Summary Report for the Peer Exchange on Data Transferability

Held December 16, 2004
Washington, DC
Keck Center of the National Academies

Sponsored by:
Federal Highway Administration Travel Model Improvement Program and

TRB Committees on Urban Transportation Data and Information Systems (ABJ30), Traveler Behavior and Values (ADB10) and National Transportation Data Requirements and Programs (ABJ10)

Prepared by the Volpe National Transportation Systems Center
U.S. Department of Transportation

Table of Contents

I. Introduction

The following report summarizes the results of a peer exchange on data transferability organized and sponsored by the FHWA Travel Model Improvement Program (TMIP) and co-sponsored by the TRB Committees on Urban Transportation Data and Information Systems (ABJ30), Traveler Behavior and Values (ADB10), and National Transportation Data Requirements and Programs (ABJ10). The exchange brought together representatives of state and local departments of transportation (DOTs), metropolitan planning organizations (MPOs), academics, and transportation consultants. It consisted of presentations on data transferability topics, followed by a discussion of data transferability issues structured around a set of questions prepared prior to the meeting. A common theme throughout the day was that a great deal of further research is needed to progress the field of data transferability so that more widespread data sharing can occur. To this end, the group produced a set of research topics that would be beneficial to data transferability, shown in Appendix A.

II. The Travel Model Improvement Program Peer Exchange

The Travel Model Improvement Program (TMIP), begun in 1992 under the sponsorship of the U.S. DOT and the Environmental Protection Agency, provides resources to planning agencies in their efforts to improve their travel analysis techniques. TMIP has three goals:

1. Help build the institutional capacity of planning agencies to perform technical analysis,
2. Support development of analytical methods that respond to the needs of planning and environmental decision making processes, and
3. Support mechanisms to ensure the quality of technical analysis used to meet local, state, and federal program requirements.

The goal of the current peer exchange was to lay down the issues associated with data transferability for travel demand modeling and identify ways the FHWA can help improve the transferability of data across the country. The gathering, held on December 16, 2004 in Washington, DC, brought together fifteen people with experience with or a need for transferable data. Tom Rossi of Cambridge Systematics moderated the peer exchange, assisted by three representatives from TMIP. A representative from the U.S. DOT Volpe National Transportation Systems Center assisted in documenting the exchange proceedings. A list of all participants is provided in Appendix B.

To begin the meeting, four participants made presentations on their past and current research relating to data transferability. These presentations can be found on the TMIP web site at http://tmip.fhwa.dot.gov. The presentations were followed by a panel discussion of issues relating to and affecting data transferability. At the end of the day, the group developed a set of research topics that the FHWA could pursue to enhance knowledge about data transferability.

The peer exchange had three objectives:

1. To present existing research on data transferability
2. To discuss issues in data transferability
3. To write scopes of work for research which would help progress the field of data transferability

III. Background on Data Transferability

This peer exchange assembled experts in travel demand modeling and data collection and analysis to discuss and debate the issue of data transferability for four-step travel demand models and activity-based models. The term “data transferability” is used loosely to refer to using data (or products derived from data such as models, equations, and parameters) for modeling situations other than those for which the data were collected. Data transfers can be used to forecast into the future (temporal transfer) or can be applied to another region (spatial transfer). It might mean applying nationally collected data to local areas, applying regional survey data to other regions, or using data collected in the past to model the future.

Data transferability is important because primary data collection for household surveys for travel demand modeling is expensive and time consuming. Planning agencies are eager to minimize their costs by borrowing or transferring data from other regions, to the extent possible. Also, emerging practices like activity- and tour-based models and micro-simulation approaches require additional and more specialized data on travel and activity patterns.

The primary goal of this meeting was to discuss ways that data collected in one region can be used for travel analysis in other regions. A broader objective was to identify an action plan for research and develop project scopes for advancing the concepts of data transferability. This, in turn, will support improvement of travel models; facilitate the implementation of activity, tour-based and micro-simulation models; and stretch limited data collection resources. Another objective was to solicit input for the National Household Travel Survey (NHTS) Transferability project to be conducted by Oak Ridge National Laboratory and to review the work of others in the field.

IV. Presentations

National Household Travel Survey Data for Benchmarking and Transferred Data
Nancy McGuckin

Nancy McGuckin is one of the authors of *NCHRP Report 365, “Travel Estimation Techniques for Urban Planning.”* This report reviews the fundamentals of the four-step travel demand process and provides transferable parameters for use by travel demand modelers. This publication updated *NCHRP Report 187, "Quick Response Urban Travel Estimation Techniques and Transferable Parameters,"* which became obsolete when microcomputers replaced the large computers previously used in travel demand modeling.

The team that updated the NCHRP publication was tasked in 1994 to look at existing work in travel demand modeling by surveying planning agencies, reviewing existing manuals, and

collecting other relevant data. The team discovered that there exists a tremendous amount of difference in the way surveying and modeling was done from one model to the next. For example, definitions varied and weighting techniques were different. Thus, the NCHRP 365 team determined that it was best to use the 1990 National Personal Transportation Survey (NPTS) for calculating household trip rates by metropolitan area size and trip purpose, vehicle occupancy and trip purposes by time of day. *NCHRP Report 365* outlined step-by-step the best practices for four-step travel demand modeling, and walked the user through a case study by transferring parameters, many of which derived from the 1990 NPTS.

The target market for data transferability is small areas that are not able to collect their own data. However, it is believed that even large areas are not collecting the data that fully explain travel, evidenced by the fact that models are still able to explain only about half of the differences in, for example, household trip purposes. More research must be conducted before data transferability can be fully understood and implemented. The first step in this research is to collect data on more variables, which would allow for study of additional relationships in travel demand modeling.

Can Nationwide Personal Transportation Survey Data be Used by Small Communities?
Pat Hu, Center for Transportation Analysis - Oak Ridge National Laboratory

The NPTS is a national survey. Using it indiscriminately to represent an area smaller than the U.S. can result in unreliable estimates. To address this weakness, Oak Ridge National Laboratories (ORNL) developed a system to use NPTS data to estimate travel behavior for smaller areas.

The method uses predictor variables (variables shown to have significant explanatory power) for daily household trips, income limits, and area type classifications to form homogenous census tract clusters. The census tract data can then be used to bridge to the NPTS data using household size and number of vehicles. Vehicle trips, vehicle miles traveled, person trips, and person miles traveled for each cluster can then be calculated using NPTS estimated rates for household size and number of vehicles.

This census tract clustering approach was tested using the 1995 NPTS and applied to four regions and states, New York, Massachusetts, Oklahoma, and Baton Rouge. The tract-level results were compared to NPTS add-on data for each location or to independent survey data. For these particular add-on surveys, the performance of the ORNL census tract method was better than the competing methods that used larger geographic definitions such as metropolitan statistical area (MSA), census division, or census region to differentiate household travel demand.

The tool they have created for calculating cluster-based estimates is available at http://nhts.ornl.gov/2001/index.shtml.

The new ORNL project will use the 2001 NHTS dataset and see if the results are similar and if they can be improved.

Travel Related Inputs Model for Mobile 6.x (TRIMM)
Mohan Venigalla, University of Illinois at Chicago

Dr. Venigalla presented his research on using NPTS data as inputs to MOBILE6. When modeling regional emissions, modelers prefer actual local data over default MOBILE6 parameters because of the important implications for air quality attainment status. The team at George Mason University was able to derive some of the local variables from NPTS that MOBILE6 requires. The subsamples are limited to large geographic areas, such as census region, MSA size, and state (if sample size is sufficient). It found that in general, NPTS data look very similar to MOBILE6 default data. However, there are significant regional variations in several variables of the NPTS that have been found to affect emission factors. Because some data gaps were found and data analysis was complex, Dr. Venigalla's team created a tool, Travel Related Inputs Model for Mobile 6.x (TRIMM), which automates the process of using NPTS data to generate MOBILE6 inputs.

TRIMM mines the NPTS data and provides as many MOBILE6 inputs as possible from the data. It then allows the user to compare each variable to those of other geographic regions and to MOBILE6 defaults, and gives the user the capability of choosing which variables to write to the MOBILE6 input format.

It was noted that NHTS data do not have location or distance information, so GPS survey information would be extremely useful in the next NHTS survey.

Data Transferability: Idealism or Realism?
Kouros Mohammadian

Dr. Mohammadian presented the results of his literature review on data transferability.

The main goals of transferability are:

- To use data collected in one context in a new context.
- To reduce or eliminate the need for a large data collection effort for model development in the application context.

Dr. Mohammadian's literature search found two existing methods for transferring travel data. The first group of studies addresses the transferability of travel demand models either in a spatial context or temporally. A second group of studies addresses transferring of data, of which the Institute of Transportation Engineers (ITE) trip generation data are the most commonly transferred. Both of these methods involve aggregate data transfer.

Dr. Mohammadian and colleagues at the University of Illinois at Chicago (UIC) are researching a third method of data transfer—data simulation. Data simulation can be used at a local level where NHTS and NPTS sampling are small and inadequate. Data simulation combines local socio-demographic data for individuals and households from sources such as a census with probability distributions of activity and travel patterns from other travel surveys such as NHTS to simulate local travel survey data. Essentially, data simulation expands any existing data (with all the inherent biases) to a wider population. The team at UIC plans to develop a framework to

facilitate transferability of household survey data for calibrating and validating travel forecast models. To achieve this goal, they will be researching more detailed classifications of NHTS and census data using advanced clustering schemas since homogeneous groups will improve the simulation.

V. Peer Exchange Discussions

Topics for discussions were developed by the FHWA and prioritized by the panel members prior to the meeting. Thirty minutes were allotted for discussion on each topic.

Question 1

What are your ideas about data transferability for travel demand/activity models? How do you define "transferability" spatially and temporally?

Panelists agreed that there are several **"layers" of transferability:**

- A conceptual layer, which consists of the modeling structure or mechanisms
- The parameters layer
- The outcomes layer (e.g. trip rates)

Data are typically transferred because agencies lack resources—money or skills—for adequate data collection for developing models unique to local circumstances. However, data transfer also requires resources. Unfortunately, variables are often transferred with little understanding of the limitations or insufficient analysis. Although it is always possible to transfer data from one model to another, it may not always be technically valid to do so. It is not clear how often inappropriate transfer of data occurs.

Models vary across regions and, subsequently, so do their data requirements. Smaller areas may only require **aggregate numbers or average values** for transfer, such as aggregate trip rate, while other models—typically those done for large urban areas with complex tour-based models—require extremely disaggregated data. Planning agencies in larger regions are particularly concerned with whether the determinants of **travel behavior** (e.g. household composition, disposable income, proximity to shopping opportunities) **are really being captured** and **whether models from different areas are similar enough for transferability** to and from other regions. Although data such as trip rates, trip generation, and process tables are frequently transferred, there is little research that shows that such transferability is actually valid.

Participants felt that data transferability guidelines would be helpful for the entire travel demand modeling community for preventing technically invalid data transfers while encouraging proper data transfer. **Standards for transferability of data** would lay out criteria and guidelines on what data are transferable, define a correct method to conduct data transfer, and provide a method for measuring whether data transfer was performed successfully and correctly (beyond data matching).

Some areas of travel demand modeling and data transferability will require more research before a "correct" or recommended method can be identified. One area that needs more research is

identifying **pre-conditions** that signal the appropriateness of data transfer, e.g. the existence of data on the availability of transportation systems and alternatives or transportation system measures, such as data on transit or pedestrian friendliness.

Some participants related their own experiences with models as anecdotal support of the concept of model transferability. For example, a participant who is a consultant to many large MPOs said that from his experience in several cases two models or regions that initially look substantially different, and therefore do not appear to be good candidates for transferability, actually become sufficiently similar or almost identical when properly scaled. An important task for the travel demand modeling community is to **validate the current state-of-practice of data transferability, both spatially and temporally**. This is discussed in more detail under question #3.

Research that helps to explain travel behavior will also facilitate data transferability analyses. For example, it maybe possible to **determine whether certain variables or types of variables are more transferable than others**. Some participants speculated that variables which dominate travel behavior may be more transferable that other variables. This is discussed in more detail in question #3.

Research should also be conducted to determine **whether the model constant is capturing some of the regional characteristics.** If so, explicitly breaking the characteristics out would facilitate transferability. Such research would require combining data from multiple regions, which is discussed under question #7. Participants suggested adding the following variables to capture regional variability: seasonal weather, condition of the transportation system (e.g. level of congestion, road condition), and land use variables (e.g. distance to employment, entertainment, and retail stores).

Land use variables are very important in travel demand modeling. However because regions typically have different definitions of the variables, they are generally not directly transferable. The first step to capturing the differences is to introduce quantifiable variables into the model rather than using central business district (CBD) dummies.

Travel demand modelers are generally interested in transferability of three different types of variables:

- Headline variables – variables which are easy to categorize, such as household size
- Underlying variables – household composition, presence of children of a certain age
- **Variables which represent hypothetical measurements of demand** – a household score representing maintenance required for children

This third type of variable requires more work to generate. However, when identified and defined correctly, it can be highly explanatory and capture a great deal of travel behavior. If these hypothetical measurements of demand are properly identified, it is possible that, in addition to improving travel demand modeling across the board, models would become more transferable.

However, modelers should consider the utility of obscure variables since they must be available or be estimated for the region or time period where the transferred model will be applied.

Therefore, the additional **variables should be limited as often as possible to universal variables, particularly those that are obtainable through census data**.

Question 2

For which types of applications does data transferability already occur, and how has the transfer been achieved? Was it successful? What applications have data/parameters that are not typically transferred currently, but might be difficult to estimate originally due to future data limitations? What are some new and different applications for which transferability of data/parameters might work?

Data transfer is often applied when forecasting the effect of **policies such as carpool lane requirements or infrastructure such as new modes**. For example, **pricing and toll** policies are being considered in most new travel demand models, even though very few cities have real pricing data. Although it is useful to transfer data from other regions for modeling these new effects, it is important to also collect local stated preference data to capture regional characteristics.

Another prerequisite for successful data transfer is that the **source data set and the target data set be comparable**. To determine if two data sets are comparable, one should combine the data sets and perform usability and reasonability tests, such as testing whether a variable works the same way before and after the data set combination. Although one goal of transferring data is to save money, agencies must invest some resources into **analyzing data before transferring** it to another application. One common mistake is to overlook scaling of the data. This could involve, for example, using one definition for low, medium, and high population densities in one region but different definitions in another region. For example, 2000 persons per square mile in Manhattan might be low density, but this same value might be high density in Kansas City. This becomes even more difficult for qualitative variables such as "pedestrian accessibility." It would be beneficial to have an outline of some **basic requirements for testing data comparability.**

A new application of transferability is to the **emerging market of statewide models.** Statewide models will require additional data collection on rural, small urban, and long distance trips, creating additional opportunities for data transfer between states and regions. Subregional models are also growing, although they will require analysis methods that can be done with minimal resources.

Another important potential use for transferred data is for **air quality modeling**. This typically uses a great deal of data from state departments of motor vehicles (DMVs). One participant suggested validating the DMV data with NPTS data and reconciling any discrepancies.

The participants noted that very few agencies have the resources to conduct a household survey every ten years, and most small agencies are lucky to have the opportunity to conduct any surveys at all. To get a sense of what data are currently available, participants were asked to list household surveys of which they are aware. Table 1 contains the list of surveys.

Table 1 - Agency Household Surveys

Agency	Last household survey	Comments
Capitol Region Council of Governments, CT	1976	Trip generation model adjusted based on 1976 data collection effort.
Chicago	1970 and 1990	Would require $20M to do another survey
Iowa DOT	early 1970s (Des Moines add-on)	
Salt Lake City	1970s and another more recent one	
Oregon DOT	1996	
New York Metropolitan Transportation Council	1963, 1995 and 2001 NPTS add-ons, 1996, and one scheduled for 2005/6	
North Central Texas Council of Governments	Dallas-Fort Worth home interview in 1964, true household survey in 1984 and 1996, next one will be 2007/8	
St. Louis	Every 10 years	
San Francisco	Surveys regularly on the scale of 15,000 households recent one, 1990, 1981, 1965	
Montreal	Every 5 years	
Tampa	1996, 2000/1, 2005/6	

Question 3
What types of data/parameters can be transferred or should not be transferred?

Some participants felt that, although **temporal transferability is currently used regularly, its validity has not been sufficiently studied** by modelers. A **controlled study of temporal transferability** would help the industry learn how to model temporal changes such as increases in trip rates. Trends over time should also be analyzed to determine whether the context of the data is changing. For example, changes in **household characteristics and structure** (e.g. change in the number of multi-worker households, smaller households, increase in the amount of eating out, extra stops for coffee) may not be captured in existing household survey data. Large but unmeasured changes in the context will affect the temporal transferability of the travel behavior modeling. This type of study is difficult to perform retroactively because one cannot go back in time to collect additional information from previous years. These studies need to be planned well in advance.

Disaggregate trends also affect temporal transferability. The reasons behind changes in travel behavior are important to modeling. Detailed travel time expenditure and travel time budget information is available in a paper by Polzin, Chu and Toole-Holt, “Case for Moderate Growth

in Vehicle Miles of Travel A Critical Junction in U.S. Travel Behavior Trends" available at http://nhts.ornl.gov/2001/articles/moderateGrowth/moderateGrowth.html. In this paper, the authors use NPTS and NHTS data to show that there is a 1.9 minute per day annual increase in travel time. Questions such as whether people are spending more or less time traveling are complicated by the "multi-tasking" characteristics becoming increasingly common for travelers. For example eating or talking on the cell phone while traveling may make travel time less of a disutility. It has been shown that non-work travel has been increasing steadily while work travel has stayed fairly constant. The growth in vehicle miles traveled (VMT) per person has been linear, but this may change as speeds in urban areas deteriorate.

Temporal transferability is also affected by the **changing urban form**. Specifically, low density areas are beginning to form their own structure. Some of this effect is currently being captured in travel models in the density variables, but the variable also needs to be interactive. If this trend continues, these low density areas will depend less on large urban areas.

The most difficult component to transfer is the destination choice model due to the high level of calibration needed. If one were to **rank components of the travel model by their ease of transferability**, trip production and tour-generation seem easier to transfer (as long as urban service variables are available) due to the ease of calibration. The second most transferable component would be the time-of-day choice model, as it does not require geographic constants.

Panelists all believed that there are probably **key or core variables** in travel demand modeling that are transferable and that there are **context sensitive variables** that are less transferable. However, it is not yet clear which variables belong to which group nor which variables can be transferred in a valid manner. Some participants suggested that for the time being, until further research is performed, the **most stable variables could be transferred with reasonable confidence**. This would allow certain steps of the modeling process, such as trip generation, to be transferred so that resource-limited agencies could concentrate on other modeling issues. An example of a variable which has proven to give good results is the rule that in-vehicle time is about half of the value of out-of-vehicle time. A great deal of existing data is available, allowing the transferability of this variable to be tested.

One variable of particular interest is travel time reliability. Some attempts have been made to better understand travel time reliability, but a more solid understanding of the effect of the variable would immediately add explanatory power to models and make them more transferable.

There are currently some **elements for which modelers have little understanding and which are difficult to transfer**. For example, the observed positive travel distance utility in Los Angeles is probably unique to this region. Also, cross-cultural variables currently are not very transferable because not enough is known about which variables are the most explanatory, such as whether a speaks English, the number of years a person has been in the U.S., the level of assimilation, and location of residence. An intrametropolitan variable is also most likely not transferable. Finally, constants should not be transferred as they represent factors that the data do not explicitly explain.

The participants also discussed the **Federal Transit Administration (FTA) New Starts guidance on travel demand modeling**, which can also be considered a data transfer**.** Participants emphasized that since modelers everywhere are attempting to adhere to these guidelines, it is important that they contain valid ranges for coefficient values and that the context for these guidelines be clearly conveyed by FTA to the modeling community. Examples of these guidelines are: in-vehicle time must be between the values of -0.2 and -0.3; auto and transit highway in-vehicle coefficients must be the same; or the acceptable in-vehicle to out-of-vehicle time ratio must be between 2 and 3.

Auto occupancy values are not transferable from area to area, unless the system configurations are similar (e.g. both cities have similar employer sponsored vanpool activity). An auto occupancy equation or model may be more transferable. A good model would account for the differences in configurations and allow for transferability. Large MPOs use mode choice instead of auto occupancy, with auto occupancy used for carpooling. Capturing auto occupancy by time has been found to be valuable. An alternative to transferring auto occupancy in a trip assignment model would be to include "carpool" separate from "drive alone" and "transit" in a mode choice model and treating high-occupancy vehicles (HOV) as a sub-division of mode choice, since it has been shown to be nested with transit (as it is a strong alternative to transit). With this model structure, in regions with transit, transit will be used more heavily. In regions without transit, then a high HOV usage will be forecast.

Question 4

How are data/parameters transferred in current applications? What are correct methods for data/parameter transfer?

In most cases, agencies simply borrow variables from the best source they can find. At this point, it is still not clear how to determine whether it is acceptable to transfer data without any supporting data. However, there are some general guidelines for good data transfer.

Ideally, modelers should **have data from both the source and recipient areas** to determine the suitability of the data transfer for each specific case. It is very important that there is a basic understanding of the source and recipient circumstances before transferring data, such as understanding of what type of errors are associated with the source data. Before transferring out-of-vehicle time, modelers should also carefully examine the way it is specified in the source and recipient models, as there are many different ways that it is specified.

Modelers should also perform a "goodness of fit" test to determine whether data can be transferred. However, a goodness of fit test is not sufficient on its own since a simple goodness of fit does not mean that the coefficient makes sense. The sign and value of the coefficient should also be reasonable. Verifying the reasonableness of assumptions, parameters, and results is important to any travel demand modeling process, especially when transferring data. To aid users in identifying transferability, it would be beneficial to come up with a set of supplementary model specification tests for transferability.

Two camps of opinion exist for how to model **travel cost**. One camp believes that perceived data should be used for travel cost as this is what the traveler uses to make decisions. The second

camp believes that perceived data is not consistent enough for modeling since people's perceptions are often wrong. Therefore, this second camp believes that travel costs should be calculated using network data. These two methods of modeling travel cost could result in significant model differences.

Transferability between the four steps within a single model is currently assumed, but this issue requires further research. For example, the North Central Texas Council of Governments (NCTCOG) found that the value of time was different for transit and toll users and therefore must be calculated separately. It believes that people with toll tags are not thinking about the cost of the trip and neither do people who use a monthly transit pass. Therefore, the derived value within NCTCOG's model that each vehicle-hour is worth $10 is not a true value of time and should not be transferred to another model without understanding its origins. A clarification of the different values of time would be a good first step. For example, a $1 transit fare is not the same as a $1 toll because sensitivity to travel time and cost is not the same for all modes.

Agencies should be careful when transferring results from models written using different software. The **software being used** for modeling can affect the resulting value of the coefficients. For example, some software does not consider fares when performing the skims in the path modeling process. EMME/2 allows for specific boarding penalties that it uses to compute the cost of a path. Therefore, given the same network inputs, different software may produce different path assignments. This leads to the promising research question of whether a model could be adapted for use on several different software programs. An often debated topic is whether a **universal model** could be created. Such a model would require that the exact same procedure be applied to different sets of data, including the skims. A universal model could be applied to multiple regions and answer many transferability questions.

Data are also sometimes transferred as distributions instead of averages. This method would not work for four-step modeling, but could be used in simulations and in an activity-based model. For example, the variable "free parking," where some people are reimbursed for parking and others have to pay for parking, is better represented as a distribution than as an average.

In general, average **trip rates are used for attraction and production models.** Most attraction models assume that attractions and productions for different retail businesses are similar. However, empirical evidence suggests that there can be large differences in trip rates depending on the nature of the business. A suggestion for accounting for this difference is to use the North American Industry Classification System to classify commercial destinations.

Special generators are also sometimes transferred, but they too are not well understood. Smaller areas can use special generators because it is easy for them to deal with a few special cases, but large areas need to avoid complicating the model with too many special generators.

Question 5

What are the implications in using transferred data (e.g. need to use same input variables)?

Data itself are not generally transferred; rather, the estimates generated from data are transferred. Currently, **data are not transferable because most data do not account for the context in which they are gathered.** For transferability to be successful, modelers must understand the context in which transferred data were gathered and the context in which models and parameters were estimated. Lacking an understanding of the context, data are often transferred incorrectly or inappropriately.

Context may account for a large proportion of information. For example, no travel data can be transferred without network O-D information. Even differences in the survey design will affect the data. Without context, utility functions for the data cannot be developed and the data cannot be weighted since the weighting factor is part of the survey records. The **data generation process could be standardized** to include the required context so that data from different regions can be pooled and exchanged. Standardization should include standardization of the surveys, building of the skims, the preferred sources for land use data, and how the data are to be summarized.

Question 6

What can be done with existing datasets such as NHTS, past travel surveys, etc.? What new data should be captured in future surveys and data collection efforts?

A list of "missing mysterious variables" is needed so the NHTS can consider collecting additional variables next time around. The following variables were discussed:

- Anything that captures the information process. For example, some research in Europe is underway on a "social network" effect that is not seen or included in the data. European researchers are finding that there is a convergence on personal propensities but also that people are affected by how and what everyone else around them is doing. Behavior may be dependent on not only traditional demand modeling variables, but also on what the person's circle of acquaintances is doing. Are there variables that could capture some of this "social network" effect?

- Variables that capture the fundamental decision making process: the hows and whys of a decision. How did you come to this decision, why did you choose this mode for the trip? How did you choose the path? Why did you choose the path that you did? What were the most important factors in your decision? This type of information could be obtained by asking people to rank cost and time as factors in their decision. This would ideally not be stated preference data but process data about revealed preference for specific trips. This information could help to better understand the differences between how households make choices compared to how individuals make choices.

- Sequence questions: Which decision was made first: destination, TOD, mode, etc?

- Technology information (internet)

- Data on the difference between stated preference and actual travel behavior.

- Variables that could relate home-based-work trips to other trip purposes, since-home-based work data are available in census data. This is a topic that should be researched.

- Auto occupancy for trip purpose. Occupancy rates in a region can be gathered from state accident reports and adjusted, but it is difficult to estimate occupancy based on trip purpose. Currently, auto occupancy by trip purpose exists pretty uniformly in NCHRP 365, which was calculated using 1990 NPTS results. An improvement for the next generation of surveys would be to mark transit availability into census tracts so that it would be easy to model transit availability.

Further **research is also needed on how to model non-existing alternatives or alternatives that are physically available but not actually practical alternatives,** such as a transit trip requiring multiple transfers. One suggestion on how to capture the real-world viability of alternatives when performing the travel survey is to use the following method:

1. Perform behavior data collection; find out why the traveler made the trip using the method he/she did.
2. Quickly shift the question to ask "Would you have rather done it another way?"

This would capture which alternatives were actually considered by the person in the decision making process.

One caveat is that process information needs a process model, which does not currently exist. The NHTS cannot gather the data needed for process models until the models are sufficiently developed.

Question 7

What types of models could reasonably and correctly be applied using default national parameters? What information would be needed to estimate these national parameters?

A useful exercise would be to **identify three MPOs with good existing survey data or to gather the exact same data**, collected the same way, for three MPOs, and then perform the skims to calibrate a sophisticated four-step or activity-based model. This would allow for analysis not only of which parameters are consistent at a national level but of whether the right questions are being asked for data transferability. Agreeing on the questions, contract management, and network coding for the participating MPOs would be a difficult task. All decisions would have to be agreed upon at each step. Fewer participating MPOs would obviously make this easier. This project may be ideal for smaller areas as it may be possible to collect fewer data than would be needed for a single MPO to perform a complete survey.

Although four-step models are easier to estimate, **activity-based models may have a better chance of being transferred**. Unfortunately, small areas cannot support activity modeling and most areas are not ready for it yet. Therefore, it would be wise to design a survey to allow either model to be estimated from the data. If it can be shown that activity-based models are more

transferable between regions, the movement towards activity-based models may pick-up. However, rigorous testing of activity-based sensitivities is necessary to ensure that activity-based models are sensitive and logical in the same way that the four-step model is.

Appendix A. Research Topics/Scopes

An important output of the peer exchange was a set of research topics and scopes for advancing the concepts of data transferability. The research topics were generated through group discussion. Subgroups of panelists developed the detailed scope for each topic. Scopes for the research topics listed below are presented in this appendix.

A. Identifying needs and approaches for standardization of travel model input data

B. Use of standardized metadata in improving the documentation and transferability of Spatial and travel model data

C. Analysis of temporal stability and dynamics in activity-travel behavior

D. Part 1: Regional impact on travel behavior
Part 2: Drivers of travel behavior
Part 3: Facilitation of travel data and model transferability

E. A guidebook that outlines data transferability issues and guides a user step-by-step through evaluating data transferability

F. Simulation of household activities and travel behavior data

G. Employment data and transferability issues in modeling

Research Topic A. Identifying Needs and Approaches for Standardization of Travel Model Input Data

Problem Statement

The future of travel modeling is certain to be impacted by the costs and difficulties of primary survey data collection and processing. As costs of household travel surveys have increased and limited resources have made availability of such surveys problematic, the travel modeling community has sought approaches to better utilize what data are collected and to improve both temporal and geographic transferability of data and models.

A large part of developing a travel model is developing inventories of local characteristics of the area's development and transportation system. These inventories are based on a number of different data sources, both local and national in origin. Local sources tend to be characterized by nonstandard, one-of-a-kind type attribution and categorization. Moving towards the development of standards for the associated data collection and keeping holds the promise of improving temporal and spatial comparability of such data, as well as establishing recognized best practices in terms of scale and categories of data collected and used. Developing common standards to guide disaggregation of spatial information will also provide a better basis for comparing findings and models from different areas and assess their transferability.

Proposed Research

The purpose of the proposed research is to investigate the feasibility and potential value of standardizing survey data collection and categories for transportation network and area characteristic information. Also, the researchers will be charged with developing a prototype/template of the specific attributes and data categories that can benefit from standardization based on a survey of current practices used throughout the U.S. The research objective will be to identify how creating a set of standards or guidelines can: 1) improve the efficiencies for data use in different applications; and 2) increase data comparability and transferability both for different actors within a region and across different geographic regions.

Specific tasks to achieve this objective will include:

1. Develop and perform a survey of current data collection and keeping practices for transportation models nationally. The survey should necessarily include a broad range of travel model users;
2. Identify types and categories of information that could potentially benefit from standardization. This process should be sensitive to the feasibility of standardization given local needs and established practices;
3. Develop a prototype/template identifying the most promising categories and attributes for standardization, required subcategorization, and detailed definitions. Apply this model to a sample set of modeled areas and identify issues to be resolved with such application;
4. Produce a project report discussing the potential benefits to be obtained from standardization; the pros and cons of standardization for specific attributes and categories; identified obstacles to temporal and geographic data comparability; and providing recommendations to practically phase and implement standardization as an industry practice.

The research is intended to provide:

- The necessary background information to inform a transition from current to recommended practices regarding more standardized data collection and keeping procedures for travel modeling inputs and data;
- An objective assessment of the potential benefits and associated costs of moving towards such standardization;
- Specific identification of areas and categories where such practices could prove beneficial;
- A prototypical implementation model as a practical guideline to implementing such practices;
- Documentation which is usable for both demonstrating the case for moving towards standardization and to guide trial and large scale implementation of recommended practices;
- A summary plan for phased implementation.

Duration

Including survey activity, it is anticipated that this work could be completed in seven to nine months. Review and finalization of the project report would be independent of this time frame.

Estimated Cost

The anticipated funding requirement for this effort is $135,000 to 150,000. This amount includes all consultant costs including direct costs associated with survey administration and a total of three progress meetings (including travel) with the project review panel.

Research Topic B. Use of Standardized Metadata in Improving the Documentation and Transferability of Spatial and Travel Model Data

Problem Statement

The transportation community is data dependent for all its research and developments. As we look forward into the future of transportation, the problem solving and implementation of models for all processes will get more data intensive. It will become necessary to have a more accurate representation of results. In today's world there is widespread data transfer and data sharing among and within organizations; it is imperative that the data have be clearly understood with respect to what information it relates to and the source that was used to compile it.

Standardization of metadata description and of the means of integrating it with associated data resources holds promise for improving data reuse and transferability. Moving towards the development of standards associated with addressing metadata needs and requirements for data integrity holds the key to improving data distribution and transferability. Developing common standards to guide various metadata developments for data ranging from spatial, traffic, and socio-economic data will also provide a better basis for comparing and tracking where and what the data represent and how to best use them for the necessary processes. Metadata can potentially provide the necessary information about the original survey and data development environment and statistics on data collection and analysis for subsequent users to better assess its value for reuse in other areas.

Proposed Research

The purpose of the proposed research is to investigate the possibility and value of standardizing metadata for traffic and travel modeling. This can be based on the existing metadata standards used for spatial data. The definition of metadata standards is well developed for geographic information systems and spatial datasets through a number of ongoing national and international efforts. These efforts, along with investigating work being done for other disciplines, would serve as a departure point for this effort. Assessing and cataloguing the state of such parallel activities along with a survey of model data users in the U.S. would be used to develop a comprehensive list of the types and categorization of metadata that may be required to address needs for data transferability. Researchers would also be tasked with developing a series of metadata templates to assist in assessing needs and potential value.

Specific tasks to achieve this objective will include:

1. Investigate current uses and standards for metadata in geographic information and other relevant disciplines;

2. Develop and perform a survey of metadata needs and practices for traffic/travel model data distribution nationally. The survey would be oriented towards determining the types of metadata information that could help users assess data transferability. The survey should necessarily include a broad range of transportation users using various data sets for travel modeling purposes;

3. Identify types and categories of information that could potentially benefit from standardization and description through metadata. This process should consider the additional resource requirements that metadata development would put on users;

4. Develop a series of metadata templates oriented towards supporting specific types of datasets and for each type identifying the most promising categories and attributes for common metadata, required sub categorization, and detailed definitions of metadata elements. Apply this templates to specific cases and identify issues to be resolved with each application;
5. Produce a project report discussing the potential benefits from standardized metadata; the pros and cons of standardization for specific attributes and categories; identified obstacles to development and standardization; and provide recommendations to practically phase and implement metadata standards as an industry practice.

The research is intended to provide:

- The necessary background information to inform a transition from current to recommended practices regarding more standardized metadata practices;
- An objective assessment of the potential benefits and associated costs of moving towards such standardization;
- Specific identification of areas and categories where such practices could prove beneficial;
- A prototypical implementation model as a practical guideline to implementing such practices;
- Documentation that is usable to both demonstrate the case for moving towards standardization and to guide trial and large scale implementation of recommended practices;
- A summary plan for phased implementation.

Duration

Including survey activity, it is anticipated that this work could be completed in seven to nine months. Review and finalization of the project report would be independent of this time frame.

Estimated Cost

The anticipated funding requirement for this effort is $145,000 to 160,000. This amount includes all consultant costs including direct costs associated with survey administration and a total of three progress meetings (including travel) with the project review panel.

Research Topic C. Analysis of Temporal Stability and Dynamics in Activity-Travel Behavior

Problem Statement

The transferability of travel survey data and/or the model parameters that are obtained from them is an issue that many urban areas around the country need to address in their travel demand model development and transportation planning processes. Many areas conduct household travel surveys sporadically while others simply do not have the resources to conduct their own household travel surveys. As a result, many areas look to transfer survey data and/or model parameters estimated using travel survey data from other areas of similar size, density, population characteristics, and transportation system characteristics.

While issues related to the spatial transferability of data are yet to be fully understood in a wide variety of contexts, issues related to the temporal transferability of data are understood even less. Most travel survey data and/or model parameters estimated using travel survey data are used to forecast travel demand and impacts of alternatives and policies, often 20 or more years after the data were collected and model parameters were estimated. How valid is it to use travel survey data and/or model parameters from one time point to forecast many years into the future? To what extent are activity-travel behavior and demand model parameters stable or dynamic over time? If behavior and/or model parameters change substantially over time, then how can data and/or model parameters be transferred or adjusted for travel demand forecasting? It is very possible that the assumption that behavior and relationships among variables remain static 20 or more years into the future has been the bane of travel demand forecasts until today. It is critical that we develop a body of knowledge that speaks to the temporal transferability, stability, and dynamics in activity-travel behavior and demand model parameters.

Research Objectives

The overall objective of the research effort is to study the temporal dynamics and stability in activity-travel behavior and travel demand model parameters. Specifically, the research effort would be concerned with determining the extent to which activity and travel behavior characteristics and model parameters change over time and the implications of these dynamics on travel forecasts. The research effort would propose methods for ensuring robust and meaningful temporal transfer of data and relationships that are implicit in all forecasting studies.

Proposed Research

The proposed list of tasks for this study includes:

1. **Compilation of data sets** - There are several data sets that can be used to analyze the temporal transferability of data and model parameters. A few metropolitan areas around the country (e.g., San Francisco) have repeated cross-sectional activity/travel survey data that can be used to analyze dynamics in behavior. The NPTS and NHTS constitute repeated cross-sectional survey data for the entire country. In addition, the Puget Sound Transportation Panel (PSTP) survey data constitutes an excellent source for analyzing dynamics in behavior at the disaggregate level while controlling for individual specific effects. This task would entail compiling a series of longitudinal data sets that would facilitate the analysis of temporal dynamics in behavior.

2. **Analysis of dynamics in activity-travel characteristics** - In this task, a comprehensive analysis of temporal dynamics in activity-travel characteristics would be undertaken. Depending on the nature of the survey data (e.g., activity-based, trip-based, or time-use based), changes in activity-travel characteristics will be examined in detail. Changes in trip rates by purpose, mode split, trip chaining characteristics, journey to work characteristics, activity durations, trip length distributions, travel durations, vehicle occupancy, and other activity-travel characteristics will be analyzed using robust statistical methods. The changes in activity-travel characteristics will be examined in the context of changes in socio-economic characteristics and changes in other demographic and transportation system characteristics so that correlations can be drawn that relate changes in behavior to changes in exogenous factors.

3. **Analysis of dynamics in activity-travel model parameters** - Forecasting using travel demand models estimated using travel survey data collected at one point in time involves making the implicit assumption that model parameters and relationships among variables will remain stable over the forecast time period. This is a rather strong assumption and may lead to serious errors in forecasting. It is imperative that a thorough analysis be undertaken to examine the extent to which travel demand model parameters remain stable or change over time. A series of basic travel demand models will be estimated using the data sets compiled in Task 1. These may include linear regression and/or count models of trip generation, logit models of mode choice, friction factors, and any other models that would be considered appropriate for examining stability in travel demand model parameters. These models would have to be estimated on pooled data sets and separately on each data set to study whether model parameters are significantly different from one time point to another. Dynamic models of travel demand can be estimated using the PSTP data set to determine the extent to which time-specific and individual-specific effects are significant.

4. **Methods for temporal transfer of data and model parameters** - If the results of Tasks 2 and 3 indicate that activity-travel behavior and/or model parameters change over time, then it is important that robust methods for the temporal transfer of data and parameters be developed. There are numerous methods by which demand characteristics and model parameters can be updated. Bayesian updating approaches, Monte Carlo simulation techniques, and synthetic travel data generation methods are but a few of the potential ways in which travel survey data and model parameters can be transferred in a rigorous way over time. In this task, specific methods for temporally transferring and/or adjusting demand characteristics and model parameters will be identified and tested. Differences and improvements in forecasts that are brought about by the use of such updating methods (if any) will be determined in this task.

Duration
24 months

Estimated Cost
$250,000

Research Topic D. Part 1: Regional Impact on Travel Behavior

Problem Statement

Background: context for the problem and past/current approaches

The current state of the art and practice in travel modeling is characterized by critical limitations in explaining travel behavior variability. The vast majority of operational models are specific to the regions where they are developed and their transferability to different regional conditions has never been fully explored. Several published research works on cross-regional comparisons have been limited to aggregate statistics and have not come to definite conclusions regarding the nature of regional impacts on travel behavior and transferability of the models. Although significant differences in travel patterns between regions were revealed, it was not explored whether the observed differences could be fully attributed to differences in socio-economic mix or the urban/travel environment (quantifiable within a travel model by means of explanatory variables), or whether there was a qualitative regional effect on travel habits that could only be quantified by means of geographic constants.

Statement of the problem

Most travel models lack full explanatory power and are frequently over-specified by geographic and/or other constants that offer an acceptable statistical fit to the observed travel behavior. However, such over-specified models do not properly predict traveler responses when demographic, land-use, or network conditions change.

Implications of the problem

Over-specification by "flat" geographic constants hampers spatial and temporal model transferability. The most difficult model to transfer is the destination choice (trip distribution) model due to the high level of calibration needed that normally results in a large set of geographic K-factors. The second model that is difficult to transfer is the mode choice model where mode-specific constants stratified by geography play an important role. Trip (tour) generation models as well as time-of-day choice models are relatively easier to transfer (as long as urban type/density variables are available) due to a simpler calibration procedure. However, even these models frequently require geographic constants.

Proposed Research

Research objectives

Proper modeling of the regional impact on travel behavior and elimination of "flat" constants-based structures will significantly improve the explanatory power as well as potential transferability of travel models. This research is intended to create a cross-regional database and explore regional impacts on travel behavior in the context of operational models. The research results will show how much of the regional impact on travel behavior can be explained by measurable factors and how much may still require region-specific parameters and constants.

Methodology/approach to studying the problem

The research is focused on exploration and quantification of the regional impacts on travel behavior, with the basic intention of building a transferable model. To overcome the problem of expensive and time-consuming collection of suitable new data for several regions, the proposed research will rely on a pooled database of available household surveys. Existing datasets will be

intensively reviewed with the emphasis on getting household data in combination with consistent level-of-service skims and land-use data attached to the survey data records.

The main travel-related choices such as household car ownership, daily activity-travel patterns of trip (tour) generation, mode choice, destination choice, and time-of-day choice will be structurally compared across several regions in terms of the observed statistics and impact of various explanatory variables through statistical estimation of choice models. This analysis will give insights into similar and different aspects of travel behavior across regions and explain the reasons for differences. Based on the analysis and limitations of the existing datasets, the research will also identify what needs to be collected in future surveys. However, it appears that significant improvement of current modeling practice will be possible by using already conducted surveys.

Specific tasks to accomplish the research

The scope of work will include the following research objectives:

- Identification of two (or more) dissimilar regions with existing household travel surveys and development of a combined multi-regional database for a study of the regional impact on travel behavior.
- Statistical analysis of cross-regional similarities and differences in travel behavior including estimation of region-specific and combined travel models with a wide set of variables.
- Identification of generic and region-specific factors and variables with the emphasis on substantiation of model transferability rules.

Multi-regional database

The research team will identify two (or more) dissimilar regions with good survey data and good land-use/network/other data. The following regions are potential candidates for consideration because extensive household travel surveys have been conducted over the last ten years and network processing procedures and land-use data at the traffic-analysis-zone level has been developed:

- Atlanta Regional Commission (ARC)
- East-West Gateway Council of Governments (EWGCOG—St Louis)
- Metropolitan Transportation Commission (MTC--San Francisco/Oakland)
- Mid-America Regional Council (MARC--Kansas City)
- Mid-Ohio Regional Planning Commission (MORPC--Columbus)
- New York Metropolitan Transportation Council (NYMTC)
- North Central Texas Council of Governments (NCTCOG—Dallas/Fort Worth)
- Oregon DOT
- Puget Sound Regional Council (PSRC--Seattle)
- Southeast Michigan Council of Governments (SEMCOG--Detroit)

For example, New York and Mid-Ohio represent two very different regions where extensive household travel surveys are available (with all complementary components like network level-of-service variables and zonal land-use data already attached) and activity-based micro-simulation models have been developed. However, the research team will consider other regions and make a final selection after analysis of the data quality and availability.

The selected regional surveys will be converted into a uniform format with a careful consideration of compatibility of the applied coding conventions and categories for variables. The data processing technique will be standardized and documented so that data from different regions could be pooled and exchanged in future. This will include standardization of the household surveys, building of the skims, the preferred sources for land-use data, data processing techniques, etc.

Cross-regional similarities and differences

The proposed statistical analysis will include two major parts:

- Analysis and cross-region comparison of the observed outcomes of travel behavior (number of trips, number of tours, activity episodes, time allocation by various types of activities, etc), with a special emphasis on determining the most stable components (invariants of travel behavior).
- Design and estimation of a single advanced four-step model (with activity-based model components) to cover two (or more) dissimilar regions. In particular such sub-models as car-ownership choice, trip (tour) generation, mode choice, destination choice, and time-of-day choice will be designed and estimated in two different ways:
  - A set of models for each region.
  - A set of combined models estimated on the pooled dataset.

The results of this analysis will help determine whether regional behavior can be separated from generic variables and captured in variables that explain regional differences. Potentially transferable model components and variables will be determined through comparing utility functions estimated for the same travel and socio-economic segment for different regions. From the current practice of trip generation, mode choice, and destination choice models it is known that the least transferable model components are constants and generic coefficients; at the same time more market-segment-specific coefficients and variables normally exhibit more similarities and potential for transferability. The proposed research will explore potentially transferable variables and coefficients in detail.

Another aspect of transferability to be explored relates to complex regions like the New York Metropolitan Area, which has an internal diversity in travel behavior and conditions that is probably greater than any other metropolitan area. The developed NYMTC model needed various adjustments of constants especially for mode and destination choice components across geographic segments (for example significant transit dummies were needed to explain transit-oriented choice for trips to and from Manhattan). An analysis of ways to explain these differences by means of socio-economic, level-of-service, and urban density variables would be a useful contribution to exploring transferability.

Generic and region-specific factors and variables

There is general agreement that the conventional set of explanatory variables must be extended in order to more fully explain observed travel behavior and eliminate numerous non-behavioral constants that are applied in travel demand models. The proposed research will undertake a search for (and detailed analysis of) additional explanatory variables. Many of the new variables may already be available in travel surveys or can be derived from network processing procedures, land-use data, or other information sources.

For example, traditional mode choice variables include:

- Travel time and cost components,
- Number of transfers for transit
- Household car ownership / sufficiency
- Household income
- Person age and driver license possession
- Area-type constants

Additional variables that may influence mode choice include:

- Individual GIS-based walk time and pedestrian conditions for transit and non-motorized modes
- Reliability in terms of transit schedule adherence
- Probability of having a seat for transit
- Comfort and convenience in transit cars (air conditioning, possibility of reading/using laptop)
- Commercial and information services on transit stations and P&R lots
- Frequency and location of stops on the way to and from the primary destination
- Road and personal safety (crime rate and public image associated with the area of transit station/line)
- Probability of having a parking place for auto and P&R
- Parking constraints, search, and conditions
- Individual parking costs (including free parking and discount parking eligibility)
- Driving conditions/road type
- Travel time uncertainty/variability (probability of delays)
- Individual car availability for the person and given travel tour taking into account broken cars (on one hand) and rented cars (on the other hand)
- Joint travel arrangements with the other household members
- Person-type, gender, age, and income-specific time and cost perceptions (VOT)
- Non-linear effects corresponding to marginal impacts of time, cost, and other variables as functions of trip length
- Quantifiable density variables instead of CBD dummies and other geographic constants

In a similar way, traditional destination-choice models for trip distribution include:

- Mode choice logsum or particular time/cost/distance variables
- Zone attraction variable based on the employment/land use mix

Additional variables that may influence destination-choice include:

- Bottleneck facilities (river crossings, bridges, tunnels)
- Statutory borders (states, counties, municipalities, school districts)
- Social frictions (income incompatibility, social/ethnic clusters)
- Special sensitivity to transit-accessible destinations by the non-driving population (children under 16, households without cars)
- Household composition and activity patterns that limit the spatial domain for activities (e.g., having a preschool or school child at home)
- Cognitive maps based on the spatial domain of the household and person with the pivot points corresponding to the most frequently visited locations (residential, work, school)
- Land-use data at a disaggregate level (not generally used because of computational complexity and data unavailability)
- Attraction characteristics and special trip generators that take into account the size and profile of individual attractions (e.g., most models use aggregate zonal attractions based on 3-4 crude employment variables)

Widening the list of explanatory variables should eventually allow for a full exclusion of flat mode-choice constants and distribution K-factors that dominate the current models. The expected result is a better transferability of modeling structures.

Implications of results of the research
The results of the proposed research will be used for improvement of the current modeling practice. These improvements will relate to the model structure, variables, and segmentation as well as to the model estimation techniques. The results will also be helpful for smaller MPOs who cannot afford large-scale surveys. In addition, the results will serve as a practical benchmark for comparison of models developed in different regions.

Duration
15 months

Estimated Cost
$225,000

Research Topic D. Part 2: Drivers of Travel Behavior

Problem Statement

Background: context for the problem and past/current approaches

The current state of the art and practice in travel modeling is characterized by critical limitations in explaining travel behavior variability. The vast majority of operational models are specific to the regions where they are developed and their transferability to different regional conditions has never been fully explored. The prevailing practice for such important model components as mode choice and destination choice (trip distribution) is based on a limited number of explanatory variables (household size, number of workers, income, car ownership, travel time and cost) and a few large travel segments (home-based work, home-based other, and non-home based).

Statement of the problem

Most travel models lack full explanatory power and are frequently over-specified by geographic and/or other constants that offer an acceptable statistical fit to the observed travel behavior. However, such over-specified models do not properly predict traveler responses when demographic, land-use, or network conditions change.

Implications of the problem

Over-specification of travel models by "flat" geographic constants and the absence of a real explanatory mechanism for travel behavior hampers model transferability in both spatial and temporal terms.

Proposed Research

Research objectives

Understanding and modeling of the real drivers and factors determining travel behavior and elimination of the "flat" constants-based structures will significantly improve the explanatory power as well as potential transferability of travel models. This is a very ambitious and long-term task that includes numerous aspects of travel model improvement. The proposed research is intended to identify the most important "breakthrough" directions and consolidate the already acquired experience in transportation and other fields.

Methodology/approach to studying the problem

The existing body of research on travel behavior in adjacent fields (like marketing research) will be surveyed in order to identify methods and models applicable for travel modeling and having potential to be included in operational travel models. This synthesis will consider model structures and methods, as well as practical data collection techniques to support these model structures. This will result in recommendations to improve model structures and surveys to better explain mechanisms of travel behavior. The improved travel models will be more generic and transferable since they will have more flexibility and sensitivity to external factors.

Specific tasks to accomplish the research

The current research scope is intended to resolve several issues of primary importance and outline further practical steps to be undertaken in this direction. The scope of work will include the following research objectives:

- Identification of available techniques for analysis and modeling of mechanisms and underlying processes of travel behavior (including the possibility of "borrowing" suitable methods from market research and related fields).
- Identification of the process data collection methods and recommendations for realistic extensions of conventional travel surveys.
- Formulation of approaches to enhance structure and transferability of travel models based on a better accounting of causality in travel decision-making (as opposed to the simple adjustment of constants to match the observed outcomes).

Analysis and Modeling of Mechanisms and Underlying Processes of Travel Behavior

What are the (external and internal) drivers of travel behavior variability and dynamics? To answer this question we should be able to qualitatively explain the mechanism and causality of the decision-making process and support this understanding with data for estimation of quantitative models. The current generation of conventional four-step models provides only limited help in this analysis since the model structure is based exclusively on the observed outcomes and cannot accommodate mechanisms of behavior.

A very general and important set of interrelated questions will be considered within this research:

- What do we know about travel-related decision-making and its complexity?
- What new data might help us understand the process better?
- How should our data collection methods improve to collect these data and understand the decisions?
- How might we then use what we have learned in models (e.g. new variables, new model structures, etc.)?

It is generally agreed that though prevailing travel demand modeling methods have been traditionally oriented towards replication of aggregate outcomes rather than explanation of individual travel-related decisions, there are some other fields with similar problems where certain useful techniques for understanding and modeling individual behavior have already been developed and applied successfully. In particular, market research methods and some other related fields that deal with decision-making and behavioral aspects will be explored and recommendations made regarding possible "borrowing" of methods. This will also include comparison of survey and modeling techniques and identification of methods applicable for travel behavior analysis.

The general issue of travel behavior will be considered in the context of developing operational models. The following aspects for improvement of travel models will be considered:

- Decomposition of complicated travel behavior into a set of manageable operational models that can be supported by data;
- Better understanding of the difference between individual and household decision-making and the mechanisms of intra-household interactions;

- Proper sequencing of travel-related decisions and identification of exogenous and endogenous factors for each decision-making step;
- Substantiation of the main variables that explain the "majority" of travel behavior variation;
- Understanding the choice scope and context for each particular choice dimension (car ownership, trip frequency, mode choice, destination choice, time-of-day choice), including a process of learning alternatives and formation of the sets of alternatives actually considered by persons and households.

Process Data Collection Methods

The question of what are the main variables for explaining travel behavior is closely related to the scope of variables being collected in the current travel surveys. To support better behavioral models it is necessary to extend the conventional set of outcome-based explanatory variables (so-called "what?" variables) and include some new variables in household travel surveys that relate to the decision-making process (so-called "why?" variables).

The proposed research will identify the creative survey work required to gather process data to use in the model and will determine the extent to which these new data add explanatory power to models. A need to collect additional process data to better understand the drivers of travel behavior is not that onerous if a proper focus is kept on the most important and realistic aspects, with the primary emphasis on variables useful for operational travel models.

In this respect, the causality-based approach represents an important step in moving from outcome-based to process-based approaches. We have to explore the causality of travel-related decision-making changes better. In particular, better sequencing of choices (mode, destination, time-of-day) and better understanding of the actually considered choice sets is needed. This requires certain stated-preference and attitudinal extensions to the conventional household travel surveys.

Further research is also needed on how to identify and model non-existing alternatives that are physically available but not actually practical (e.g. a transit trip requiring four transfers), along with actually considered alternatives in the choice framework. One suggestion on how to capture the real-world viability of alternatives when performing the travel survey is to use the following method:

1. Perform behavior data collection, i.e. ask how something was actually done.
2. Quickly shift the question to ask "Would you have rather done it another way?" This would capture which alternatives were actually considered by the decision maker.

The causality-based approach is oriented to proper sequencing and conditioning of the decision-making steps. In addition to "what" happens because of the combination of explanatory variables, the causality-based approach offers insights into the "why" sequence of decisions and events that led to the modeled "what."

Introducing causality and proper sequencing requires adding specific questions to the household surveys that would refer to the order and conditionality of decisions as well as the formation of

the choice set. In particular, for each visited activity location and the corresponding choice of destination, mode, and time of day, the following set of questions can be added to surveys:

- Was this activity scheduled or undertaken because of occasionally saved time in the course of the day?
- Were the destination, mode, and time-of-day choices made simultaneously or was there a certain order of conditional choices? Which of these choices are usual and stable over time and which are subject to change?
- If the actually chosen alternative was not available, what would be the second-best choice?
- Is there any predetermined area from which the location choice was made (like shopping on the some shopping street in the town or visiting the closest cinema theatre), or was the choice of location based on some unique properties of the location not associated with the surrounding area (like visiting Madison Square Garden or Carnegie Hall in New York)?

An important direction in improvements to survey techniques is associated with attitudinal and stated preference (SP) extensions to the conventional revealed preference (RP) surveys. A conventional household survey still represents the major source of information for travel model estimation. It includes detailed household and person information as well as a full description of the actual daily activity-travel patterns of all household members. It constitutes an ideal basis for additional attitudinal and SP-type questions that would be put in the actual context. It is much better than a stand-alone SP survey where normally one of the trips/activities is taken out of the daily context and then different questions about hypothetical alternatives are pivoted off the observed choice.

However, addition of attitudinal and SP questions represents a practical problem since the existing household surveys are already at the border of the length and complexity that can be tolerated by the interviewed persons. Thus, it is important to make these extensions easy and natural, and not time-consuming. This can be achieved by pre-prepared sets of answers from which the interviewed person could choose one and use the open question only if needed. These extensions are not intended to replace SP surveys, but to result in a better understanding of the actually made choices and their sequencing, as well as how the choice sets were formed. There are several examples of extensions of this sort that could be added to the conventional household surveys:

- For mode and destination choices, there can be a question about whether the mode/location was usual or occasional.
- For mode and destination choices, there can be a pre-prepared set of answers on the "why" question. For transit mode choice, it could include answers like "auto was not available," "travel time is better," "parking is a problem," etc. For auto choice, it could include answers like "had to drive a kid on the way," "transit was not available," "poor transit service," etc. For location choice for non-mandatory activities, it is important to distinguish between choice of "the closest location for this activity" and choice of "special pre-planned location for a particular activity."
- For time-of-day choices, there can be a pre-prepared set of answers on questions of how the schedule was actually built like "usual schedule for this activity," "planned in advance," "occurred in the course of the day because of necessity," "was added in the

course of a day because of the saved time." It may also be possible to ask respondents to order activities in the schedule by their schedule priority. For mandatory activities, it may be possible to ask respondents if there was any schedule adjustment to accommodate other activities in the schedule.

Enhance Structure and Transferability of Travel Models

The general logic and mechanisms of travel behavior could be very generic and transferable. However, it is necessary to develop quantitative approaches to explain and model behavior. Observed outcomes of the same behavior mechanisms and processes can be very different in different regions. Simplified models that do not give insights into travel behavior and deal directly with outcomes are generally not transferable. Essentially transferability means that the impact of the same particular factor (variable) on some particular outcome and/or process is similar across geographical segments.

Although conventional four-step models are easier to estimate, activity-based models that attempt to explain travel behavior have a better chance of being transferred. One of the constructive technical ways to determine transferable variables is comparing utility functions estimated for models in different regions. From the current practice of trip generation, mode choice and destination choice models it is known that the least transferable model components are constants and generic coefficients; at the same time, the more market-segmented coefficients and variables normally exhibit more similarities and potential for transferability. The proposed research will include a survey of the estimated models from different regions and comparison of the model components and impacts of different variables.

Implications of results of the research

The results of the proposed research will be materialized in terms of practical recommendations for improving travel models and surveys in a coordinated way. These improvements will introduce a better behavioral background into travel modeling structures and procedures applied in practice. As a result, models developed and estimated in different regions will become more comparable and transferable.

Duration

12 months

Estimated Cost

$200,000

Research Topic D. Part 3: Facilitation of Travel Data and Model Transferability

Problem Statement

Background: context for the problem and past/current approaches

The current state of the art and practice in travel modeling is characterized by critical limitations of the existing approaches in terms of explaining travel behavior variability. In particular, the absolute majority of developed operational models were specific to a single region and their transferability to different regional conditions has never been fully explored. It has been generally adopted by the travel modeling community that transferability of the conventional models is problematic since the basic aggregate statistics (trip generation rates, average trip length, mode shares, etc) vary significantly from region to region. There has been a very little attempt to explore this issue in more detail and facilitate transferability by means of more disaggregate analysis of travel behavior, finding invariants as well as explaining variation by means of quantifiable variables.

Statement of the problem

As the result, the majority of developed operational models were developed based on the local data and adjusted to the particular regional conditions. The model structures are not easily compatible and are difficult to synthesize or even compare.

Implications of the problem

Most of the developed regional models are unique to the region. Each MPO has to develop its own model and support it by an expensive local survey. There is a vary limited practice of transferring model components from region to region even in terms of conceptual model structure not talking about data and model coefficients.

Proposed Research

Research objectives and methodology

Understanding different aspects of transferability will help take full advantage of the already existing surveys and developed models. The current research scope is intended to explore different aspects of transferability and facilitate transferability of the best operational models. The scope of work will include the following research objectives:

- Identification of travel behavior invariants in terms of different activity units (trips, tours, activity episodes, and time allocation), decision-making units (household, person type) and time frameworks (day, week, month, year);
- Identification of the most transferable model structures and ways to facilitate transferability including practical guidelines for data / model transfer from region to region.

The scope of work will build upon and reference prior research:

- Identification of generic and region-specific factors and variables with the emphasis on substantiation of model transferability rules;

- Identification of available techniques for analysis and modeling of mechanisms and underlying processes of travel behavior (including the possibility of "borrowing" suitable methods from market research and related fields);
- Identification of the process data collection methods and recommendations for realistic extensions of conventional travel surveys;
- Formulation of approaches to enhance structure and transferability of travel models based on a better accounting of causality in travel decision-making (as opposed to the simple adjustment of constants to match the observed outcomes).

Transferability Concept Revisited

Transferability of data and models is a very open-ended topic. Transferability has actually never been really explored since most of the research and model-development projects have always been bound to a particular region. There are several "layers" of transferability that have to be explored:

- Transferability of data
- Transferability of model parameters
- Transferability of model structures

Transferability of data in a sense of observed outcomes of travel behavior (for example, trip rates) is rarely successful since outcomes of travel behavior are functions of numerous (external) regional characteristics; However data can be transferable is a sense that data collected in one region can help estimate a model in other regions if the potential scales and biases are taken care of. In this sense, data is always more transferable than models since any model has a simplification/aggregation bias that can hamper transferability. To take advantage of surveys from the other regions an appropriate statistic technique should be developed and tested. This technique should be able of pooling several regional surveys together and estimating a model specified with the maximum number of generic parameters while a limited set of region-specific parameters/constants should provide a reasonable fit for each region.

Transferability of model parameters assumes that the models have the same basic structure; this type of transferability can be successful but it normally requires an extensive segmentation and scaling of models.

With respect to transferability of model structures and underlying behavioral mechanisms, it is believed that we will be able to achieve a significant level of transferability, especially with the new generation of activity-based models that address behavioral processes in addition to observed outcomes.

To transfer travel survey data, parameters, or behavioral processes, the behavior must necessarily be well understood and the causality must be quantifiable. These needs lead towards state-of-the art survey work to identify decision processes that explain why choices are made, as well as towards databases and model structures sophisticated enough to describe realistic choice contexts throughout an entire urban area and within individual households.

Invariants of Travel Behavior

Research of the invariance, or constants, of travel would assist with determining whether there are additional variables which are not currently being included in travel demand modeling that can provide mode explanatory power and facilitate transferability, for example:

- What are people's lower and upper limits / tolerances in terms of the number of trips, number of tours, activity episodes, and time allocation by various types of activities and person/household types?
- Is there a variability or persistence of travel by individuals per household/person type by day, day of week, or by month?

There are research works that show that though cross-regional differences in travel behavior in terms of number of trips can be significant, there is a great deal of uniformity (and potential transferability) when travel behavior is measured different units such as number of tours, activity episodes, or duration of activities (time allocation). While conventional four-step models are built exclusively on trips, the new generation of activity-based models incorporates all behavioral units (trip, tour, activity episode, and duration) in a balanced way. These additional dimensions form a better potential for transferability.

Research previously proposed to better understand travel decision-making processes and recommend possible modifications to survey techniques could provide a valuable base from which to begin this work on understanding invariants. Existing activity-based surveys offer a significant opportunity to explore invariants, but enhancing these surveys to explore things such as which choices were actually considered and whether the activities were routine or not would benefit the research into invariants.

Transferable Data and Model Structures

Data transferability guidelines developed as the result of the proposed research would be helpful for the entire travel demand modeling community by preventing technically invalid data transfers while encouraging proper data transfer. This will include substantiation of standards for transferability of data and laying out criteria on what data is transferable, define a correct method to conduct data transfer, and provide a method for measuring whether data transfer was performed successfully and correctly (beyond simple matching of aggregate statistics).

A critical issue to be addressed is whether differences in travel behavior can be separated from generic variables and flat constants and captured in variables that explain these differences reasonably well. Significant regional differences in travel patterns may be a function of urban design, level-of-service and socio-economic variables. The ability to transfer operational models hinges on the ability to explain within the models why travel behavior varies significantly in different contexts.

The ability to adjust operational models for regional and temporal variation is also essential for transferability. More detailed model specification, scaling of variables, and segmentation can make models more generic and transferable. This includes various aspects like segmentation by trip length (or scaling of travel times and distance perceptions by the average regional time or distance) and income (capturing differential value of time effects), etc. Model specification tests for transferability will be developed.

Another option that could be tested and ought to be addressed is using aggregate "field" variables and social/spatial networks that can replace geographic constants and reduce proliferation of these constants in models. This captures the underlying information process that is missing in conventional models. For example, some research in Europe is underway on a "social/spatial network effects" that are not included in conventional surveys. They have found that people are affected by what everyone else around them is doing. Thus, travel behavior may be dependent not only on traditional person/household variables, but may also depend on what the person's circle of acquaintances (people from the same neighborhood or from the same social circle) is doing.

The practical experience with regional travel models has shown that even carefully estimated (with disaggregate data) models require additional calibration with numerous adjustments of parameters to match external target data (traffic and transit counts, journey-to-work census statistics, etc). The process of model calibration is generally more complicated and time-consuming than the original estimation. In this sense, if extensive disaggregate data of good quality are not available starting with a default set of model parameters is a reasonable approach. Also, to facilitate transferability, reasonable upper and lower limits (reasonable ranges) can be established for the model parameters.

Implications of results of the research

The results of the proposed research and data/model transferability guidelines will be used for improvement of the current modeling practice. They will help substantiate more generic model structures that can be used as defaults or starting points for development of regional models. Data and/or model transfer is essential when forecasting the effect of policies such as carpool lane requirements or infrastructure such as new modes that do not currently exist in a region. This will also be helpful for MPOs that cannot afford large-scale surveys and/or extensive model development.

Duration

18 months

Estimated Cost

$300,000

Research Topic E. Create a Guidebook That Outlines Data Transferability Issues and Guides a User Step-by-Step Through Evaluating Data Transferability

Problem Statement

One of the basic requirements for a good travel demand model is the availability of quality data. The option of using data transferred from previously developed travel demand models in different regions is an attractive alternative for transportation planning agencies. Data transferability is important because initial data collection related to household characteristics and travel patterns is often too expensive and time consuming. Planning agencies are eager to minimize their costs by borrowing or transferring data from others, to the extent possible. However, transportation professionals should use caution when incorporating borrowed data into their model. Not all data and parameters can be successfully transferred from one model to another. Additionally, emerging practices like activity and tour-based models and micro-simulation approaches require additional and more specialized data on travel and activity patterns. This significantly increases the need for the transportation professional to understand what data and information is a good candidate for transferability and those that are not.

Borrowing data or model parameters from somewhere else is an enticing alternative to transportation modelers as it could reduce or eliminate the need for a large data collection effort to support model development and application. This practice can save a significant amount of time and expense, which is always an issue. Unfortunately, a lack of understanding of data transferability and the knowledge of the opportunities and issues of using transferred data prevent transportation modelers from completely understanding the impact to the operation of the model. Variables, coefficients and other modeling components are often transferred with very little understanding or insufficient analysis of the actual effect to the model.

The greater interest to transportation professionals is to determine if another regions model or any number of its components are similar enough to their own that transferability of that information will result in an acceptable model and indicate good modeling practice. Thus, a guide is needed that will assist transportation modelers in determining if data, model structure, model components, variables or coefficients from other areas are indeed transferable to their model. As well as providing caveats and direction on how to accomplish this successfully.

Proposed Research

It is proposed to develop a web-based guidebook that outlines data transferability issues and guides a user through a step-by-step process for evaluating the transferability of specific data and understanding the options and factors that need to be considered. The guidebook would discuss general issues related to transferability, for example, what is meant by transferability, data management issues, and how to evaluate data quality. The guidebook will also provide guidance to address questions a planner may have on transferability, such as:

- What types of data/parameters can be transferred or should not be transferred?
- How are data/parameters transferred in current applications? What are correct methods for data/parameter transfer?
- What are the adjustments recommended (or required) and what are the local data needs to perform the adjustments?

- What are the implications in using transferred data (e.g. need to use same input variables)?
- What results should a modeler expect from using specific types of transferred data? (average trip lengths, VMT and VHT levels, number of trips, etc.)

The discussion of transferability will address issues related to the use of transferred survey data, model structure, model components, utility functions, variables, or coefficients. Experiences with different types of transferability will be provided as appropriate.

The deliverable will include a web-based tool that allows a user to go through a checklist that will assist in the understanding of the options available and factors that need to be considered to use transferred data or model parameters. Features of web-based applications will be engaged throughout the web-based guidebook to provide interactive notes, explanations and resources, including related documents and web sites, as needed. The guidebook may provide direction or links to resources where specific data and parameters available for transfer can be accessed. A maintenance plan to keep the web-based guidebook up-to-date will be provided as part of the deliverable.

The web-based guidebook of model transferability will serve as a knowledge base and resource center for transportation planners and modelers to understand the options available to them other than conducting a new survey and developing new models for their study areas. The products of this proposed research, listed below, would provide a cost-effective solution to help transportation agencies to find alternatives to serve their modeling needs with limited resources.

1. Research and gather related materials.
2. Proposed framework of the web-based guidebook with detailed site map
3. Web-base guidebook
4. Maintenance plan with a user-friendly tool to update the web-based guidebook

Duration
12 – 18 months

Estimated Cost
$100,000

Research Topic F. Simulation of Household Activities and Travel Behavior Data

Problem Statement

Mid-size and large metropolitan areas with populations of over 50,000 are required to conduct transportation planning. Household travel survey data that are necessary to develop and maintain travel demand models are expensive instruments. Household travel and activities data required to support the planning process can become outdated. Furthermore, emerging modeling techniques like micro-simulation are becoming available to planners. These require much richer data for calibration and validation. However, such rich data do not exist in most metropolitan areas, making it very difficult to move these emerging modeling techniques to practice. As a result, many questions are emerging with respect to development of a robust procedure for simulating household travel surveys data. Research work has been underway since the late1990s to develop procedures for simulating household travel survey data (Greaves and Stopher, 2000[1]; Stopher, Bullock, and Rose, 2003[2]). However, these works mainly focused on simple travel characteristics like trip rate, trip purpose, mode choice, etc. More recently, Stopher et al (2004)[3] introduced preliminary procedures to simulate uncomplicated trip tours and tour characteristics. There is more research required to determine whether simulation of the travel data is feasible and to identify potential approaches and develop a framework for simulation of household travel survey data.

Proposed Research

The main objectives of this study are to:

- Determine whether simulation of household travel data is feasible;
- Identify potential approaches and develop a framework for simulation of household travel survey data;
- Develop and validate a robust methodology to synthesize household travel survey data using local socio-demographic characteristics in conjunction with a national source of travel data;
- Evaluate the performance of travel-demand models estimated with these synthetic data compared to models or statistics developed in other contexts.

Proposed list of tasks for this study include:

1. **Data preparation and analysis -** The main purpose of this task is to identify important descriptive parameters that uniquely characterize household or individual travel behavior. This includes a detailed literature review and comparison of household travel

[1] Greaves, S.P. and P.R. Stopher (2000) Creating a Simulated Household Travel/Activity Survey – Rationale and Feasibility Analysis, *Transportation Research Record No. 1706*, pp. 82-91.

[2] Stopher, P.R., P. Bullock, and J. Rose (2003) Simulating household travel data in Australia: Adelaide case study, *Road and Transport Research*, Vol. 12, No. 3, pp. 29-44.

[3] Stopher, P.R., S. P. Greaves, and M. Xu (2004) Using Nationwide Household Travel Data for Simulating Metropolitan Area Household Travel Data, Paper presented at the *National Household Travel Survey Conference: Understanding Our Nation's Travel*, November 1-2, 2004, Washington, D.C.

characteristics across areas at different boundary levels and identify socio-economic, geographic, and land use factors that may influence travel decisions.

2. **Classification analysis -** Using parameters derived from Task 1 and different data mining and analytical methods (e.g., factor analysis, clustering analysis, machine-learning, genetic algorithms, etc) an attempt will be made to categorize geographic units into homogeneous clusters. Households from a larger level travel data (e.g., NHTS) will be assigned to these clusters and cluster-specific and geographic-specific travel statistics will be estimated.

3. **Data simulation** - The aim of this task will be to develop and validate a methodology to simulate household travel survey data in conjunction with a larger source of travel data and to further develop a general framework to facilitate data simulation and, more specifically, to determine how travel survey data can be simulated for smaller areas where local survey data are not available. In addition to basic travel characteristics required to develop traditional four-step models (e.g., trip rate, mode, trip purpose, etc.) the model should support data required to develop advanced travel demand models (activity-based and microsimulation models). This includes new simulation procedures to simulate trip tours, characteristics of the tour, duration, duration of each segment of the tour, stops between segments, travel modes, start times, and simulation of the destinations for primary tour purpose and intermediate stops. The destination simulation can probably utilize data obtained from GPS-enabled travel surveys. Several key assumptions need to be explored and examined including what distributions are observed in travel behavior and whether they exhibit different rates in different clusters. Specific questions to be answered include how to improve the estimations of probabilistic distributions (e.g., Bayesian approach) or what are the best data simulation techniques (e.g., Monte Carlo simulation)?

4. **Travel model calibration and validation** - The purpose of this task is to determine the reliability of the simulated data and to evaluate whether travel demand models estimated with the simulated data perform as well as models developed using actual survey data.

Duration
24 months

Estimated Cost
$250,000

Research Topic G. Employment Data and Transferability Issues in Modeling

Problem Statement

One of the main data components of a travel demand model is the availability of accurate and detailed employment information. Employment data are used in the modeling process to determine work related trips. Acquiring quality data with the required attributes is often one of the major impediments in travel demand model development. Depending on the region, various sources exist for employment data; however, the data sets need to be thoroughly assessed prior to their use. Common issues surrounding employment data sets are the age of the data, the frequency of collection, the geographic level of data aggregation, and the attributes that are included in the data set, such as the level of industrial classification. These can vary greatly for each source of data. Many sources of data are aggregated to protect the privacy of businesses and individuals. Modelers should have a good understanding of what specific information is needed to generate trips for their particular model. Some processes, such as activity-based modeling, require more data than others. Unfortunately, the modeler may find certain data sets are not as good as others only after a fair amount of time has been spent working with the data.

The greatest interest to transportation professionals is to determine if specific employment data sets are indeed transferable to their travel demand model efforts. Do they include the specific attributes needed to produce trip information and at the desirable level of geography? This largely depends on the sophistication of the trip generation process the modeler is using. Also of importance is the comprehensiveness of the data. Are all businesses included? How much checking of the data needs to happen before it can be used? What information is present with the data to allow the geographic location of the employment center? Are there address ranges or geographic coordinates?

Thus, a research study is desired to aid transportation professionals in the identification of common sources of employment data and assess the quality and transferability of the data with respect to the attributes included, frequency of capture, and overall accuracy and comprehensiveness relative to the need of their particular model and trip generation process used.

Proposed Research

It is proposed that a research study, culminating in a report, be undertaken that identifies the majority of available sources for employment data and assesses each one in respect to its ability to provide the necessary information for various travel demand modeling processes, including the four-step process and activity-based modeling. The research will provide guidance and address questions a planner may have on the transferability of employment data to their type of model, such as:

- What are common employment data sources?
- What attributes are included in the employment data?
- How often are these data collected?
- How comprehensive are the data?
- How accurate are the data?
- To what level are the data aggregated?

- To what level of industrial classification are the data assembled?

The research would consist researching what employment data sources are commonly used and addressing each of the above questions. The research could consist of a survey to collect this information.

Research Tasks:

1. Research state of the practice.
2. Identify common employment data sources.
3. Provide assessment for use in application to various travel demand models.
4. Prepare research report.

The implications of the research is that modelers can save time and effort by having an easy-to-use guide for finding the right data sources. Guidance will also be provided for the caveats and use of each data set and its respective level of transferability.

Duration
12 – 18 months

Estimated Cost
$100,000

Appendix B. List of Participants

Names	Affiliation	Telephone	E-mail
Carlos Arce	NuStats	512.306.9065	carce@nustats.com
Richard Arnold	Oregon DOT	503.986.4218	Richard.ARNOLD@odot.state.or.us
John Britting	Wasatch Front Regional Council	801.363.4250	jbritting@wfrc.org
Ken Cervenka	North Central Texas Council of Governments	817.695.9266	kcervenka@nctcog.org
Kuo-Ann Chiao	New York Metropolitan Transportation Council	212.383.7212	kchiao@gw.dot.state.ny.us
Ed Christopher	FHWA	708.283.3534	edc@berwyned.com
Brian Gardner	FHWA	202.366.4061	brian.gardner@fhwa.dot.gov
Edward F. Granzow	CH2MHill Engineers	530.546.3113	egranzow@ch2m.com
Pat Hu	ORNL	865.946.1349	psh@ornl.gov
Ging Ging Liu	Volpe National Transportation Systems Center/Cambridge Systematics	617.494.2039	liu@volpe.dot.gov
Tom Maziarz	Capitol Region Council of Governments, CT	860.522.2217	tmaziarz@crcog.org
Nancy McGuckin	Oak Ridge National Laboratories	202.248.8640	n_mcguckin@rocketmail.com
Phil Mescher	Iowa DOT	515.239.1629	Phil.Mescher@dot.state.ia.us
Kouros Mohammadian	UIC	312.996.9840	kouros@uic.edu
Elaine Murakami	FHWA	206.220.4460	elaine.murakami@fhwa.dot.gov
Ram Pendyala	University of Southern Florida	813.974.1084	pendyala@eng.usf.edu
Thomas Rossi	Cambridge Systematics	617.354.0167	trossi@camsys.com
Jim Ryan (not present)	FTA	206.366.0954	james.ryan@fta.dot.gov
Mohan Venigalla	George Mason University	703.993.1630	mvenigal@gmu.edu
Peter Vovsha	PB	212.613.8807	vovsha@pbworld.com

Appendix C. Data Transferability Resources

Greaves, Stephen. "Simulating Household Travel Survey Data." Presented at International Conference on Travel Survey Methods. Punta Renas, Costa Rica, August 2004. http://www.its.usyd.edu.au/isctsc/costarica_papers/resource/B8%20-%20Resource%20Greaves.pdf

Xu, Min, P. Stopher, and S. Greaves. "Using Nationwide Household Travel Data for Simulating Metropolitan Area Household Travel Data." Paper presented at TRB Conference on 2001 National Household Travel Survey, Washington, D.C., November 2004. http://www.trb.org/conferences/nhts/Xu.pdf

Ruescher, T., R. Schmoyer, and P. Hu. "Transferability of Nationwide Personal Transportation Survey Data to Regional and Local Scales" draft paper submitted to XX in 2001. http://npts.ornl.gov/npts/1995/doc/transfer.pdf

Urban Travel Demand Forecasting Use website http://nhts.ornl.gov/2001/presentations/index.shtml (this is list of presentations only)

Martin, WA, McGuckin, NA. "Travel Estimation Techniques for Urban Planning." NCHRP Report 365. Transportation Research Board Barton-Aschman Associates, Inc., Washington, DC, 1998. http://nationalacademies.org/trb/bookstore

Wilmot, Chester G. and Peter R. Stopher. "Transferability of Transportation Planning Data." TRR 1768, pp 36 –42. National Academy Press. Washington, D.C. 2001.

Venigalla, Mohan, S., Chalumuri. "Vehicle Activity and Personal Travel Inputs to Emission Models." Final report to the Transportation and Environment Research Program (TERP), Office of Planning, Environment, & Realty (HEP), FHWA. George Mason University, Fairfax, VA, March 2004. http://www.edthefed.com/xferability/TRIMM-FinalReport.pdf

Beckman, R.J., K.A. Baggerly and M.D. McKay. "Creating Synthetic Baseline Populations," *Transportation Research A*, Volume 30A, Number 64, 1995, pp. 415-429.

TRANSIMS: Transportation Analysis Simulation System, Version: TRANSIMS - 3.0." Volume 3 (Modules), Chapter 2 (Population Synthesizer), January 8, 2003. http://www.transims.net/Transims-Docs-3_0/Ver_3-0_Vol3-Ch2-PopSynth-08Jan03.pdf

Appendix D. Agenda

Peer Exchange on Data Transferability
Thursday, December 16, 2004
Keck Center of the National Academies
Transportation Research Board
500 Fifth Street, NW, ROOM 206
Washington, DC 20001

AGENDA

8:00 AM	CONTINENTAL BREAKFAST
8:30	**Introduction and Goals for the Peer Exchange** – *Ed Christopher, FHWA and Brian Gardner, FHWA*
8:45	**Presentations** – *Tom Rossi, Cambridge Systematics (moderator):* NCHRP 365 – *Nancy McGuckin* 1995 NPTS Transferability – *Pat Hu, ORNL* TRIMM – *Mohan Venigalla, George Mason University* Other transferability research including Stopher/Greaves – *Kouros Mohammadian, University of Illinois at Chicago*
10:00	COFFEE BREAK
10:15	**Facilitated Discussion Part A**: **Topics 1, 2, and 3**– *Tom Rossi, Cambridge Systematics (moderator); Elaine Murakami, FHWA (recorder):*
11:45	LUNCH
1:00 PM	**Facilitated Discussion Part B**: **Topics 4, 5, 6 and 7**– *Tom Rossi, Cambridge Systematics (moderator):*
3:00	BREAK
3:15	**Breakout Session**: **Developing research scopes of work** – *Tom Rossi, Brian Gardner, Elaine Murakami, Ed Christopher (note takers)*
4:45	**Conclusion and Next Steps** – *Ed Christopher, FHWA and Brian Gardner, FHWA*