A Practical Policy Sensitive Activity-Based Model

Yoram Shiftan, Technion – Israel Institute of Technology
Moshe Ben-Akiva, Massachusetts Institute of Technology

3rd Israel- Netherlands Workshop in Regional Science

Jerusalem, November 4-5th
A Practical Policy Sensitive Activity-Based Model

Yoram Shiftan
Department of Civil and Environmental Engineering
Technion – Israel Institute of Technology
Israel
Phone: +972-4-8292381
Fax: +972-4-8295708
E-mail: shiftan@technion.ac.il

Moshe Ben-Akiva
Department of Civil and Environmental Engineering
Massachusetts Institute of Technology
USA

Abstract

The growing complexity in activity and travel patterns resulting from various social and economic changes, together with growing congestion and negative externalities and the need to estimate changes in travel behavior in response to new policies have motivated the development of activity-based model as a new tool to understand travel behavior and forecast transport demand.

Today, there are a number of such models in operation, and new ones are being developed. One of the main challenges facing the travel demand modeler is the trade-offs between behavioral realism and complexity. This paper proposes an approach that on one hand captures the key behavior aspects and policy sensitivities, and on the other hand, is practical and requires reasonable computational resources so that it can be widely used for decision-making.

The paper analyzes the three main model elements that should be consider regarding the trade-off between behavioral realism and computational practicality: the model structure, data and application. Drawing on examples from an ongoing model development for Tel-Aviv and from existing U.S. models, the paper shows that quite a lot of behavioral realism and policy sensitivity can be achieved with a reasonable level of model complexity.

Keywords
Activity based models, policy sensitive, behavioral realism, computational complexity
1. Introduction

Activity-based modelling treats travel as being derived from the demand for personal activities. Travel choices, therefore, become part of a broader activity scheduling process based on modelling the demand for activities rather than merely trips. The explicit modelling of activities and the consequent tours and trips enables a more credible analysis of response to policies and their effect on traffic and air quality.

A variety of research methods have been used to study activity behaviour including duration analysis, limited dependent variable models, structural equation models and computational processes models (for a good review see Pas, 2002). There are two main approaches for practical activity-based models: the discrete choice or random utility maximization (RUM) approach and the simulation or rule-based approach. The focus of this paper is on the discrete choice modelling approach of utility maximization which is currently the more common approach in practical models. For a good comparison of the two approaches see among others Jovicic (2001) and Bowman and Ben-Akiva (1997).

Based on the RUM approach, Ben-Akiva et. al. (1996) proposed a practical activity-based modelling approach. They proposed a comprehensive travel demand-modelling framework that captures the mobility, activity and travel decisions of individuals and households; and a corresponding system of models that can be used for planning and policy analysis. Applications have followed but in an effort to enhance behavioural realism and to make them sensitive to a wide spectrum of today’s planning and policy needs they have reached a significant level of complexity that put their practical use, which is their main objective, at risk.

The growing complexity in travel patterns and the need to estimate changes in travel behaviour in response to new policies called for a better understanding of travel behaviour. This includes issues such as: How travel behaviour is affected by new information and communication technologies? How does land use and growth management affect travel behaviour, how much travel is induced as a result of new infrastructure, and how do travellers respond to auto restrain policies? Understanding such effects is essential for a better design of new policies and is the main motive behind the development and advance of activity-based models to a complex level. This paper looks at these various research advances together with the current practice of actual models used or under development by metropolitan planning organization and discusses the trade-offs between behavioural realism and complexity practicality in actual real world models. There is no question regarding the need and the importance of the continuing and growing research on activity –based modelling and their contribution to our understanding of travel behaviour, the question is how much of all this advanced research should go into practical models, so they can be widely used.

Figure 1 shows how the move from the trip-based (four-step) models to more advanced models increase behavioural realism together with computational complexity. As the figure shows the cost in terms of model complexity increases exponentially as we advance the models, while the benefits in term of behavioural realism are increasing at a decreasing rate. Figure 2, shows the same concept, but the computational complexity curve has been
converted to computational simplicity, so both curves represent model benefits. We are seeking a model of good behavioural realism with simple computational complexity. As Figure 2 shows there is some optimal level that provides the maximum benefits of the model, and this level is not one with ideal behavioural realism.

This paper tries to answer the question how much of the process of household and individual activities scheduling do we really need to incorporate in our models for the purpose of travel demand forecasting and policy analysis and what is the “right” trade off between behavioural realism and model complexity to optimize the model benefits.

A key aspect affecting model complexity that reflects the trade-off between behavioural realism and computational practicality is the model structure. The model structure should be compatible with data availability for estimation and with the level of detail and computational effort suitable for application. The paper analyzes each of these three related components: model structure in Section 5, data in Section 6, and model application in Section 7 and consider how these elements are affected by trade-off between behavioural realism and policy sensitivity on one hand and computational practicality on the other hand in each of them. Before getting into these details we discuss the ideal behavioural realism that activity-based models can achieve (Section 2), the policy issues that they should deal with (Section 3) and the barriers for their implementations (Section 4).

2. The Ideal Behavioural Realism

For activity-based model to have the required behavioural realism they need to be theoretically sound, and at a sufficient resolution to explain policy impacts. The ideal activity-based model should consider activity participating along continuous time dimension capturing time use and allocation behaviour with explicit consideration of constraints by the spatial, temporal, and social dimension, accounting for inter-dependency among individual in the households, among trips, and trip chaining. To better understand activity behaviour there is a need to analyze also the context of the activities including the why, when, with whom, and the duration and sequence of those activities (see Bhat and Koppleman, 1999 and Goulias et al, 2004). It also requires a detailed understanding of how households and individual acquire and assimilate information about their opportunities for activity participation and travel options, how this information is used to determine time allocation for activities and travel, and whether the attributes of activity episodes are determined jointly or sequentially (Bhat and Koppelman, 1999).

Some researchers have claimed that for a good understanding of activity-participation there is a need to understand the evolution of activity schedules from intentions to final outcomes for a weekly period (Lee and McNally; 2003) and that our understanding of the scheduling process can be advanced by observing household scheduling behaviour in realistic planning conditions (Doherty and Miller, 2000) and have developed special computerized survey instruments to records such scheduling behaviour. Doherty (2005) used a computer-based multi-day diary that recorded when respondents plan an activity, change its attributes or cancel it.

Finally, as Ben-Akiva et. al. (1996) showed the importance of the integration with lower level
decision, such as parking choices or route choice and with higher level decisions, such as residential location, work location, and car ownership. For a good discussion of “Behavioural realism in urban transportation planning models” see the paper with that name by Ben-Akiva et al (1998) who raised additional issues such as allowing for a randomly distributed value of time and the use of latent variables.

In term of ideal behaviour realism, activity based models should have few elements that are not appearing in most practical model and are discussed in the next sections.

2.1 Household Interactions

In keeping a simple structure most existing activity-based models of transport demand are based on individual activity-travel choice and not on household activity-travel choices. Although it has been realized that activities in multiple person households need to be coordinated and sometimes synchronized in time and space, current models neglect this behavioural realism in trying to keep the models simple and practical.

Recently, several potentially productive areas of time use research have been identified, including the allocation of tasks, resources and possessions to household members, and joint activity engagement by household members. Zhang et al. (2005) developed of a household task allocation and time use model based on a multi-linear group utility function that also allows quantifying the relative influence of the household members. Earlier attempts to deal with the issue include Wen and Koppelman (1999) and Goulias (2000). Goulias (2000) showed that proper accounting for the interaction between the household members and for within-person correlations over time is necessary to obtain more consistent estimates of activity participation and travel behaviour. Scott and Kanaroglou (2002) developed an approach that incorporates in its framework both interactions between household members and activity setting (i.e. independent and joint activities).

The Mid-Ohio Regional Planning Commission (MORPC) model accounts for explicit modelling of intra-households interactions and joint travel with particular interest at modelling the share ride as a travel model (MORPC); however, it is based on sequential modelling of households’ members by a predetermined order by person types. The Atlanta (ARC) model currently under development tries to estimate all household members’ activities simultaneously (Bradley and Vovsha, 2005).

2.2 Integration with higher and lower level decisions

In advanced models we move more and more to disaggregate models. With the move to activity-based model all choices related to the daily activity pattern and related choices are estimated with disaggregate models. The question is how much higher and lower level decisions are being estimated by disaggregate choice models and are integrated with the activity-based models or are they using some aggregate short cuts.

Figure 3 shows the system of models used in travel demand analysis with the activity base model as one of its elements. The upper level of the figure, above the double line, shows aggregate applications while the lower level shows disaggregate applications. Full arrows show the current practice of most activity-based models while dashed arrows show
additional integration at a disaggregate level with higher level decisions, such as land use and auto ownership, and with lower level decisions, such as parking choice and route choice.

2.2.1 Integrated Land Use and Activity Based Models

There is a rich literature on the importance of integrated land-use and transportation modelling for a better understanding of the effect of various transport policies. This is the higher level in Ben-Akiva et al (1996) framework; however, little has been done so far in the integration of activity-based model with land use models. Initial efforts are shown by Dong et. Al. (2006) and by Miller et. al. (2004).

2.2.2 Integration with Lower Level Decisions

From the lower level, below the activity scheduling, most activity-based models developed so far have done little to account for route choice behaviour and the effects of that behaviour on activity participation, duration and scheduling patterns. Rather, most model use traditional aggregate assignment models thus not utilizing all the benefits of activity-based models.

3. Policy Sensitive

As discussed above one of the main motivation for the development of activity-based model is to provide a model that would be sensitive to current emerging policies. Therefore, in designing these models, it is important to think what type of policies the model should be sensitive to. The most important policies that raise the need for activity-based models are:

- Demand management
- Land use policies
- Information and communication technologies
- Transit improvements
- Latent demand.

3.1 Demand management

Most commonly these days, emphasis is put on evaluating the response to different demand management strategies such as auto restrain policies, mainly parking restrictions and congestion pricing. Activity-based models are of significant importance to analyze such measures mainly because of their ability to predict how the whole daily activity pattern may change as a result of a specific measure. For example, a new congestion pricing may make a commuter change his or her mode from drive alone to transit; however, because the person no longer drives to work there may be other adjustments in his daily activity schedule. For example, using transit he may not be able to stop on the way back to buy groceries. Therefore, upon returning home, the person may take the car and drives to a nearby store. Only an activity-based model can deal with these types of responses.
3.2 Land-Use Policies

Land-use policies include mixed development, concentrated development in centers or corridors and pedestrian-friendly site design. Land-use measures can affect all activity and travel decisions; therefore, only an activity-based model can be sensitive to such measures. The importance of activity-based modelling to model responses to land-use measures lies in its ability to consider adjustments in the daily activity pattern considering time and space constraints.

3.3 Information Communication Technology (ICT)

ICT are rapidly being developed and introduced to daily life offering various business and personal opportunities such as telecommute, teleshopping, information and access to opportunities and travel information. The effects of ICT on daily activity pattern and travel behaviour are far more than affecting specific trips such as the commute trip and only an activity-based model can try to catch some of these effects. One of the main contributions of the Portland model is its ability to distinguish between in-home and away-from-home activities and to make the trade-off between work and any other activity at home or away from home.

3.4 Transit Improvements

Transit improvements are a major concern in an attempt to increase transit mode share and reduce traffic congestion. The advantage of activity-based modelling to predict responses to transit improvements is in its ability to model mode switching more accurately by considering the feasibility of such a response based on one’s daily activity pattern and adjustments to the daily schedule resulting from such a change. For example, if a person has to drop a child off, at a day-care center but not before a given time, and then be at work not after a given time, the person may find it unfeasible to switch the home-daycare-work auto half-tour to an auto tour of serving the child and a transit trip to work; nevertheless, all other variables show that the person would be willing to make such a switch under specific transit improvements.

3.5 Latent Demand

Latent demand is not a policy per se but it is a critical issue in policy making regarding transportation investment. The benefits from a new highway can be significantly biased by not taking into account induced demand. Travel speeds will be much lower and it has been shown that this overestimation of benefits outweigh the benefits to the new users not considered by the traditional model (Williams and Yamashita, 1992). The ability of activity-based models lie in its integrated approach including log sum variables that bring level of service variables up the structure till the daily activity pattern model, thus one’s activity participation may be a function of the accessibility to opportunities.
3.6 Summary

In the consideration of trade-off between practicality and behaviour realism emphasis should be on these aspects of behavioural realism that are important for the specific policies of interest. In the Tel Aviv model, for example, emphasis is put on parking pricing and supply and congestion pricing. For this purpose a special parking survey was conducted and a parking demand and supply and a detailed time of day models are being estimated, while compromising some other features such as limiting the number of activity types to five, and the number of stops in a tour to three (See below for more detailed on the Tel-Aviv time of day model.).

A parking model was developed on the basis of the parking survey conducted in 2001 (see below more on the survey) and on additional efforts aimed at the characterization of TAZ from the point of view of parking difficulties. These models estimate parking search time, walking time to the destination, and parking costs at each destination zone which are important determinants of travellers’ choice behaviour. However, walk time was found not to vary significantly among zones.

In San-Francisco the supply, cost, and availability of parking were developed from a variety of sources, including parking surveys, a small sample stated preference survey, parcel data and aerial photographs of on-street parking. The San-Francisco model also developed a policy variable to measure the potential impacts of improved pedestrian system and expected growth that will likely impact future travel demand, (Outwater and Charlton, 2006).

4. Implementation Barriers

In developing practical activity-based models it is important to understand the various barriers to their implementation. Models should be technically and financially feasible to develop, apply and maintain, which required also a reasonable run time. Activity-based models require advance disaggregate model estimation at a level that does not exist in most MPOs. While the estimation of such models is usually done by outsourcing to consultants, the application and use of the models also requires modelling skills that many MPOs lack.

Barriers to the implementation of more advanced models include the lack of willingness of planning agencies to invest into better data collection to support such models. Quite a lot can be done with current practice of travel diary surveys; however, a lot of data processing effort is required to convert the trip data to tours and day patterns. Finally, there are not yet off the shelf software to process the data and apply such models and their application require custom software.
5. Model Structure

5.1 From trip-based to activity-based models

Initial development in the advance from the traditional trip-based also known as the four-step paradigm to activity-based model started in the implantation of tour-based models (see for example the Stockholm model (Algers et al., 1995), the Boise model (Shiftan, 1998) and the New Hampshire statewide model (Rossi and Shiftan, 1997)) capturing the behavioural interactions across trips within tours (defined as a sequence of trips that start and end at the same location usually the home) but not across tours. Therefore, if a person stops for shopping on the way from work to home it will not affect his probability regarding making a special shopping tour once he get back home. The activity-based model also known as the day-pattern approach captures also the interactions across tours. To date it is quite agreeable that the advantage in behavioural realism of activity-based models outweigh the extra complexity in such models. However, also within the activity-based approach there is quite a variety in the level of behavioural realism they capture and in their level of complexity which continuously increases from early applications like the Activity Mobility Simulator (Kitamura et. Al, 1996) that was applied in Washington, DC. (Kitamura et. al. 1995), models in the Netherlands (Gunn and Van der Hoorn, 1998), in Denmark (Algers et al., 1995), in Germany (Ruppert, 1998) and in Italy (Cascetta & Biggiero, 1997), to more recent applications including San Francisco (Bradley et al., 2002; Jonnalagadda et. al., 2002 ), New York, Columbus, Ohio, Atlanta (Bradley and Vovsha, 2005), Dallas Forth-Worth, Sydney and the Dutch Albatros model (Arentze and Timmermans, 2001).

5.2 Daily and tour activity patterns

In designing the daily and tour activity patterns model the modeller has to consider various aspects of the daily patterns including the activity purposes, the number of tours per day and the pattern of each tour regarding the number of stops, the definition of a primary destination, inclusion of work sub-tours (tours that start from work and come back to work in the middle of the day) and more. In an effort to represent all possible daily activity patterns, model structures have increased in complexity to include large number of alternative daily patterns that with interactions among the various choices make the model application highly computational demanding and resulting in applications that take days to run.

Due to the large number of attributes in the daily activity patterns and the large number of alternatives for each attribute it is impossible to model all these alternatives jointly. A common and recommended simplification in the model structure is to decompose the model to three levels and make the distinction between the daily activity pattern, tour level models and trip level models as shown in Figure 4. The daily activity pattern predicts either the overall daily structure or the characteristics of the main activity of the day. Given the daily activity pattern, it is common to include in the tour level models the tour structure and mode, destination and timing of the main activity. In many cases the location, mode and timing of trips to intermediate stops are applied at the trip level after all other tour-level models are predicted, conditioned on the tour level choices but without feedback from the trip level models back to the tour level models. Some models like Portland have work-based subtour which is an intermediate level between the tour and trip level. This approach can significantly simplify the application of the models as will be explained in the Application Section.
It is important that the model will cover a good percentage of the cases that appear in the data. In practice coverage of 90-95% is considered good. For example, the Tel-Aviv model considers up to two tours per person, a primary tour and a secondary tour. The analysis of the data showed that only 1.6% of the people made three tours, 0.3% made four tours and only 0.1% made five or more tours (Shiftan et. al. 2004). Therefore, by capturing up to two tours per person we cover 98% of the population accurately and the percentage of the tours covered is even higher. It is also important when considering the number of stops per tour that there will be a good coverage in term of VMT. For example, in the Boise model the number of stops per tours was limited to four, but for travellers with more stops, the chosen four are those that make the largest VMT so ignoring the other have only a marginal effect of the total VMT estimate (Shiftan, 1999).

The classification of activity patterns is an important research topic in activity analysis as shown for example by Joh et al (2002). Table 1 shows the definition of main activity type for various models. Through these examples different approaches to deal with modelling the day activity pattern are discussed.

Table 1: Main Activity Types in Few Models

<table>
<thead>
<tr>
<th>San Francisco</th>
<th>Tel Aviv</th>
<th>Portland</th>
</tr>
</thead>
<tbody>
<tr>
<td>HB Work</td>
<td>Work</td>
<td>Subsistence at home</td>
</tr>
<tr>
<td>HB Education</td>
<td>Education</td>
<td>Subsistence on tour</td>
</tr>
<tr>
<td>HB Other</td>
<td>Shopping</td>
<td>Maintenance at home</td>
</tr>
<tr>
<td>HB Secondary</td>
<td>Other</td>
<td>Maintenance on tour</td>
</tr>
<tr>
<td>WB Subtour</td>
<td>Home</td>
<td>Discretionary at home</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Discretionary on tour</td>
</tr>
</tbody>
</table>

In the San-Francisco model the full day model predicts by one nested logit model (full information):
- The purpose of the primary home-based tour (work, education, other, or none)
- The trip chain type of the primary home-based tour (one or more stops before, after, neither, or both)
- The number of home-based secondary tours (0,1,2+)

There are overall 48 out of home activity patterns, 16 primary tour patterns (4 trip chain type for each purpose plus 4 work base subtour options for the work purpose) time 3 categories of secondary tour frequencies.

The Portland daily activity pattern (shown in Figure 5) determines the purpose of the person’s primary activity of the day, and whether it occurs at home or on a tour, allowing it to capture trade-offs between at-home and on-tour activities. The primary activity is one of six alternatives as shown in Table 1. If the primary activity is on-tour, the activity pattern model also determines the trip chain type for that tour (tour type), defined by the number and sequence of stops in the tour. Four tour types available for all tour purposes are: simple (no intermediate activities), one or more intermediate, activities on the way from home to the
primary destination, one or more intermediate activities on the way from the primary destination to home, and intermediate activities in both directions. For work/school tours, additional four types are defined as above with the addition of a work-based subtour. Simultaneously with primary activity and primary tour type, the activity pattern model predicts the number and purposes of secondary tours. There are six alternatives: no secondary tours, one secondary tour for work or maintenance, two or more secondary tours for work or maintenance, one secondary tour for discretionary purpose, two or more secondary tours for discretionary purpose, two or more secondary tours with at least one for work or maintenance and at least one for discretionary purposes. Since not all of the tour types apply to all of the primary activity types, there are 8+1+4+1+4+1 = 19 possible combinations of primary activity/tour types. Each of the six secondary tour alternatives is possible for all primary activity/tour types, so the model has a total of 19 x 6 = 114 alternatives.

The Tel Aviv activity patterns are defined by 4 primary out of home activities and for each activity there are 4 primary tour patterns: with one or more stops on the way to the primary activity, with one or more stops from the primary activity, stops on both ways, and no stop. For each of these 16 combinations, there is an option to have a secondary tour or not, while each secondary tour has a similar structure as the primary tour resulting in 16 alternatives of primary tour plus an alternative of no primary tour resulting in a total of 272 (16*17) out of home activity patterns.

In the Tel Aviv model it was decided that before modelling the details of the full day patterns it is more important to first model the destination, mode and time of the primary destination and only conditioning on these decisions to model the probability to make additional stops and a secondary tour. This structure is shown in Figure 6. Figure 6 shows the model structure of the main tour, only after the main tour is determined a similar model structure is used to estimate if there is a secondary tour and all the details of that tour. The logic is, for example that the duration of the main activity may have an affect on the propensity to make a secondary tour, while in the San Francisco and Portland models, decisions regarding destination, mode and timing are made only after the full day structure is determined. A drawback of the Tel-Aviv model is that the full day activity pattern is not estimated simultaneously with full information. These examples show that there are many ways to model activity patterns and more research is required on what make a better behavioural realism, which may be different for different areas. In designing the model structure there is a need to consider the trade-off between more patterns but estimated without full information versus fewer patterns with full information.

The MORPC model also first predict mandatory activity patterns including their time of day, mode and destination, and only next details of secondary tours are modelled given the residual time window. Finally a trip-level model is estimated to predict stop frequency, trip mode choice and destination (MORPC). A similar approach is used in FAMOS where mandatory activities are predetermined at a higher level module as part of the Household Attributes Generation System (HAGS).

Bhat (2004) in developing the CEMDAP model defined the work start and end time as temporal pags on which the worker’s complete activity-pattern rests. These pegs, along with the commute duration, determine the departure time to work and the arrival time at home from work. Accordingly, the first set of models determines the individual’s decision to
participate in mandatory activities and other activities are models next.

5.3 Destination Choice Models

Given some areas have a large number of TAZ, it is common to sample zones for model estimation, for example in the San Francisco model the number of alternative was 40 (Jonnalagadda et al, 2002) which is similar to the number used in Boise, Idaho (Shiftan, 1999) and New Hampshire (Cambridge Systematics, 1998). In Portland the destination and mode choice of the primary activity are modelled simultaneously as a nested structure. A sample of 21 zones was used for each tour drawn from the full set of 1244 zones and estimated together with nine alternative modes.

In Tel-Aviv it was decided to use the full 1200 zones for destination choice models. It should be noted that using the full set of zones as alternatives doesn’t complicate the estimation task per se and provides more efficient estimates. However, data preparation, specifically for secondary destinations where the level of service refers to the additional travel time that the second destination imposes on the already determined tour from home to the main destination requires a calculation of three dimensional matrices of size of the number of zones, and these can be cumbersome to calculate both for estimation and application. None of the models estimate two or more destinations simultaneously, rather all models estimate the main destination first, and additional destinations one by one, given previously determined destinations.

5.4 Time of Day Models

Time of day is a very important element from a policy point of view. With the increasing interest and importance of various pricing policy by time of day such as congestion pricing or various parking policies, it is of significant importance to have more detailed time of day models, and not just making the choice between peak and off-peak travel as common in most models. Ideally the time component should be model continuously; however, this is probably easier in the simulation/rule-based type models than in the discrete choice type models. Even with only few time periods, the time of day model is a an element that can complicated the model quite a lot with a large number of alternatives given multiple activities and the need to predict start and end time for each activity.

Most activity-based models use the two-level approach for time of day modelling, where the timing of the main activity is predicted first at the tour level, and the timing of other stops are predicted at the trip level in the remaining time window. This approach was used in a similar way both in San-Francisco and in Portland. These models first predict the period when the traveller leaves home to begin the primary tour simultaneously with the period when the traveller leaves the primary destination to return home. There are five periods in the day, early, a.m. peak, midday, p.m. peak, and late. Excluding overnight tours, there are 15 possible combinations. This approach is also being implemented in the Tel-Aviv model for the activity level, but given the importance of a more detailed time of day for congestion pricing and parking policies a much more detailed time of day model is being developed in Tel Aviv at the trip level. It was decided that a detailed time of day analysis at the trip level is more important than further development with less details at the full activity time of day model.
The detailed time of day model in Tel Aviv is based on a model Cambridge Systematics has developed for the U.S. Federal Highway Administration (FHWA) to advance the practice of forecasting person travel demand by time of day. The time of day choice model is based on demographic characteristics of travellers as well as the transportation level of service of the periods, which represent congestion and pricing levels. Since travel time data are available from model skims for only a few time periods (three in the Tel Aviv case), a key aspect of this approach is in developing a model that estimate travel time for all time periods used in the choice models. The basis for developing this model is to relate the reported travel time given in the household survey to the three model travel time skims and to various other network variables using ordinary least square regression. More specifically the model relates the ratio of reported speed to network free flow speed to various explanatory variables such as network delay (derived from peak and free flow speed), trip distance, and origin and destination area types. A cyclic function of time is used to ensure that the travel time corresponding to a given departure time will be the same 24 hours later.

The developed models are flexible in structure, and their main features can be summarized as follows:

- Time of day is modelled at a fine level of resolution using half-hour time periods. Compared to the use of a few time periods, the use of more refined time periods allows for substitution between these more disaggregate time periods.
- Congestion pricing strategies specific to a wider range of time segments during the day can be evaluated in a more flexible manner. Peak period pricing can be tested for half-hourly or hourly increments or any other combination. For example, the time period between 7:30 and 9:30 which may extend the definition of the AM peak period can be studied for peak pricing.

The model is applied only for auto trips to capture peak spreading and make the model sensitive for congestion pricing, these are non-issues for transit. Given that this is an auto model at the trip level its location in the model structure is after trip mode choice and before traffic assignment. This makes the location (and purpose) of all stops on a tour to be known (or modelled) prior to time of day modelling. Thus, it is assumed that the travel time from one activity to the next activity is known once the timing of the activity is known (that is, when an individual chooses a certain time to pursue an activity, s/he is also choosing the corresponding travel time.)

The MORPC model also had a detailed time of day for the tour level time of day choice model at a resolution of one hour, resulting in 190 hour-by-hour departure-arrival time alternatives. However, given there are four network simulations; there are only four different level of service variables for the different periods.

To summary, current activity based models have made a big step from trip-based models in analyzing simultaneously the start and end time of the primary activity and determine the timing of other activities in the remaining time window, however, there are still far from implementing the detailed time use allocation at a daily level that is currently mostly at the research level. We recommend the use of a two-tier approach model, while the tour level model can capture the behavioural time constraints on individuals by simultaneously estimating departure from home and from the main activity, while a more detailed model
implemented at the trip level with detailed time resolution can support the analysis of various congestion pricing policies and their impact of auto trip time shift.

5.5 Mode Choice Models

Most mode choice models in activity-based models consisted of two level models, models for the tour mode choice and models for the trip mode choice. The tour mode choice model determines the primary tour of the mode, while the trip mode choice models determine the mode for each individual trip made on the tour, based on the mode chosen for the tour. In some cases the trip level model only allows deviation from the tour main mode such as in the Tel Aviv model that include a more detailed list of modes already at the tour level. In other cases, like in the San Francisco model the trip level model allows also for further detailing of the modes, for example the trip based model will predict if the share ride mode predicted in the tour main model is Share-ride of 2 or 3+ people, and if the transit mode is one of four transit modes: local, MUNI, premium and BART, and between various combination of walk and auto access and egress modes to a total of eight different transit modes.

In Portland, based on the data analysis that showed that only 3% of the tours have a change in mode from trip to trip within the tour, usually with auto drive alone in one direction and auto drive with passenger in the other direction which are basically both drive modes. Therefore, the Portland model predicts only the main mode of the tour and assumes that all trips within the tour use the same mode. A set of rules was developed to translate all possible mode combinations to the nine main modes shown below. As mention above in Portland the tour main mode chose is nested under the destination choice model and estimated simultaneously.

The Tel-Aviv data showed that 91% of all tours are single mode tours, and out of the other 9% little more than half are combinations of bus and auto passengers. Therefore it was decided that the definition of the main mode tour is not crucial as long as the trip level allows for deviation from the tour main mode. The tour main mode was defined as the mode leaving home and allows for the whole array of modes. The trip level model is a nested model with a higher level choice between deviation from the main mode or not, and if the person deviated the lower level determines the other mode conditioned on the main mode of the tour.

Table 2 shows the modes that appear in the tour main mode choice models in few activity-based models. This table shows differences in definitions of modes among models. For example SFCTA differentiate between driver and passenger, but for driver didn’t make the distinction between drive alone and share ride, while CEMDAP made the distinction between drive alone and share ride but for the share ride didn’t make the distinction between driver and passenger while Portland made the full distinction between drive alone, share ride driver and share ride passenger.
Table 2: Modes in the tour level models of different model systems

<table>
<thead>
<tr>
<th></th>
<th>CEMDAP Commute</th>
<th>Portland</th>
<th>San Francisco</th>
<th>Tel Aviv</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drive alone</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Driver</td>
<td></td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Share ride</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Driver</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Passenger</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Transit</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transit Walk</td>
<td></td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Premium Drive</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Park &amp; Ride</td>
<td></td>
<td>+</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Kiss &amp; Ride</td>
<td></td>
<td>+</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Premium Bus</td>
<td></td>
<td>+</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Premium Walk</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Transit Drive</td>
<td></td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bus Drive</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Park &amp; Ride</td>
<td></td>
<td>+</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Kiss &amp; Ride</td>
<td></td>
<td>+</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bus Walk</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Taxi</td>
<td></td>
<td></td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>Walk/Bike</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Walk</td>
<td></td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Bike</td>
<td></td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
</tbody>
</table>

6. Data

The perfect activity-based model calls for the collection of very detailed time use data including activity diaries of all household members during a period of time, including in and out of home activities, detailed travel information, land use data and transportation level of service data. The data should also include spatial and temporal constraints and opportunities, and interactions in time and space as well as interactions among household members. The question is what of all these data is required for a good practical policy sensitivity model. In this section we discuss some of the main data issues in activity-based models.
6.1 The Activity Travel Diary

One of the main considerations in developing activity-based models is the type and length of the travel diary. One of the main questions regarding the diary is do we collect a one-day diary of a multiple-day diary? Obviously the Mobidrive data collected in Germany (Axhausen, 2000) over a six week period may be cumbersome for an MPO to use for developing a practical activity-based model. This relates to the question how important it is to incorporate inter-day and inter-week interactions to accurately model a single day’s activity-travel pattern which is our desired final practical product.

We believe that these various interactions are very important from a research point of view and future directions, but that quite a lot can be achieved with simpler travel diaries and one day diary are sufficient for the current practice. Even when diaries were longer, in many cases such as in Tel Aviv that includes three days and Portland that includes two days, the actual models do not deal with across-day interrelations.

Collecting detailed activity and travel data is problematic and imposes significant burden on respondents resulting in all types of problems and errors, such as under reporting, specifically of midday trips. Therefore for practical models it is suggested to keep these surveys to the minimum level of required complexity in terms of questionnaire design and invest additional resources in quality control.

Some surveys include in-home activities such as the Portland one as well as those conducted in Dallas-Fort Worth, Texas and the Research Triangle, North Carolina. The Portland model uses this information and its activity model patterns distinguishes between in-home and out-of-home activities. This is an important feature to capture trade-off between in home and out of home activities that is of increasing interest with the growing concern of the negative externalities of transport. However, research should be conducted regarding what level of in-home activities should be captured. Obviously it is impossible to record all in-home activities and guidelines to what types of in-home activities should be included in the diary should be developed.

There are a lot of issues in designing the activity survey in addition to the design of the model structure. For example, as discussed above we may define few activity purposes in the model structure, however, the survey may include many more purposes that are aggregated for the model development task. Another example is the use of open questions requiring the respondent to fill in his response versus close form requiring him to choose from a list of potential responses. These and other issues are dealt widely in the survey and data collection literature.

6.2 GIS and GPS

GIS data have been used in transportation planning but haven’t got much attention in time-use research. A possible reason as suggested by Bhat and Koppelman (1999) is that almost all GIS systems do not accommodate a temporal dimension. However, GPS include this time dimension and when combined with activity-travel survey can make some progress in this direction. They still impose significant burden on the modelling and application efforts introducing a new level of data complexity by its possibility to provide second by second
location information, that need to be dealt with, in addition to the cost implications. This complexity requires tools that can reduce and synthesize this type of information into meaningful data.

Linking GPS data with land use data at the parcel level can provide another layer of rich data to support the analysis of activity and travel data but again at a level that is too detailed for an MPO to work with for practical applications. GPS also provides opportunities to complement travel diaries to minimize under reporting of short and infrequent trips as observed in many diaries. GPS can also eliminate the need to report some aspects of the diary, mainly location and time, and the respondent can concentrate on fewer items such as mode, purpose, and occupancy that can not be collected from the GPS. While GPS has a potential in contributing to activity based models and maybe even in simplifying data collection, further research and development is needed before this can be a practical and easy tool to use for the advance of activity-based models.

6.3 Combining Data Sources

Activity-based models contain a larger number of alternative choices and a greater number of unknown parameters than tour or trip-based models. Therefore, it is critical that maximum use is made of travel survey data and combining them various other data sources such as stated preference data and auxiliary intercept surveys for an efficient use of the data. Integrated analysis of disparate data sources and the integrated application of different modelling methods and approaches that best fits each data set is an important feature to rigorously accounts and reconcile complicated travel behaviour characteristics as shown by Ben-Akiva and Morikawa (1990).

In the Tel-Aviv model for example, a combination of existing RP and SP data with new intercept survey and SP data was used. The new tour-based stated preference survey collected revealed preference data on respondents’ daily activity pattern, and then presented stated preference experiments regarding how these actual tours would be affected in response to parking restrictions, parking pricing, and congestion pricing. Potential alternatives included changes in mode, changes in the time of day and chaining of trips.

Given the importance of parking policies, special effort was made in Tel-Aviv to collect meaningful parking data that will used to estimate parking supply and demand models. These data include:

- Parking inventory (amount of parking places along the streets and in open parking lots, both legal and illegal; in-building parking was not included)
- Parking occupancy for street and open parking during the day (number of cars parked, by hour)
- Parking occupancy for selected in-building parking lots
- Interview with drivers, who park in the area (trip purpose, arrival time, search time, walk time, payment, personal data).

The tour level mode choice model estimated for Tel-Aviv is a good example of an efficient use of existing and new data as well as revealed and stated preference. This model (see Figure 7) consists of two set of revealed preference data and two sets of stated preference data. The revealed preference data include the National Travel Habit Survey and an extension of it
conducted specifically for the development of this model in communities adjacent to a rail corridor to extend the number of rail trips in the survey. The stated preference data include a previous stated preference survey conducted for the development of the Tel-Aviv mass rapid transit system and the new driver tour-based stated preference described above capturing sensitivity to auto restrain policies.

The lack of more detailed data collection and specifically more detailed activity and travel diaries are a barrier for the research advance of time use data and activity participation. However, quite a lot can be done with the current practice of data collected and overall, the San Francisco and Tel Aviv models while using a rich variety data, were developed with the same data sources that are used for traditional trip-based models. The one extension recommended on top of these two questionnaires is the addition of limited main in-home activities to the diary as was done in the Portland survey.

7. Model Application

Application of activity-based model is a complicated task and usually the one that most constrains the level of behavioural realism achieved in such models. The application consists of few elements including the activity simulation program, the population synthesizer and the transportation networks and assignment procedure.

7.1 The Activity Simulation Program

The application of activity-based models for forecasting usually employs a ‘sample enumeration’ or a ‘micro-simulation’ approach with a representative sample of households or synthetic population. In sample enumeration the probabilities across all possible alternatives are added across all individuals in the sample. In the micro simulation approach, the probabilities are used in a Monte Carlo way to predict specific choices for each individual in the sample. The key difference is that the sample enumeration approach enumerates all possible combinations of model outcome and multiplies probabilities, while the Monte Carlo approach predicts a single outcome per person drawing randomly from the model probabilities. See Bradley et. al. (1999) for more on the differences of these two approaches).

There are various options for short-cuts and run time reductions. For example in applying the Portland model using the sample enumeration approach (Bradley et. al., 1999) the following short-cuts were made:

- Running the model with only 10% of the sample (see also below).
- The destination choice and stop location models were applied to only a sub sample of the 1244 possible zones, usually 20 zones selected randomly using stratified sampling based on distance and employment levels.
- The application of the work-based subtour and intermediate stop locations were applied at the zonal level using sample enumeration, and therefore also couldn’t use logsum variables (see Bradley et al, 1999 for more details).

The Monte Carlo simulation introduces some random sampling error into the forecast, however, this decreases as the number of households simulated increase and therefore it is
recommended to use large samples and even as much as the size of the population. On the other hand, by simulating choices for a specific individual, all his characteristics can be retained to provide a wealth of information for other purposes, such as equity consideration. Outwater and Charlton (2006) specified this advantage as the reason for choosing this approach for the application of the San-Francisco model. The micro simulation approach is also preferred as it makes more sense to use a sub sample of destination in applying destination choice models since we are simulating a single choice by a single individual (Bradley et al, 1999).

The consideration for the specific method to apply involves a trade-off between computer run-time, geographic coverage and the accuracy of the results. Sample enumeration was used in some of the tour-based models in combination with Monte-Carlo simulation. The problem with sample enumeration is that the more levels there are in the model systems, the more costly it is to store in memory the probabilities of all the possible combinations (multiplying the probabilities from the different levels) for sample enumeration. For example in the Boise model, sample enumeration was used at the high level models where relatively few alternatives were available such as the tour purpose and patterns, but once the model went down to models with many alternatives like the destination choice models, Monte Carlo simulation was used to avoid book keeping of large number of probabilities resulting from multiplying probabilities from the different models (Shiftan, 1999). With the move to activity-based models the number of alternatives has significantly increased, making this book keeping more cumbersome and the trend with the applications so far seems to be of Monte Carlo simulation.

The model application process may require excessive time and computational resources. The main elements in the running time of the application are the model structure and the sample size. Eliminating insignificant interactions or linkages among sub-models may be considered at this stage as a way of reducing run time (see next section on logsums). The required sample size depends on the nature of the application and a flexible application procedure can be used to reduce run time when a lower level of resolution of the outputs is required, by allowing the user to select an appropriate sample size.

The literature on model application usually does not report running time. Bradley et al (1999) report running time for the Portland model using both sample enumeration approach and Monte Carlo simulation. Running the model system on the full sample of 600,000 households using the Monte Carlo simulation approach took 32 hours on a 400 MHz Pentium II computer, while running the sample enumeration approach of the same model on only 10% of the sample took little more time, 32 hours. As Bradley et al (1999) report, 75% of that time was needed to run the zonal enumeration to calculate the distribution on intermediate stop locations between every OD pair in the region showing the advantage of the micro-simulation approach from a running time point of view. Although it is more important to use larger sample size with micro-simulation approach it seems that the shorter run time compensate for this disadvantage, so we recommend the Monte-Carlo simulation approach to be used with large samples, and as said above ideally equal to the size of the population in the study area. Most recent applications including MORPC and the generic application of CEMDAP also use micro simulation approach.
7.2 The Use of Logsums

Logsums are not a separate element in the application rather they are an integrated part of the activity-based model system and their simulation application. However, they impose a major computation complexity in model application and therefore deserve a special discussion. The computational complexity is a result of the need to first calculate the utilities of every combination of alternatives from the bottom of the structure going up the tree before calculating probabilities on the way back down. As indicated by Bradley et al (2002), adding logsums variables to the model add a good deal of complexity to the model estimation and application process, as well as requiring a great deal more computer time to run. Therefore it is common to make various shortcuts and assumptions to reduce this complexity. For example in the San Francisco model the program first applies the work tour mode choice model (at the highest level of the model) to calculate a mode choice accessibility logsum across all modes to each alternative work location, however, since the tour type is not predicted yet at this point, an assumption is made that it is an am peak-pm peak work tour with no intermediate stops in each direction. In the San Francisco model logsums are also calculated and used from the main mode choice models to the primary destination choice models for non-work tours (because work destinations are modelled as the highest decision in the tree).

Logsums are important elements of the activity-based model and ones that allow a lot of its advantageous features by allowing the connection from low level models to higher one, a connection that also makes the consideration of latent demand possible. Given their importance on one hand and their computational complexity on the other hand they require a careful thinking of how much and which logsums variables to include in the model. It is recommended to estimate the models with as many logsums variables possible but then to carefully consider which are the most important ones to retain for model application, which are the ones that can affect specific policy analysis and the treatment of latent demand.

7.3 Population Generator

Another aspect of the application is the population generator. One of the main issues in the population generator is the dimension of the marginal distributions that define the number of segments in the population that are controlled for. Table 3 shows these dimensions for some of the models and the variables that are used as controls for the marginal distribution.

Table 3: Dimensions of marginal distributions

<table>
<thead>
<tr>
<th>Model</th>
<th>San Francisco</th>
<th>Portland</th>
</tr>
</thead>
<tbody>
<tr>
<td>HH size and number of workers</td>
<td>9</td>
<td>4</td>
</tr>
<tr>
<td>Household income</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Age of head of HH</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Total combinations</td>
<td>9<em>4</em>3=108</td>
<td>4<em>4</em>4=64</td>
</tr>
</tbody>
</table>

Most models found household size and number of workers to be good variables to distinguish
between important households life-cycle groups.

The sample size generated by the population generator and used in model application directly affect the run time as discussed above. Portland uses a sample of 0.6 Million households and 1.5 million inhabitants, matching the actual population for the application base year of 1994. However, to save run time, many simulations were run with a partial sample usually 10%. In designing model applications, it is recommended to provide the user with an option to use the fraction of the complete sample to use in each run. In this case, initial analysis or sketch planning levels can be conducted with a smaller sample, and only final analyses would be conducted with the entire sample to achieve better accuracy. Bhat (2004) showed a policy testing with a sub-sample of 1,000 households for the DFW area.

7.4 Networks

Another aspect of the complication of the application is the number of traffic analysis zones (TAZ). The smaller and more of them, we have a more refine level of spatial variables that allow for a better spatial resolution, but every operation has to be performed on more TAZs and thus run time is increased. San Francisco has 1,728 TAZs in the metropolitan area, MORPC has 1,805 TAZs, Portland, 1244, and Tel-Aviv has a little over 1200. Overall it seems common in the best practice to use between 1,000 and 2,000 zones. It may make sense to consider a two level zone system, as some applications, such as alternative transit alignments really require the fine resolution system, but other applications, such as an area wide taxing policy, man not require that level of resolution. Finally, as shown in the Destination Choice Model section some applications use the entire zone system, but sample a subset for some specific applications.

8. Conclusions

Before getting into detailed design of an activity-based model it is important to define the planning needs and the purpose of the model, and more specifically what are the important policy issues that the model needs to be sensitive to, and design the model in respond to these needs.

It is also important to understand that in activity-based models it is the application procedure that drives the complexity of the model. Therefore, in designing the model structure, one must keep in mind the application. Given more behavioural realism can be achieved in estimation than in application, it is also recommended to design and estimate a model with a one or two level higher complexity than what can actually is reasonable to apply. In this way, the important features and linkages can be identified and kept, while other can be removed in application to obtain a reasonable complexity level and run time.

Quite a lot of the elements of activity-based models can be estimated and applied with data and efforts that do not exceed much of those needed for traditional trip-based. The Tel-Aviv and San-Francisco model were estimated with the same data that are needed for trip-based models and the San Francisco model was developed and implemented in a relatively short
period of time (a bit over a year). These models already provide a big step towards better policy sensitivity comparing to trip-based models.

With the development of more advanced model systems such as the MORPC one and the Atlanta currently under development, there is a need for future research to test their actual contribution to demand forecast and a better policy sensitive analysis. The interrelations among household members and among days are clearly important from a behavioural realism point of view, but the magnitude of their complexity in practical models raises the question is their contribution justify this extra level of complexity. For now, we recommend that MPOs shall move ahead with simpler activity-based models and not wait till they will be able to implement the perfect behaviour realism model. As research advance and various tools, such as the generic computer software to apply such models (Bhat et.al, 2004), are being developed it is expected that better behavioural realism will be easier to implement and will eventually be part of practical models.

The following sections summarize the main recommendations for the tree main elements of activity-based modelling: model structure, data and model application.

8.1 Model Structure

There are endless options for various model structures and there is a lack of research pointing to what make a better structure. However, few general recommendations can be made based on the discussion in this paper.

It is recommended to use a two level structure for mode choice and time of day decisions and potentially also for destination choice, in which the main decisions are modelled at the tour level while secondary decisions are modelled at the trip level give the tour-level decision. More details and refined decisions can be modelled at trip level, specifically, it is recommended to develop a detailed time of day model for auto trips with a resolution of half hours to support the analysis of various congestion pricing policies.

Logsums are important element for capturing behavioural realism, and various logsums variables should be tested in estimation. However, due to their significant contribution to model complexity, only significant and important logsum variables should be retained for application.

8.2 Data

Much can be done in activity-based model with the same data used for trip-based models and a one day diary seems sufficient for a good practical activity-based model. Given the complexity of the diaries, it is more important to invest extra effort and resources in quality control. One recommended extension to most travel diaries in practice is the inclusion of main in-home activities, but further research is required on how to define these main activities.
8.3 Application

We recommend the Monte-Carlo simulation approach to be used with large samples, ideally equal to the size of the population in the study area. Various short-cut should be considered in application to reduce complexity and running-time. For example, it is reasonable to apply destination choice models on only a sub sample of destination to reduce running time, since we are simulating a single choice by a single individual. There still a need for further research regarding the sensitivity of these type of short-cuts. There is also a need to further investigate how much random sampling error is introduced into forecasts by using the Monte-Carlo simulation approach and how good is its geographic coverage.

8.4 Concluding Remarks

To better answer the questioned posed in this paper we need to compare predictions and policy forecasts from complicated models that capture the full spectrum of behavioural realism with simpler more practical models. However, unfortunately, opportunities to perform such comparison are rare and this is an important future direction for research that will assist in developing guidelines indicating when a more realistic process model is warranted and when a simpler more practical model will suffice. While the research on activity-based models already has a history of few decades and had advanced a lot in the last decade, practical activity-based models are only now starting to be operational. This is an important time to test these models, use them for policy analysis and recommend on how to better move activity-based models to a wider practice.

9. References


MORPC (undated). Mid-Ohio planning commission transportation modelling system: overview and summary. Available at www.drcog.org/documents/MORPC.pdf


24


Figure 1: Behavioral Realism and Computational Complexity in Travel Demand Models
Figure 2: Benefits from Behavioral Realism and Computational Simplicity in Travel Demand Models
Figure 3: Combined Aggregate and Disaggregate Model System
Figure 4: A Three Level Model System
Figure 5: The Portland Daily Activity Pattern Model
Figure 1  The Model Structure

Figure 6a: The Tel Aviv Model Structure
Figure 1  The Model Structure (continued)

Figure 6b: The Tel Aviv Model Structure (Continue)
Figure 7: The Tel Aviv Combined RP-SP Model Structure