SUMMARY OF WORKSHOP ONE: ACTIVITY AND TIME USE DATA NEEDS, RESOURCES AND SURVEY METHODS

MARTIN LEE-GOSSELIN
Laval University, Quebec

JOHN POLAK
Imperial College, London

Objectives

The workshop set out to address the three key questions posed by Peter Stopher in his opening address to the conference. In the context of the workshop these were interpreted as:

i. What elements of the activity-based perspective to data collection are both desirable and immediately available and what steps are needed to bring these elements into practice in the short term?

ii. What are the potential areas of application of activity-based data collection methods over the next 2 to 5 years and what opportunities and constraints are likely to influence the evolution of practice within this time frame?

iii. What steps are needed in terms of dissemination, training and research in order to promote the integration of activity-based methods into the mainstream of transportation analysis?

The workshop participants comprised practitioners at MPO, State and Federal level, consultants specialising in data collection, as well as those with a more general background in transportation modelling, academics and other researchers. Whilst the level of interest in activity-based methods was high, few practitioners had direct experience of applying these methods themselves and so understandings and expectations regarding the current status and future potential of the methods varied widely.

A particular difficulty surfaced in the opening session when it was discovered that a significant number of those present were unfamiliar with the fundamental differences between activity-based and trip-based survey instruments in current use. As the second session was on the succeeding day, it was possible, thanks to the quick action of several consultants and practitioners, to assemble and present a sample of recent instruments, which then served as examples in the remaining discussions.

General Issues

In addition to the broad issues raised by Peter Stopher, the contributions by the discussants Kay Axhausen and Ken Cervenka highlighted two further important issues that provided a backdrop to the deliberations of the workshop.

- The need to recognise that the complexity of the data which we would wish to collect and use is increasing (due to both new modelling requirements and the increased complexity of the policy environment). This raises important issues regarding assessing and improving the quality of data collection and about the optimal trade-off between quality and quantity in data collection.

- The need to recognise that the priorities and practices of transportation planning authorities vary greatly as do the resources and expertise available at the local level. New methods developed and training/dissemination activity undertaken should be sensitive to these variations. This means finding ways of facilitating some authorities in making a transition from 4-step to activity-based modelling whilst also providing mechanisms to support other authorities that currently undertake little or no formal modelling work in adopting the activity-based approach anew.

General Characteristics and Status of Current Activity-Based Data Collection Methods

Drawing on the diversity of background and experience of the participants, and the examples of instruments presented, the workshop attempted to clarify the characteristics and status of existing activity-based methods. The view was that activity-based data collection methods are defined as much by the use to which the data are put as by the method of collection *per se*. However, particularly in the case of travel demand data, the workshop was able to identify some broad characteristics that were felt to distinguish activity-based approaches. These included:

- A focus on the relationship between travel and the activities generating the demand for travel. This is a subtle but important difference. For example, whereas in a conventional (trip-based) approach a respondent might be asked to recall all the trips they made yesterday and the purpose of each trip, in an activity oriented approach they would be asked to recall the activities in which they participated and how they travelled between these activities.

- Developing naturally from the focus on activities, is a strong emphasis on issues of the sequencing and scheduling of behaviour. The activity based perspective emphasises the overall structure of activity/travel relations, both spatial *and* temporal.

- The emphasis on timing in turn leads to a concern with the dynamics of behaviour in terms of the relationship between different elements of behaviour within a day, the relationship of behaviour on different days and the effect of a wide range of changes in factors (socio-demographic, economic, technological or regulatory) on behaviour in the longer term.

- A final feature identified was an emphasis on household and institutional constraints as factors influencing travel behaviour. Just as trip making is viewed as part of a wider pattern of activity participation, so the behaviour of the individual is placed within the wider context of household decision making, taking account of the consequent inter-personal and institutional constraints placed upon freedom of action.

Elements of Activity-Based Models

In order to further explore data needs, Ken Kurani and John Polak laid out the main elements of activity-based models. These were summarised as:

- Activity generation/participation:
  - activity opportunities
  - time-use

- Activity scheduling:
  - timing, sequencing, duration
  - planning

- Activity execution:
  - network events
  - re-planning

- Dynamics:
  - within day
  - day-to-day
  - weekly
  - longer term

- Cognitive:
  - perceptions
  - learning

- Policy sensitivity:
  - information systems
  - pricing/tolling
  - regulation
  - non-transport measures
  - non-motorized transportation
  - land-use effects and impacts
  - in-home/out-of-home opportunities
  - derived analyses (environment, QOL...)
  - etc.

Specific Types of Activity-Oriented Data Collection Method

A number of different types of activity-oriented data collection method, relating both to demand and supply data were identified and discussed.

Demand Data

- *Activity diary surveys:* These are the activity-based counterpart of traditional travel diaries and are relevant to a similar range of modelling issues, principally the description of baseline population behaviour and the development of cross sectional models. They involve respondents reporting activities and travel rather than just travel and may also involve the collection of a limited amount of information regarding activities performed in the home. Experience suggests that activity diaries can be more effective at recovering mobility information than conventional trip diaries, especially information on short journeys and journeys made by non-motorised modes (which are often overlooked in trip diaries). A number of examples of activity/travel diary instruments have been used in the US in recent years and the benefits of the approach are beginning to be more widely (but by no means universally) accepted.

- *Panel surveys*: These are methods in which the same individuals are surveyed at two or more points in time, possibly, but not necessarily, by means of an activity diary. Panel surveys are principally of value in the assessment of impacts of new policies (before and after studies) and for the development of disaggregate dynamic models of travel demand (such as microsimulation models). The organisation and administration of a panel survey can be complex, but the data can be uniquely valuable, especially if the *turnover* of individuals who participate in different types of activity and travel has policy implications. A small number of panel surveys are currently in progress in the US, and there is more extensive experience of this method in some European countries.

- *Event-based surveys:* A variant of the panel survey and one that is especially suited to the investigation of long term demand issues is the event-based survey. In this approach respondents are tracked over a longer period of time, but rather than being surveyed at regular intervals (as is normally the case with panel surveys), then are only required to report key changes in circumstances as and when they occur (such as the acquisition or disposal of a car or a relocation of residence). Although experience with this type of survey is very limited, it was regarded as being potentially of considerable interest, especially in relation to modelling the relationship between demographic/land use change and mobility.

- *Stated Response surveys:* Most of these surveys, labelled "Stated Preference", involve presenting respondents with two or more hypothetical activity/travel situations and inviting the respondent to indicate which situation they would prefer or which they would choose. Some surveys on small samples involve more elaborate choice simulation exercises which are used to collect information on choice processes and the origins of choice-sets. The advantage of these approaches is that the researcher has control over the form of the hypothetical situations that are presented and therefore can explore new policy measures or complex behavioural processes that it would not be possible to address directly in the real world. Both these applications were regarded as being of considerable relevance to the development of activity-based models. The disadvantage of this approach is a concern over the validity of the results obtained and their transferability from a hypothetical context to the real world. Stated Response data is increasingly used by practitioners in the US, but some participants expressed concern about validity, transferability or even whether agencies who asked for such data really understood their limits. However it was noted that for some types of issue, Stated Response may be the only feasible approach and that methods exist to enable the 'blending' of Stated

Response and conventional Revealed Preference data, which can go some way towards addressing concerns over validity.

- *Passive data collection:* Developments in telecommunications and mobile computing are beginning to make it possible for certain types of behavioural travel data to be collected passively by means of remotely linked monitoring devices attached to vehicles and conceivably also to individuals. Accurate geocoding of trip ends and routes using Global Positioning Satellite systems, and real-time data transmission, are among the possibilities. Active research projects are currently underway in the US, Canada and Europe aimed at developing and testing appropriate systems.

Supply Data

- *Activity opportunities:* The main issue discussed in relation to supply data was how best to address the need that it was envisaged would arise (as a result of the development of highly detailed and disaggregate microsimulation models of travel demand) for much more detailed information on the spatial and temporal pattern of activity opportunities. This was seen as a major issue that can be only partially addressed through the enhancement of current digital databases. It was also regarded as raising difficult institutional issues concerning the relationship between the public and private sectors.

- *Network topology and performance:* In the same manner that detailed microscopic demand models were anticipated to give rise to new data requirements, so it was envisaged that the application of highly disaggregate traffic flow simulation models to larger and larger networks would give rise to a growing requirement for detailed and up-to-date network inventory data. Moreover, as both roadside and in-vehicle ITS systems are developed to play a larger role in some areas, so the scope of the inventory information required increases. Although this was not regarded as being as significant a problem as that posed by the activity opportunity data, a number of participants drew attention to the high costs associated with establishing and maintaining comprehensive network inventory data, and to the vulnerability of monitoring systems to funding cuts even where they have been established.

Progress Towards Consensus

The discussion revealed a tension between the preconceptions of practitioners and researchers. Understandably, many practitioners wanted a clear briefing on applicable and affordable new modelling and data collection techniques that would better equip them to deal with the increasingly complex policy issues. Perhaps equally understandably, researchers whilst quite clear about the benefits of activity-based approaches, preferred to emphasise the outstanding research issues and tasks.

Even though there was, in principle, more potential for consensus on the issue of data requirements and collection methods, we realised early in the workshop that there was little conventional wisdom on the strategies for selecting the optimal mix of data collection to feed the development of activity-based models. We need to draw particular attention to this insofar as there is a tendency to focus mostly on Revealed

Preference data, notably activity-travel diary methods, to the exclusion of other data sources.

It was unfortunately not possible to arrive at a clear plan for the short (0-2 year) or medium (2-5 year) term horizons, or to address all the dissemination and training aspects of Peter Stopher's charge to workshops. However, we took a position on some important, concrete issues: (i) the most important data collection "unknowns and partly knowns" which deserved urgent research, development and innovation in order to bring activity-based data collection into practice, and (ii) some priorities for coordinating the process of activity-based travel forecasring implementation from the data perspective. These are briefly summarised below.

Priorities for Research and Development - the List of "Unknowns and Partly-knowns"

These issues were considered under two headings: the content of the information required and the techniques and methods appropriate to collecting such data.

Content

- *Information on in-home activities:* Current activity-travel diary methods at most provide only a very crude specification of in-home activities. Yet for a variety of reasons (e.g. growth of virtual environments, teleworking and other forms of flexible employment) the substitution/complementarity between in-home and out-of-home activities is likely to become an increasingly important issue. It is not clear how much detail it is necessary or feasible to seek on this topic.

- *Level of detail in activity-based surveys vs level of detail in activity-based models:* The emerging activity-based modelling tools are highly disaggregate and in principle make large demands in terms both of the description of population characteristics and understanding of behavioural response mechanisms. However, it is not necessarily the case that all dimensions of these data are required simultaneously. For example, the ability to synthesise populations from marginal distributions of population characteristics means that it may be possible to significantly reduce the data requirements for model application. This is one of a number of areas in which the workshop identified the need to establish close cooperation between model developers and data collection specialists.

- *Dealing with the long term:* Long term aspects of travel behaviour are not well understood and few existing datasets contain appropriate information. It would therefore seem desirable to focus new effort in this area.

- There was a recognition of an increasing need to understand how and why people make the travel choices they do, not simply what choices they make. That is, there is a felt need for better "*cognitive" as well as "behavioural" data*. Stated Response and related methods are one of the few ways of addressing this need. However, issues of validity and transferability must be seriously addressed.

- At a more practical level, we need to have a better understanding how to deal with *"non-forecastable" explanatory variables* within a forecasting context.

Techniques

- *Non-response*. Problems of non-response and incomplete response are already serious issues in conventional travel surveys. The scope for these problems to increase in magnitude as we move towards potentially more complex survey instruments is substantial. Moreover, a peculiar advantage of data collection around activity patterns is the inclusion of all traveller segments. In short, activity-based models face more *complex* non-response and response bias problems than do trip-based models. It therefore appears urgent to better understand both how to reduce the incidence of non-response and, given that it will nevertheless continue to occur, how best to detect it and to deal with the consequences. More work is needed on the merits and problems of such strategies as the use of aide-memoires in telephone surveys, re-contacting respondents, rostering between responses from members of the same household, imputation, etc. Also, the workshop recognised growing support for an more "open", detailed reporting of non-response and incomplete response for different *subgroups* of interest. This implies that an adequate set of descriptors of both respondents and non-respondents is recorded, and it was noted that these may not be the same variables or classes which are required for the analysis of survey results.

- *Event-based methods to reduce respondent burden:* One aspect of attempting to reduce the problems of non-response is reducing respondent burden, and in this connection it was consider potentially worthwhile to investigate further the scope for "event-based" reporting strategies within behavioural surveys, especially those addressing longer term issues.

- *Integrated data strategies*. It is clear that no single data collection method is capable to furnishing all the required information. Rather, there is a need to integrate data from different sources. Techniques for approaching the problem of merging data from a diverse collection of sources are available, but are generally not well known within the transport research community. Further work in this area would therefore seem justified.

- *Talking the respondent's language.* Limited research has been undertaken into how respondents understand (or otherwise) the concepts that are presented to them in typical travel behaviour surveys. Elements of survey grammar are often driven by the grammar of a model. However, this may not be the most effective means of extracting sound data. Therefore research into respondents' comprehension and understanding of typical survey grammars would appear desirable. The ethics of passive data collection must also be addressed, although it was recognised that the problems in this area may have been exaggerated, and that the work most needed is on the institutional aspects of responsible data management (e.g. for the production of transparently anonymous disaggregate data files) and on the field testing of appropriate ways of obtaining informed consent.

- *The role of Stated Response methods.* These methods appear in principle to have a major role to play in the development of the behavioural response models that will be embedded in microsimulations of travel demand. There was felt to be a need to clarify the role of such techniques and in particular,

as pointed out above, to seriously address the issues of validity and transferability.

- *Exploiting technology:* Technologies are advancing rapidly and this opens up new opportunities for the collection and collation of data. The potential for individual passive monitoring has already been discussed and it would seem highly desirable to reinforce the existing research initiatives in this area. There is also important progress that could be made in terms of the secondary use of aggregate data sources such traffic monitoring system and automatic tolling systems.

Implementation Issues

The workshop identified three broad strategies that it believed would be of value in advancing the dissemination and implementation of activity-based methods.

- *Cooperation* between research into model development and research into data collection methodology. The need for research in these two areas to proceed jointly was regarded as being of paramount importance.

- *Test Bed:* The facilitate the co-development of modelling and data methods, the idea was put forward of "integrated sites" with unusual attention paid to the collection of both demand and supply data. Such integrated sites could act as test beds for the emerging data and modelling developments.

- *Justification:* Practitioners need both a rationale for shifting to from 4-Step to Activity-Based Travel Forecasting and a straightforward explanation of the differences in forecasts likely to be produced by the two types of model. The workshop noted that these should focus on the anticipated superior performance of activity-based approaches in estimating the effects of *simultaneous changes* in both the population and the supply characteristics. Ideally, more effort should be made to run both types of model in parallel in a small number of test cases, thus allowing planning agencies to weigh an increased amount of "real" evidence.

Other Issues Identified As Important

At the end of the workshop, participants also identified three areas as meriting discussion, but which did not get covered sufficiently in the time available to become the subject of recommendations:

- *More on uses of secondary or "support" data sources:* In addition to aggregate sources from transport systems (such as traffic monitoring system and automatic tolling systems, mentioned above), the workshop would have liked to examine the potential roles of such external sources as electronic directories to improve geocoding, or credit-card transaction databases.

- *The secondary analysis* of important existing travel data, as well as its "intelligent archiving" (meaning that the accumulated knowledge about the performance of data is preserved in a database system, together with the database itself).

- *Planning area boundaries:* How extensive an area should be covered by activity-based data

collection and the model which it feeds? Are MPO boundaries the only realistic definition of the appropriate universe?