

## **Data Structures, Sampling and Survey Issues**

**A.J. Richardson**

The Urban Transport Institute  
PO Box 363  
Alexandra, Victoria, 3714  
AUSTRALIA

**J. Wolf**

GeoStats  
530 Means St, NW, Suite 510  
Atlanta, Georgia 30318  
U.S.A.

Richardson, A.J. and Wolf, J. (2001). "Data Structures, Sampling and Survey Issues – Report of Workshop M6 ". 9<sup>th</sup> International Association of Travel Behaviour Research Conference, Gold Coast, Australia.

### **Introduction**

The Workshop on Data Structures, Sampling and Survey Issues took as its starting point a comment made by one of the Keynote Speakers at the conference. In discussing the perceived complexity of many Stated Preference survey designs, Jordan Louviere asked “what does it matter if we annoy the respondents, so long as we get some data for modelling?” Many of the participants in the workshop disagreed strongly with the sentiment behind this question, and argued that data obtained from “annoyed” respondents would not be of much use to model builders. It was felt that the principles of quality survey design should not be ignored simply because the end purpose of the data collection is to build models of one sort or another.

However, the Workshop participants were reminded that this was not a workshop at a conference on Survey Methodology, but a workshop on survey issues at a conference on travel behaviour research. To that end, the paper authors were asked to address two specific questions when presenting their papers:

- What major aspect of Travel Behaviour does your survey method address?
- How does your Survey Method enable that aspect to be better analysed?

Following on from these presentations, the workshop participants used the limited remaining time to identify major developments in the area over the past five years and to identify areas of future research.

## The Workshop Papers

The workshop contained nine paper presentations. Four were on the topic of general methodological advances, four were on the topic of the application of technology to surveys, and one was on the issues involved in the use of large volumes of microdata on travel behaviour.

Anable (2000) considered the issues involved in the collection of data for modelling the demand for leisure travel. Her paper highlighted the significant contribution of leisure travel to current patterns of travel and presented the hypothesis that the distinct characteristics of 'discretionary' travel warrant a radically different understanding of the factors influencing this behaviour. As a result, she argued the need for new approaches to data collection by explaining the rationale behind a survey of National Trust visitors to conservation and heritage sites in the United Kingdom which utilises psychological theories of behaviour and statistical segmentation analysis to develop mobility management policies in this sector.

It was considered that the survey techniques identified by Anable were particularly useful in addressing the following aspects of travel behaviour:

- Leisure travel and other discretionary travel
- Issues of car dependency
- Travel awareness and travel adaptation
- The formulation of strategies to influence travel behaviour.

The advantages of the methods identified by Anable in addressing the above travel behaviour issues were summarised as:

- The psychological basis was seen as being more applicable to leisure travel than the more conventional economic bases
- The methods highlight the role of perception in travel decisions
- The methods emphasise the linkage between travel and activity patterns.

Armoogum (2000) described the sampling issues involved in surveys of long-distance travel. Because of the relative rarity of long-distance trips, a number of specific issues arise in sample designs for long-distance travel surveys. For the calculation of overall mobility rates or distances travelled, it is important to collect information on both high and low frequency travellers. However, for the estimation of O-D matrices it is more important to collect information from high frequency travellers, since low frequency travellers contribute very little information to an O-D matrix. Armoogum described a method for optimising a sample design when faced with such conflicting objectives.

It was considered that the sampling techniques identified by Armoogum were particularly useful in the context of the following aspects of travel behaviour:

- Long-distance travel
- Other surveys of “rare” travel behaviour
- Surveys with conflicting sampling objectives.

The advantages of the methods identified by Armoogum in addressing the above travel behaviour issues were summarised as:

- The optimised sample is more cost-effective than other sampling designs
- The optimised sample can reduce the variance in estimates of model parameters
- By directly the sample towards those who are most relevant to the survey objectives, the method can reduce the annoyance felt by respondents.

Pratt (2000) described the concept of “piggybacking” onto existing surveys for obtaining new perspectives on travel behaviour by, firstly, adding questions to existing periodic surveys and, secondly, phrasing the questions in objective terms so that the responses can be compared across data sets. The need for piggybacking is caused by the fact that large samples and lengthy questionnaires are often seen as being necessary to capture the variety of travel behaviours. However, cost, respondent burden and other barriers often preclude surveying large samples, and, typically, questions needed to illuminate the factors that lead to trip choices are sacrificed to limit the survey to a reasonable length. In such circumstances, adding on to existing surveys may be a more effective way of identifying emerging travel behaviours. Pratt described her experiences with piggybacking in the context of several major US surveys such as the Nationwide Personal Transportation Survey, the American Housing Survey, Current Population Survey, National Longitudinal Surveys of Labor Market Experience, Survey of Income and Program Participation, the Characteristics of Business Owners and the Dallas-Fort Worth Household Activity Survey 24-Hour Diary.

It was considered that the piggybacking techniques identified by Pratt were particularly useful in the context of the following aspects of travel behaviour:

- Work-at-home behaviour, which is not usually picked up by conventional travel surveys
- Rare activities and travel behaviour, which may not be picked up by relatively small-sample travel surveys
- Emerging behaviours, which are not yet established enough to justify a special survey
- Travel behaviour which has multiple dimensions, such as demographics, income, employment and transportation patterns

The advantages of the piggybacking methods identified by Pratt in addressing the above travel behaviour issues were summarised as:

- Adding standardised questions to several surveys provides the opportunity for multi-dimensional data merging
- Emerging behaviours can be identified by adding “screening” questions to an existing survey, to determine the extent of the emerging behaviour and to see whether a special survey of that behaviour is warranted
- Adding selected questions to ongoing surveys of large samples provides a useful source of information for monitoring trends in travel behaviour.

Stopher and Wilmot (2000) described the development of a time-use diary based on the concept of the familiar “day-planner” used in many personal diaries. Rather than have respondents describe their daily activities (and travel) in chronological order by ticking boxes about each activity, the “day-planner” diary allocates space for each hour of the day and allows respondents to record their activities (and travel) in relatively free-format fashion within each of the hours. Stopher and Wilmot reported that it was desired to improve the methods in such a way as to increase response rates, reduce perceived burden of the survey procedures, and increase the respondent-friendliness of the survey process. Stopher and Wilmot described their experience in developing this method in the context of several small-sample pilots and surveys in Louisiana.

It was considered that the day-planner diary techniques developed by Stopher and Wilmot may be particularly useful in the context of the following aspects of travel behaviour:

- Situations where total activity patterns, including out-of-home activities, are needed for a full understanding of travel behaviour
- Situations where it is desirable to record secondary activities, in addition to primary activities.

The advantages of the method described by Stopher and Wilmot were seen to be:

- The ease with which respondents could add travel and activities which they had initially forgotten
- Reduced burden for respondents by using a format with which many would be familiar (this, however, may be balanced by the need for respondents to write responses in full text, rather than ticking boxes).

In addition to the above four papers on general aspects of survey methodology, another four papers concentrated on the application of relatively new technologies to specific aspects of travel survey design.

Asakura and Hato (2000) described the possibility of using the positioning function of mobile phones for travel and activity surveys. They describe how the positioning function is applied to determine the location of a traveller in time and space. A pilot survey in Osaka City is described, in which ten individuals participated for two weeks. The use of mobile phone positioning was compared to GPS positioning as a means of obtaining time/space tracking information.

It was considered that the mobile phone positioning techniques described by Asakura and Hato may be particularly useful in the context of the following aspects of travel behaviour:

- The time/space monitoring of passenger travel patterns
- Tracking “lost miles” involved in activities such as parking search behaviour.

The advantages of the method described by Asakura and Hato, particularly when compared to GPS tracking surveys, were seen to be:

- Many people already possess mobile phones, and would not need to carry an extra piece of equipment such as a portable GPS system
- Mobile phones have generally better reception than GPS, and work reasonably well within vehicles and buildings, under trees and in “urban canyons” (on the other hand, the positioning accuracy of mobile phones is generally not as good as GPS)
- There is the possibility of using the 2-way communication capabilities of mobile phones during the survey, to give and receive information from respondents.

Lee et al. (2000) (presented by Craig Rindt) described the development of an interactive survey technique (REACT!) designed for deployment and retrieval over the Internet. The survey is based on CHASE (Computerized Household Activity Scheduling Elicitor) developed previously by one of the authors. The main objectives of CHASE were to explore a household’s activity agenda, from which all activities are drawn, and to track the entire process of when and how activities from the agenda are added, deleted, and subsequently modified in a week-long period. These were accomplished through a household interview, self-completing data entry of a weekly activity schedule through the CHASE program, and a follow-up interview. The CHASE program was installed on laptop computers rotated among surveying households to record weekly household activity schedules. REACT! is designed to allow respondents to download the program over the Internet, use their own computers to enter activity diaries and then upload them to a server. At the time of presentation, however, REACT! was being sent to respondents on a CD-ROM, and only the upload was being done via the Internet.

It was considered that the REACT! program described by Lee et al. may be particularly useful in the context of the following aspects of travel behaviour:

- To understand the process by which activities (and travel) are scheduled, modified and executed
- To understand the nature of non-planned activities and travel behaviour.

The advantages of the Internet as a survey distribution method proposed by Lee et al., particularly when compared to non-Internet surveys, were seen to be:

- Many people already have access to the Internet at home (at least in some Western societies) and are comfortable using the Internet as a medium of communication (however, attention still needs to be paid to the biases introduced by the use of the Internet)
- The method may be useful for surveying special niche markets, where biases due to non-availability of the Internet may be minimised
- More rapid turnaround times may be achieved, where the survey is both downloaded and uploaded using the Internet (or done on-line interactively).

Thorpe, Law and Nelson (2000) outlined a survey in which palm-top computer technology was used to implement a self-completion electronic travel diary as part of a travel awareness program in which feedback about recorded travel patterns was given to respondents in an attempt to bring about behavioural change. Feedback from respondents about the palm-top surveys indicated that the majority felt comfortable with the method.

While the use of the palm-top technology was described only in the context of a specific travel awareness program, it was considered that the survey technique described by Thorpe, Law and Nelson may also be useful in the context of the following aspects of travel behaviour:

- Tourism and activity site surveys, where details need to be captured in-field
- Traveller tracking surveys, when the palm-top technology is combined with GPS.

The advantages of the palm-top survey method proposed by Thorpe, Law and Nelson, particularly when compared to pen and paper surveys, were seen to be:

- Easier recording of more complex data
- Automatic time recording for activities and time stamping of data entry
- Increased capabilities in the future, especially when combined with other new technologies such as GPS and online access to public transport information systems.

Wolf et al. (2000) gave a comprehensive overview of the use of automated data capture systems, especially in the context of metropolitan surveys in Atlanta in 2000. In particular, the paper addressed three different types of automated data capture; passive in-vehicle GPS systems, electronic travel diaries with GPS, and a

comprehensive in-vehicle data collection system with both GPS technology and an engine-monitoring device.

It was considered that the automated data capture systems described by Wolf et al. may be particularly useful in the context of the following aspects of travel behaviour:

- The recording of the time/space aspects of travel surveys and physical activity surveys (which record the level of physical exercise that travellers obtain as a function of their trip-making patterns)
- The detailed recording of vehicle use, as an input to the measurement of emissions from vehicles
- Understanding the details of route choice, under normal and emergency conditions
- Performance monitoring of transport systems
- Patterns of commercial vehicle usage.

The advantages of the automated data capture systems described by Wolf et al., were seen to be:

- The automatic recording of complex and highly detailed data
- The ability to use GPS traces to check the validity of information collected by means of traditional travel surveys
- The ability to record travel details for high frequency tripmakers, such as commercial travellers, with minimum effort on the part of the respondent
- The ability to obtain detailed O-D information
- The ability to conduct surveys over extended periods with the same respondent

Many of the papers in the workshop, especially those describing the use of new technologies for travel surveys, highlighted the fact that new survey methods would enable more complex data to be collected than had been possible in the past. In this context, the final workshop presentation by Lee-Gosselin et al. (2000) was a sobering reminder of the issues that need to be considered when dealing with the plethora of microdata that can now be collected on travel behaviour.

## **Major Developments**

The presentation and discussion of the nine workshop papers occupied the majority of the available time for the workshop. However, following the presentation and discussion of the papers, workshop participants were invited to nominate the major developments they had observed in survey methods for travel behaviour research over the past five years (whether or not they had been addressed by the workshop papers). The major developments identified in the workshop included:

- Declining response rates  
A common feature of surveys in most countries was the decline in response rates, in which households (or people) as a whole failed to respond to the survey.
- Increasing non-contacts  
A major reason for non-response was identified as the inability to contact households in the first instance. This was becoming a particular problem for telephone surveys, where a wide variety of call-screening devices are now available. However, it was also a problem for face-to-face household interviews, where household security systems were preventing interviewers from making personal contact with the potential respondent.
- Correcting for non-response  
Given the increasing levels of non-response, it was becoming more accepted that serious efforts need to be made (at the survey design stage) to devise ways of correcting for various forms of non-response. It was also recognised that different survey methods had different biases from survey non-response.
- Over-surveying of the population  
One of the suggested reasons for increasing non-response is a general over-surveying of the population. While little can be done at a global level to prevent this, serious consideration needs to be given as to whether a new travel survey needs to be performed, whether older data might suffice, or whether some travel questions might be piggybacked onto other general purpose surveys.
- Consideration of respondents' needs  
One way of increasing cooperation and response from potential respondents is to be more mindful of their needs and preferences. Choosing survey methods that match their lifestyles, and designing questions with their preferences and limitations in mind, are ways in which greater response and higher quality data may be achieved.
- Increased interest in data  
At the same time as data was becoming more difficult to collect, it was generally agreed that there was increasing interest in the collection of data. After a hiatus in the 1980s, the number of large-scale travel surveys increased substantially during the 1990s.
- Increased complexity of travel and activity patterns  
One of the reasons for the increased interest in data collection was the realisation that travel and activity patterns were changing and becoming more complex towards the end of the century. Data collected in the 1970s and 1980s was

becoming increasingly outdated as a representation of travel and activity patterns in the late-1990s.

- **Increasingly detailed data needs for modelling and policy**  
The increased complexity of models and the wide variety of policies to be addresses using survey data has resulted in increasingly detailed data needs. Activity and time use data, stated preference data, and relationships between travel data and social and environmental data are now commonplace requirements.
- **The effects of privatisation**  
One of the reasons for the greater demands for higher quality travel data has been the increased involvement of the private sector in transport operations. Data for the design and analysis of private sector road-tolling systems, and for the prediction of revenue from privatised public transport systems, has given rise to needs for increased levels of accuracy and precision in travel data. The private sector has much higher expectations for the quality of travel data than does the government sector.
- **The increasing number of electronic aids for surveys**  
As reflected in the papers presented in the workshop, there is a greater emphasis on the use of a variety of electronic aids for the design, conduct and analysis of travel surveys. GPS, GIS, mobile phones, laptops, palmtops, dataloggers and the Internet are just some of the new techniques being integrated into travel surveys. One caution, however, is that the use of these electronic aids should not blind survey designers to the fundamental principles of good survey design. No amount of electronic wizardry, for example, will overcome the poor selection of a biased sample.
- **Increased number of communication channels**  
Technological developments have also given us an increased range of communication channels. The traditional options of face-to-face, telephone and written communication have been supplemented by fax, email, internet, mobile phone and a range of other niche modes of communication. While this may be seen as offering new ways of communicating with people for the purpose of conducting surveys, it also imposes new barriers. For example, modern phone systems have many new ways of preventing communication by means of call-screening, while many people have also converted entirely to mobile phones (for which there is currently no convenient directory system).
- **Collection of route choice data**  
As a result of the use of various tracking devices (such as GPS and mobile phones) which provide high quality time/space trajectories, a greater interest was

being placed on the collection of detailed route information, which could be used in the development of improved models of route choice.

- **Information overload**  
While new technology can enable the collection of a multitude of new types of data, this can also lead to an overload of information available to the analyst. New methods of data manipulation, reduction and analysis must be developed to deal with the quantities of data now becoming available.
- **The diversity of survey instruments being used**  
The combination of new technological developments and refinements in more traditional methods has meant that the survey designer now has a far greater range of options from which to choose.
- **Mixed-mode surveys**  
As well as a greater range of survey options being available, it was now becoming more common for mixed-mode surveys to be designed, rather than relying on a single type of survey. The traditional alternatives of face-to-face interview, phone interview and written questionnaire are now being combined in various ways, and supplemented by other techniques such as GPS traces of personal and vehicular travel.
- **Increased use of non-conventional data sources**  
As the relationship between travel and other activities is being explored, it is becoming more common to use non-conventional sources of data to assist in travel analyses. For example, electronic phonebooks, combined with databases on the geographic distribution of employment, can provide a good foundation for the mapping of trip attractors for use in destination choice modelling.
- **Interest in time-use data**  
A particular data source that is becoming increasingly popular in travel behaviour modelling is the group of time-use surveys that have been conducted around the world. While these surveys have mostly been used for many different types of social policy analysis, they also provide a detailed description of the underlying activity patterns that give rise to travel behaviour.
- **Increased consideration of travel substitutes**  
One reason for the increased attention being paid to time-use data sets is that more policy attention is being paid to the role of potential travel substitutes (e.g. working at home, use of the Internet etc). Traditionally, however, travel surveys have only concentrated on out-of-home activities and have ignored at-home activities. Time-use data sets provide a means of exploring the relationship between out-of-home and at-home activities in more detail.

- Increased diversity of languages in western societies  
A common problem faced in many countries is the increased diversity of languages spoken in the region. This poses special problems for the design of surveys that are aimed at a multi-lingual society.
- Behavioural change studies involving surveys  
An emerging trend arising in several countries is the role of surveys in increasing the awareness of respondents to their own travel patterns. Travel surveys are sometimes the first time that someone has consciously thought about their own travel patterns. As a result, they may spontaneously change their travel patterns after becoming aware of them. For example, they may now walk to the local shop after realising how many times they drive to the local shop. Several projects are now short-cutting the traditional process of collecting data, building a model, thinking of a policy, modelling the policy, implementing the policy, and monitoring the outcomes. Instead, they now collect the data, provide feedback to the respondent, then monitor the outcomes.

## Future Research

Based on the paper presentations and discussions and on the major developments seen over the past five years, workshop participants were asked to nominate the areas of future research that they would like to see undertaken over the next few years. They were asked to do this by nominating the titles of papers they would like to see presented at the next IATBR conference. While some of these titles may be somewhat optimistic, they probably capture the spirit of the workshop participants. The suggested titles include:

- “Do Attitudes predict Behaviour?”
- “Interfacing Data on Travel Demand and Supply”
- “The Demographics of Survey Non-Respondents”
- “Lifestyles of the Non-Contactable”
- “Travel Patterns of Survey Non-Respondents”
- “Applications of Data Mining in Travel Behaviour Research”
- “Increasing Mobility, while Decreasing Vehicle-Kilometres of Travel”
- “The Flexible Survey Instrument”
- “Bridging the Gap between Intentions and Behaviour”
- “Privacy Considerations in Advanced-Technology Travel Surveys”
- “How to Design, Maintain and Use the Ultimate Transport Panel Survey”
- “Agreed Standards for Travel and Activity Surveys”
- “Data when you need it: Guidelines for the Timely Availability of Travel Data”
- “Re-arranging the Deckchairs: Survey Designs in a Changing World”

Time will tell which of these papers come to fruition at the next IATBR conference.

## **Workshop Participants**

Jillian Anable, Carlos Arce, Jimmy Armoogum, Yasuo Asakura, Cyril Bouvier, David Bunch, Ed Chandra, Heather Contrino, Umberto Crisalli, Andre de Palma, Pierre-Olivier Flavigny, Jacqueline Golob, Andrew Harvey, Heinz Hautzinger, Simon Kingham, Susan Lapham, Karina Loxton, Glenn Lyons, Elaine Murakami, John Peachman, Joanne Pratt, Tony Richardson (Chair), Craig Rindt, Peter Stopher, Neil Thorpe, Henk van Evert, David Wagner, Cherill Watkins, Jean Wolf (Rapporteur), Elvie Yates,

## **References**

Anable, J. (2000). "New issues, new methods: a segmentation approach to leisure travel demand". M6 Workshop Paper, IATRB 2000, Gold Coast, Australia.

Armoogum, J. (2000). "Sampling schemes for long-distance surveys". M6 Workshop Paper, IATRB 2000, Gold Coast, Australia.

Asakura, Y. and Hato, E. (2000). "Analysis of travel behaviour using positioning function of mobile communication devices". M6 Workshop Paper, IATRB 2000, Gold Coast, Australia.

Lee, M. S., Sabetiashraf, R., Doherty, S., Rindt, C. and McNally, M. (2000). "Conducting an interactive survey of household weekly activities via the Internet: Preliminary results from a pilot study". M6 Workshop Paper, IATRB 2000, Gold Coast, Australia.

Lee-Gosselin, M., Doherty, S., Theriault, M. and Noe, N. (2000). "Avoiding drowning: new strategies for the collection and analysis of microdata on emerging changes in travel behaviour". M6 Workshop Paper, IATRB 2000, Gold Coast, Australia.

Pratt, J. (2000). "Piggybacking on existing surveys: a methodology for obtaining new perspectives on changing travel behaviors". M6 Workshop Paper, IATRB 2000, Gold Coast, Australia.

Stopher, P. and Wilmot, C. (2000). "Development of a prototype time-use diary and application in Baton Rouge, Louisiana". M6 Workshop Paper, IATRB 2000, Gold Coast, Australia.

Thorpe, N., Law, M. and Nelson, J. (2000). "Raising transport and travel awareness through feedback from an electronic self-completion travel diary". M6 Workshop Paper, IATRB 2000, Gold Coast, Australia.

Wolf, J., Guensler, R., Washington, S. and Frank, L. (2000). "The use of electronic travel diaries and vehicle instrumentation packages in the year 2000 Atlanta regional household travel survey: final test results and implementation plans". M6 Workshop Paper, IATRB 2000, Gold Coast, Australia.