beneficial to the use of passive data sources in travel forecast modeling. Additionally, research to evaluate the ability to obtain demographic and personal data directly from the traveler and even the traveler's household or link to other sources of these data is an important activity in creating the ability to use passive data items in travel forecast modeling.

## CHAPTER 3: SURVEY OF MODELERS

### Background

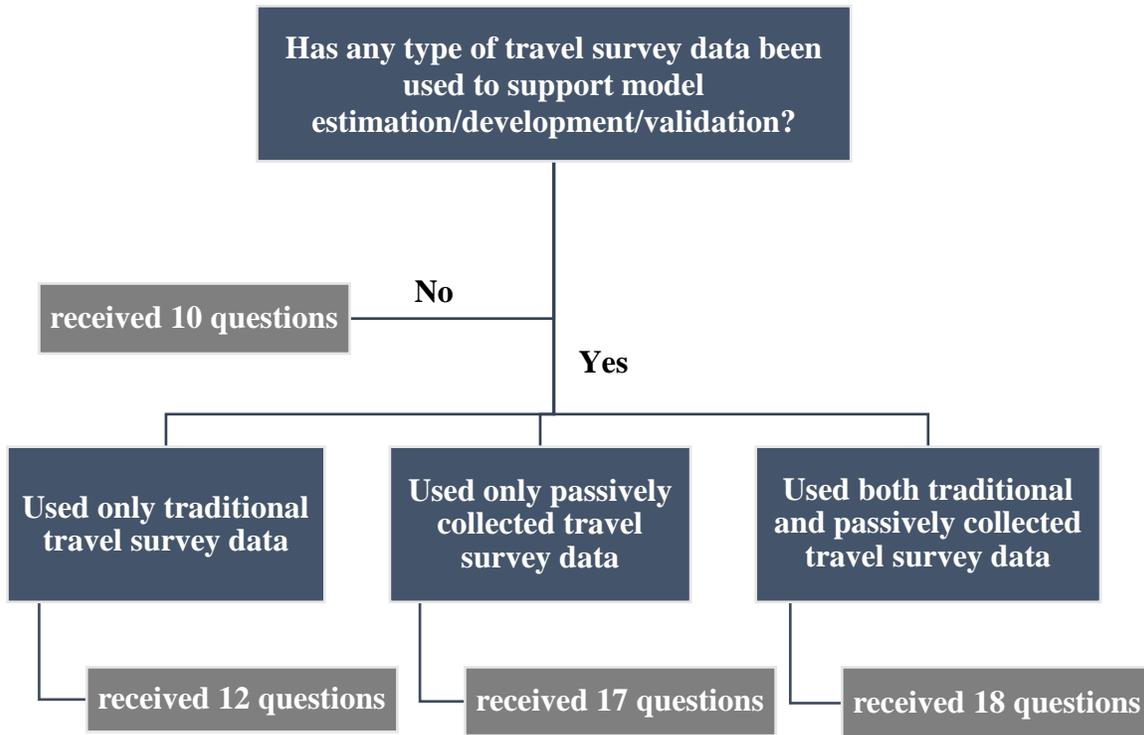
The research team developed a web-based survey to gather information and better understand the use of passively collected travel survey data in travel demand model development. The survey was designed using LimeSurvey, an online survey tool, and made available on the Internet.

In accordance with the objectives of the study, the target audience for the survey was those in the public or private sector engaged in the use of travel survey data for developing and applying travel demand model. Whether they had or had not used passively collected survey data in model development, individuals were asked to consider completing the survey.

The survey was launched on May 22, 2014, and ran for two weeks until June 5, 2014. The research team disseminated information about the survey to the FHWA's Travel Demand Modeling Improvement listserv of more than 1,600 members (during the period of the survey) and to over 340 people identified as responsible for travel demand modeling and/or travel surveys in MPOs. In addition, several other researchers or practitioners (who were known to conduct research in the passive data field) were targeted and informed about the survey. The survey responses were kept completely anonymous, and respondents were not personally identified unless they were willing to provide contact information at the end of the survey for a follow-up interview. No compensation was provided for participation in the web-based survey.

The survey was comprised of 25 questions; however, some questions were conditionally offered depending upon participants' responses related to their use of travel survey data. Those indicating experience with travel survey data, and passively collected data in particular, were asked additional follow-up questions (see Figure 1). The questionnaire took approximately 5–7 minutes to complete.

In the survey, traditional survey data referred to travel behavior data collected through direct or indirect interaction via CATI, mail-based surveys, or direct interviews. Passive data were defined as information related to the time, location, purpose, mode, or route of personal travel collected without traveler interaction, including cell phone, GPS, and Bluetooth data.



**Figure 1. Distribution of Survey Questions.**

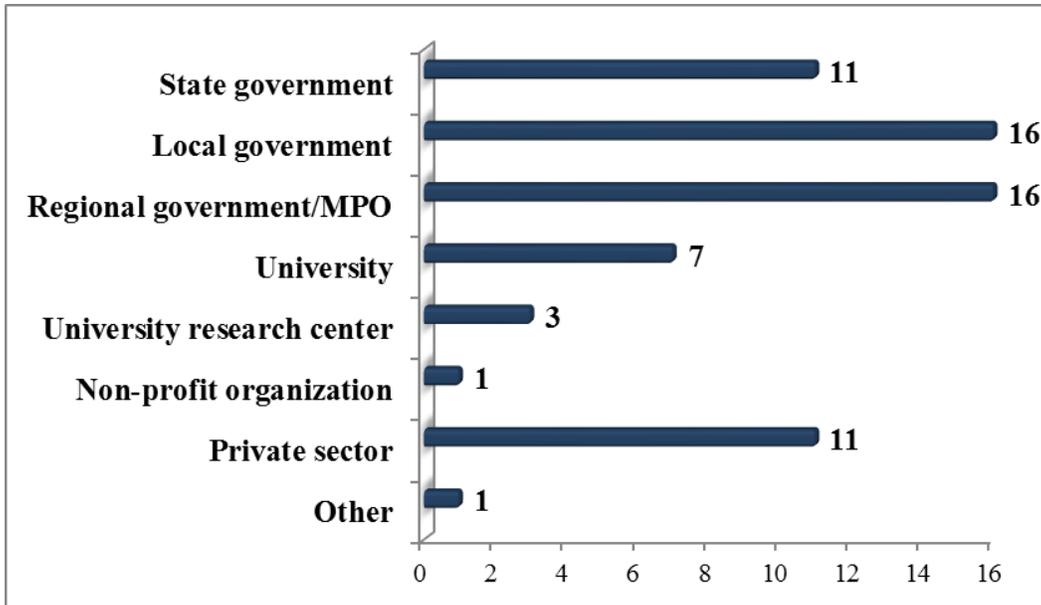
Following the survey, several follow-up interviews were conducted to gain further insight into the use of passive data. Researchers conducted follow-up interviews using two approaches:

- At the end of the survey, respondents were asked whether they were willing to be contacted for any follow-up questions/discussions. Additional information was gathered via email from those who agreed to be contacted.
- In-person interviews were conducted to get hands-on information from users with early experience using passive data in modeling.

The following sections provide information on the web-based survey response rates, general characteristics of respondents, a summary presentation and discussion of the survey results, and discussion of the follow-up interviews.

### **Survey Response Rates and Respondents' General Characteristics**

A total of 110 individuals accessed the survey online. Of those 110, 64 completed the entire survey, 30 partially completed it, and 16 failed to provide any response. As presented in Figure 2, approximately two-thirds of survey respondents (65 percent) worked in state, local, or regional government, while nearly all others were affiliated with the private sector or universities/university research centers.

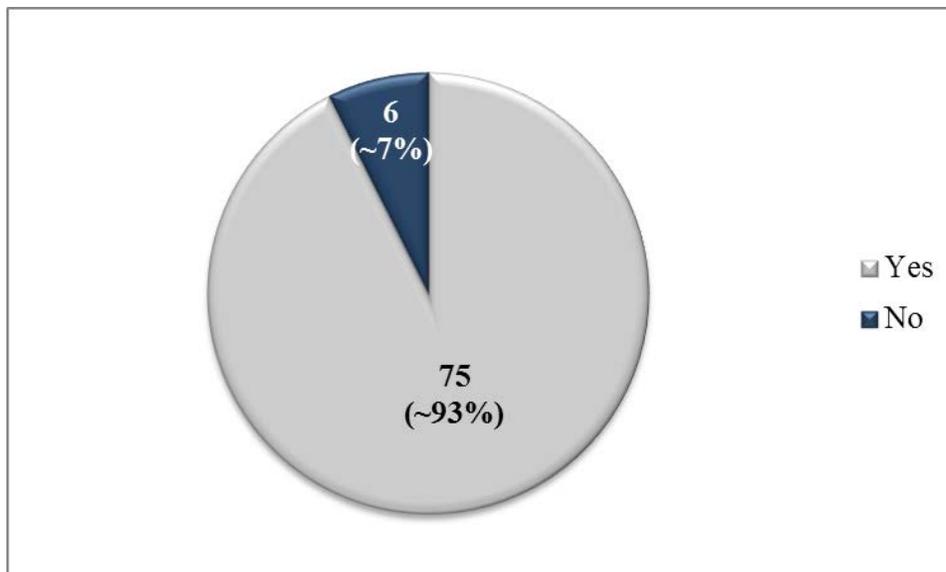


**Figure 2. Organizational Affiliation.**

## Summary of Responses

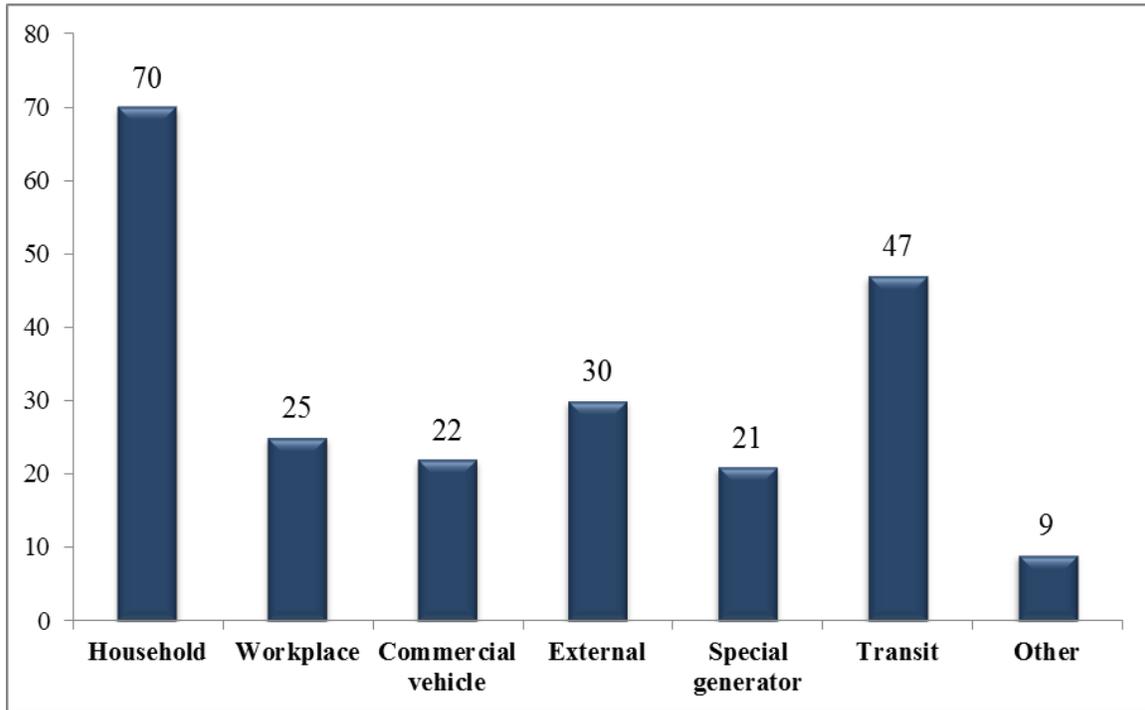
### *Experience with Travel Survey Data*

Most respondents (93 percent) indicated that their region had employed some form of travel survey data to support model development, estimation, or validation (Figure 3). Of the six that answered otherwise, two respondents mentioned cost being a factor, while another who had experience working with various MPOs reported that none of these MPOs used travel survey data in their modeling.



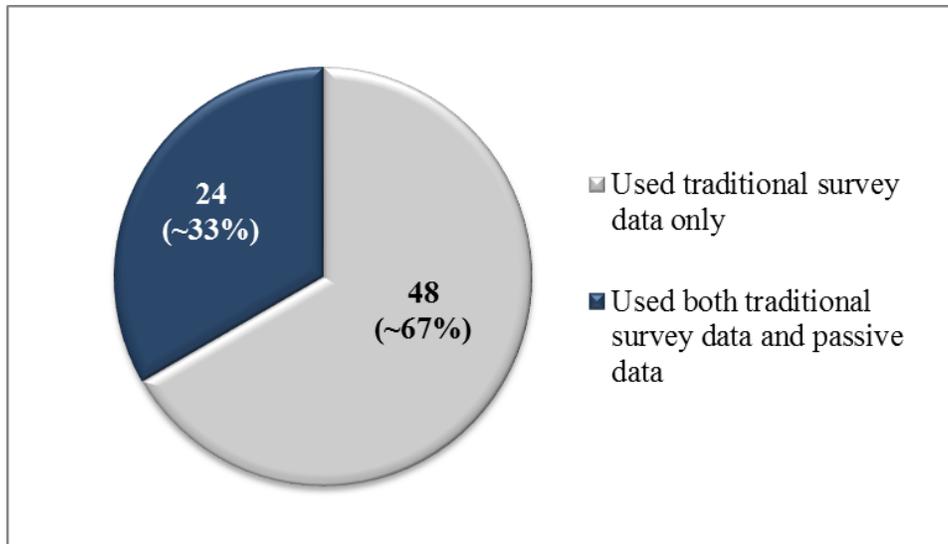
**Figure 3. Has Any Type of Travel Survey Data Been Used to Support Model Estimation/Development/Validation?**

Survey participants had the most experience with traditional survey data, and in particular household travel surveys (n=70) and transit surveys (n=47). Other common sources of traditional travel survey data were external, workplace, commercial vehicle, and special generator surveys (Figure 4). Responses provided in the “other” category included roadside interviews, airport passenger surveys, university surveys, bike/pedestrian surveys, surveys of long trips, casino trips, truck movement, visitors, toll user groups, and airport passengers.



**Figure 4. Type of Traditional Travel Survey Data the Region Has Used.**

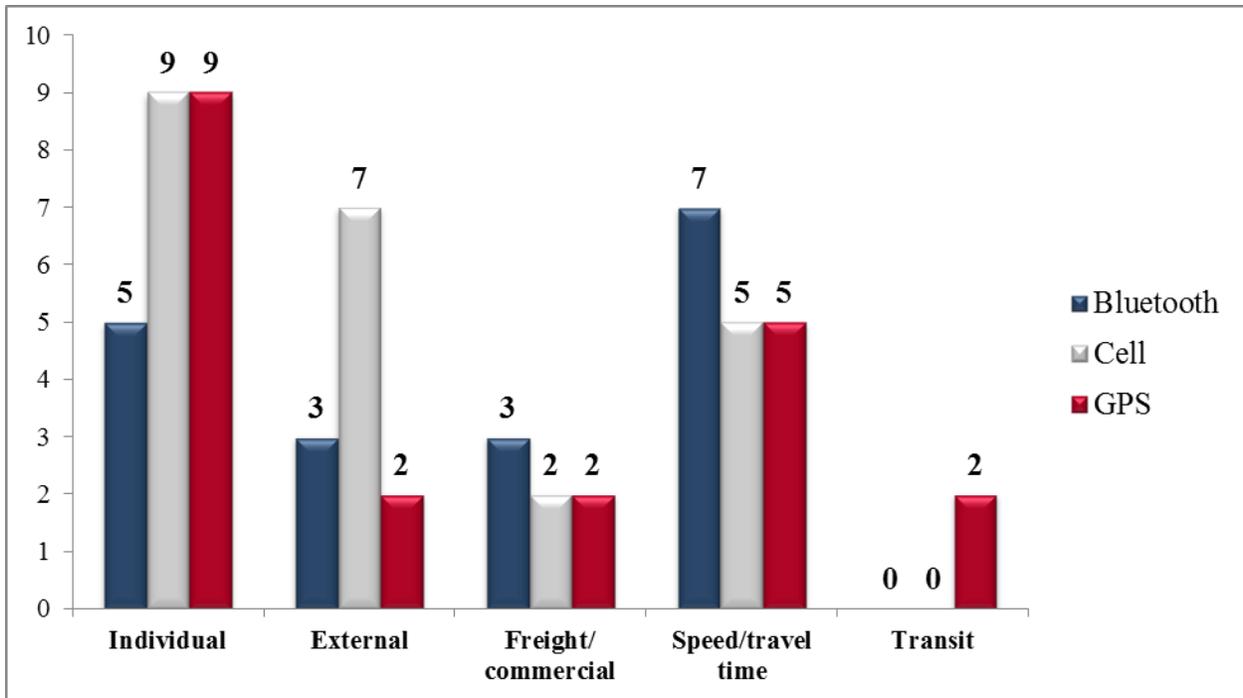
It appears that passive data are beginning to gain a foothold alongside traditional data sources. Among those that had experience with travel survey data for modeling purposes, one-third had used both traditional survey data and passive data, while the rest relied solely upon traditional data sources (Figure 5).



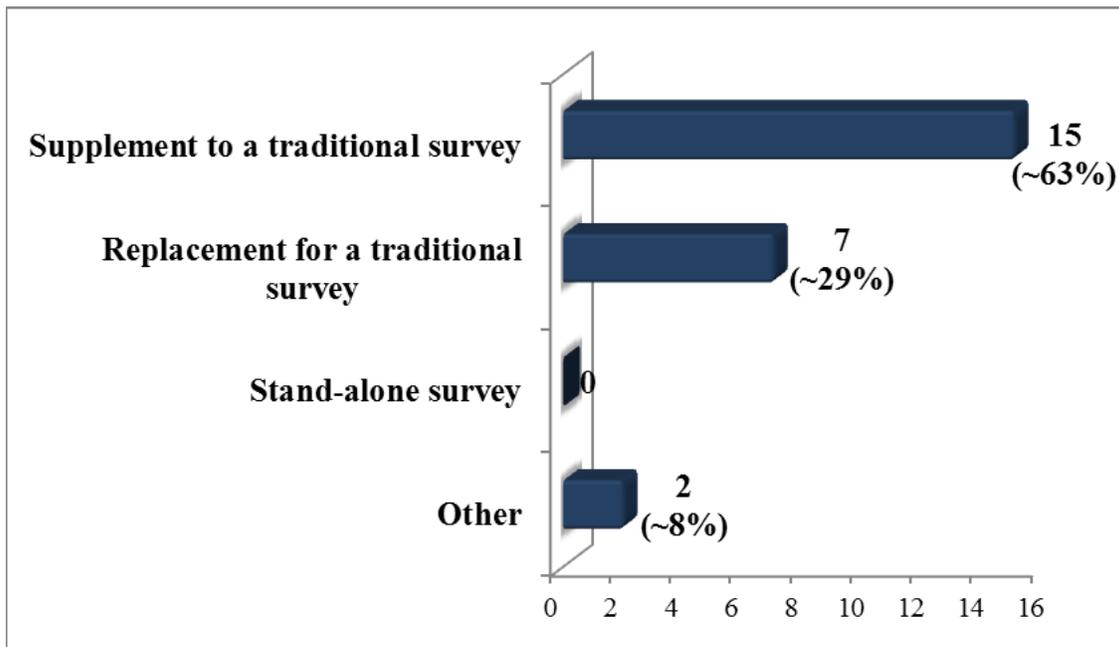
**Figure 5. Type of Survey Data the Region Has Used to Support Modeling.**

*Passive Data Usage*

This subsection reports the responses of the 24 individuals that indicated having used passively collected data. Figure 6 presents the various types of passive data used, with individual, speed/travel time, and external survey data being the most prevalent. Cell phone data were commonly used across nearly every data type, while individual and speed/travel time data were the primary uses of GPS and Bluetooth data. Passive data were often used as a supplement to traditional survey data (63 percent) or sometimes as a replacement (29 percent), but never as a standalone data set (Figure 7).



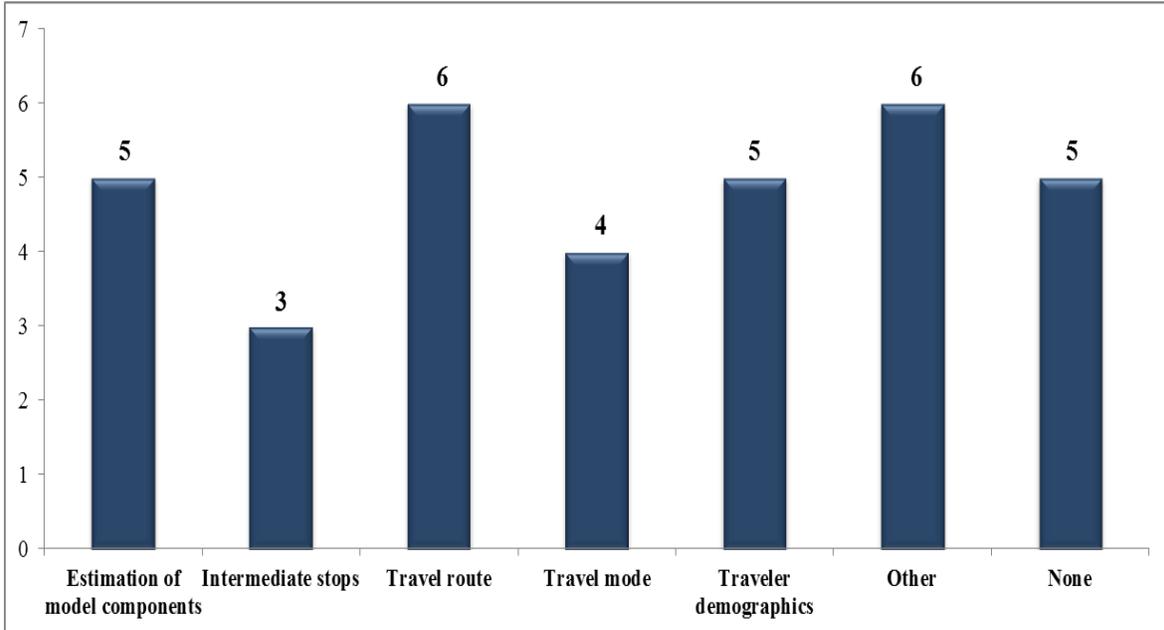
**Figure 6. Type of Passively Collected Travel Survey Data That Has Been Used.**



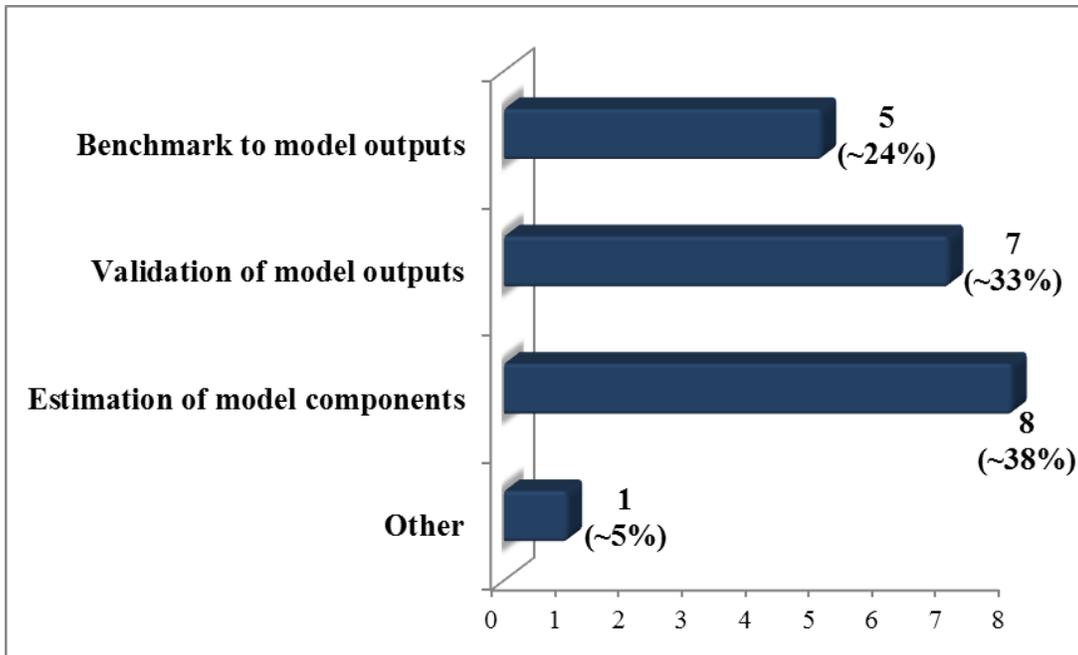
**Figure 7. Survey Context in Which Passively Collected Travel Survey Data Have Been Used.**

A number of travel/trip attributes were imputed or developed through fusion with other data sources (Figure 8). These included route, demographics, and travel mode; trip purpose was also frequently cited in the “other” category. In regard to travel models, these data were used for the

estimation of model components (38 percent), for validation of outputs (33 percent), and as a benchmark to model outputs (24 percent) (Figure 9).



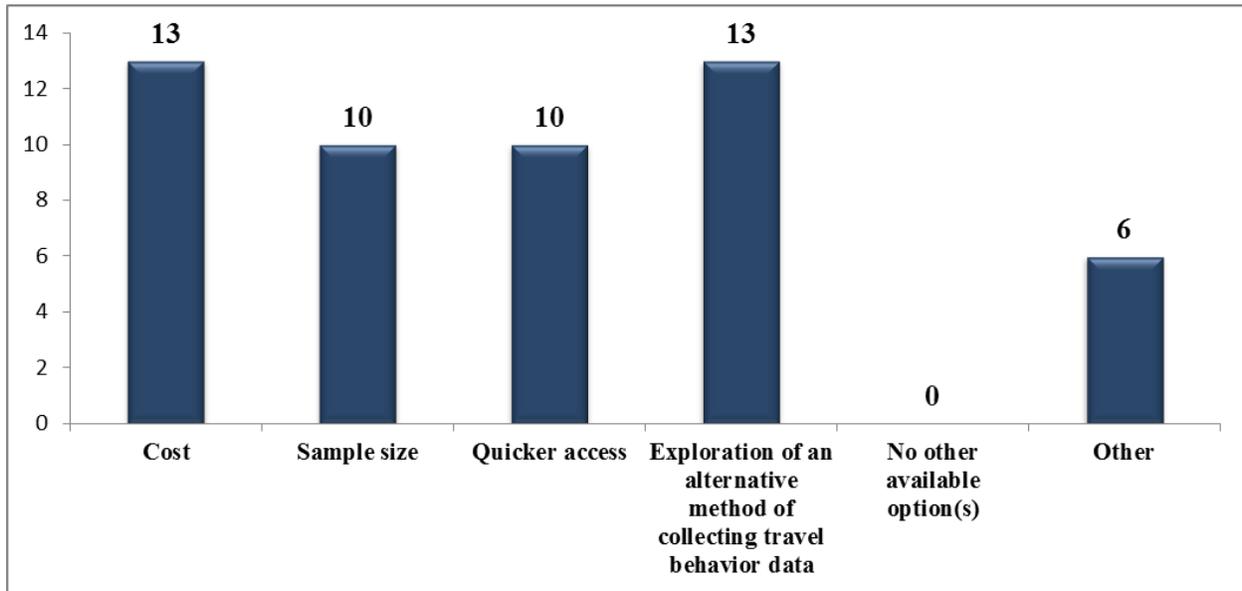
**Figure 8. Attributes of the Passively Collected Data That Have Been Imputed or Otherwise Developed through Linkage/Fusion with Other Data Sources.**



**Figure 9. Primary Use of Passively Collected Data for Travel Modeling.**

Just over half of passive data users selected “cost” and “exploration of an alternative method of collecting travel behavior data” as reasons for choosing to use passively collected data

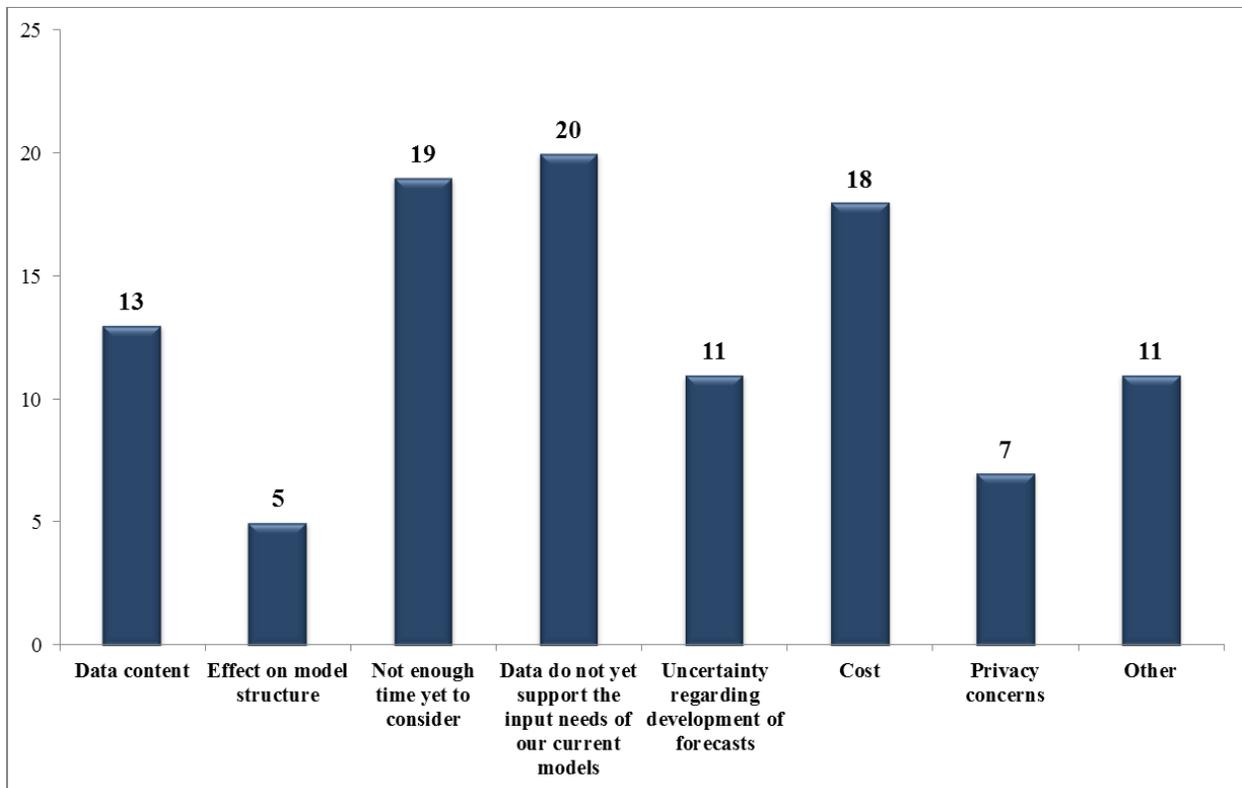
(Figure 10). “Sample size” and “quicker access” were selected by 42 percent of users. Other reasons added by participants included “to reduce respondents’ burden” and “better speed data.”



**Figure 10. Reasons for Choosing Passively Collected Data.**

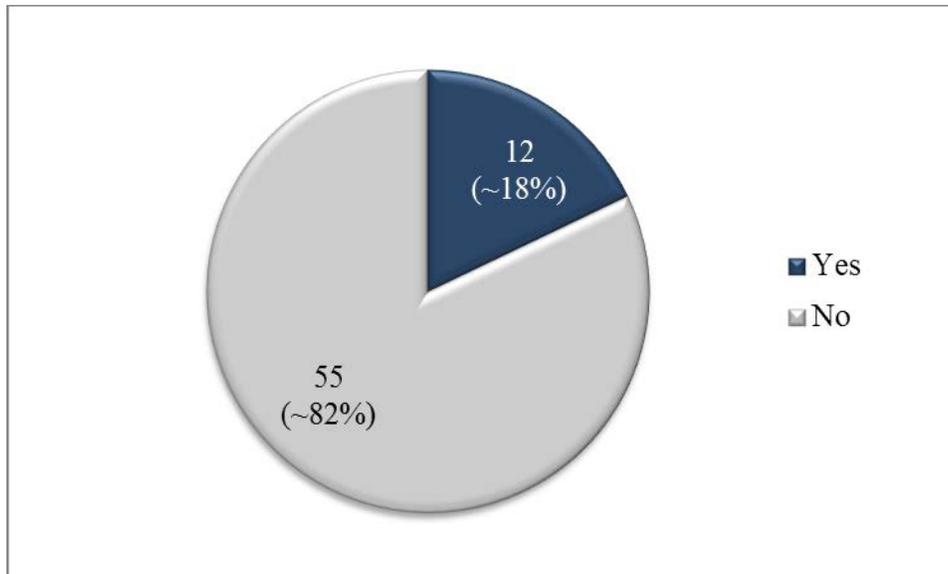
#### *Passive Data Concerns*

Users of traditional survey data only and non-survey data users were asked to select reasons why they had not made use of passive data in travel modeling. The most commonly reported reasons were: “data do not yet support the input needs of our current models,” “not enough time yet to consider,” and “cost” (Figure 11). Interestingly, cost was also cited by users of passive data as one of its primary benefits (Figure 10). It is possible that despite the relatively low cost, many respondents viewed passively collected data as an additional expense because they did not believe they are capable of acting as a standalone data source. Indeed, one respondent stated that passive data “can only supplement [traditional data], yet the cost of traditional [data] eats up the budget before passive data are added.” Concerns regarding data content, development of forecasts, privacy, and passive data’s effect on model structure were also chosen as reasons for not using passive data. Additional reasons added by respondents included sampling bias, accuracy, and lack of knowledge.



**Figure 11. Reasons/Concerns That Influenced the Decision to Not Use Passively Collected Data.**

All respondents were asked to provide their opinions on whether passively collected data need to provide the same information in the same way as traditional data. Eighty-two percent of participants did not agree with this statement (Figure 12). Several of these responses emphasized the notion that passive data are only suitable as a supplement to traditional survey data because they are not capable of providing the same information. As expressed by one participant, passive data could “serve as a supplement [but] cannot provide the detailed information required to develop a modern transportation model.” Another common theme was the idea that the current modeling paradigm will need to adapt to modern data collection technologies, rather than the other way around. Respondents stressed the “need to fit our forecasting tool to the observed behaviors of our traveling public” and that travel models must be “disabused of the objective that matching a household survey is the only measure of validity.” Only seven respondents believed that passive data should mirror traditional survey data, with one suggesting that “credibility will always be an issue” with passive data.



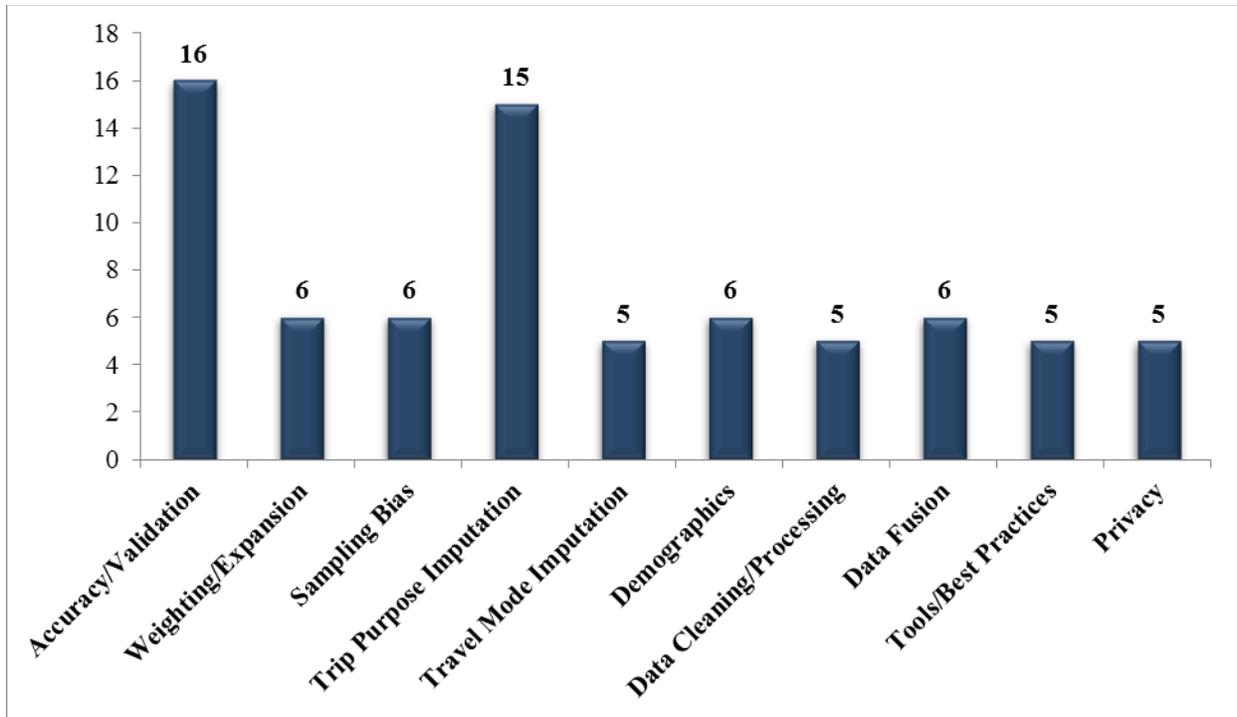
**Figure 12. Do Passively Collected Data Need to Provide the Same Information in the Same Way as Traditional Data in Order to Be Considered for Use in Model Development?**

#### *Areas of Research*

Finally, participants were asked about potential areas of research that would encourage their use of passive data. A number of research topics were identified, along with existing barriers and concerns. Responses were categorized and are presented in Figure 13. Sixteen respondents mentioned uncertainty regarding the accuracy, data quality, and reliability of passively collected data. In particular, several were interested in research validating passive data against traditionally collected data. Another respondent had experience using cell phone data to model at the county level but was unsure of their suitability for finer-scale studies.

Six respondents brought up weighting and expansion of passive travel data as an area for further study, including techniques for doing so and studies assessing the accuracy of these techniques. Tied to this idea were concerns over sampling bias, given that passively collected data likely draw from a different population than traditional survey data. One respondent wondered, “What markets are over/under represented [?] Are low-income missing? Are trip chaining and group travel correctly represented?”

Despite the large quantities of passively collected data available, one of the main drawbacks of passive data appears to be reliance on simplified trip purpose definitions. A number of respondents (n=15) suggested that they would value research into trip purpose imputation techniques. One commented that “even though cell phone data can now provide limited trip purposes, it will be nice if we can learn the full range of trip purposes from the data.” Determination of travel mode and demographic characteristics for passive data were identified as other potential areas of research. Pertaining to cell phone data, one participant noted, “No demographic information is included, which limits the usefulness of the data.”



**Figure 13. Potential Research Areas for Passively Collected Data.**

Technological barriers also appeared to discourage the adoption of passive travel data. Respondents cited the need for documentation and validation of data cleaning/processing and fusion techniques. Five others reported a desire for research resulting in best practices or tools for practitioners. These suggestions included:

- “User friendly software to understand what the data are telling us.”
- “Some best practices or lessons learned from an MPO that has actually used these data.”
- “An off-the-shelf, standardized travel demand model/land use model to consume the passively collected data.”

Others simply acknowledged their lack of knowledge regarding the use of passive data (e.g., “I am not that familiar with the technology” or “unable to comment as my knowledge in this area is limited”). Privacy was also brought up as a factor that should be considered when using passive data. As one respondent asked, “How does the public feel about their travel patterns being monitored passively? How can we assuage their fears without skewing the data?”

Passively collected data options are growing more prevalent, but respondents appeared to have reservations regarding their use. Although many recognized their potential, passive data are currently viewed as a supplement to traditional survey data rather than a standalone data set. Data quality, trip purpose imputation, and lack of expertise were among the primary concerns mentioned by survey respondents. Several other factors were also mentioned, including sampling bias, data weighting and expansion, privacy, demographics, data cleaning, and tools for practitioners. It appears that further research and experience will be needed before more practitioners feel comfortable incorporating passive data into their travel models.

## Follow-Up Questions

### *Email Interviews*

At the end of the web-based survey, individuals were asked whether they were willing to be contacted for any follow-up questions/discussions. In total, 25 respondents expressed interest in being contacted if needed. Among these respondents, eight individuals indicated that they had used passively collected data and were willing to provide additional details. The eight respondents were asked additional follow-up questions to develop a further understanding regarding the nature of their passive data use. All eight of the follow-up contacts preferred to correspond by email.

The eight respondents were first asked whether the model for which passive data were used was a trip-based or tour-based model. All indicated that passive data were used for trip-based models. Two of the eight also used passive data in the context of activity-based models.

Four of the eight respondents indicated that they used passive data for the estimation of model components. Three of these four individuals used the data in some aspect of the estimation of trip matrices, either parameters used to develop trip length distributions or trip demand matrices for types of travel not represented by their travel model, such as weekend or visitor/tourist demand. All four individuals indicated that their use of passive data for modeling involved the estimation of several components of their region's activity-based model.

In three cases, passive data were used to validate components of regional travel models. The data were used to validate model output origin-destination (OD) patterns, flows, shares, trip lengths, and travel times. In one case, the passive data were used to make adjustments to model-based tables of trip patterns. Two of the eight respondents used passive data to benchmark model components with the passive data being used to compare model trip patterns to passive-data-derived trip patterns. One participant planned on comparing travel patterns from one form of passive data (Bluetooth) to another (cell phone data).

The research team was also interested in discovering actions passive data users took to facilitate their use of the data within their travel models. Specifically, the team was interested in any simplifications made to the target models to accommodate passive data, and whether the users processed, merged, or joined data from secondary sources. Four of the respondents indicated that some model simplifications were performed in order to integrate passive data. One of the simplifications involved the collapsing of trip purposes to more general purposes to mirror those available with the passive data. The only other simplification specifically mentioned by respondents was the aggregation of traffic analysis zones (TAZs) outside of a corridor of interest. This was done not necessarily to accommodate a limitation of passive data, but more so to streamline the analysis of travel patterns and demand for a corridor of interest.

Four respondents had merged or joined passive data to secondary data sources. Two reported using various street layers to facilitate the development of route traces from the passive data. Another reported comparing passive data to traffic counts and helicopter imagery to expand the passive data to represent the full traffic stream. The fourth respondent had performed additional

processing of a passive data set to obtain desired line diagrams and trip counts by purpose and time of day.

The final question dealt with comparisons of passive data to traditional data sources. Three respondents reported they had compared passively collected data with either observed and/or traditionally collected data. Two of the three respondents stated that passively-derived trip length information had been compared to that of traditional sources. In one instance, longer-distance trips appeared to be overrepresented by passive data as compared to household survey and roadside interview data. Another user's experience suggested that passive data resulted in shorter trip lengths compared to journey-to-work data. In addition to comparing trip lengths, one respondent conducted a comparison of trip rates developed from a traditional household survey and passive data. Passive data trip rates were surprisingly similar to those generated by the traditional survey, even though the passively collected data represented year-round data, while the traditionally collected data omitted the summer months. Lastly, one respondent compared passively-derived data to modeled volume and OD data. The passively-derived data outperformed the modeled data for OD estimation, but compared less favorably to modeled volume counts.

In summary, the majority of passive data were used for validation or benchmarking of modeled travel pattern and trip length data from trip-based travel models. Half of the respondents performed some form of simplification, either in terms of trip purposes or TAZ geography. Slightly fewer than half of passive data users processed or merged passive data with a secondary data source to facilitate its use.

### *In-Person Interviews*

The research team also conducted two in-person interviews with modeling practitioners who had used passive data to obtain further information regarding their use.

One practitioner attempted to use a passive data set to calibrate the region's travel demand model for use in a freeway corridor study and develop inputs for dynamic traffic assignment (DTA) analysis. The practitioner found that the geographic scale and time scale of the data presented challenges, particularly for DTA. The practitioner did use the passive data in conjunction with some of the travel demand modeling analysis and obtained results that indicated the data were most useful at the regional level. Based on these experiences, the practitioner suggested that research should be devoted to providing expanded methods for using passive data to supplement traditionally collected data sources. The practitioner also believed that efforts to promote the exclusive use of passive data for modeling should not be prioritized. Rather, given passive data's robustness for disaggregate modeling, tools should be developed to obtain passive data through up-front interaction with survey subjects who could proactively consent to allow the recording of their travel behavior.

Another practitioner described research comparing passively-derived OD, trip length, and assignment data to model-based trip pattern information. The work demonstrated differences between the trip tables, but results of the passive data trip table assignment compared favorably to the assignment of modeled demand. The practitioner also relayed the successful development of external (to a region) demand using passive data. As with the first practitioner interviewed,

this second practitioner felt that the most promising current and near-term future uses of passive data were as supplements to traditionally collected travel data. The practitioner stated that passive data could be used in cases where no traditionally collected data existed or where traditionally collected data were of questionable accuracy. The practitioner also remarked that research into the use of passive data for modeling target subsets of demand, such as external or visitor demand, could be beneficial to the modeling community.

## **Summary**

Roughly a third of the survey respondents reported experience having worked with passively collected data, primarily to estimate travel/trip attributes and model components, validate model outputs, or perform benchmarking. Cost and sample size were among the key benefits of passive data identified by users, but respondents also expressed several concerns. These included data accuracy, trip purpose imputation, sampling bias, privacy, and a lack of knowledge. Comments by participants suggest that further research, greater familiarity with passive data, and improved data quality will be required to overcome these concerns.

Follow-up interviews with a subset of respondents who made use of passive data in travel modeling revealed that passive data were used in the estimation of model components and the validation of model outputs for trip-based or activity-based models. The results of comparisons between passively derived and traditionally derived trip attributes were mixed, and half of the respondents found it necessary to simplify their target models to accommodate passive data limitations. In-person interviews with two respondents indicated that passive data currently lack the critical descriptive data required to meet the needs of modeling practitioners as a standalone data source. In its current form, passive data was seen as generally suitable only as a supplement to traditionally collected data.

## CHAPTER 4: WORKSHOP

### Background

A critical aspect of this research effort was a workshop among data collection and modeling researchers and practitioners geared toward the identification of future research topics and work tasks for the expanded use of passively collected data in modeling. The workshop involved the presentation of the results of the literature review and web-based survey, discussion of their significance relative to future research directions, and the identification of research topics for further development.

The half-day workshop was conducted on June 19, 2014, at the TTI offices in Austin, TX, with participation from TTI researchers with expertise in areas such as travel surveys, travel modeling and forecasting, transportation planning, and transportation data. Also attending the workshop was an MPO modeling director who had worked with passive data in modeling and had been involved at the national level in research-funding panels that review and make recommendations for modeling- and data-collection-oriented research studies. A list of workshop participants is located in the appendix of this report.

The workshop opened with a presentation from the project team members, who briefly covered the background of the project, the results of the literature review, and discussion of the documented proceedings of the FHWA Cell Phone Data and Travel Behavior Research Symposium held on February 12, 2014. The workshop then moved to a presentation and discussion of the results of the web-based survey. The remainder of the workshop was spent discussing passive data gaps and issues, and the potential for the use of passive data and other non-traditionally collected data in travel modeling. The following sections present details of these sessions.

### Workshop Sessions

#### *Project Team Presentations*

#### *Discussion of Previous Research Findings*

The project team presentations started with the description of the documented prior research efforts in preparation and enhancement of cell, GPS, and Bluetooth travel data as well as direct use of such data for travel modeling purposes. In summarizing the literature review for the participants, the project team noted that integrated or hybrid approaches that combine traditional and passive data currently appear to be the preferred method of use of passive data. These approaches are attractive for their ability to blend quantitative information required of disaggregate behavioral modeling with extensive trip data. In closing the discussion on the literature review, the project team discussed future research needs to increase stand-alone usage of passive data. Issues related to data fusion, validation of results, and the need for continuous data streams suggest possible directions for future research.

### *FHWA Cell Phone Data and Travel Behavior Research Symposium*

Some of the workshop participants had attended the FHWA Cell Phone Data and Travel Behavior Research Symposium on February 12, 2014. The next segment of the workshop correspondingly opened with a discussion of symposium topics and materials, which confirmed many of the findings of the literature review in terms of current practice and future research needs.

With regard to areas of research, symposium topics fell into two general categories:

- The collection and preparation of passive travel data for modeling.
- Travel modeling itself.

In terms of passive data, the topics dealt with:

- Capturing full or only partial trips.
- Trip chaining.
- Origin, destination, route, and mode precision.
- Evaluation of data linking and imputation methods.
- Validation of data.
- Methods for extracting behavior data.

The modeling discussion at the symposium covered two subjects:

- Balancing the use of traditionally collected data and data from new/emerging technologies in model development.
- Consideration of changes to model development methods to accommodate new data collection methods and new types of data.

### *Web-Based Survey*

The information presentation portion of the workshop concluded with a presentation of preliminary findings of the web-based survey of modeling practitioners and researchers. A review of literature<sup>3</sup> on documented research in preparation and enhancement of cell, GPS, and Bluetooth travel data as well as direct use of such data for travel modeling purposes was used to structure the data gathering of the web survey. Information regarding respondents' use and/or considerations concerning the use of passive travel data in modeling was presented. Workshop participants primarily discussed three issues:

- Workshop participants wondered if the frequent estimation of model components using passive data was an indication of its ability to be used in as a stand-alone data source. Participants remarked that many respondents did not perform any imputation or development of data items. Some believed that this could be a reflection of respondents using the data in a stand-alone fashion to support modeling needs.

---

<sup>3</sup> Technical Memorandum: Emerging Data Collection Techniques for Travel Demand Modeling: A Literature Review, Lee, Sener, Mullins, September, 2014.

- Discussing the issue of whether or not passively collected data needs to provide the same information in the same manner as traditionally collected data, workshop participants noted that passive data represent a much larger number of observations and are much more continuous than traditionally collected data. Even given potential biases, participants felt that passively collected data can be useful in overlooked areas of travel demand analysis or areas with thin or old data sets.

### *Open-Ended Discussion*

The workshop concluded with a discussion of the information presented thus far to facilitate the development of research ideas for the expanded use of passive data in modeling.

Concerning the literature review, there was discussion concerning the fact that many of the passive data being used in modeling come from third-party sources, often making it impossible to analyze raw travel data. Workshop participants also expressed doubt as to whether standardization was necessary or achievable given the way in which resellers process and repackage data before they are used by modelers. Participants mentioned that even the terms used to describe mobile phone positioning data are not standardized. What is *active* versus *passive* can mean different things depending on the technology behind the data. Participants noted that researchers working with passive data should properly document the sources and methodologies used in modeling.

The discussion of the literature review, FHWA symposium proceedings, and survey closed with a dialog about considerations when using passive data. The dialog dealt with the possibility of working with smaller data sets that offer longitudinal views of travel behavior. There were also suggestions that the modeling community should be researching the development of different modeling forecast inputs and considering new questions these data can answer. Participants discussed the possibility that the modeling paradigm will need to change to fit the different types of data being collected with mobile technology. In general it was suggested that research should focus on case studies in small-scale settings before attempting large-scale modeling. Participants also wondered why some respondents believed that passive data should provide the same level of information as traditional data sets.

### *Future Research Directions*

Several topics were brought up and discussed dealing with the fundamentals of passive travel data collection and other forms of big data. Among these was concern about the loss of access to data given increased privacy concerns, and the importance of demonstrating the utility of these data to counteract potential efforts to limit or curtail access. Alternately, some felt that privacy concerns would lessen over time. This led to a discussion about exploring other types of passive/remotely collected data, such as consumer data, as well as future innovations beyond current passive data techniques. Methods such as remote sensing or technologies that build on the developing communications technologies between vehicles and infrastructure were mentioned.

There was also discussion regarding the potential for the development of big data processing capabilities beyond the sphere of those who currently process such data. There was speculation

that future competition, potentially from the original sources of passive travel data, could offer an alternative to current data providers not currently available to end users such as travel modelers and planners. This discussion acknowledged the challenges of decreased reliance on current data providers, including the need to interact or partner with those in engineering and computer science fields.

Several issues raised during the discussion had to do with processing and using passive travel data. These issues ranged from developing a better understanding of the biases in passive travel data, to tying passive data to demographics in order to demonstrate their utility for modeling at the urban level.

Another set of topics fell into what could be termed applications research. These topics dealt with the development of tools to allow practitioners to take passive data products and derive commonly used outputs/parameters, and to focus on areas where traditional data are poorly-suited or nonexistent. Freight modeling was raised as a prime example, where a fair amount of modeling has been done but in many cases with compromises and simplifications due to the lack of detailed information. Participants felt that GPS data offered opportunities in the near-term for improved freight modeling that should be the focus of applications research. The workshop discussion on future research directions concluded that research in this area needs to be focused on finding a compromise between bigger but shallower data and modeling, which requires higher levels of detail.

Following the workshop, there was additional discussion on research needed to advance the use of passive data in modeling. In the opinion of one workshop participant—who had been involved at the national level in research-funding panels that review and make recommendations for modeling- and data-collection-oriented research studies—the most appropriate research may involve case studies, proofs of concepts, and small implementations of new data in model development.

## **Summary**

A workshop was conducted with a group of travel survey, data, and modeling researchers and practitioners to identify potential passive data research avenues and to discuss the implications of passive travel data. Following presentation of the literature review, symposium topics, and web-based survey results, an open-ended discussion was held. The discussion highlighted several potential areas of research, many of which also emerged in the web-based survey. With the development of new data sources and advances in big data and vehicle-to-infrastructure technologies, researchers expected the availability and usage of passive data to expand in the future. Conversely, some were worried about a potential loss of access to passive data due to growing privacy concerns. A common theme that also emerged from the survey was the notion that researchers should develop innovative applications for passive data, instead of simply trying to replace traditional data sources. Workshop participants also expressed interest in the development of practical tools to allow practitioners to more easily analyze and implement passive data. As was conveyed in the surveys and follow-up interviews, a greater understanding of passive data's limitations and biases will be required before many will feel comfortable in its use. In general, it was believed that near-term research should focus on smaller case studies and proofs-of-concepts before embarking on large-scale modeling.

## CHAPTER 5: FUTURE RESEARCH AND TESTING

### General Recommendations

Based on the collective findings of a literature review and survey of travel modeling practitioners and researchers, data from passively collected travel data sets are being used with greater frequency in travel demand modeling. While passive data are still not prevalent enough to be considered part of standard practice, results of research directed toward increased use of passive data in modeling suggest that these data have started to gain the attention and acceptance of practitioners. In particular, passive data are now being frequently used for modeling of trip and traffic demand. Most commonly these data are being used to model travel pattern characteristics such as origins and destinations of travel and travel time, distance, and routing.

The results of the web-based survey indicate that the modeling community is beginning to turn to passive data to compensate for some of the shortcomings of traditional travel surveys. These shortcomings include increasing cost and non-response bias, as well as the inability to provide sufficient data on hard-to-reach population groups. Passive data also offer a viable alternative to other sources of travel data such as the decennial Census, American Community Survey, and Census Transportation Planning Products. These data are either too generalized for many travel models, suffer from small sample sizes, or are not attuned to the specific needs of locales.

Survey and workshop participants were particularly interested in research and research applications that:

- Investigate and evaluate passive travel data sets for additional data item content to be used in travel modeling.
- Demonstrate the use of passive data through model applications testing.

With respect to evaluation of the ability to develop additional data items for use in modeling, research to evaluate the ability for passive data to provide data as described and listed in Chapter 4 would be of most immediate benefit, interest, and need. Such research would include evaluation of methods to address sampling, bias, and data expansion characteristics related to passive travel data.

Other research that would be of benefit would involve evaluation of the ability and methods to use passive data to model underrepresented travel demand markets such as low income, low auto ownership, and visitors. Also, passive data contain information that potentially support modeling of the variability atypical demand scenarios such as weekends, holidays, and special events analysis of the variability of regional and sub-regional. Applications research testing and case studies of modeling of this type of demand with passive data should be considered.

The pursuit of research and applications funding could come from national, state, or local partners. Based on our findings and dialog with individuals participating on national-level research panels, the identified research needs involving passive data enhancement for modeling and application of passive data for modeling are somewhat ahead of national funding availability. Many of the advancements in the development and use of passive travel data for

modeling have come at the state and local levels, with several efforts being self-funded studies by researchers or research institutions.

For this reason, a near-term strategy in this area should focus on the state and regional levels for applications research opportunities, particularly via existing relationships with departments of transportation (DOTs) and MPOs. The results of this project suggest that modeling and data collection professionals at the state and local levels need timely, inexpensive, and robust data sets for use in the maintenance and development of travel models, and procedures to support their travel demand analysis needs. In particular it will be important to focus on case studies, proofs-of-concepts, and small implementations of new data in model procedure development. On a longer-term scale, researchers should seek opportunities to assist DOTs and MPOs in the investigation of continuous travel data collection programs featuring passively collected data in conjunction with, or as a supplement to, large-scale episodic travel survey programs.

At the national level, NCHRP studies (NCHRP 8-36, NCHRP 8-89, and NCHRP 8-95) and a recent TRB symposium represent federally funded efforts to evaluate prospects and a future for increased use of passive data in transportation planning and travel modeling. Researchers should look to the results of the completed studies and closely follow on-going such as NCHRP 8-36 and 8-95 because they could establish future research directions at the national level. In anticipation of expanded federal research opportunities, researchers should continue to present their work at national conferences and establish strong relationships with organizations such as TRB and NCHRP.

## **Research Areas**

Based on this research effort, two priority research areas were identified: evaluation of passive data for modeling and targeted modeling applications testing of passive data.

### *Evaluation of Passive Data for Modeling*

#### *Topic A: Identification and Evaluation of Methods to Obtain Demographic Characteristics of Smartphones Users*

This research would focus on the need for demographic and personal information of travelers. The effort involves identification of cost effective methodologies to obtain demographic data directly from smartphone users to link with the travel data flowing from the smartphone itself. Included in this research would be an evaluation of factors, methods, and incentives that would motivate smartphone users to provide and/or allow collection of household and personal demographic characteristics. The efforts would include the following work.

- Identification of variables that are desired to obtain from smartphone users.
- Evaluation of techniques to obtain desired variables, including incentive-based techniques.

*Topic B: Merging of Secondary Demographic Data Sources to Mobile and Smartphone-Based Passive Data*

A companion research topic to Topic A above would be evaluation of methods to merge cell and smartphone data with data from secondary demographic sources. In a limited fashion, linking of trip origin and destination and routing data with secondary data sources such as census data has been accomplished. This research would seek to identify additional data sources and evaluate their content for suitability as inputs to travel demand models. The efforts would include the following work.

- Inventory potential sources of data for merging with passive travel data.
- Identify variables among the data sources that would provide data for use in travel modeling.
- Investigate and develop methods for merging data with travel behavior data collected from mobile/smartphones that are most appropriate for modeling purposes

*Topic C: Use of Automated/Connected Vehicle (AV/CV) Technology to Supplement Mobile/Smartphone Travel Data*

Even in the early stages of AV/CV presence, there may be data items that would provide inputs related to the purpose and amount of travel and the relationship of trips to each other, such as:

- Inventory of travel demand related data items among current AV/CV implementations
- Identification of specific data elements that would be pertinent for use in travel modeling.
- Identification of methods to extract and merge AV/CV data items with mobile/smartphone travel data for use in modeling.

*Topic D: Passive Data Sampling Confidence*

Sample sizes and sampling rates that are used in traditional survey data collection are based on desired confidence intervals for modeling data items such as trip rates, trip time and distance, and the times-of-day of travel. There is currently no basis for knowing or establishing statistical confidence for modeling data items built with passive travel data. This research would investigate a replacement or substitute for confidence intervals for modeling data items created with passive travel data.

*Model Applications Testing of Passive Data*

The following research topics provide opportunities to take advantage of some of the inherent benefits of passive data. The topics primarily target aspects of travel demand modeling that are not well-served by traditional data collection methods, and thus are underdeveloped or almost completely lacking in capable tools. This research would focus specifically on applications using case-study or proof-of-concept approaches.

*Topic A: Use of Cell Phone-Based Travel Data to Develop Estimates of Weekend Travel Demand*

The research would seek to evaluate the utility of using cell phone travel data to adjust weekday travel model-based estimates of travel to represent weekend travel demand. The efforts would include the following work.

- Investigation of the variation in travel demand among weekday and weekend travel of cell phone-based travel data.
- Identification of common variables among travel model and cell phone data set for use in adjusting travel model to represent weekend travel.
- Development and evaluation of methods to use common variables to adjust model-based travel using cell phone data metrics to represent weekend travel.

*Topic B: Evaluation of the Use of Passively-Collected Data in the Development of Visitor Travel Demand Models*

This research would evaluate sources of passive travel data to determine their potential suitability for development of visitor travel demand through the following actions

- Identify travel data variables that would be needed to model visitor travel.
- Evaluate passive travel data sources and determine availability and usefulness of variables within these sources to develop weekend travel estimates.
- Identify and evaluate techniques to incorporate such variables into current travel models or to develop new models of weekend travel.

*Topic C: Evaluation of Cell Phone Data for Long Distance Models of Individual Travel*

Long distance travel models are difficult to build because of the dearth of behavioral data. This research will investigate the suitability of cell phone data for use in estimating and modeling long-distance travel by individuals through the following efforts.

- Identify cell phone travel data variables that could be used for estimating the amount frequency and routing of long-distance travel.
- Evaluate modeling techniques that could make use of these data items for estimating inter-regional travel.

*Topic D: Developing Future Year Travel Forecasts with Mobile Phone Data*

This research would seek to evaluate the feasibility of growing/forecasting observed mobile phone-based travel data by performing the following activities.

- Obtain mobile phone based origin-destination trip matrices for a region or major travel corridor within a region as well as any background characteristic data for this demand.
- Identify variables in mobile phone travel data sets for which future year forecasts exist or could be developed.
- Identify methods for developing and/or obtaining forecast values of variables from mobile phone-based travel data set for a forecast year for which travel model-based data is available.
- Apply methods for developing forecasted values to observed mobile phone-based travel data to create forecast trip tables.
- Evaluate performance of methods for forecasting mobile phone-based trip tables and traffic volumes.

## CHAPTER 6: CONCLUSIONS

This multi-phase project examined the current state of passive data usage in travel demand modeling from the perspective of researchers and practitioners, and put forth recommendations for future research directions. The study began with an initial literature review that identified recent studies and applications assessing the capabilities of passive data collection technologies, including mobile phone positioning, GPS, and Bluetooth. In the second phase of the project, a web-based survey was developed to characterize the views of transportation professionals regarding the use of passive data, and a workshop was held among TTI researchers to present preliminary project findings and to discuss possible research directions. Overall, this research provided several insights:

- Findings from the literature revealed much research into the enhancement of passive data collection methods, but more limited research in the enhancement of passive data for use modeling and applications of passive data in modeling. What applications-oriented research that was identified involved use of passive data to either supplement the modeling of trip patterns and characteristics, or expand passively obtained travel and traveler data characteristics. Although passive data are often unreliable as a standalone data source, hybrid approaches that integrate passive data with traditional travel data appeared to be especially promising.
- Feedback received through the survey, interviews, and workshop efforts indicated that many researchers and practitioners see passive data as having a valuable supplemental role in travel modeling, but are also aware of its shortcomings and limitations. Many also recognized the usefulness of passive data, independent of traditional travel data, given their superior robustness in several dimensions.

Findings from the literature review, survey, interviews, and workshop formed the basis for recommendations for future research avenues and partnerships. In general, data and modeling practitioners identified two areas of basic research and applications research for increasing the use of passive travel data in modeling:

- The first area focuses on evaluation of modeling data items for which passive data are a potential source. Several research areas evaluating the use of passive travel behavior data to replace the use of traditional survey data in the development of model components were identified. Expansion of available data items, either through the use of technology linked to passive data devices or through linkage to third-party commercial data sets, was seen as an additional area in need of research.
- The second area of research assesses the usability and utility of passive data within the current travel modeling context. Many modeling practitioners, wary of implementing these new data sources, expressed a desire for successful examples of applications-oriented research. Such research could also serve to enhance the data-gathering and data-processing techniques of third-party data collectors. In addition, the general consensus from the workshop and survey was that applications research should demonstrate the capability of passive data to model aspects of travel demand that are currently unserved or underserved by current modeling with traditional travel data.

## APPENDIX: LIST OF WORKSHOP PARTICIPANTS

	Name	Affiliation	Program Name	Contact Info
Project team members (TTI)	Andy Mullins	Research Engineer	Travel Forecasting Program	a-mullins@tamu.edu
	Ipek Sener	Assistant Research Scientist	Travel Forecasting Program	i-sener@tti.tamu.edu
	Rich Lee	Student Researcher	Travel Forecasting Program	r-lee@ttimail.tamu.edu
Workshop participant (Atlanta Regional Commission)	Guy Rousseau	Program Manager	Travel Surveys & Transportation Model Development	GRousseau@atlantaregional.com
Workshop participants (TTI)	Byron Chigoy	Research Specialist	Transportation Planning Program	b-chigoy@ttimail.tamu.edu
	Chris Simek	Assistant Research Scientist	Public Engagement Planning Program	c-simek@tamu.edu
	Ed Hard	Research Scientist & Program Manager	Transportation Planning Program	e-hard@tamu.edu
	Jim Benson	Senior Research Engineer	Travel Forecasting Program	Jimmie-Benson@tamu.edu
	Kevin Hall	Research Scientist	Travel Forecasting Program	Kevin-Hall@tamu.edu
	Mark Ojah	Research Specialist	Transportation Planning Program	m-ojah@tamu.edu
	Michael Martin	Assistant Transportation Researcher	Transportation Planning Program	m-martin@tti.tamu.edu
	Praprut Songchitruksa	Associate Research Engineer	System Reliability Division	Praprut@ttimail.tamu.edu
	Shawn Turner	Senior Research Engineer & Division Head	Mobility Division	Shawn-Turner@tamu.edu
	Stacey Bricka	Research Scientist & Program Manager	Mobility Management Program	s-bricka@tamu.edu
	Tom Williams	Research Scientist & Program Manager	Travel Forecasting Program	t-williams@ttimail.tamu.edu