Directions for Coordinated Improvement of Travel Surveys and Models

Eric Petersen,
RAND Europe,
petersen@rand.org

Peter Vovsha,
PB Consult, Parsons Brinckerhoff Inc,
vovsha@PBworld.com

A number of recent studies have pointed out the gap between academic interest in activity-based models and the relative scarcity of activity-based models implemented for regional and state-wide planning agencies. The issues that hinder the adoption of activity-based models range from the difficulty in obtaining resources to re-estimate existing models as well as staff to run more complicated models to theoretical concerns over the variability involved in micro-simulation. This paper focuses on the data requirements to support the estimation of an activity-based model and will lay out the minimum requirements as well as desirable features to be included in future household surveys. The underlying message of this paper should reassure planning directors that the basic surveys required to build an activity-based modeling application are similar to those required to update and revalidate a conventional model, although certain extensions are desirable. Focusing on more limited improvements to conventional surveys does not represent a digression from moving toward activity-based models, but rather offers a useful intermediate "stop" on the way, taking practical advantages of what can be already done today on in near future. For modelers wishing to explore the cutting-edge of activity-based modeling, the paper will also examine two promising areas of research: attitudinal and SP extensions to conventional surveys. The paper will conclude with a survey of existing household surveys from large metropolitan regions in North American and Europe and examine their suitability for supporting activity-based models.

For a long period of time, the structure of travel surveys was limited by the considerations of supporting the development of conventional 4-step models. One of the major deficiencies of 4-step models is the matrix structure of the trip distribution and modal split sub-models that severely limited the model segmentation and the number of explanatory variables that could be used. The surveys were actually much "richer" than the models and it was not clear why travel surveys should be made even more complicated (and more expensive to collect). Shifting to the micro-simulation modeling paradigm has lifted this technical limitation on model segmentation, allowing for richer, more complex models, in turn, fueling the desire for better data.

There are several following directions in which travel demand models and the corresponding surveys can be significantly improved:

1. Widening the range of explanatory variables used in models and collected in surveys.
2. Better understanding and modeling of causal linkages across various dimensions of travel behavior.
3. Adding attitudinal and SP extensions to conventional RP surveys

Each of these points is described below in details. These three directions are not independent and actually are closely intertwined. It is also important to note that model improvement can proceed in incremental steps rather than requiring dramatic improvements in all three areas simultaneously. In a resource

constrained environment, the most practical approach may be to carry out a survey for a convention model, but we would strongly recommend that the standard surveys be enhanced with some of the additional variables described in the following section to allow for more advanced model improvements in the future.

1. Explanatory variables

Conventional travel demand modeling has developed in an environment that stressed economizing on explanatory variables as much as possible to avoid extensive model segmentation. This led to a standard approach that was expressed in a limited set of variables like household size, number of workers, car ownership, income group, etc that are indeed important for travel behavior but are not nearly exhaustive.

Zonal "attractiveness" was measured by a limited number of employment variables stratified into 3-4 major categories like industrial employment, office employment, and commercial employment, etc. In a similar way, level of service variables by different travel modes were limited to average time and cost components that could be skimmed by the existing network simulation procedures.

New modeling frameworks open a constructive way to add variables and explanatory power to travel models. We believe that considerable improvements can be made within a conventional decision-making framework by adding additional explanatory variables. Of course, in order for these variables to be available to the modeling process, they must be present in the surveys.

Below is a list of traditionally used variables and new variables that could add significant explanatory power to such important travel models as mode and destination choice (trip distribution) taken as examples:

- Mode choice
  - Traditional variables:
    - Average travel time and cost,
    - Number of transfers
    - Household car ownership / sufficiency
    - Household income
    - Person age and driver license possession
    - Area-type constants
  - New variables:
    - Travel time uncertainty (probability of delays)
    - Reliability in terms of transit schedule adherence
    - Parking constraints, search, and conditions
    - Individual parking cost including free parking and discounted parking eligibility
    - Driving conditions / road type
    - Probability of having a seat for transit
    - Probability of having a parking place for auto and P&R
    - 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
    - Individual car availability for the person and given travel tour taking into account cars in disrepair (reducing the household's typical car availability measure) and renting cars (which supplements it)
    - Joint travel arrangements with the other household members
    - Individual GIS-based walk time and pedestrian conditions for transit and non-motorized modes
    - Road and personal safety / crime rate / public image associated with the area of transit station / line
    - 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
        - Comfort and convenience in transit cars / possibility of reading / using laptop / air conditioning
- Destination choice
    - Traditional variables:
        - Mode-choice log-sum or particular time/cost/distance variables
        - Zone attraction variable based on the employment /enrollment mix
    - New variables:
        - 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 of non-driving population (children under 16, zero car households)
        - Household composition and activity patterns that limit spatial domain of activity (for example, presence of child at home)
        - Individual attraction characteristics and special trip generators that take into account size / profile of the individual attraction (we are going into more and more detail on the household / person side but still have terrible aggregate zonal attraction variables based on 3-4 crude employment variables
        - Cognitive maps based on the spatial domain of the household and person with the pivot points corresponding to most frequently visited usual locations (residential, work, school).

The variables listed above have already been examined in different research and modeling frameworks and contexts. These are measures which can be quantified and added to a survey instrument. What is needed is to move these research achievements into practice for travel surveys and models. In particular, widening the range of explanatory variables should eventually allow for the removal of flat mode-choice constants and distribution K-factors that dominate the current models and "explain" most of the observed variability.

An important but under-researched area is the examination of long-term trends in travel behavior. Travel behavior obviously undergoes a significant evolution that is not captured by static travel demand models. There have been only several attempts to capture long-term trends in VOT estimates with the corresponding consequence for the choice model coefficients. {How does this relate to data – the need to provide data in a sustained way over time?}

2. Causal linkages

In our view, focusing on causality represents a constructive intermediate stage between a fairly standard outcome-based approach and the new process-based approach. The difference between outcome-based, cause-based, and process-based approaches can be illustrated by the following example of location choice for shopping.

The conventional outcome-based approach would try to explain the chosen location by means of the location characteristics (size, distance from home, accessibility by different modes) and person/household characteristics (person type, gender, age, car ownership, presence of children, etc) in a single choice framework where all location, person, and household attributes would be blended in the utility function and all other locations (zones) would be considered as available alternatives.

The cause-based approach would be focused on formation of the available choice set under the given conditions of the person that are considered as prior in the causal chain and prove that these conditions

indeed were fixed in the decision making at the time of the making the modeled decision (available time window, car availability, usual spatial "domain" of the person) and then formulation of a choice model that would take maximum advantage of the causal/conditional variables along with the conventional variables. The cause-based approach is oriented to proper sequencing and conditioning of decision making steps in an overall static environment.

The decision-process-based approach would be focused on both causal and chronological aspects of the decision making associated with the modeled event. Ideally, this would include a historical sequence of preliminary decisions regarding the time and location for the modeled shopping activity including probably numerous corrections and adjustments until the final decision was made and the corresponding activity was implemented.

The described three approaches are not actually alternative. It can be easily seen that they are sequentially inclusive. All factors, variables, and observed statistics pertinent to the conventional outcome-based approach are still relevant for the cause-based approach as well as causality is still a part of the decision-making screening. However, in addition to "What" happens as a result of the combination of explanatory variables, the cause-based approach offers insights into the "Why" sequence of decisions and events that led to the modeled "What." The decision-process-based approach makes additional step further in mapping the whole "How" chronology of the decision-making that was build up around the modeled event. The modeling complexity and amount of information needed to these approaches grows exponentially from "What" to "Why" and then to "How."

Chronological peculiarities of individual decision-making are less important for large-scale models and frequently lead to complicated multi-stage procedures with numerous feedbacks that are difficult to convert into operational models. Understanding of casual linkages is a simpler task though it is a limited view on travel behavior. It may significantly improve the structure of the travel model system and sequencing of the modeled choices and associated decision-making steps.

The cause-based approach to surveys pragmatically serves the existing static structure of choice models and helps improving it. It is not a substitution to a full-fledged process-based approach; it is a simplification that is practically helpful in a short term. It may also be helpful in the longer term as well, however, since the knowledge and understanding acquired in causal analysis may be of great value for the subsequent process-based analysis.

Introducing causality and proper sequencing in a static framework 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 TOD, the following set of questions can be added to either RP or SP surveys:

- Was this activity preliminary scheduled or undertaken as a result of occasionally saved time in the course of the day?
- Was the destination, mode, and TOD choice 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 locations choice was made (like shopping on the some shopping street in the town or visiting the closes cinema theatre) or the choice of location was based on some unique properties of the location not associated with any area around (like visiting Madison Square Garden or Carnegie Hall in NY)?

Introducing casualty into the modeling framework should naturally reduce a tendency of using simplified models of compensatory utility maximization and work in favor of more elaborate decision making chains with partially non-compensatory rules (eliminations).

3. Attitudinal and SP extensions to the conventional RP surveys

For the foreseeable future, standard RP household survey will remain the major source of information for travel demand model estimation. The most satisfactory surveys are those that essential form travel diaries with a full accounting of all daily activity-travel patterns for all household members. This type of survey constitutes an ideal basis for additional attitudinal and SP type questions added to reflect each traveler's actual situation, and is much better than a stand-alone SP survey where normally one of the trips / activities is taken out of the daily pattern context and then different questions about hypothetical alternatives are pivoted off the observed choice.

However, the addition of attitudinal and SP questions to the household survey represents a practical problem since existing household surveys are generally already at the upper limit in terms of length and complexity that can be tolerated by the interviewed persons. Thus, it is important to make these extensions as easy, natural, and short as possible. it is worth noting these extensions are not intended to replace SP surveys with extensive SP games; they are mostly intended to better understanding of the observed choices, their sequencing, and the way in which 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 location choices, there can be a question if the mode/location was usual or occasional;
- For mode and location choices, guided questions on the reasons "why" behind the choice of a specific mode or destination;
- For departure and arrival time choices, there could be a pre-prepared set of answers on question how the schedule was actually built such as "Planned in advance" or "Occurred in the course of the day out of necessity." A different set of questions might be asked at the end of the survey about the schedule priority of all activities and whether any schedule adjustment took place in order to accommodate some other activities in the schedule.

While we recognize that not all agencies will have the budget to support such extensive surveying, it is also the case that activity-based models can make the biggest advances in the exploration of the sequencing and scheduling of activities both for at the individual and household level. Yet making these advances requires data that has not conventionally been collected in the context of travel demand surveys and may require new innovations in data collection technology. It might well be worth treating these SP extensions as a pilot study or only collecting the additional data on a subset of the households.