Resource Paper

Data Needs, Data Collection, and Data Quality Requirements of Activity-Based Transport Demand Models

Theo Arentze and Harry Timmermans
Urban Planning Group/EIRASS
Eindhoven University of Technology
The Netherlands

Frank Hofman and Nelly Kalfs
Ministry of Transport, Public Works and Water Management
The Netherlands

ABSTRACT

Recent policy changes and methodological advances have led to new modeling approaches of increasing complexity in transportation research. Some of these approaches require new kinds of data. Moreover, the increasing complexity of these models often also implies that more detailed data are required, leading to increased demands on respondents. This paper focuses on activity-based models. It identifies the data required to estimate these models, discusses possible approaches to collect the data and formulates quality requirements for these data. Observations in this paper are largely based on two projects conducted for the Dutch Ministry of Transport, Public Works and Water Management. One project involved an examination of the validity, reliability and data quality of alternate ways of collecting diary data. The other, still ongoing, involves an activity-based model, called ALBATROSS.

INTRODUCTION

Traditionally, transport models were largely used to predict or assess transport demand and capacity as a function of the distribution of dwellings, jobs and facilities and the characteristics of the transport system. The conventional four-stage methodology of transport modeling involved the prediction of transport demand generation, destination choice, mode choice and route choice. The explanatory variables of such models were largely confined to household characteristics, attributes of the destinations, and attributes of the transportation system. Moreover, the theoretical underpinnings of conventional models were relatively weak. Traditional gravity and entropy-maximizing models were basically an attempt to describe observed traffic flows in terms of statistical principles or concepts adopted from physics, but lacked any convincing behavioral foundation. Discrete choice models, developed in the 1970s and 1980s, are better in this regard, if one is willing to accept that individual mode, destination and route choice behavior is the result of utility-maximizing behavior or, at the very least, that the principle of utility-maximizing behavior constitutes an adequate representation of individual decision-making.

Regardless of one's position, recent changes in policy, theory development and modeling in general have led to the development of new classes of transport demand models that have in common the element of increased complexity. It has been realized that observed traffic flows are a manifestation of people's activity patterns and that these are influenced not only by transport policy, but also by spatial and socioeconomic policies and the changing institutional context. Consequently, the explanatory variables of transport models, which traditionally were largely confined to attributes of the transportation system, have become inadequate. To give a few examples, conventional models do not incorporate any mechanism to predict the impact of task allocation on transport demand within households. Likewise, they often do not contain any component to assess the effects of institutional change, such as changing opening hours of stores, or different policies for flextime arrangements on transport demand. At best, modelers can develop scenarios or submodels, external to their main core transport model, typically leading to potentially incoherent modeling systems.

To avoid these problems, some modelers have argued in favor of activity-based models, which attempt to predict which activities are conducted, where, when, for how long, with whom and with which transport mode, if any. In principle, these models have the advantage that the various decisions impacting transport demand are treated in an integrated manner. However, this potential advantage comes at the cost of increased data needs, as one no longer requires data only on traffic flows, but also on activity patterns, described in terms of the various choice dimensions of interest.

Transport policies themselves have also changed and become more complex. The traditional policy of adding capacity has been augmented by better management and use of the existing capacity. Consequently, there is a need to predict users' reactions to information systems, route guidance systems and the like. This means that traditional models that predict the choice of a particular alternative (mode, destination, route) are no longer adequate to address these changing needs, as the changing policies require predictions on the timing of decisions, and the duration of activities. These changes in transport policies have therefore induced efforts to develop and apply model types such as hazard and ordered logit models, which were largely unexplored in transportation research until recently. As these models are based on duration data, data different from those traditionally collected in transport studies are required.

Another factor contributing to the increased complexity of transport demand models is the realization that the principle of utility-maximizing behavior may be inadequate to explain the complex relationships between the spatial environment, the transportation system, the institutional context and individual preferences, motivations and adaptive behavioral strategies. Across many different disciplines, including transportation research, an increasing number of researchers have started to examine the heuristics that individuals use to cope with a changing environment. These heuristics are often context-dependent. In terms of modeling efforts, this shift in the theoretical underpinnings of transport demand models means that a single probabilistic equation is replaced by a set of qualitative, context-dependent, logical expressions, again substantially increasing the complexity of the model. As these heuristics may be difficult to derive from real-world observations of traffic flows, this shift in theory has resulted in a need to explore qualitative methods of data collection.

Finally, developments in modeling in general have cleared the way for specific model developments in transportation research. Where a major breakthrough in discrete multivariate analysis in the 1970s stimulated the development and use of multinomial logit models in transportation research in the 1970s and 1980s, recent advances in the analyses of panel data and event history data have led transportation researchers to apply models based on panel and event history data as opposed to cross-sectional data. As the difficulty and effort involved in collecting panel and event history data are of a different magnitude from those for collecting cross-sectional data, the increased sophistication in modeling has generated increased problems of data collection.

Thus, recent policy changes, theoretical shifts and methodological advances have led to new modeling approaches of increasing complexity in transportation research. This increased complexity often requires different kinds of and more detailed data. The purpose of this paper is to identify the specific data needs of these models, to identify possible methods for collecting these data and to discuss some of the quality requirements. The discussion will focus in particular on activity-based models. Of course, ultimately, the data needs and data quality requirements are dictated by the specific model that one wishes to develop. Our observations regarding activity-based models in this paper are largely based on two projects conducted for the Dutch Ministry of Transport, Public Works and Water Management. One project involved an examination of the validity, reliability and data quality of alternate ways of collecting diary data. The other, still ongoing, project aims at developing an activity-based model called ALBATROSS.

To this end, the paper is organized as follows. In four sections, we will first briefly discuss the principles underlying activity-based modeling. Then, we will discuss the data needs, possible methods of data collection and data quality requirements. Finally, we draw some conclusions.

ACTIVITY-BASED MODELS

Activity-based models describe which activities people pursue, where, when, for how long, and with whom; which transport modes are involved, if any; and how these activities are scheduled, given the locations and attributes of potential destinations, the state of the transportation network, the aspects of the institutional context, and the personal and household characteristics. Many different theoretical perspectives and modeling techniques have been suggested over the years (e.g., Carpenter and Jones, 1983; Jones, 1990; Ettema and Timmermans, 1997). Each of these modeling approaches has its specific data needs. Although these approaches can be classified in many different ways, the following classification serves our purposes.

The *constraints-based approach* primarily aims at identifying feasible activity schedules, given various space-time and institutional constraints. Simulation models have been developed and applied to demonstrate that particular activity patterns can no longer be performed as a result of policies, such as closure of schools and shorter opening hours of service facilities (e.g., Lenntorp, 1976; Huigen, 1986). The input to these models consists of activity programs, which describe a set of activities of a certain duration that are performed at particular times. These activity programs are typically derived from observed activity patterns. The space-time environment is defined in terms of locations, their attributes, available transport modes and travel times between locations by transport mode. To examine the feasibility of a certain activity program, a combinatorial algorithm is used to generate all possible activity sequences. The feasibility of each sequence is tested by (i) checking whether the interval between the end time of the previous activity and the start time of the next is sufficient to perform the activity and cover the associated travel time; (ii) checking whether the activity can start after the earliest possible start time and be finished before the latest possible end time; and (iii) checking whether conditions about the sequencing of activities are not violated. The number of feasible activity schedules in this modeling approach is often used as a measure of flexibility of the time-space environment.

A second modeling approach consists of *utility-based models*. Their focus is on predicting activity models from principles of utility-maximizing behavior. Although some of the earlier attempts used Monte Carlo techniques to simulate individual choices at the successive stages of the activity chain (e.g., Kreibach, 1979; Sparmann, 1980; Swiderski, 1982; Schmiedel, 1984), or developed aggregate models for homogeneous groups with the same location and activity chain (e.g., Kutter, 1984; PTV, 1987; Fellendorf et al., 1997), discrete choice models became the dominant model type in this tradition. Daily activity patterns, rather than single trips, became the basic behavioral unit. For example, Adler and Ben-Akiva (1979) assumed that individuals evaluate a number of complete, one-day activity patterns and choose the pattern that maximizes their utility. A multinomial logit model is assumed to predict the choice of activity pattern, taking into consideration such variables as the number of sojourns and tours made for various purposes, the total travel distance traveled by particular modes, the destinations visited, the time spent on various activities, and the sociodemographics. In a more recent study, Ben-Akiva and Bowman (1995) suggest breaking down the activity scheduling process into a number of partial decisions, which are embedded in a hierarchical nested logit structure. Perhaps the most general model in this utility-based tradition is STARCHILD (Recker et al., 1986a, 1986b). In this model, activity participation is formulated as an individual constrained-choice process subject to the outcome of activity generation and allocation. The generation and allocation of activities occur continually over a multitude of time frames; however, the execution phase is conveniently conceptualized as a daily pattern when the actual participation and scheduling choices are implemented. After several preprocessing steps, the ultimate choice is formulated as a problem of maximizing the utility of a daily activity pattern from the derived choice set.

A third activity-based modeling approach is the *rule-based approach*. These models involve a set of IF <condition> THEN <action> rules, which are used to simulate activity patterns. The rules may refer to the priority of activities, the spatial proximity of activities and the implied waiting times of the schedule. The result is a sequence of activities performed at specified locations, and the route followed between the locations. An example is SCHEDULER (Garling et al., 1989), although it should be noted that the model is not fully operational, but primarily a conceptual framework for understanding the process by which individuals organize their activities. Following Hayes-Roth and Hayes-Roth (1979), individuals and households are assumed to attain certain goals. Activities are defined as the means that the environment offers to attain these goals. Individuals and households hold beliefs about how instrumental activities are in attaining their goals. These beliefs dictate their preferences for activities. Choice of participation in activities is determined by these preferences in conjunction with prior commitments, constraints and the like. Scheduling is defined as the process of deciding how to implement a set of activity choices during a defined time cycle. It entails an interrelated set of decisions made by the individual, interactively with other (household) individuals, concerning who will participate in the activities, where, when, and for how long, and how to travel between locations where the activities can be performed. The model assumes that the feasibility of the schedule is a primary goal, and the minimization of travel a secondary goal. The input to the model consists of a list of activities that have to be performed, a map of the environment with the available activity sites, and the available time slots for various activities. Another example of a rule-based model is described in Vause (1997). ALBATROSS also belongs to this category.

This classification of these modeling approaches is necessarily arbitrary. Hybrid forms have received increasing attention lately. For example, SMASH (Ettema et al., 1994, 1997) has been developed to combine rule-based models and discrete choice models. It conceptualizes the process of activity scheduling as a sequential process of activity sequence formation. In every step, the schedule, which is empty at the beginning of the process, can be adjusted by one of the following basic actions: (i) adding an activity to the schedule; (ii) deleting an activity from the schedule; (iii) substituting an activity from the schedule with an activity from the agenda; or (iv) stopping the scheduling process. Time-dependent utility functions are derived for each of these options. The option, chosen at each step, is the one with the highest utility at each step. By repeatedly applying one of these basic actions, the schedule is constructed and adapted until a satisfactory schedule is created. The schedule only includes information about the locations and the sequence of the activities. In every planning step, the rule-based component creates all possibilities to perform the basic actions. For instance, in the case of substitution, all activities in the schedule can be replaced by all activities on the agenda, that can be inserted at every position in the sequence. Of all possible options, the action that gives the highest utility is performed. Another example is AMOS (Kitamura, et al., 1995), which uses neural networks to derive the rules that drive the simulation of activity patterns.

DATA NEEDS

Activity-based models thus differ widely in terms of their theoretical underpinnings, their modeling approaches and their complexity. In fact, this field of research represents a set of different modeling approaches as opposed to a single, completely specified model. Moreover, each modeling approach can be further developed to include new elements. Consequently, it is impossible to provide a definitive answer to the question of which information is required to estimate and apply activity-based models. The answer will depend on the specific modeling approach of interest and the application area. Figure 1, therefore, only lists the typical data needs of constraints-based models, utility-based models and rule-based models.

It reflects our understanding of the dominant current emphasis of these modeling approaches, rather than an ideal situation.

As we have noted, the purpose of a full activity-based model is to predict which activities will be conducted, where, when, for how long, with whom, and with which transport mode. Consequently, this general description dictates the kind of data that are required. Thus, first of all, one needs data on *activities*. As we will indicate later, there is some empirical evidence suggesting that data on activities should be collected at a relatively detailed level; at the very least, one needs data on the activities that are conducted out of home. The more detailed the model, the higher the breakdown of these activities should be. These data are required for all modeling approaches.

This is also true for the second dimension, is the spatial dimension or *location* of activities. Assuming that actual traffic flows are important, geocoded data on land use and the location of various facilities where particular activities can be conducted are required. Because activity-based models typically predict which specific locations will be chosen out of a set of possible locations, data specifying the attractiveness of the possible locations will also be required. Since activity-based models involve many different types of activities, the definition of the attractiveness of locations also has a wide range of variables. For example, in the case of shopping locations, variables such as floor space, price, number of parking lots, atmosphere, variety, and service, are important. Many of these variables, however, are less relevant for specifying the attractiveness of locations for leisure activities. Thus, many of the attractiveness variables need to be activity-specific. Some activity-based models, such as the constraints-based models and the rule-based models, highlight the influence of constraints on behavior. These kinds of models, in addition, require data on the identified constraints, such as opening hours of stores, flextime, legal issues, etc.

The third dimension concerns *time*. Activity-based models incorporate the timing of travel decisions. Consequently, data on start and end times, and hence duration, of activities and related travel should be collected. Constraints-based and rule-based models tend to incorporate the timing decision in more detail, and hence require a more detailed recording of the timing dimension compared with utility-based models.

The next dimension concerns *"with whom"*, although it should be noted that we are not aware of any model that actually incorporates this dimension. If this dimension is incorporated in the modeling effort, however, its data implications are straightforward: one needs to collect data on who is participating in the various activities/episodes, at least for household members.

The final dimension, that of *transport mode*, implies that data on which transport mode is used to conduct the various activities should be collected. As one often will try to predict the choice between competing modes, data on variables reflecting the attractiveness of the modes should be collected, as well. One may, for instance, think of costs, frequency, comfort, travel time, and other commonly used descriptors of attractiveness.

The dimensions mentioned above are the principal choice dimensions underlying activity-based models. Depending on the specific modeling approach, data on each of these dimensions will be required. In particular, all of the three identified modeling approaches require data on activities, location and transport mode.

Constraints-based and rule-based modes, in addition, require (more detailed) information about the timing dimension and constraints, and the latter modeling approaches may also need data on travel company. The conceptual frameworks underlying all activity-based models recognize that activity participation and engagement are often related to personal and household characteristics, and characteristics of the physical environment, the institutional context and the transportation system. Hence, in most applications, data on such characteristics need to be collected. Figure 1 gives a few examples. Moreover, some scholars have discussed the option of linking activity-based models to models of longer-term decisions, such as residential choice, choice of work place, vehicle replacement choice, etc. If one is pursuing this idea, additional data need to be collected.

Special attention is required for task allocation within households and substitution. Although almost none of the existing activity-based models take these components into consideration, they are nevertheless important in understanding activity patterns. Thus, if one is incorporating these components in the model, data on task allocation and substitution are required.

DATA	CONSTRAINTS-BASED MODELS	UTILITY-BASED MODELS	RULE-BASED MODELS
CHOICE DIMENSIONS			
ACTIVITIES	X	X	X
LOCATION	X	X	X
TIME			
- START TIME	X		X
- DURATION	X		X
- END TIME	X		X
WITH WHOM			X
TRANSPORT MODE	X	X	X
EXPLANATORY FACTORS			
INSTITUTIONAL CONTEXT			
- OPENING HOURS	X		X
- LEGAL ISSUES			X
- POLICIES	X	X	X
SPATIAL ENVIRONMENT			X
- LOCATION	X	X	X
- LAND USE	X	X	X
- ATTRACTIVENESS		X	X
TRANSPORTATION SYSTEM			
- NETWORK	X	X	X
- TRAVEL TIMES	X	X	X
- FREQUENCY		X	X
- COSTS		X	X
- # OF TRANSFERS		X	X
- COMFORT		X	X
OTHER			
- TASK ALLOCATION			X
- SUBSTITUTION			X
- SOCIODEMOGRAPHICS	X	X	X

FIGURE 1 Data needs of activity-based models.

DATA COLLECTION

Having defined the information needs of activity-based models, the next question is how to collect that data. Figure 2 lists various possible methods of data collection that can be used to collect information about the choice dimensions and explanatory factors underlying activity-based models of transport demand. It shows that diaries, interactive computer experiments and conjoint experiments are the most viable methods to collect data on activity patterns. We will discuss how these methods can be used.

The experience with the latter two methods is still limited. Moreover, comparative methodological research is completely lacking. In contrast, the use of diaries has become more widespread, and several studies have collected empirical information, allowing us to discuss the pros and cons of specific operational decisions, such as type of diary, time horizon, and mode of administration. In particular, we will discuss the potential advantages of diaries as opposed to questionnaires, and the consequences of various implied operational decisions. Following this, we will briefly discuss interactive computer experiments and conjoint experiments, but this section is not meant to be exhaustive. A more extensive discussion is provided elsewhere (Ettema et al., 1996).

DATA	METHODS OF DATA COLLECTION
ACTIVITIES	■ Activity or trip diary ■ Interactive computer experiments ■ Conjoint experiments
LOCATION	Geocoded survey of potential destinations
TIME - START TIME - DURATION - END TIME	Element of diary or other chosen method to record activities
WITH WHOM	Element of diary or other chosen method to record activities
TRANSPORT MODE	Element of diary or other chosen method to record activities
FUNCTION OF	
INSTITUTIONAL CONTEXT - OPENING HOURS - LEGAL ISSUES - POLICIES	Fieldwork and scenario development
SPATIAL ENVIRONMENT - LOCATION - LAND USE - ATTRACTIVENESS	Existing databases, combined with remote sensing/field work
TRANSPORTATION SYSTEM - NETWORK - TRAVEL TIMES - FREQUENCY - COSTS - # OF TRANSFERS - COMFORT	Existing databases, combined with remote sensing/fieldwork
OTHER - TASK ALLOCATION - SUBSTITUTION	■ Questionnaire ■ In-depth interview ■ Decision table or other knowledge acquisition method ■ Full assignment conjoint experiment

FIGURE 2 Possible methods of data collection.

DIARIES

In this section, we will discuss the potential advantages of collecting diary data for estimating activity-based models, and the various operational decisions that should be considered when collecting such data. This discussion is based on desk research, supplemented with our recent experiences in collecting four different sets of diary data in the Netherlands. The following operational decisions will be discussed: questionnaire versus diary, type of diary, time horizon, recall period, frequency, timing, type of interval, form of administration and form of instrument. It should be emphasized that this discussion may lead one to believe that these operational decisions are largely independent, but they are not. Researchers are typically faced with trade-offs and, moreover, the decision made for a particular dimension may preclude particular decisions on other dimensions. Before discussing these operational decisions, we will first discuss the use of diaries as opposed to questionnaires focusing on an average day.

QUESTIONNAIRE VERSUS DIARY

The questionnaire has long been the dominant form of data collection in transportation research. It has been argued, however, that a questionnaire format with a focus on an average day may result in an under-reporting of trips. In particular, there is significant accumulated evidence that travel surveys especially under-report off-peak, non-home-based trips of short duration (e.g., Meyburg and Brög, 1981; Koppelman, 1981; Robinson, 1985; Dijst, 1993).

Thus, Stopher (1992) argued that an activity diary outperforms a travel diary and travel surveys in that short, non-home-based trips are no longer under-reported. This is consistent with the findings of Clarke et al. (1981), who reported that the activity diary indicated a significantly higher level of trip-making than the travel survey (13-16% higher). Mandatory trip purposes did not vary significantly between the methods of data collection. It is the discretionary trips that showed the greatest differences. Similar differences in degree of reporting have also been found in time-use studies (e.g., Niemi, 1993).

Nevertheless, diaries are not perfect, as exemplified by the study conducted by Golob and Meurs (1986), who found a systematic under-reporting of walking trips and an under-reporting of walking segments on trips involving more than one mode in diary data. People tend to overlook short, non-vehicular trips with increasing incidence during the diary period. Murakami and Watterson (1992) reported similar effects for the Puget Sound Transportation Panel.

There is also evidence on differential nonresponse by sociodemographic groups. For example, Roveri (1992), discussing experiences from the Italian time-use study, concluded that the level of nonresponse for time diaries is higher than for conventional questionnaires. Dowling and Colman (1995) reported a higher nonresponse rate for lower-income groups in San Francisco, while Sen et al. (1995) found for the Chicago region that managerial and professional workers had higher response rates than blue-collar workers. Larger households were found to have lower response rates than do single-member households. Finally, households with vehicles were more likely to respond. Thus, the available literature seems to suggest that the diary is likely to outperform the questionnaire in terms of the validity of trip and activity data. Diaries are also a richer source of information that allows additional kinds of analysis.

However, collecting diary data is quite demanding for respondents, which may result in lower response rates and differential nonresponse, and hence in higher costs and possibly bias. The latter can be a very serious problem: respondents who cannot afford the time to fill out the diary might also be the people who travel a lot.

When the objective of the study, however, is to build an activity-based model of transport demand, detailed information about people's activities and related travel is required. It is difficult to see how such data can be collected through the use of questionnaires asking respondents to report on the total number of trips during some time horizon. Diaries seem to be the best choice, as they represent the only instrument that tries to capture the relationships between activity choice, location choice, timing, travel company and transport mode in a systematic fashion. It is important, therefore, to construct diaries of good quality. In the next sections, we will discuss the operational decisions that are involved in this effort.

TYPE OF DIARY

Once one has decided to use a diary format, one then has to choose between a trip-based diary, a full activity diary or an out-of-home-only activities diary. These formats not only differ in terms of the amount of information collected, but also in terms of the format used to lead the respondents through the data collection process. In the case of trip data, the leading question relates to the trips made, and all other information is derived from or linked to trips. In contrast, the leading question for the other two types of diaries relates to activities, and all other collected information is associated with the activities the respondent says to been engaged in. Full activity diaries report all activities; out-of-home activity diaries only record out-of-home activities.

Clarke et al. (1981) pretested these various options. They found that no respondent had any apparent difficulty in recalling yesterday's trips in the trip diary format, although some acknowledged difficulty with trips made the day before. Omissions were either of all trips in one trip-purpose category, or short-duration trips. Some activities involving local travel, notably walking to local shops, tended not to be popularly construed as trips. Many disclaimed yesterday's trips as atypical, and there was some reluctance to detail 'unusual' trips or their attributes. Trips over the week were recalled in largely unordered sequences.

The results for an out-of-home activity diary format suggested that this format was accepted more readily than the trip diary format. All respondents recalled and discussed patterns of shopping activities in particular detail, and patterns of social activities were quite readily described. However, many people omitted one or more categories of activity entirely. Some respondents did not construe leisure as a distinct category, and it appeared that many nonwork, trip-dependent activities shared attributes of several categories, which made it difficult to identify them uniquely (e.g., shopping activities comprising shopping, social and recreational attributes). Similarly, multipurpose travel arrangements tended to be obscured by this format.

The full activity diary format generated the initial questioning of the relevance of such information to a travel study, but nevertheless, all respondents cooperated. This format demanded a longer interview time than others, and generated more discussion. It also produced more causal accounts of travel behavior than did the other formats. Conditions at the time of the trip were reported in more detail, because the adjacent activities were recalled quite fluently and in sequence. In no case did subsequent questioning reveal omitted trips.

Thus, Clarke et al. (1981) concluded that the use of a format in which participants discussed travel in the context of the day's activities seems to provide the closest correspondence with the natural storing of information and the planning of activities. Conversely, the conventional trip diary would seem the least satisfactory means of eliciting travel information. Asking respondents what trips were made during the period of time under investigation neither ensures that all travel is recorded nor defines the researcher's notion of trip to the respondent. A format that uses activity rather than trip diaries will substantially increase trip rates. Increases of up to 13 percent have been reported between activity diaries and traditional travel surveys (Stopher, 1992). On the other hand, Kalfs and van der Waard (1994), comparing seven Dutch diaries, concluded that an activity diary will not necessarily result in a more accurate and more detailed reporting of trips, unless it incorporates particular features to provide high-quality data.

Thus, it seems that the choice between a trip diary and an activity diary is largely determined by the goals and objectives of the study. If data are to be collected for a conventional four-step model, a trip diary might be sufficient. It is likely to give better information than a traditional questionnaire. However, activity diaries may provide more reliable data, as they imply a more natural way of storing information. If the purpose is to build an activity-based model or to assess the likely consequences of policies on activity rescheduling and hence transport demand, an activity diary seems the preferred choice. The question then remains whether this should be a full activity diary or an out-of-the-home activity diary. The latter type of diary may be cost-saving, but also limits the analysis of possible substitution of activities. On the other hand, a full activity diary is more demanding for the respondents.

TIME HORIZON

This operational decision concerns the question of whether diary data should be collected for the past or for the future. In principle, respondents may be asked to recall trips or activities, or even activities longer ago, or be asked to fill out the diary for a day in the future. Often, the latter option is referred to as 'leave-behind', as it typically involves an interviewer's leaving behind the diary for the respondent to fill out, after explaining the diary. In contrast, the recall format involves asking respondents, with or without previous notification, to report their activities performed during a given previous day. When the interviewer meets the respondent or contacts the respondent by phone, all events of the previous day are systematically reviewed in order to elicit from the respondent's memory the whole sequence of activities and trips, while establishing also the time of the day at which the consecutive events took place, the location of the activities, the persons in whose presence they took place, etc.

There is some empirical evidence that higher-quality data are obtained from leave-behind than recall diaries (Harvey, 1993). The latter approach appears to yield 5 to 10 percent fewer diary entries. On the other hand, the costs of the leave-behind diary are from 1.5 to 4 times higher. Juster (1986) argued that although leave-behind diaries appear to be of higher quality than recall diaries, the difference does not appear to justify this cost difference. We disagree with this recommendation if higher- quality data are required for the analysis. Even though it is likely that the respondent will fill out the diary at the end of the day, respondents are likely to be more conscious about their activities and trips, which is likely to improve the reliability of the data that are collected, unless respondents would change their behavior to reduce the time required to complete the diary.

FREQUENCY

This term is used in the time use literature to describe how long the diary should be kept. This issue has been subject to considerable debate in the literature. One group of scholars argues that the demands are so high that reliable results can only be obtained for one- or two-day diaries. Others (e.g., Scheuch, 1972; Goodwin, 1978; Lawton and Pas, 1995) have argued that one- or two-day diaries are not very valuable in that they do not capture multiday cycles in activity patterns, as evidenced in a large number of empirical studies (e.g., Goodwin, 1978; Hanson and Huff, 1982a, 1982b; Barnard, 1984; Harrison, 1986; Huff and Hanson, 1986, 1990; Pas and Koppelman, 1987; Mahmassani et al., 1991). In the time-use literature, the overriding opinion seems to be that one or two days per respondent for general studies are to be preferred, while for specialized studies, longer periods may be appropriate. The basic choice in this regard, for a given budget, is between sample size and number of days per respondent. Increasing the number of days per respondent reduces important dimensions of measurement error and marginal costs (Gershuny, 1992), and increases the usefulness of the data for subgroups and model development. However, it also increases nonresponse.

Rydenstam (1995), for example, reporting about the Swedish pretest of the harmonized European time-use survey, mentioned that the tested three-day diary caused heavy burdens on the respondents. Some families with children regretted that they had agreed to participate in the pretest. Moreover, the youngest children, in particular, had great difficulty in keeping their diaries, sometimes even causing quarrels in the family. Bagatta (1995), in a pretest of the same diary in Italy, arrived at the same conclusion.

On the other hand, it should be remembered that nonresponse does not necessarily imply behavioral bias. In fact, there is some evidence that there is no serious behavioral bias from nonresponse (Gershuny, 1990). The problem is that we cannot really know whether nonresponse implies behavioral bias. Hence, the nonresponse should be reduced as much as possible, without pushing respondents too much, as they may be inclined to quickly fill out the diary just to avoid future reminders.

Another important concern of multiday diaries is that diaries covering more than two days may be more prone to bias. Tendencies to skip certain types of trips or activities and record inaccuracies may increase over time, although it should be noted that the empirical evidence on such temporal biases in the transportation and time-use literature is still scarce and nonconclusive. Ampt and Richardson (1994) found that the number of reported daily trips slightly increased. In particular, they found that the second day had approximately 30 percent more information on trip-making than was obtained from the first day. If this higher reporting is the result of selection bias, which they also evidenced, it might point at a significant problem. There is evidence of reduced data quality beyond the second day (Clarke et al., 1981; Golob and Meurs, 1986). On the other hand, Gershuny et al. (1986) reported that data quality was not significantly reduced throughout a diary week, suggesting that data quality may not be systematically related to how long a diary is kept, at least within particular time horizons.

Gershuny (1992) argued that a one-day diary gives a good estimate of the prevalence of the activity of the population as a whole, but leaves open the question of differential participation rates. If we observe that a certain proportion of the sample engages in a particular activity on the diary day, we do not know if they engage in the same activity on the non-diary days or not. Thus, one-day diaries seriously limit the kind of analyses one can conduct, and may even lead to misleading interpretations and conclusions.

It is of the utmost importance, therefore, to examine the possibilities of data fusion, combining different data sources. It is clearly preferable to have a record of the respondents' activities over much longer periods than just the 24 hours of a day. However, it is extremely difficult to induce people to record their daily activities with due precision for a prolonged period of time, or to submit themselves to repeated interviewing over a number of days. A 48-hour frequency is perhaps the maximum achievable, unless resources suffice to offer substantial monetary incentives and build a substantial supervisory network. It would be interesting to examine the possibilities of computer-assisted panel research, asking respondents to fill out the diary once every so many days, and using personal patterns to reduce the interviewing time.

TIMING

Another aspect of the implementation of diaries concerns the choice of which days of the week to choose and how many hours to capture. As everyday activities are strongly subject to periodic variations depending on the days of the week, it is critical to spread the diaries equally over all seven days of the week. In order to eliminate as completely as possible extreme seasonal variations, diaries should also be spread over the year, assuming sufficient budgets.

This argument is supported by several empirical findings. Lyberg (1989), for example, found that a mail/telephone omnibus survey carried out by Statistics Sweden in 1982/83, using a yesterday approach, yielded an under-reporting of weekend days. Kinsley and O'Donell (1983), who tested convenient (i.e., respondent- selected) versus designated-day diaries, found that the designated-day diaries were less likely to overestimate hours spent out of home because they are more representative. However, they were more likely to underestimate the number of activities on some days, in comparison with convenience days. This may be because respondents have less time and therefore spend less time on recalling and reporting their activities.

Juster (1985) argued that a designated-day approach will enhance the representativeness and seems practical, given that overall, there is relatively little quality deterioration due to the recall decay. There seems to be no correlation between the propensity to respond and the designated day of the week. This finding suggests that a designated-day approach is appropriate and needed to obtain representativeness in terms of days of the week.

RECALL PERIOD

If a recall format is used, the next decision concerns the length of the period between the day for which the diary is reported and the day it is recorded. Juster (1985) concludes that the recall period appears to affect data quality differentially, depending on the day being reported. He concluded that there is little if any deterioration for Friday-to-Sunday diaries, and with less confidence, that there was only approximately ten percent deterioration for Monday-to-Thursday diaries, with recall periods examined as high as four to seven days. Similarly, Clarke et al. (1981) indicated that respondents generally had no apparent difficulty in recalling yesterday's trips, but some acknowledged having difficulty with trips made the day before. Therefore, the empirical evidence suggests that a recall period of two days is to be recommended.

OPEN VERSUS FIXED INTERVALS

Researchers also have to decide on the accuracy with which start and end times of activities are recorded. Respondents can be requested to mention start and end times of their activities (open interval) or report their main activity during some fixed (e.g., 15-minute) interval. Two counteracting considerations affect this choice. First, time slots should be such that reliable information can be given. Fixed intervals with longer time periods may cause an under-recording of those activities that take less than the chosen interval. Second, the choice of open- versus fixed-interval diaries should be considered in the context of a user-friendly format, probably leading to a preference for the less demanding and error-free fixed-interval diaries.

These assumptions were empirically confirmed by Lingsom (1980). Although open diaries are burdensome for respondents and coders, longer, fixed time intervals encourage respondents to concatenate activities, and activities of short duration tend to disappear altogether. Moreover, the use of fixed time intervals inhibits response from the more conscientious respondents, who become aware of the problems of enforced concatenation and suppression of activities (Gershuny, 1992). Time estimates, although not necessarily error-free, will more naturally reflect the duration of activities. Because short activities will not be reported in fixed-interval diaries, one should expect a tendency to over-report the duration of other (non-travel) activities. Consequently, one might observe a tendency that space-time constraints cannot be met, as there is less time left for travel. This will have an immediate impact on the outcomes and parameters of activity-based models, provided there is a systematic bias in under- and over-reporting the duration of particular activities.

FORM OF ADMINISTRATION

As with any survey, diaries can be administered by mail, personal interviews, and phone. The telephone mode of administration is likely to cause most difficulties. As diaries tend to be complex and take considerable time, it is difficult to imagine that respondents can provide such detailed information over the phone, that is, beyond simple retrieval of the information that they have generated before. The empirical evidence, however, is far from conclusive. For example, Klevmarken (1982) found no difference in time use between personal and phone interviews in a pretest of the Household Market and Non-Market Activities Study in Sweden.

The results of Lyberg's study of the Swedish Time-Budget Survey (Lyberg, 1989) suggested that the self-completed diaries provided more detailed reports than the interviewer-administered recall diaries. The estimates of time spent on different activities, however, were very similar for these two modes of data collection. The interviewer-administered diaries seemed to underestimate the percentage of participants in different activities, and overestimate the time spent among the participants. Meyburg and Brög (1981) also found that face-to-face interviews yielded less-valid estimates of travel time compared with mail questionnaires. Moreover, the diaries completed in the presence of the interviewer yielded 15 percent fewer trips. They argued that this finding is caused by unconscious mistakes of the interviewees, conscious mistakes of the interviewees who are unwilling to provide certain information, and the influence of the interviewer who attempts to complete the interview as quickly as possible.

There is also some evidence on the contrary. Libs and Lehart (1994) found in the Multi-Media Time-Budget Survey that, when using self-completed questionnaires, the number of regular readers was significantly lower than in face-to-face interviews. Stopher (1992) compared the completion rates between mail and telephone. Using stringent rules of completeness, he found a rate of 23.5 percent for the mail-back retrieval, and only 19 percent for the telephone. He did not find significant differences between the forms of administration in terms of person trip rates for home-based work, home-based nonwork and non-home-based trips.

Other scholars supported the efficacy of self-completion and mail-back questionnaires once contact has been made. Thus, it seems that self-completed diaries tend to result in a more detailed reporting and thus have much to offer, as long as appropriate action is taken to guarantee the required response rate.

The form of administration may also impact response rates. Ampt and Richardson (1994) compared six diary instruments that differed in terms of form of administration. One involved the telephone to establish initial contact, four established initial contact by mail, and the final instrument was based on personal initial contact and a personal interview to retrieve the diary data. The findings indicated that the instrument based on personal contact resulted in a response rate of 65.5 percent, whereas the lowest response rate was observed for the instrument that used the telephone to establish initial contact. The response rates for the self-completed mail diaries varied between 50.9 percent and 66.7 percent, depending on the type of diary. In another study, Ampt (1989) compared a self-administered and a personal interview for the collection of 24-hour travel diaries in Australia. As expected, she found the personal interview to provide higher response rates.

Given the available evidence, it seems that a mixed format that capitalizes on the best features of the various forms of administration may be preferred. Personal contact is likely to increase the response rate. Self-administered mail can then be used to realize a relatively high degree of reporting.

FORM OF INSTRUMENT

A final operational decision concerns which instrument (paper-and-pencil or computer) to use to collect diary data. Most diaries to date have used the paper-and-pencil format. If respondents are asked to keep a log of their activities and/or trips, the booklet is probably the most convenient way of keeping track of events, although recently introduced palmtop time-management systems may also be considered. If, on the other hand, a recall format is used, computer-assisted instruments offer an alternative, although the use of such electronic diaries is still scarce. The study conducted by Verweij, Kalfs, Saris and de Pijper (1987) is an exception in this respect. They used an electronic diary for time-use research and concluded that the electronic diary yielded detailed information of high quality. Similar results were obtained in related studies (Kalfs, 1992, 1993, 1994, 1995; Kalfs and Saris, 1997).

It seems to us that, in general, the pros and cons of computer-assisted technologies for collecting diary data are the same as those for surveys. Computer-assisted forms have the advantage of data quality control. This advantage can be appreciated when considering the anomalies and errors found in Italian time-use surveys (Panizon and Venturi, 1992). Close to 27 percent of all errors and anomalies were caused by skipping transfer activities (sequence of activity and/or place from outdoors to indoors, and vice versa). Another 17.8 percent was due to missing information about other people. The third reason was not reporting the hours. Hence, these three most important contributors of errors and anomalies account for approximately 56 percent and can easily be remedied in a computer-assisted instrument. It is straightforward to build an expert-based system that would systematically check for these anomalies and prompt the respondent for additional information. Arentze and Timmermans (1998) developed such a system, called SYLVIA, for diagnosing and repairing the activity diaries collected to build the ALBATROSS model. In their case, the computer program was used once the data were already collected, but there is no reason why this technology cannot be used during the interviewing process.

The next most important factor was inconsistency between activity and place. This seems to be more difficult to solve on line, as the diary instrument would have to be linked to a geographical information system. Although there is no reason in principle why this cannot be done, at the very least, it would slow down the interviewing process.

Other errors and anomalies mentioned included the same activity in consecutive records, out-of-own-town activities, missing information about the day of the interview, inconsistency between main and secondary activity, codes out of range, beginning hour later than ending hour, first or last activity of the day not equal to sleep, secondary activity that should be main activity, less than 1 activity per day, fewer than 15 activities in the diary, inconsistency between activity and age, activities of children without an adult being present, and main activity equal to secondary activity. All of these, and others, can be easily detected by computer-assisted forms of data collection.

Rydenstam (1995) reported similar sources of error in a pretest of the Swedish time-use survey. Respondents often failed to separate travel from activities, and some activities were simply forgotten. To avoid these errors, one needs instructions, but these are often difficult to provide. Oral instructions may involve information overload and have the additional disadvantage that substantial differences between interviewers may occur. Likewise, extensive written instructions will not be read. In his pretest, very few respondents read or even glanced at the instructions before filling out the diaries. The instructions were more used as a reference book in case of uncertainty.

All these reasons and practical experiences favor the use of computers, budgets permitting. There are no reasons why checks on the types of errors mentioned above cannot be incorporated into an intelligent computer-based diary, avoiding close to 90 percent of all errors reported for the Italian time-use survey. In some cases, the use of computers can also reduce the costs of the data collection process. One should not forget that cleaning the data after the fact is often time-consuming and hence costly.

INTERACTIVE COMPUTER EXPERIMENTS

Up to this point, we have discussed the potential use of computers as an alternative means of recording and checking data and guiding the interviewing process. However, computers can also be used in a more interactive way, seeking respondents' reactions to scenario-based or rule-based conditions. This is an interesting alternative way of data collection, which is more difficult to accomplish with self-administered paper-and-pencil formats. As this topic is discussed elsewhere at length, we will discuss it only briefly in this contribution.

An example of such an interactive computer experiment is MAGIC, specifically developed by Ettema et al. (1993) for activity-based analysis and model development. In addition to offering a tool for the registration of activities and related travel, destination and socioeconomic information, MAGIC can also be used to systematically measure, code and store information on respondents' reactions to planned or unplanned changes in the attributes of the transportation network, destinations or institutional context, and to changes in temporal, spatial, institutional and/or coupling constraints that might impact on activity performance and scheduling behavior.

The system starts by asking questions about the respondents' cognition of the environment. That is, respondents are requested to indicate where they typically perform a set of predefined activities, the alternative locations they know where they can perform these activities, the transport mode they use, the required travel time, etc. The first part thus results in a description of respondents' cognitive representation of the urban environment and the transportation system. Next, respondents are requested to indicate their typical activity patterns for the day of interest. The computer is used to test the consistency of the patterns. The information that is provided at an earlier stage of the experiment is automatically used at this subsequent stage. Then, the truly interactive aspect of the system is used. Having provided their activity pattern for the day(s) of interest, respondents are informed about possible changes to their environment, or their personal situation. For example, opening hours of shops may change, or they may have to work fewer hours. They are then requested to (re)schedule their activities under such changed conditions. Results can be interpreted as strategies of scheduling activities under changing conditions, and thus offer clues as to which heuristics or principles are used. The value of this additional interactive option lies in its ability to systematically obtain data on respondents' choices and reactive behavior, required to estimate activity-based models, that otherwise would be difficult, impossible or too costly to collect.

MAGIC has been successfully applied in projects in The Netherlands and Sweden. Respondents generally had few difficulties in understanding the interactive tasks, using the computer screens and providing the requested information. Thus, at least in terms of face validity and general experiences with the application of this computer program, it seems safe to conclude that interactive computer experiments are potentially valuable for collecting activity-related information on individuals' scheduling and rescheduling behavior, required to estimate activity-based models or for the ex ante evaluation of policies.

As for all stated measurement procedures, it remains an issue whether behavior observed under quasi-experimental conditions is systematically related to behavior in the real world. We cannot be sure whether respondents' choices under these quasi-laboratory conditions reflect their daily behavior or are merely a response to a computer game. In particular, if the experimental task is quite different from their everyday decisions, it may be difficult to grasp and understand the specific situation and provide reliable and valid responses. This measurement procedure may also be questioned if a high degree of detail is requested. A solution might be to focus on the basic principles and try to validate the principles that underlie the response patterns in these hypothetical situations, as opposed to the specific patterns themselves. On the other hand, there may be situations where real-world observations are simply not available. The question then is whether these interactive computer experiments provide a better understanding of activity scheduling decisions than expert knowledge.

CONJOINT EXPERIMENTS

Although conjoint analysis (or stated preference techniques, as they tend to be called in the transportation literature) has gained increased popularity, attempts at building conjoint-based activity models are still scarce. Existing attempts are limited in that only particular components of a full activity model are addressed. For example, Timmermans (1988) examined the issue of multipurpose behavior, and later elaborated on this approach (Timmermans and van der Waerden, 1993; Timmermans, 1996). In these cases, the options of trip-chaining are defined a priori. Fractional factorial design is used to build choice sets, and respondents are requested to choose the alternative they like best. Responses are then aggregated and analyzed through some a priori assumed-choice model. The results of this analysis reflect the contribution of the underlying factors in predicting the probability that a particular trip structure will be chosen. The original conjoint experiments developed along these lines were quite simple, while the later ones allowed the researcher to quantify substitution effects by using mother logit models (Timmermans and van der Waerden, 1993), as opposed to multinomial logit models.

This modeling approach allows one to calculate the utility of the choice alternatives, but not to break this down into its constituent components. In another project, Dellaert et al. (1996, 1997) developed a methodology for such portfolio choices. The starting point of this project was that the assumptions underlying conventional stated choice analysis may not be valid for portfolio choice. Conventional stated choice models assume that there is only one evaluation process that simultaneously determines the utility of all activities in an activity pattern. However, when consumers make choices between activity patterns, different activities in the pattern may be evaluated at different stages. For example, a series of interrelated subchoices (e.g., timing, activity type and destination) may be made to generate the overall activity pattern. This implies that additional assumptions are required to model how these subchoices are interacting and are integrated to arrive at the choice of a complete activity pattern.

Different conceptualizations will lead to different formal representations. The utility functions of different subchoices may have their own error components. Also, consumers may change their preferences when the number of alternatives free to choose from is reduced. Dellaert et al. (1997) describe the properties of a joint logit model, a heterogeneous logit model and a set of separate logit models to represent different situations.

The experimental designs needed to support stated choice experiments that allow one to estimate these different models then depend on the assumptions one is willing to make. These designs should support estimates of (i) interaction effects between preferences for activities undertaken at different moments in time; (ii) differences in error components between activity choices at different stages in the activity pattern; and (iii) variations in preferences that may occur over different stages of the activity pattern. The authors propose an activity pattern stage-dependent experimental approach that combines a set of experimental designs for different types of pattern choices. A set of N conditional subdesigns is constructed in which, for each stage n, only the activities of the activity pattern alternatives vary in each choice set. In addition, one subdesign is constructed that varies the attributes of all activities. To estimate interactions, the design has to be constructed in such a way that interactions between activities in different periods can be estimated independently from the main effects of those periods. Potential differences in preferences for activities when undertaken at different moments in time can be determined by allowing these activities to reoccur in several periods in the design.

Although this data collection method allows one to apply principles of conjoint experiments to activity patterns analysis, this approach is in fact rather simplistic in that many of the choice dimensions (e.g., transport mode, sequencing, accompanying persons) of a full-scale activity-based model of transport demand are not addressed. The projects described in this section should be considered as small- scale projects to test some of the model specifications and design strategies that are required to develop a full-scale conjoint model of activity patterns.

Over the last year we have been working on a more general conjoint-based activity model (Wang et al. 1997). This involves, first, a breakdown of the activity pattern generation and scheduling problem into an activity participation and an activity scheduling subproblem. The first subproblem is meant to indicate which activities will be conducted on a particular day, as a function of contextual variables. The probability that a particular program, defined as a set of activities, will be carried out can then be calculated. Because sequencing decisions will necessarily depend on the number and kind of activities that are performed, separate experiments need to be designed for these combinations. Although the construction of such designs does not cause any specific problems, the fact that separate designs are required for specific activity programs implies a labor-intensive task. If it is assumed that the actual scheduling also depends on space-time constraints and the attributes of the destinations, then these factors need to be varied in the experimental design. For most situations, the construction of such experiments is feasible. There is, however, one issue that requires special attention. Because one needs to vary several factors simultaneously, the degrees of freedom left are sometimes low, or in some cases, infeasible space-time patterns may result.

Thus, the construction of experimental designs to develop conjoint-based models of activity scheduling behavior requires considerable effort and expertise. The results of our projects, however, indicate that the estimated utility functions are consistent with the findings of qualitative research, given support to the face validity of the approach. It is fair to say, however, that the task of building a full-scale activity model derived from conjoint experiments is overwhelming. The problem is that separate designs have to be constructed to estimate the various components of the model. Although considerable progress in finding appropriate specifications and design strategies has been made, it is difficult to imagine that a full model can be developed using common methods of data collection.

In fact, it seems worthwhile to investigate the potential of interactive, condensed-time computer simulations to collect the necessary data. This would involve creating environments that meet the conditions of the underlying experimental design and invite respondents to make a series of interrelated choices. This is, however, a costly option, as respondents need to come to the laboratory to complete the interactive computer simulations. Currently, the development of a virtual reality system for conjoint analysis is on its way (Dijkstra and Timmermans, 1997a, 1997b), but the system has not yet been tested in larger scale experiments. With increased computing power, such systems could be accessed through the Internet. If one then would have a panel, very interesting new means of data collection would emerge.

It seems that, at the moment, the contribution of conjoint experiments to building activity-based models of transport demand is primarily restricted to estimating particular relationships in the overall model, or estimating respondents' reaction to transport or other policies. Conjoint experiments allow one to control the covariance structure of the data, which has particular advantages in theory testing. In many cases, however, techniques other than those commonly used in stated preference analysis in transportation research are required. Many decisions are context-dependent, hence context-dependency should be included (Oppewal and Timmermans, 1991). Many decisions are related, implying that models of sequential or portfolio choice seem the natural choice. Respondents also have difficulty in conceptualizing a completely different environment. Hence, design and model strategies that would focus on changes in their *current* environment have the potential advantage of increased reliability and validity.

Finally, the choice of activity patterns to some extent involves group decisions. Conjoint models of group decisions are therefore more likely to capture the nature of the decision-making process than the currently used conjoint models, which are derived from the responses of individuals.

KNOWLEDGE ACQUISITION

As we have seen, one of the approaches in activity-based modeling is the rule-based approach. The quintessence of this approach is to represent activity-scheduling decisions in terms of qualitative rules. This focus leads to the question of how such qualitative rules should be derived. One theoretical option is to derive such rules from observational data. This means that rules should be found that describe observed activity patterns. We are, however, not aware of any transportation literature along these lines. In fact, this is the very reason why Kitamura et al. (1995) suggested the use of stated preference techniques and neural networks. In terms of data collection, the elicitation of rules from observed activity patterns is not causing a new challenge, as diary data could be used.

Another and, in fact, dominant option in the literature on production systems is to use techniques of knowledge acquisition to find the rules and simulate activity-scheduling behavior. Although never expressed in these terms, the early gaming simulations such as CARLA can be considered as an example of such an approach. These games were used to identify the conditions, constraints and mechanisms that influence people's activity-scheduling decisions. If this qualitative information would be captured in terms of IF <…>, THEN <…>, ELSE rules, a simulation system could be developed.

Another possibility is the use of decision nets (Timmermans et al., 1987). Compared with the games or similar methods such as protocol analysis, decision nets constitute a more structured interview. In particular, the decision net method involves, first, identifying the factors that influence a particular decision. Next, for each of these factors, one examines its specific nature. Respondents are requested to indicate, first, the level of each factor at which one would no longer consider the choice alternative. For example, if we assume that the amount of time left to go shopping is one factor influencing an activity-scheduling decision, one would ask the respondent at what minimum time duration he/she would decide not to go shopping anymore on a particular day. Once this critical level is known, the respondent is asked about his/her decision if the time is less than this threshold. In principle, respondents can provide three answers. First, they may decide not to consider the choice alternative anymore. In this case, one has a rejection-inducing dimension. A second answer might be that the alternative will still be considered if some other condition is met. This is the situation of a compensatory dimension, in the sense that a low evaluation of one condition or factor might be compensated by other factors. Finally, respondents might give an indication of relative preference.

The construction of such decision nets involves considerable time. Interview sessions of a few hours are not an exception. Moreover, there may be some concern about whether respondents can actually rationalize their decisions. To some extent, this is also true for stated preference techniques. However, in the case of stated preferences, one is primarily concerned with the ultimate responses to experimentally varied conditions, and not with the specific rationale behind such decisions. Regardless of this issue, the verification and validation of decision nets are difficult, as the standardized interviewing protocol does not include the mechanisms required to prove the exclusivity and exhaustiveness of the rules that can be derived from the decision nets.

Decision tables are an alternative representation, and have the advantage of formal logic. A decision table can be defined as a table that represents the exhaustive set of exclusive conditional statements to represent the problem domain. It consists of conditions and actions. Every column of the table represents a decision rule. Given the complexity of the activity-scheduling problem, a single decision table will be insufficient. Therefore, a system of decision tables will be required, in which the separate tables are linked. The ALBATROSS model is based on this representation format.

In terms of data collection, the construction of decision tables requires personal interviews and hence is rather labor- and cost-intensive. Moreover, the involved rules tend to be deterministic and do not have an associated error theory. Consequently, small changes in the conditions may lead to the rejection of alternatives. Work on fuzzy decision tables, however, is on its way (Wets et al., 1996a, 1996b).

FIELD WORK AND GEOGRAPHICAL INFORMATION

All the methods of data collection discussed above concern the measurement of activities or the mechanisms that are supposed to drive the underlying decision-making processes. However, activity-based models typically include the relationships between the spatial and transportation environment, the institutional context and the participation in and duration of activities. Consequently, one also needs to collect data on the environment and the institutional context. Although commercial databases are often available, they tend to differ in their definitions and, at least in The Netherlands, they are very costly. Hence, additional fieldwork will often be required. A variety of sources is typically needed to gather information about opening hours, attractiveness of locations, land use patterns, travel times, characteristics of the transportation network, etc.

Geographical information systems offer the framework to geocode all this information. Although there usually are no major problems, putting all relevant information in a GIS represents a considerable amount of work. Different sources of information are often based on different operational definitions, classifications always seem to differ, the spatial resolution varies from source to source, etc. The inclusion of the diary data themselves also represents a challenge, as this information is time-dependent and geographical information systems do not have an explicit time dimension. Hence, some efficacy may be lost.

DATA QUALITY REQUIREMENTS

In our discussion to this point, we have focused on data needs and issues in methods of data collection. We have been asked to comment also on the data quality requirements. Compared with the literature on data collection methods, the literature on data quality is scare and often not well-conceptualized. In fact, data quality is a multidimensional concept that has received different meanings in different disciplines. Although it is not intended to discuss the multidimensionality of the concept in travel analysis at any particular length, a few remarks are required to better position our discussion of the consequences of increasing model complexity on data quality requirements. In particular, we will introduce a checklist of data quality requirements that might help modelers to decide how much attention the various elements of the data collection need, and in case of limited resources, how best to allocate these to the various data sources that are required to build an activity-based model of transport demand.

A first requirement is that of *reliability*, which can be defined as the consistency with which an instrument or question measures whatever it does measure. The aspect of reliability comes in many different forms when collecting data on activity patterns. For example, when detailed activities have to be classified into activity classes, different interviewers may use different schemes. Likewise, the choice of fixed versus open time intervals may lead to differences in reported time taken to conduct particular activities. Parts of the diary may be incomplete, or errors may occur. Many other factors are likely to influence the reliability of the collected data. Data quality can be improved by developing detailed and consistent coding and classification instructions.

Perhaps the most important aspect of reliability is the number of errors in the diaries. Durations may not sum to 24 hours, travel may be included in activity duration data, the activity may be inconsistent with the land use, respondents may not report the location of an activity if they have reported it earlier in the diary, etc. If such errors are not detected and corrected, the activity-based model may capitalize on such unreliable information. Hence, it is of the utmost importance to detect such errors and clean the diaries. As we have noted before, we have developed SYLVIA for performing such analyses. This computer program uses a set of rules to detect missing information and potential anomalies. In addition, the kind of errors is cross-tabulated against the kind of diary and sociodemographics, allowing the researcher to analyze whether the reliability of the diaries is related to sociodemographics. The system also contains rules to repair the diaries. Because some of these rules may be conflicting, and the repair of one error may lead to the introduction of another, algorithms have been developed to guarantee the reliability of the diaries along more dimensions. Not all potential errors can be detected and repaired by general rules. Therefore, the system also allows the user to select the diaries with a particular error and repair these interactively or delete them altogether.

Data should not only be reliable, they should also be valid. Indeed, *validity* is perhaps the foremost aspect of data quality; it is a multidimensional concept by itself. In general terms, validity refers to whether the selected variables measure the concept they intend to measure and encompasses a wide range of notions related to conceptualization, sampling, and choice of the right measurement instrument. Various kinds of validity could be mentioned for these various aspects, but for the present discussion, the term refers to whether the data truly measure representative activity patterns along all dimensions identified by the model.

A third aspect of data quality that we distinguish is *level of detail*. This requirement is meant to indicate whether the operational definitions of the choice dimensions and explanatory variables are consistent with the aim of the model-building process. For example, the timing of activities may be coded in terms of broad categories (morning, afternoon, evening), or in, for instance, 5-minute intervals. This level of detail dictates the level of possible detail in the predictions of the activity-based model. If the model is supposed to predict activity patterns and related demand across the day, this should be reflected in the level of detail in the data. Another example is the level of detail in the locational data. If space-time constraints are critical in the model, the locational data should be available at the desired level of resolution. If the model is supposed to check the feasibility of an activity schedule, more detail is required compared with a model that is not based on such feasibility checks.

Another aspect of data quality is *consistency*, which is defined in this paper to indicate whether the same definitions and classifications have been used across the data sources. If the researcher has control over the data collection process, consistency should not be an issue. However, in many cases, activity-based models need to be based on different sources of information, often collected for general purposes, and certainly not with a specific model in mind. Different sources may use different classifications of activities, different classification of land uses, etc. In this case, it is important to develop schemes that link the various classifications.

Related to consistency is the aspect of *completeness*. It refers to the notion that one should always be able to classify every unit of observation. Specific activities should all be related to the activity classes that are used in the model-building process: all time should be accounted for, one should be able to systematically link all land-use categories to the activity classes, etc. Again, to improve data quality on this aspect, a tested set of instructions is very important. In our own work, we checked whether all respondents, activities, destinations, time periods and transport modes could be subsumed under all relevant classifications. Also, we checked whether the classifications could be linked.

Finally, we distinguish four aspects of data quality that are related to the efficient and effective use of data. First, data should be easily *accessible*, that is, available on modern media, using the same representation format. Second, because the world is constantly changing, and one needs to incorporate this in the application of the model, the databases should be *dynamic*, suggesting that they should be updated immediately when change occurs. Third, we identify the aspect of *flexibility*. The modeler will often have to operate at different spatial levels of resolution, or at different time frames. For example, while the module on activity generation might involve broad time frames, a module for checking the feasibility of activity schedules is likely to be defined in terms of minutes. The data should allow this flexibility, which in practice implies that a consistent set of codes for various levels of spatial resolution and time frames should be maintained. Finally, we believe that *integration* is important, expressing the idea that different data sources should have common elements to allow the user to link the various data sources. Various common elements may be envisioned, but in particular, geocoding seems critical in building activity-based models of transport demand.

CONCLUSION AND DISCUSSION

Recent developments in transport modeling have stimulated the interest in building activity-based models of transport demand. As Jones (1995) has argued, activity-based approaches offer a broader perspective and richer set of concepts from which to study travel behavior than the traditional view of travel as the movement of vehicles, people and goods over transport networks. Because a richer set of concepts is involved, activity-based models involve a considerably greater complexity than traditional gravity, entropy and discrete choice models of transport demand. Moreover, and as a result of this increased complexity, the data needs and data quality requirements of activity-based models are considerably higher.

In this paper, we have discussed some of the issues related to data needs, data quality requirements and methods of data collection that are pertinent to the further development of activity-based models. We have argued that, from a modeler's perspective, the data collection process should be derived from the specification and objectives of the model. This implies that, in many cases, available data, collected for general purposes, may not be sufficient to develop a particular kind of model, or the complexity and hence potential use of the model may be prohibited by the nature and limitations of available data.

Starting from this perspective, we have briefly reviewed part of the available literature on methods of data collection and have made some recommendations for future data collection. In addition, we have briefly commented on some specific aspects of data quality that we felt were important in building an activity-based model. As illustrated by our real-life decisions, we feel that, often, a mixture of different methods of data collection presents the best of both or a multitude of worlds.

In preparing this contribution, it became abundantly clear that, with a few notable exceptions, the transportation community has not spent that much effort on systematically examining (under controlled conditions) the linkages between model building and data collection. If it is realized, however, that a model can never be better than the data it is built on, methods of data collection should be as sophisticated as the models they support.

ACKNOWLEDGMENTS

The authors are grateful to John Bates, Juan de Dios Ortuzar and two anonymous referees for their constructive comments on the draft paper. Their feedback has considerably improved the readability of the paper, and forced us to further articulate some thoughts and notions discussed in the paper. Any errors remain the responsibility of the authors.

REFERENCES

1. Adler, T., and M. Ben-Akiva (1979). A theoretical and empirical model of trip chaining behavior. *Transpn. Res. Rec*. 13B, 243-257.
2. Ampt, E. (1989). Comparison of Self-Administered and Personal Interview Methods for the Collection of 24-hour Travel Diaries. Proceedings of the WCTR, pp. 195-206.
3. Ampt, E., and A. J. Richardson (1994). The Validity of Self-Completed Surveys for Collecting Travel Behavior Data. Proceedings of the 22nd European Transport Form Conference, Seminar H: Transportation Planning Methods, Vol. II, pp. 77-88.
4. Arentze, T., and H.J.P. Timmermans (1998). SYLVIA, A System for the Logical Verification and Induction of Activity Diaries, EIRASS, Eindhoven University of Technology, Eindhoven, The Netherlands.
5. Bagatta, G. (1995). European Time-Use Survey Pretest Italy, Report ISTAT.
6. Barnard, P. O. (1984). Use of an Activity Diary Survey to Examine Travel and Activity Reporting in a Home Interview Survey. Proceedings ATRF Conference, Adelaide, pp. 273-302.
7. Ben-Akiva, M., and J. Bowman (1995). Activity-Based Disaggregate Travel-Demand System with Daily Activity Schedules. Paper presented at the Workshop on Activity-Based Analysis, Eindhoven, The Netherlands.
8. Carpenter, S., and P. Jones (1983). *Recent Advances in Travel Demand Analysis*. Gower, Aldershot.

9. Clarke, M., M. Dix and P. Jones (1981). Error and uncertainty in travel surveys. *Transpn.*, 10, 105-126.
10. Dellaert, B. G. C., A. W. J. Borgers and H. J. P. Timmermans (1996). Conjoint choice models of joint participation and activity choice. *Int. J. of Res. in Mark.*, 13, 251-264.
11. Dellaert, B. G. C., A. W. J. Borgers and H. J. P. Timmermans (1997). Consumer activity pattern choice: development and test of stage-dependent conjoint choice experiments. *J. of Ret. and Cons. Services,* 4, 25-39.
12. Dijkstra, J., and H.J.P. Timmermans (1997a). Exploring the Possibilities of Conjoint Measurement as a Decision-Making Tool for Virtual Wayfinding Environments. Paper presented at CAADRIA 97, Taiwan, April 17-19, 1997.
13. Dijkstra, J., and H.J.P. Timmermans (1997b). The Application of Conjoint Measurement as a Dynamic Decision-Making Tool in a Virtual Reality Environment. Paper presented at CAAD Futures Conference, Munich, August 4-6, 1997.
14. Dijst, M. (1993). Gewijzigde Opzet van Dagboekonderzoek Succesvol, OSPA, TU-Delft, Delft.
15. Dowling, R. G., and S. B. Colman (1995). Effects of increased highway capacity: results of household travel behavior survey. *Transpn. Res. Rec.*, 1493, 143-150.
16. Ettema, D. F., A. W. J. Borgers and H. J. P. Timmermans (1993). A simulation model of activity-scheduling behavior. *Transpn. Res. Rec.*, 1413, 1-11.
17. Ettema, D. E., A. W. J. Borgers and H. J. P. Timmermans (1994). Using Interactive Computer Experiments for Identifying Scheduling Heuristics. In: *Preprints of the 7th International Conference of the Association for Travel Behavior Research*. Santiago, Chile.
18. Ettema, D. E., A. W. J. Borgers and H. J. P. Timmermans (1997). SMASH (Simulation Model of Activity-Scheduling Heuristics): empirical tests and simulation issues. *Transpn. Res. Rec.*, to appear.
19. Ettema, D. E., and H. J. P. Timmermans (1997) (eds.). *Activity-Based Approaches to Travel Analysis*. Pergamon, Elsevier Science Limited, Oxford.
20. Ettema, D., H. Timmermans and L. van Veghel (1996). Effects of Data Collection Methods in Travel and Activity Research, European Institute of Retailing and Services Studies, Eindhoven University of Technology, Eindhoven, The Netherlands.
21. Fellendorf, M., T. Haupt, U. Heidl and W. Scherr (1997). PTV Vision: activity-based demand forecasting in daily practice. To appear.
22. Gärling, T., K. Brännäs, J. Garvill, R. G. Golledge, S. Gopal, E. Holm and E. Lindberg (1989). Household activity scheduling. In: *Transport Policy, Management and Technology Towards 2001: Selected Proceedings of the Fifth World Conference on Transport Research,* Vol. 4, pp. 235-248. Western Periodicals, Ventura, CA.
23. Gershuny, J. (1990). International comparisons of time budgets. In: *Zeitbudgeterhebungen* (Von Sweitzer et al., eds.). Metzler-Poeschel, Stuttgart.
24. Gershuny, J. (1992). Time-Budget Research in Europe, EUROSTAT, Doc E3/IS/5/92.

25. Gershuny, J., I. Miles, S. Jones., C. Mullings, G. Thomas and S. Wyatt (1986). Time budgets: preliminary analyses of a national survey. *The Quarterly J. of Soc. Affairs*, 2, 13-39.
26. Golob, T. F., and H. Meurs (1986). Biases in response over time in a seven-day travel diary. *Transpn.*, 10, 163-181.
27. Goodwin, P. (1978). Intensity of car use in Oxford. *Traffic Eng. & Control*, 514-517.
28. Hanson, S., and J. O. Huff (1982a). Assessing day-to-day variability in complex travel patterns. *Transpn. Res. Rec.*, 891, 18-24.
29. Hanson, S., and J. O. Huff (1982b). Systematic variability in repetitious travel. *Transpn.*, 15, 111-135.
30. Harrison, B. (1986). Electronic road pricing in Hong Kong. *Traffic Eng. & Control*, 13-18.
31. Harvey, A. S. (1993). Guidelines for time use data collection. *Soc. Indicators Res.*, 30, 197-228.
32. Hayes-Roth, B., and F. Hayes-Roth (1979). A cognitive model of planning. *Cogn. Sc.*, 3, 275-310.
33. Huff, J. O., and S. Hanson (1986). Repetition and variability in urban travel. *Geo. Analysis*, 18, 97-114.
34. Huff, J. O., and S. Hanson (1990). Measurement of habitual behavior: examining systematic variability in repetitive travel. In: *Developments in Dynamic and Activity-Based Approaches to Travel Analysis* (P. Jones, ed.), pp. 229-249. Avebury, Aldershot.
35. Huigen, P. P. P. (1986). Binnen of buiten bereik?: een sociaal-geografisch onderzoek in Zuidwest-Friesland. Amsterdam/Utrecht: KNAG/Universiteit van Utrecht, Nederlandse Geografische Studies 7.
36. Jones, P. M. (1990). *Developments in Dynamic and Activity-Based Approaches to Travel Analysis*. Avebury, Aldershot, England.
37. Jones, P. M. (1995). Contribution of Activity-Based Approaches to Transport Policy Analyis. Paper presented at the Conference on Activity-Based Approaches, Eindhoven, The Netherlands, May 25-28, 1995.
38. Juster, F. T. (1985). The validity and quality of time-use estimates obtained from recall diaries. In: *Time, Goods and Well-Being* (F. J. Juster and F. P. Stafford, eds.), pp. 63-91. Institute for Social Research, University of Michigan, Ann Arbor.
39. Juster, F. T. (1986). The effect of recall period on the quality of time diary data. In: *Time-Use Studies: Dimensions and Applications* (A. Dagfinn, A. S. Harvey, E. Wunk-Lipinsky and I. Niemi, eds.), pp. 149-175. Central Statistics Office of Finland, Helsinki.
40. Kalfs, N. (1992). Time-budget research: three data collection modes. In: *Time-Use Methodology: Towards Consensus*, pp. 283-285. ISTAT, Rome.
41. Kalfs, N. (1993). Hour by Hour: Effects of the Data Collection Mode in Time-Use Research, Ph.D. thesis, NIMMO, Amsterdam.
42. Kalfs, N. (1994). The construction of a new electronic diary. In: *Fifteenth Reunion of the International Association for Time-Use Research* (N. Kalfs and A. S. Harvey, eds.), pp. 246-260. NIMMO, Amsterdam.
43. Kalfs, N. (1995). Effects of different data collection procedures in time use research. *Transpn. Res. Rec.*, 1493, 110-188.

44. Kalfs, N., and W.E. Saris (1997). New data collection methods in travel surveys. In*: Activity-Based Approaches to Travel Analysis* (D.F. Ettema and H.J.P. Timmermans, eds.), pp. 243-262. Pergamon, Oxford.
45. Kalfs, N., and J. van der Waard (1994). Kwaliteit van Gegevens van Tijdbestedings-en Verplaatsingsdagboeken, Colloquium Vervoersplanologisch Speurwerk, Deel 1, pp. 889-907.
46. Kitamura, R., R.M. Pendyala, E.I. Pas and D.V. Reddy (1995). Application of AMOS: An Activity-Based TCM Evaluation Tool Applied to the Washington, D.C. Metropolitan Area. Proceedings of the 23rd Summer Annual Meeting, PTRC, London, pp. 177-190.
47. Klevmarken, N. A. (1982). Household Market and Non-Market Activities: A Pilot Study, University of Goteborg, Department of Statistics.
48. Koppelman, F. S. (1981). Nonlinear utility functions in models of travel choice behavior. *Transpn.*, 10, 127-146.
49. Kreibach, V. (1979). Modeling car availability, modal split and trip distribution by Monte Carlo simulation: a short way to integrated models. *Transpn.*, 2, 153-166.
50. Kutter, E. (1984). Integrierte Berechnung Städtischen Personenverkehrs. Technische Universität Berlin.
51. Lawton, K., and E. Pas (1995). Household Travel Surveys: New Concepts and Research Methods. Paper presented at the Conference on Household Travel Surveys: New Concepts and Research Needs, Irvine, CA, March 12-15.
52. Lenntorp, B. (1976). Paths in space-time environment: a time geographic study of possibilities of individuals. Lund: The Royal University of Lund, Department of Geography. Lund Studies in Geography, Ser. B. Human Geography, No. 44.
53. Libs, M., and L. Lebart (1994). Analysis of media-contact frequency as measured by self-completion questionnaire and face-to-face interview. In: *Proceedings of the RC-33 Sessions on Nonresponse and Measurement Error* (J. van der Zouwen and E. de Leeuw, eds.). Amsterdam.
54. Lingsom, S. (1980). Open and Fixed International Diaries: A Pilot Study on Time Use, Statistisk Sentralbyra, Oslo.
55. Lyberg, I. (1989). Sampling, Nonresponse and Measurement Issues in the 1984/85 Swedish Time Budget Survey. Paper presented at the US Bureau of the Census Fifth Annual Research Conference, Washington, March 19-22.
56. Mahmassani, H. S., S. G. Hatcher and C. G. Caplice (1991). Daily Variation of Trip Chaining, Scheduling and Path Selection Behavior of Work Commuters. Proceedings, Sixth International Conference on Travel Behavior: Methods for Understanding Travel Behavior in the 1990s, Quebec, pp. 29-45.
57. Meyburg, A. H., and W. Brög (1981). Validity problems in empirical analyses of non-home-activity patterns. *Transpn. Res. Rec.*, 807, 46-50.
58. Murakami, E., and W. T. Watterson (1992). The Puget Sound transportation panel after two waves. *Transpn.*, 19, 141-158.
59. Niemi, I. (1993). Systematic error in behavioral measurement: comparing results from interview and time-budget studies. *Soc. Indicators Res.*, 30, 229-244.
60. Oppewal, H., and H. J. P. Timmermans (1991). Context effects and decompositional choice modeling. *Pap. in Reg. Sc.*, 70, 113-131.

61. Panizon, F., and M. Venturi (1992). Data processing and control in the Italian time-use survey: analysis and results for improving data quality. In: *Time-Use Methodology: Towards Consensus*, pp. 319-335. ISTAT, Rome.
62. Pas, E. I., and F. Koppelman (1987). An examination of the determininants of day-to-day variability in individuals' urban travel behavior. *Transpn.*, 14, 3-27.
63. PTV (1987). VISUM and VISEM. Product description, Karlsruhe.
64. Recker, W. W., M. G. McNally and G. S. Root (1986a). A model of complex travel behavior. Part 1: theoretical development. *Transpn. Res.*, 20A, 307-318.
65. Recker, W. W., M. G. McNally and G. S. Root (1986b). A model of complex travel behavior: Part 2: an operational model. *Transpn. Res.*, 20A, 319-330.
66. Robinson, J. P. (1985). The validity and reliability of diaries versus alternative time- use measures. In: *Time, Goods and Well-Being* (F. J. Juster and F. P. Stafford, eds.), pp. 33-62. Institute for Social Research, University of Michigan, Ann Arbor.
67. Roveri, L. (1992). Nonresponse patterns: experience from the Italian time-use survey. In: *Time Use Methodology: Towards Consensus*, pp. 301-309. ISTAT, Rome.
68. Rydenstam, K. (1995). *The Harmonised European Time-Use Surveys*. Statistics Sweden.
69. Scheuch, E. K. (1972). The time-budget interview. In: *The Use of Time* (A. Szalai, ed.), pp. 69-87. Mouton, The Hague-Paris.
70. Schmiedel, R. (1984). Bestimmung Verhaltensähnlicher Personenkreise für die Verkehrplanung. Universität Karlsruhe, Karlsruhe.
71. Sen, A., S. Sööt, L. Yang and E. Christopher (1995). Household travel survey nonresponse estimates: the Chicago experience. *Transpn. Res. Rec.*, 1493, 170-178.
72. Sparmann, U. (1980). Ein Verhaltensorientiertes Simulationsmodell zur Verkehrsprognose. Universität Karlsruhe, Karlsruhe.
73. Stopher, P. R. (1992). Use of an activity-based diary to collect household travel data. *Transpn.*, 19, 159-176.
74. Swiderski, D. (1982). A model for simulating spatially and temporally coordinated activity sequences on the basis of mental maps. In: *Recent Advances in Travel Demand Analysis* (S. Carpenter and P. Jones, eds.), pp. 314-344. Gower, Aldershot.
75. Timmermans, H. J. P. (1988). Multipurpose trips and individual choice behavior: an analysis using experimental design data. In: *Behavioral Modeling in Geography and Planning* (R. G. Golledge and H. J. P. Timmermans, eds.), pp. 356-367. Croom Helm, London.
76. Timmermans, H. J. P. (1996). A stated choice model of sequential mode and destination choice behavior for shopping trips. *Env. & Plan. A*, 28, 173-184.
77. Timmermans, H. J. P., and R. E. C. M. van der Heijden (1987). Uncovering spatial decision-making processes: a decision-net approach applied to recreational choice behavior. *TESG*, 78, 297-304.
78. Timmermans, H. J. P., and P. J. H. J. van der Waerden (1993). Modeling sequential choice processes: the case of two-stop trip chaining. *Env. & Plan. A*, 24, 1483-1490.
79. Vause, M. (1997). A rule-based model of activity scheduling behavior. In: *Activity-Based Approaches to Travel Analysis* (D. F. Ettema & H. J. P. Timmermans, eds.), 371 pp. Pergamon, Elsevier Science Limited, Oxford.

80. Verweij, N, N. Kalfs, W.E. Saris and M. de Pijper (1987). Tijdbestedings-onderzoek middels tele-interviewing en de mogelijkheden voor segmentatie. *Massacommunicatie*, 4, 253-369.
81. Wang, D., A.W.J. Borgers, H. Oppewal and H. J. P. Timmermans (1997). A Stated Choice Approach to Develop Activity Scheduling Models. Paper presented to the IATBR Conference, Austin, Texas.
82. Wets, G., F. Witlox, H.J.P. Timmermans and J. Vanthienen (1996a). A fuzzy decision table approach for business site selection. In: Proceedings of the Fifth IEEE International Conference on Fuzzy Systems ’96, pp. 1605-1610. New Orleans, LA, September 8-11, 1996.
83. Wets, G., F. Witlox, H.J.P. Timmermans and J. Vanthienen (1996b). Choice modeling by means of fuzzy decision tables. In: Proceedings of the Fourth European Congress on Intelligent Techniques and Soft Computing, Vol.2, pp. 1335-1339. Aachen, Germany, September 2-5, 1996.

WORKSHOP SUMMARY

Juan de Dios Ortúzar
Department of Transport Engineering
Pontifica Universidad Catolica de Chile

John Bates
John Bates Services
England

INTRODUCTION

This short paper summarizes the conclusions of Workshop 9. The purpose of the workshop was to identify the types of data items required for state-of-the-art and state-of-the-practice transport planning models, to identify possible approaches to obtain these data and to formulate quality requirements for them. Finally, the workshop was enjoined to identify future research needs.

A number of specific themes were identified, and these are discussed below. In addition, however, there was some discussion about the wider role of models, not merely in respect of new departures, but in terms of different policy requirements, with the general aim to improve the quality of models. Distinctions were made between models for understanding behavior and models for advising politicians, between estimation and implementation, and between the very different data requirements for demand and supply modeling. In addition, there was considerable interest in the technological development of data collection (e.g., geocoding), and in the data requirements for models dealing with processes over time (dynamic).

We need to make forecasts; for this, we need adequate models and, for these, appropriate data. There is a general requirement for a professional approach to planning, which in turn will define a standard for the professional practice of data collection for transport models. The initial requirement is that models be taken seriously: from this, standards in models will dictate standards for collecting data. Thus, the whole issue might be viewed as an iterative process, as there are limits to what data can be collected and, thus, which kinds of problems can be addressed. In particular, if we intend to investigate model errors more seriously, this will usually require both more and better-quality data.

Technological advances open up new possibilities. Use of geocoded information should be a standard, as it is now cheap and allows collecting data very finely. In particular, GIS systems should replace zones. It is possible now to track times, routes and number of trips (e.g., using GPS devices). However, there are problems with non-household information and with measuring data for nonchosen options with a reasonable degree of precision.

More data are also needed to allow better segmentation strategies: these allow improved choice set definition, particularly in more complex cases such as destination choice. Differentiating between mandatory and discretionary trips is of particular importance: destination choice for the latter may be extremely flexible in some cases (e.g., dependent on the weather). If trip purpose is precisely defined, then the number of potential destinations diminishes greatly. An interesting possibility is to define market segments by type of policy. In another vein, problems with sample size for estimating complex models may explain some controversial results in the literature.

More research is needed in an area not well understood: how to explain the differences between the results of household surveys and volume counts on the road. Combining household and choice-based survey data seems promising in this sense, but little general practical guidance is available, despite the pioneering work of Cosslett.

Six areas of especial interest were identified for detailed discussion:

- Activity modeling
- Land use
- Freight
- Destination choice
- Parking
- Time-of-day choice (including dynamic assignment)

Finally, the workshop considered some statistical issues relating to appropriate sample size calculations.

The rest of the paper is organised under these headings: there was, however, insufficient time for a detailed discussion on destination choice. Workshop participants, whose knowledge and enthusiasm were instrumental in making it a success, are listed at the end of this paper.

ACTIVITY-BASED APPROACHES

It was agreed not to discuss this topic at length, as the resource paper (Arentze et al., 1997) concentrates on this topic in some detail, and readers are therefore referred to it directly. However, the problem of forecasting was mentioned (what is conditional forecast, given scenarios?), and also the problems of complexity and behavioral content (are discrete choice models really doing a better job than aggregate models?). More complex models need more complex data, for example, to tackle spatial and temporal constraints. For this, surveys must be collected at the home, but it is unclear how far the data collection process can be taken: many activities occur at the same time, and some of the “soft” constraints may be “harder” in certain cases. Some concern was expressed that most operational models were just extensions of discrete choice models (i.e., with activities rather than trips).

More specifically, it was agreed that two-day diaries are the most efficient way of collecting activity-based information. Although it is not possible to get full information about personal variability in this sense, at least it shows part of it. On the other hand, it was agreed that even for one-day trip diaries, use of an activity recall framework is the preferred option.

The general question was raised as to how activity-based models interact with the supply side. There seem to be no equilibrium formulations and, hence, no guarantee of convergence. It was accepted that activity-based approaches attempt to *model* choice sets rather than impute them and are, in general, more dynamic in nature. Also, they provide a scope to model the interchange of in-home activities and time spent out. However, as far as dealing with, for example, destination choice, the problem remains as large as in more traditional models.

LAND-USE TRANSPORT INTERACTIONS

Data are required about long-, medium- and short-term decisions in order to model behavior at these levels: residential location, place of activity choice (which is very important for nonfixed activities, such as shopping), frequency, mode and time-of-day choice. Precise data are not always forthcoming relating to location choice, and it is important to obtain duration information in order to explain the processes involved. Destination choice may be examined in a historical context (e.g., when did people arrive at a place, how long were they in another residence, did their children change school?), but this data may cause problems of interpretation. Other important questions relate to car use (e.g., when did people get a driver's license, if they bought a car, did it replace a previous one?).

Time of occurrence and duration of activities may be obtained from normal travel or activity diaries (i.e., it is not necessary to have a panel). At the minimum, it is recommended to add retrospective questions on residence and job changes in household questionnaires. In addition, there is a need for secondary data on (housing) stock prices and quality.

FREIGHT

There was a general consensus among workshop members that this is an underresearched area, particularly in the case of cities. There are no quality data on freight movements in urban areas. It is necessary to acquire data not merely on vehicle flows but also on commodities, and on the type of service vehicle moving them.

On the other hand, freight modeling has developed as an independent area of work in transport. Problems have even less generic qualities than in the case of passenger travel, and there are tremendous problems in building base patterns. It is important to bring the experts in this area to our meetings due to the convergence of policy needs.

DYNAMIC ASSIGNMENT AND TIME OF DAY CHOICE

In contrast to much of the behavioral analysis, the type of information required for dynamic assignment is network data in detail; this includes secondary data, measuring times and costs, etc. Data are also needed on route choice behavior (both for cars and public transport, and for regular and nonregular travellers). In this connection, it is very important to have, but rare to find, data on the nonselected elements of the choice set. In fact, choice set definition and measurements are key items of information: information about signposting and responses to variable message signs (VMS) is of interest here. In general, all the above seems to be underresearched areas.

A key issue is the collection of data for model validation, particularly in the context of changing traffic levels or conditions. Another interesting possibility is to compare traffic assignment models in different cities. The possibility of creating a generic test model data set was discussed

Time-of-day choice involves trading delays against being early or late. There is a problem of scheduling associated to each individual's preferred arrival time at his/her destination. In general, what is needed here is stated preference (SP) data, plus secondary information on entry- and exit-time restrictions, and flexibility of working hours.

PARKING

This is one of the most directly available policy variables, and it needs to be modelled more effectively (the same applies to pedestrian schemes). Parking should be integrated in the general modeling procedures (for example, the availability of parking is a determinant of mode choice), though there is evidence that results are not transferable across cultures or cities. Unfortunately, parking data collection is not simple: fieldwork on arrival/departures at parking spaces is expensive and not conclusive. There are problems with the supply side (parking capacity and duration); for example, what happens when there is no more room? This is particularly the case in private car parks. Also, duration is obviously different by purpose. For the demand side, it is possible to resort to SP data (by asking what happens when the car park is full), plus direct observations for validation purposes.

In terms of household surveys, it is necessary to ask not only how much was paid but also where the car was parked (although this is not recalled well in discretionary trips). In the case of trips by public transport by people with access to a car, it is important to know how much the parking cost would have been. In the case of surveys to employees, data on parking spaces should be sought.

SAMPLE SIZE FOR MODEL ESTIMATION AND VALIDATION

This is an area where virtually nothing sensible has been published anywhere. Only rules of thumb can be found (e.g., 30 observations per parameter), which may turn out to be absurd in practice, as they do not take account of data variability. While sample size calculations exist for pure data items, they do not for model predictions or outcomes (such as the value of time).

In model estimation, variables are retained if t-statistics are greater than standard critical values. Such “rules,” however, merely ensure that the *signs* are correct, and they assume that models are correctly specified (which they are not). Generally, therefore, our statistical procedures are ill-prepared for the tasks to which they are applied. SP data have their own share of problems (such as repeated observations).

Research is needed to define appropriate sample sizes (i.e., that guarantee a certain estimate of accuracy) for a typical transport study. An important question is whether it is possible to define a recommended procedure to assess confidence intervals.

Another important and neglected issue is that of validation samples. Several key questions about model worth cannot be assessed unless independent data are available to compare models.

REFERENCE

1. Arentze, T., Hofman, F., Kalfs, N., and Timmermans, H. (1997). Data needs, data collection and data quality requirements of activity-based transport models. *Proceedings, International Conference on Transport Survey Quality and Innovation (Transport Surveys: Raising the Standard)*, 24-30 May 1997, Grainau, Germany.

WORKSHOP PARTICIPANTS

The workshop chairman was Juan de Dios Ortúzar (Chile), the resource paper author was Harry Timmermans (Holland), and the discussant was John Bates (England). The rest of the workshop members were (in alphabetical order): Carlos Arce (USA), Patrick Bonnel (France), Michael Bryszewski (Poland), Jens Emig (Germany), Hugh Gunn (Holland), Dirk Heidemann (Germany), Norbert Klassen (Germany), Eckhard Kutter (Germany), Keith Lawton (USA), János Monigl (Hungary), Earl Ruiter (USA), Jeff Turner (UK), Wulf Wätjen (Denmark) and Manfred Wermuth (Germany). Tony Richardson (Australia) and Philippe Toint (Belgium) also made useful guest appearances.