# 30

# ON LARGE SCALE ON-GOING MOBILITY SURVEYS: THE STATE OF PRACTICE

**Paper presented at ISCTSC Conference on Travel Survey Methods, Annecy, France, 2008**

*Elizabeth S. Ampt, Sinclair Knight Merz, Adelaide, Australia*
*Juan de Dios Ortúzar, Pontificia Universidad Católica de Chile, Santiago, Chile*
*Anthony J. Richardson, The Urban Transport Institute (TUTI), Victoria, Australia*

**ABSTRACT**

Large-scale continuous mobility surveys have some advantages over less frequent (usually every 10 years) even larger-scale cross-sectional surveys; these have been well documented in previous papers (Richardson and Battellino, 1997; Ampt and Ortúzar, 2004).

In this paper we first define what we mean by "on-going mobility surveys". We then describe the state of practice in this context, briefly reviewing the state of affairs in all the cases that we are aware of. We then discuss some problems encountered in practice and offer ideas for improvement. In particular we discuss a wide range of issues that are likely to act as barriers to a high quality and sustainable implementation and suggest approaches for improvement. Issues covered include sampling frames and sampling methods, survey methods, respondent burden, weighting processes and expansion, and the increased importance of developing and maintaining field staff motivation. We also touch briefly on the practical/political issue of securing ongoing funding. Throughout, we advance some thoughts to try and explain why this method has not gained wider acceptance particularly in the Northern Hemisphere where there are more examples of travel surveys in general.

The paper also raises some ideas and issues about the way in which ongoing mobility surveys can best collect data for the environmental accounting of travel. Finally, we raise questions about the environmental impact of the survey methods themselves as a stimulus for further consideration.

**Keywords:** Continuous surveys, travel surveys, response rate, response burden, weighting, surveys and climate change

# 1. INTRODUCTION

This paper focuses on large-scale, ongoing, urban mobility surveys. Since there are various kinds of large-scale mobility surveys being carried out around the world, it is useful to commence by describing what we understand by this description. A large-scale, ongoing, urban mobility survey is typically commissioned for a particular city or region by either the city or regional authority responsible for transport planning and/or policy making. The aim of the survey is usually to describe what is currently happening transport-wise in the city or region and its immediate environs, and to use the data to plan or forecast what is likely to happen over the next years. By "ongoing survey" we mean that it is carried out every day for the whole year over several years. This definition leaves out data panels and national surveys (in Appendix 1 we list a series on national surveys conducted in the Northern Hemisphere[1]). It also leaves out cross-sectional surveys. Cross-sectional surveys repeat the same or similar surveys periodically across a city or country. On each occasion a sample of households is selected to represent the whole geographic area in question. The advantages and disadvantages are discussed in detail in Richardson et al., 2005.

[1] We are grateful to Prof. Dirk Zumkeller for having provided us with this list.

People choose ongoing surveys for several reasons. They provide data on temporal variability of travel – over seasons and over periods that can be considered different from 'normal' (e.g. periods of school holidays and other festive periods). Ongoing surveys also give an indication of changes in travel patterns in response to general changes in behaviour, e.g. responses to variations in cost of cars, public transport or petrol, modifications in land-use, and changes in social values (e.g. in response to climate change).

The core of an ongoing urban mobility survey is usually household-based, with the sample selected periodically using the repeated cross section method. This means that the same household is not being followed, as it would be in a panel survey[2], but that new samples are chosen for each period, with some households possibly being selected again by chance[3].

The surveys we describe usually collect data on daily mobility only, and concentrate on travel within the targeted urban area, thereby excluding a deliberate focus on long distance travel. This is because the uses of the data are focussed on the changes that happen and can be made in the urban area being studied. Most surveys we are considering gather 1-2 days of data[4].

While the core surveys are household based and therefore focus on private travel, most cities also choose to have parallel (though not always ongoing) surveys of commercial vehicles and public transport, some traffic flow data and measurement of level-of-service, with some level of ongoing counts of all people and vehicles for all modes of travel. The survey data is usually used to produce trip matrices (by mode, purpose, time of day) and this involves combining data from several sources, most importantly with intercept surveys (also usually not ongoing). This is summarised in Ampt and Ortúzar (2004).

We have chosen to address this type of "ongoing survey" in a research paper for this conference for several reasons. First, many of the issues raised are applicable much more broadly to applications on a country-wide level. Second, we describe many of the problems encountered and, where possible, the solutions to enable others to avoid them in the future. Thirdly, we believe that this ongoing urban mobility survey is a very valuable way of achieving the targets of many city authorities and this paper should provide a useful basis for understanding the issues.

---

[2] There is considerable literature on large scale panel surveys, e.g. British Household Panel Survey Calibrated Time Use Data, 1994-2004 http://www.esds.ac.uk/findingData/snDescription.asp?sn=5363.

[3] This depends on what is considered a period; in Santiago, for example, the sample is selected in advance for three years without repetition.

[4] There is debate about the number of days that should be selected. In 1988, Pas (1988) suggested that 2 days of travel represents about 80% of travel variability, though more recent work (e.g. Mistro and Behrens 2008) suggest that it is much less and imply that more days would be needed.

## STATE OF THE PRACTICE – WHAT IS DONE

The main features of an "ongoing survey" methodology are:

a) Data collection should aim to cover every day of the year (not just typical days and periods). This allows the survey to:

- Capture seasonal, weekly and daily variations in trip making.
- Monitor the evolution of demand and of traveller responses to changes (foreseeable or otherwise) in the transport supply and movement capacity.
- Achieve better training of the field personnel.
- Lower operating costs (rather than paying by the hour a fixed contract can be negotiated - however, this is not the most typical case).
- Develop better and more expeditious quality control procedures, as fewer data items per day/week arrive at the office.
- Obtain higher interest in the information, particularly by other potential users.

Furthermore, a continuous data collection system throughout the whole year allows intensification of the sampling rate in, for example, newly developed areas or less represented days of the week (such as weekends). These data simply receive different expansion factors.

b) The data collection effort is kept running for several years, as described below, producing matrices and models at the end of each previously agreed period.

c) The methodology uses two fundamental types of surveys to cover the most important objectives of the exercise: home interviews and intercept surveys; however the latter are not always present – particularly if there is no special interest in producing trip matrices. Each type of survey is used to maximise the achievement of the study objectives, although it is not possible to look for a strict optimization as neither the state-of-the-art nor the state-of-practice yet allow for it.

d) All the O-D information is geocoded to allow for the future use of the data under any zoning system established. This does not detract from the need to design an appropriate hierarchical zoning system for the study, on the basis of fine zones which can be aggregated into larger zones.

e) Additional information such as traffic counts, bus and Metro patronage by line, and level-of-service data for private and public transport can also be collected to validate and/or update models estimated with the household survey data.

f) Simultaneously with the survey, computerised representations of the public and private transport networks are built or updated, level-of-service data by time period is gathered (for example, using GPS devices), and land use information is collected. All this is undertaken with the aim of eventually using the survey data to estimate strategic supply-demand equilibrium models of the study area transport system.

g) Finally, matrices and models for the future can be updated taking advantage of this continuous data collection effort. Although it is possible to think about preparing partial trip matrices for specific purposes, the city matrices and models are normally updated not more often than every 12 to 18 months, depending on the type of city under scrutiny.

However, this ongoing nature also brings in certain problems; for example it is more difficult to keep the field force motivated for long periods and thus a lot of personnel replacement and re-training is necessary. Similarly, it is not possible to keep a continuous level of advertising about the survey and this has the effect of reducing respondent collaboration in periods without an advertising campaign.

## SURVEY METHODOLOGIES

While the unifying feature of the surveys described in this paper is their ongoing nature, there have been a range of different methodologies used for the conduct of the surveys depending on the demographics of the intended population, the history of travel surveys in the region, the available budget, and the experience of those conducting the surveys. The following sections describe the methods employed in three different cities to highlight the possibilities, and the experience gained in the design and conduct of ongoing surveys.

### Self-Completion

Although the Santiago and Sydney surveys, discussed below, have adopted face-to-face interviews as their primary method of data collection, this is not the only way to conduct on-going mobility surveys. This section of the paper will describe how the methods originally devised for the Victorian Activity & Travel Survey (VATS), which ran from 1994 to 2002 in Victoria, Australia, were modified over the years to arrive at the current incarnation which is being used in its successor, the Victorian Integrated Survey of Travel & Activity (VISTA). The VATS survey revealed two major issues with respect to the conduct of an on-going mobility survey. The first related to the methodology itself, while the second related to its organisational administration over an extended time period.

The first methodological lesson pertained to the self-completion nature of the survey and, in particular, to the use of reminders; these were used because previous research (Wermuth,

1985) had shown that this could double the response rate achieved in mailout/mailback surveys. However, it had also been observed that the reported trip rates fell with increasing number of reminders, and hence it had been postulated (Brög and Meyburg, 1982) that non-respondents were more likely to have trip-making characteristics like those who respond late to the travel survey. They therefore assumed that non-respondents had lower trip rates than respondents. However, a more detailed analysis of the VATS data (Richardson, 2003) showed that the reducing response rate across the reminders was due to allowing respondents to use a later Travel Day, which introduced a degree of self-selection. In fact, the mobility of the non-respondents (when their Travel Day was specified by the interviewer) was more similar to the earliest respondents (who responded to the specified Travel Day) than it was to the late respondents. The lesson from this experience was that reminders could not be relied upon to increase response rates, since they inevitably introduced a degree of self-selection into choosing a Travel Day. If a self-completion questionnaire was to be used, then other methods would be required to increase the response rate.

The second lesson from VATS was concerned with the difficulty of keeping a qualified survey team together for a long period of time. One of us (Ampt) worked on VATS for three years (1992-94) before moving. Another (Richardson) worked on VATS for seven years (1992-1998) and then also moved on. From 1998 to 2002, the survey was run by people not involved in the original design and with little experience in running a large-scale travel survey. In 2000, major changes were made to the design with many un-tested new questions added. In addition, as the University demanded "higher cost-effectiveness" several components were outsourced to market research companies, who offered lower prices. Unfortunately, this resulted in lower quality data and a loss of control in administration and documentation. As a result, the last year of the data was never satisfactorily delivered to clients, and the survey was terminated after 2002. To this day, most analyses concentrate on the data from 1994 through 1999, before the design changes were introduced in 2000.

Following the cessation of VATS in 2002, several surveys have been designed and conducted by Richardson at TUTI using self-completion methods and building on the lessons learned in VATS, i.e. Perth (2002-2005), Brisbane (2003 and 2006), Sunshine Coast (2004 and 2007), Gold Coast (2004 and 2008), Melbourne (2004 and 2005), Auckland (2006) and Christchurch (2006). However, only one of these (Perth) has been a true "ongoing" survey, in that it ran continuously for more than one year, but the others have contributed to refining the methodology which is now being implemented in VISTA (2007-08), currently in its first year, but that may well be extended to run again as an "ongoing" survey.

VISTA is still based on a self-completion questionnaire, but this is now incorporated in a substantially different overall methodology (to overcome many of the problems uncovered in VATS). The questionnaire is hand-delivered to, and hand-collected from, the survey

households. This process is also supplemented by telephone motivational calls, telephone and postal reminders, and telephone clarification calls. The process consists of 11 essential steps:

1. *Selection of sample*: At the beginning of the survey period, a sample of households is drawn from the sample frame, which is a GIS database of properties within residential areas but it could be a list of electricity connections or a list of residential addresses from a local government rates database. It is a cluster sample, based on a random sample of households within a stratified random sample of Census Collectors Districts (CCD: groupings of approximately 250 residences, as used in the Australian Census). Within each CCD, a cluster of 42 households are sampled (at an average rate of approximately 20% of residential addresses). Each week, those households designated for contact in that week are identified, and a list is made for each of the sub-regions accompanied by a GIS map showing the location of each address within the cluster.

2. *Address checking and pre-contact letter delivery*: On the Tuesday of each survey week field staff in each of the CCDs check for the existence and suitability of each listed address. At this stage, some will be classified as Sample Loss (i.e. no such address, address not findable, vacant block, business address, empty premises etc). Sample Loss addresses are replaced by selecting another one from a pre-supplied random selection of replacement addresses. Other addresses will need to be corrected (e.g. for apartments, there may need to be correction or clarification of apartment numbers, where the apartment number is either missing or mis-recorded in the sample frame). Field staff is given strict instructions about what corrections they are allowed to make. In addition, field staff is in contact with the Fieldwork Controller by mobile phone to clarify any problematic situations on the spot. Once the In-Scope addresses have been established, field staff delivers a Pre-Contact Envelope (addressed to "The Householder") to each In-Scope address (the letter is placed in the mailbox, with no contact being made with the householder at this time). The Pre-Contact Envelope contains a Pre-Contact Letter from the client and a survey brochure.

3. *Correction of sample addresses*: On the Wednesday of each week, the field corrections made on the Tuesday are entered into the Sample Address database, and a final list of In-Scope addresses is produced for the questionnaire delivery. This is accompanied by an updated map, with any comments from field staff to enable the easy and correct location and identification of households for the delivery of questionnaires (e.g. big white house with poplar tree and red letterbox).

4. *Delivery of questionnaires*: On the following Saturday and Sunday, field staff (different from those who undertook the Pre-Contact Letter delivery on the previous Tuesday) deliver the questionnaires in each of the CCDs. The questionnaires cover Travel Days from the next Monday through to the following Sunday. Every attempt is made to personally contact a household member to deliver the questionnaire in person

(since this has been found to significantly increase the response rate). Interviewers pay up to four visits to a household during the weekend to try to contact them (two attempts on Saturday and two attempts, if needed, on Sunday). If contact is made, the interviewer introduces themselves, briefly explains the nature of the survey and answers any questions about it.

If the household agrees to accept the questionnaires, then the interviewer asks how many people are likely to be in the household in the coming week, and then provides that many Travel Diaries (plus one) as part of the survey materials. They also inform the household when they will return to pick up the completed questionnaires, and where they might leave the questionnaire out for collection, if they happen not to be at home. They are also asked for a contact phone number in case we need to contact them during the survey. If contact cannot be made, a "We-Missed-You" postcard is left at the household with an envelope containing all the survey materials (including six Travel Diaries to cover most household sizes). Information is also left about the pickup of the questionnaires the following week, and instructions for where to leave the completed questionnaires if the household does not expect to be home when they are collected. If a personal refusal is received at this stage, then the field staff is instructed to first enquire whether the reason for the refusal is "transport-related" (e.g. I don't travel much, so why am I in the Travel Survey?). They then attempt to convert such refusals to an acceptance by stressing the importance of obtaining responses for all types of households. If they cannot be converted, or if the reason for the refusal is not transport-related, then the field staff immediately asks two "refusal questions": how many people are in the household and how many vehicles are in the household. This gives some information about refusals. These questions are asked anytime that a personal refusal is encountered.

5. *Motivational call*: On the day of each household's Travel Day, they are phoned by the Survey Office to ensure that they have received the survey materials, to answer any questions they might now have, and to remind them that today is their Travel Day. Phone numbers are obtained from households when survey packs are delivered to them, with a backup set of phone numbers obtained from commercial phone number databases.

6. *Pickup of questionnaires*: On the following Sunday (for households with a weekday Travel Day) and Monday (for households with a weekend Travel Day), a different set of field staff goes out to the CCDs to collect the questionnaires.

   Two attempts are made to personally contact the household (since the effect of personal contact at this stage is not as strong as in the delivery stage). As soon as possible after collecting questionnaires from the household, the interviewer performs a quick edit check to identify whether the full number of Travel Diaries have been completed and handed over. If any Travel Diaries are missing, the interviewer returns

to the household to enquire whether the Diaries were inadvertently excluded from the package of completed questionnaires. If the household is not at home, but they have left their questionnaires out for collection, then the questionnaires are collected and a "Thank-you" postcard is left to indicate that the questionnaires have been collected.

If no household member is at home, and they have not left their questionnaires out for collection, the interviewer leaves a "We-Missed-You" postcard together with a reply-paid envelope for the return of the questionnaires by mail to the survey office.

7. *Data coding*: A special program (*Speedit*) was developed by TUTI to enable data coding, including geocoding, and editing to take place in the field office as soon as questionnaires are returned. The program is readily adapted to new survey designs (especially if they are of the same general format as VISTA). The number of data coding workstations can be modified to suit the survey workload, such that completed surveys are processed within a week of being received in the office. The geocoding module has six options: (Home Address geocoding, Full Street Address geocoding [when the full address is known], Partial Street Address geocoding [when only part of the address is known, e.g. street without number], Landmark geocoding [e.g. a bank name], Cross-street geocoding, Town centroid geocoding [when the only detail is the suburb to which a person travelled) and has been designed to accept external data files of street names, suburbs, postcodes, cadastre file addresses, landmarks, cross-streets and town centroids; it can be easily modified for use in another geographic locality. All data output from the geocoding module can then be used in GIS programs such as MapInfo using the required map grid projection.

8. *Data editing*: The *Speedit* program performs both data entry and data editing at the same time. One of the advantages of this is that, because all geocoding is performed as the data is entered it is possible to run detailed editing checks involving trip distances and speeds, since these can be calculated as soon as the geocodes for successive locations are found. Thus, in addition to the usual range checks and logic checks, these more complex trip editing routines can be run immediately, thus enabling the data coder to immediately correct problems with the data or flag the data for follow-up clarification calls (which are typically performed within days of the respondent filling out the survey forms). Data entry and editing in VISTA is typically completed within a week of the questionnaires arriving in the survey office at the end of a week of Travel Days.

9. *Identification of data problems*: Following the coding and editing of data from the questionnaires, the data are merged into an integrated file for each household.

This file is then analysed looking for missing data and for data entries that fail the range checks or logic checks. These households, and the problem data, are earmarked for further investigation.

10. *Telephone clarification calls*: Those households in which problem data have been identified and not rectified, and for which a phone number is available, are then phoned to seek clarification about the problem data. The clarified information is then input into the data file. Approximately 15% of responding households are phoned during this process for VISTA.

11. *Reminder calls and letters*: About one week after the collection of completed questionnaires from households, those who have not yet responded are contacted via phone to remind them to return the questionnaire if they have completed it. Those who are not contactable by phone are sent a reminder letter. However, unlike VATS, the non-respondents are not urged to complete the survey for a later Travel Day. They are only encouraged to send back the questionnaires if they have already completed then for the nominated Travel Day.

**Face to Face**

*The Sydney Experience*

The Sydney *Household Travel Survey* (HTS) is a large scale, continuous survey of the travel patterns of residents of the Greater Metropolitan region of Sydney (Battellino and Peachman, 2004). Every year approximately 3,500 households (or 10,000 individuals) are interviewed about their travel on a particular day, with interviews spread across the whole year. Information on all travel and activities undertaken by all household members over a given 24 hour period are collected via face to face interview.

To maximise the statistical reliability of HTS results for a given year, each annual HTS estimate consists of three years of pooled HTS data weighted to the Estimated Resident Population (ERP) from the Australian Bureau of Statistics for that year. This sample size enables reliable trip estimates at the Statistical Division (Sydney wide) level in wave 1, Statistical Sub-division (regions within Sydney) level after wave 2 and the Statistical Local Area (SLA - local government area of which there are 50 in Sydney) level at the end of wave 3. Interviews are clustered within these SLAs to reduce the fieldwork costs. Each day of the year is represented with sampled households being given a nominated “travel day” for which data will be collected. All members of the selected households are interviewed. The scope of the survey population is all residents of private dwellings in the survey area. Thus visitors, tourists and residents of institutions are not included.

The sample is drawn on a quarterly basis by a process of area address listing; more details are provided by Battellino and Peachman (2004). Once the sample is selected no replacement of dwelling or “travel day” is allowed. Area address listing encounters the

expected problems associated with this method such as vacant dwellings and human error in recording addresses. Approach letters addressed to "The Householder" are sent before the interviewer attempts to make contact. The sample loss due to vacant dwellings, dwellings under construction/being demolished or non-private dwellings has been around 5% and for incorrect addresses, less than 2 %. There have been very few issues in relation to the sampling process which are a function of the continuous nature of the survey.

Questionnaire design

The questionnaires used for the continuous survey were primarily the same as those used for the Sydney 1991 Home Interview Survey: a Household Questionnaire, collecting household, person and vehicle data for the household, a 24 hour trip diary questionnaire for members of the household aged 15 years and over, and a 24 hour trip diary for members aged less than 15 years. This had the benefit of providing data which would be comparable with that collected in 1991 and to a large extent with 1981. Thus for the first time, longer term trend data on travel patterns is available. Despite the fact that there is no real information on the intervening trends between 1981, 1991 and 1997, the ability to compare the data at these points in time has proved to be very valuable in providing advice for decision makers.

It is important when conducting the survey on a continuous basis that the questionnaire remains essentially the same to ensure the comparability of the data. However there is room to progressively incorporate small improvements and updates. There is also the added benefit of being able to use the existing survey strategy to collect supplementary data on relevant issues as required, at marginal costs. There is also the potential to increase the sample size in any particular area, using the same survey strategy, to provide more statistically reliable data, for example in relation to the impact of a large infrastructure project, in a time frame shorter than would be provided by the standard sampling procedures.

Fieldwork control

At the beginning of the continuous HTS, the fieldwork was contracted to a research company on an annual basis with a rollover of contract, if certain performance indicators were met, for a maximum of three years. As the HTS is a face-to-face interview, a sizable and stable team of interviewers, especially suited and trained in personal interviewing techniques, is required. The advantage of the continuous nature of the survey for interviewers is that it provides a reliable and consistent source of work and, given the interesting nature of the topic, many find it rewarding.

Notwithstanding, the HTS team is refreshed each year and few have stayed since the first wave; the team has some 30 interviewers which dwindles to the low 20s during the wave and is refreshed with a new recruitment of up to 10 interviewers each year. It has been found that the composition of the team and the fieldwork managers/supervisors can have a significant effect on response rates.

Because of the continuous nature of the survey, there is potential for interviewer fatigue and boredom, or over familiarisation, leading to laxity in the compliance with procedures. Much effort is put in by the fieldwork company and survey management staff to ensure that the fieldwork staff is made to feel part of the team. This is done by regular team meetings, every two months, where interviewers are consulted about issues arising in the field and suggestions sought for improvements, annual retraining and debriefing days and providing information on the importance of the survey and use of the data. The performance of interviewers is also monitored by strict validation procedures which require that at least 10% of interviewer completed workloads are validated by telephone. The validation must be done within one week of the receipt of the workloads and before the data is clerically checked, entered and processed. Where a significant proportion of completed interviews by an interviewer (i.e. 10% below the average for a reporting period of one month) does not have the household's telephone number or have incorrect telephone numbers, the survey consultant validates using the face-to-face interview method by making a single visit to each household in the affected workloads. Any interviewer suspected of not following field procedures and providing incorrect data is validated at a higher rate. The team leaders also conduct monitoring of interviewers with a minimum of two supervised interviews undertaken per interviewer per year. The aim of the supervised interview is to check that all interviewers are conducting the interviews using correct procedures, are clearly explaining the survey to respondents and have a clear understanding of the survey questions. An extra supervised interview is undertaken where the results of the initial supervised interview revealed that the interviewer's performance was not at an acceptable level.

There are strict control procedures for checking that information is collected only for the nominated travel day and that respondents are contacted according to certain rules around this day. Workloads are checked weekly to obtain immediate information on the level of response obtained by interviewers so that the survey consultant is able to act at once where an interviewer is having difficulty recruiting the respondents or specific issues are identified that require special attention. In addition to the wide range of data checks that are carried out on the data following data entry, response rates, trip rates, rate of sample loss and number of calls made to households are also monitored monthly on an interviewer basis. The percentage of persons staying at home on the designated travel day should also be checked as a form of 'soft refusal' (see Madre et al. 2007). This means that any indications of potential interviewer bias can be detected, and rectified, at an early stage.

Thus, although many interviewers are very familiar with the survey, there is a constant need for monitoring of interviewers' performance and project management staff cannot become complacent assuming that the continuous process will become self-managing. However, the necessity for continual vigilance of the fieldwork is not a deficiency, or a negative side-effect, of a continuous survey, but recognition of the fact that it offers an opportunity to rectify problems that a one-off survey does not.

For example some problems were discovered with regard to the adherence to procedures by a few interviewers in the second wave of the HTS. It was then possible to disregard these interviews and resample them in the third wave. This would not have been possible if it had been a one off survey and these interviews would have been lost. The continuous nature of the survey also allows for the on-going assessment of the effectiveness and practicality of fieldwork procedures and provides the opportunity for refinement them over time. This, it is believed, has contributed to a continuation of the high quality of the data.

*The Santiago de Chile Experience*

The Santiago Continuous Mobility Survey, started in August 2001 with a first wave of 15,537 households (interviewed until April 2002), which should had been followed by roughly 5,000 households every subsequent year. As it happened, after that first year the government decided to examine the results and the process was deferred for almost two years, to be re-commenced in April 2004 for another three years (the data collection stopped in May 2007). The government is currently studying the continuation of the process.

Questionnaire design and fieldwork control

The main characteristics of the survey in this sense are very similar to those of the Sydney HTS. The design of the questionnaire was initially inspired by that of the VATS and Sydney HTS surveys, but the Santiago experience (and the evidence after a year long, very thorough, pilot study of various methods) showed that in the great majority of cases face-to-face interviewing was the best approach to obtaining data from homes at an acceptable cost in a developing country such as Chile. Notwithstanding, there are cases where the personal interview method is simply not feasible; for example, when household members travel too much and it is not possible to contact them (in spite of several attempts at different times in several days), or when the security personnel at a high income condominium or building, does not allow the interviewer to enter the premises. In these "difficult" cases, the self-completion survey method with personal delivery and recovery of the survey form provided a solution to the problem. However, it was found that only people capable of following written instructions (and indeed, reading them) were able to complete the forms in Santiago. In practice this meant that acceptable response rates were only possible in medium to high income households.

It was also found that self-completion may not be advantageous, cost wise, in a developing country. First, as the forms need to be attractive and easy to follow, this translated into much higher production costs than for personal interview material[5]. On the other hand, the sample must be over dimensioned (as response rates are clearly inferior in this method) and each household given a large number of forms (eight in Santiago) to make sure that there were enough for each household member. Finally, the cost of delivering and recovering the forms were also higher due to the over dimensioning of the sample.

In the first wave of the survey self-completion forms were used in 3% of the households (i.e. 490 cases) and only 12 had one or two missing items; this required special training of the interviewers (respondents filled the questionnaires when they "remembered", so they have to be followed closely to make sure they jotted down all their trips in the memory joggers, etc.). However and surprisingly, in the 2004-2007 wave no self-completion forms were used; although the current administrator rightly claims that they were able to achieve higher response rates among high income people than in the 2001-2002 wave, the fact remains that most probably they did not obtain information about the "difficult" cases mentioned above[6].

The field procedures have remained basically unchanged between the two periods and are almost identical to those described for the Sydney HTS. However, data validation and interviewer control have need to be more stringent, as it appears that interviewers are not as responsible as in Australia (although recruitment and training processes are similar); this behaviour could be, perhaps, attributable to the rather low wages received by interviewers (US$ 13.5 per household including transport) and other staff (e.g. the data validators in the field receive US$ 3.25 per household and the supervisors and digitisers US$ 500/month[7]). To improve on this, during the second wave of the survey 20% of all households were re-visited for validation purposes and a further 30% of the sample was re-contacted by phone.

Sample design

The first wave sampling frame had the registered homes at the National Tax Office and in the second wave the Census addresses (the last was taken in 2001). At every period, for example three years in the second wave, the appropriate number of addresses (i.e. 15,000) were generated randomly in time and space using Sobol low discrepancy sequences[8] (Bratley et al, 1992) looking at three years ahead (i.e. for a total sample of 15,000 households).

In this simple two-dimensional case, each sequence corresponds to a pairing (space and time) covering the study area in orderly form. The first column of the complete sequence would be the traditional random sample (in space) of a typical O-D survey (which is usually stratified by zone in practice). The second column is obtained by means of a linear projection between the initial and final dates of the data collection effort. In practice, it is necessary to order the pairings in terms of their "spatial" component and associate them to a

[5] In the Santiago survey each trip required two Legal size (216 x 356 mm) pages printed in two colours, apart from the pages dedicated to presentation and instructions.
[6] In fact, we have recently learned that high income zones are now represented by (high income) households with the lowest income values in that category (i.e. if the high income group is, say, over US$ 1,500/month, we have that most households in this group have incomes just over this value).
[7] The minimum wage in Chile was US$ 300/month in 2007.
[8] http://en.wikipedia.org/wiki/Low-discrepancy_sequence.

traditional spatial sample (spatially ordered[9]) so that they will be associated to an address in the sample; the pairings are then ordered by their "time" element.

The following example illustrates the method. Assume we need to generate a sample of 10 addresses in five zones (all zones have the same size[10]), to be surveyed (survey date) over two days (Saturday and Sunday). For this we first generate a two-dimensional Sobol sequence with ten elements as in Table 1.

**Table 1: Space and time Sobol numbers**

| Number | Space | Time |
|---|---|---|
| 1 | 0.5000 | 0.5000 |
| 2 | 0.7500 | 0.2500 |
| 3 | 0.2500 | 0.7500 |
| 4 | 0.3750 | 0.3750 |
| 5 | 0.8750 | 0.8750 |
| 6 | 0.6250 | 0.1250 |
| 7 | 0.1250 | 0.6250 |
| 8 | 0.1875 | 0.3125 |
| 9 | 0.6875 | 0.8125 |
| 10 | 0.9375 | 0.0625 |

We then generate a sample of 10 addresses for the five zones, ordered by their zone number; in practice this is done as follows. First, a block is selected for each observation by randomly assigning to every block in a zone a probability of being chosen proportional to its number of households. Second, for each observation a random number between 1 and the number of households in the block is chosen and this serves to choose the household in practice, i.e. the third in the block[11] in Table 2.

**Table 2: Associating zone and address to places**

| Place | Zone | Address |
|---|---|---|
| 1 | 1 | 1 |
| 2 | 1 | 2 |
| 3 | 2 | 1 |
| 4 | 2 | 2 |
| 5 | 3 | 1 |
| 6 | 3 | 2 |
| 7 | 4 | 1 |
| 8 | 4 | 2 |
| 9 | 5 | 1 |
| 10 | 5 | 2 |

Then ordering the Sobol table by its spatial component and integrating it with the sample of addresses we get Table 3. And reordering the table in ascending order by time and

[9] The ordering can be simplified using zonal codes that help finding their spatial location, as has been traditional in Santiago; in this way zones starting with the same number are closer.

[10] In practice not all zones are equal, and the number of cases to consider in each zone is the nearest integer value of the product of the number of households in the zone (from the Census) and the sampling rate (i.e. 15,000 over the total number of households in Santiago).

[11] In practice a series of numbers is generated as for each household sampled, several replacement households are also selected in each case.

distributing it according to date (i.e. 1: Saturday; 2: Sunday), we get the data shown in Table 4. Finally, eliminating redundant columns and ordering by Sobol number we get Table 5 that satisfies the objectives of being random in both space and time.

**Table 3: Ordering the Sobol table by its spatial component**

| Number | Space | Time | Place | Zone | Address |
|---|---|---|---|---|---|
| 7 | 0.1250 | 0.6250 | 1 | 1 | 1 |
| 8 | 0.1875 | 0.3125 | 2 | 1 | 2 |
| 3 | 0.2500 | 0.7500 | 3 | 2 | 1 |
| 4 | 0.3750 | 0.3750 | 4 | 2 | 2 |
| 1 | 0.5000 | 0.5000 | 5 | 3 | 1 |
| 6 | 0.6250 | 0.1250 | 6 | 3 | 2 |
| 9 | 0.6875 | 0.8125 | 7 | 4 | 1 |
| 2 | 0.7500 | 0.2500 | 8 | 4 | 2 |
| 5 | 0.8750 | 0.8750 | 9 | 5 | 1 |
| 10 | 0.9375 | 0.0625 | 10 | 5 | 2 |

**Table 4: Ordering by space and time**

| Number | Space | Time | Place | Zone | Address | Date | Day |
|---|---|---|---|---|---|---|---|
| 10 | 0.9375 | 0.0625 | 10 | 5 | 2 | 1 | Saturday |
| 6 | 0.6250 | 0.1250 | 6 | 3 | 2 | 1 | Saturday |
| 2 | 0.7500 | 0.2500 | 8 | 4 | 2 | 1 | Saturday |
| 8 | 0.1875 | 0.3125 | 2 | 1 | 2 | 1 | Saturday |
| 4 | 0.3750 | 0.3750 | 4 | 2 | 2 | 1 | Saturday |
| 1 | 0.5000 | 0.5000 | 5 | 3 | 1 | 2 | Sunday |
| 7 | 0.1250 | 0.6250 | 1 | 1 | 1 | 2 | Sunday |
| 3 | 0.2500 | 0.7500 | 3 | 2 | 1 | 2 | Sunday |
| 9 | 0.6875 | 0.8125 | 7 | 4 | 1 | 2 | Sunday |
| 5 | 0.8750 | 0.8750 | 9 | 5 | 1 | 2 | Sunday |

**Table 5: Ordering by Sobol number**

| Number | Zone | Address | Day |
|---|---|---|---|
| 1 | 3 | 1 | Sunday |
| 2 | 4 | 2 | Saturday |
| 3 | 2 | 1 | Sunday |
| 4 | 2 | 2 | Saturday |
| 5 | 5 | 1 | Sunday |
| 6 | 3 | 2 | Saturday |
| 7 | 1 | 1 | Sunday |
| 8 | 1 | 2 | Saturday |
| 9 | 4 | 1 | Sunday |
| 10 | 5 | 2 | Saturday |

A preliminary calculation, based on the spatial distribution of zones in the Santiago area (by the six macro-zones defined) and assuming that the survey would start on Thursday January 1, 2004 and end on Sunday December, 31 2006, produced the sample distribution by day and area in the three study years shown in Table 6 (note that the days between Thursday and Sunday have an additional week).

**Table 6: Sample distribution over 3 year period**

| Day | North | West | East | CBD | South | South East | Total | % |
|---|---|---|---|---|---|---|---|---|
| Monday | 289 | 464 | 346 | 107 | 478 | 453 | 2,137 | 14.25 |
| Tuesday | 283 | 481 | 326 | 99 | 502 | 442 | 2,133 | 14.22 |
| Wednesday | 292 | 444 | 342 | 105 | 507 | 447 | 2,137 | 14.25 |
| Thursday | 295 | 465 | 356 | 107 | 475 | 450 | 2,148 | 14.32 |
| Friday | 277 | 498 | 333 | 100 | 512 | 429 | 2,149 | 14.33 |
| Saturday | 297 | 460 | 346 | 100 | 482 | 464 | 2,149 | 14.33 |
| Sunday | 279 | 469 | 348 | 99 | 515 | 437 | 2,147 | 14.31 |
| **Total** | **2,012** | **3,281** | **2,397** | **717** | **3,471** | **3,122** | **15,000** | **100.00** |

Data input, digitising and checking

The rigorous procedures in place in the Santiago survey are very similar to those reported for the VISTA survey (indeed they started by mimicking the early procedures in the original VATS survey), so they will not be repeated here for lack of space. In particular, on-line checking software for data entry, which allows multiple staff to work at the same time, was designed so that the software also completes street and place names, easing the task of the digitisers and reducing errors substantially. However, in the second wave (and a little like the later experience in VATS) the strict, almost on-line, procedures for data entry and checking were relaxed and performed not every day or week, but every month and later every several months. This meant that finding out an interviewer who was performing incorrectly (for example, much lower trip rates), took three or four months to detect.

This, in turn, led to severe data losses, replacement of personnel, re-training of interviewers, much higher costs and the need to apply, in some cases, non-trivial correction factors to non compulsory trips.

**Survey Content**

The VATS survey attempted to obtain travel information from all members of the household, plus any visitors who stayed in the household on the night before the Travel Day. But while it is relatively easy to obtain demographic information about visitors, it is often difficult to obtain Travel Diary data for them, especially if they return to their own residence on the Travel Day. Also, asking for children under 5 to complete the Travel Diary merely imposes an extra burden on the parents. This is especially the case when the young children's diaries are often identical to one of the parents, with whom they travel around during the Travel Day. For these reasons, the VISTA survey does not attempt to obtain Travel Diary data for visitors (but does ask for their demographics), and does not seek

completion of Travel Diaries for children under 5 (instead reconstructs them from data provided by other household members).

In Santiago, the scope of the mobility survey did also not include all travellers in the urban area but only residents and visitors to households. People in hotels, other people in non-private dwellings (such as hospitals and military barracks) and travellers that pass through the area on the survey days, were judged to be relatively few (in comparison to residents) and also too difficult to contact. However, trips are recorded for all household members, even if they are too young to do this independently; in this case the parents have to report their travel. Also, in the case of relatively young individuals and the old and infirm, a small amount of proxy reporting is accepted.

All the on-going surveys mentioned contain questions on four essential topics; household data, persons, vehicles and trip stages for each person (in the VISTA survey only for those aged 5 and above). In most surveys, the client group inevitably wants to ask more questions than can reasonably be accommodated in a reasonably sized questionnaire. In general, the approach in these situations is to stay with a rigid maximum size of questionnaire (say, a folded A3 sheet for the Household, Person and Vehicle forms, and a 16-page A4 booklet for each Travel Diary) and to force the required content to fit within these boundaries (with a minimum font size). This forces the client to think about what information they really need, while ensuring that the respondent burden is contained within what has been found to be acceptable in past surveys.

**Sampling**

In most Australian surveys a clustered random sample of households within Census Collectors Districts (CCD) has been used. VATS was a slight exception, in that it was not clustered, because of the nature of mailout/mailback surveys. The idea was to minimise the time and cost spent by interviewers in the field in moving between sampled households. In Santiago, the situation is different as the sample is truly random in time and space; the only concession to logistics being that the days of the week are aggregated into eight week groups, where, for example, a household with a date defined for a given Monday in the period is assigned to any of the eight possible Mondays in the period to optimise field work. This also ensures that the summer period (January and February) is completely covered.

Sample frames have varied over time. Earlier ones tended to use lists of addresses from utility companies (water supply, electricity connections etc), as these were the only comprehensive lists available for use in travel surveys. More recently, the widespread use of GIS programs has meant that many agencies have comprehensive address lists available in this format. For example, several surveys conducted by TUTI in Melbourne in 2005 and in New Zealand in 2006, obtained lists of residential addresses from the local government agencies in whose areas the surveys were being conducted. This strategy, however, tends

only to work well when the number of agencies covered is relatively small. When data has to be obtained from many agencies, the effort involved in standardising the content and format of the data sets can become significant. In Santiago, as mentioned above, the sampling frame was first the registry of homes in the Tax Office, but this was inevitably a couple of years out-of-date, leading to several losses due to changes (demolitions, building of new buildings, change to commercial use) which are very rapid in developing nations. Furthermore, as this data corresponds to regularised (i.e. legal) dwellings housing actual taxpayers, we later discovered that it was biased towards higher incomes (poor people do not pay taxes nor have their dwellings regularised). In the second wave, the 2001 Census information was used (which was released late in 2002), and had no income bias, but by 2006 it had also become relatively obsolete.

In recent times, the growing availability of commercially available GIS data sets of addresses and property boundaries has opened up new opportunities, which have particular relevance for on-going travel surveys. For example, in the Sunshine and Gold Coasts surveys in Queensland and in the VISTA survey, the sample has been chosen by using three GIS data sets (the Geocoded National Address File, an overlay of planning zones, and a map of cadastral property boundaries) to choose a sample of residential addresses within the sampled CCD regions. This method is particularly appropriate for on-going surveys, because such databases tend to be updated on a regular basis by the agencies developing them for commercial purposes (such as market research and location planning).

## CHALLENGES OF THE CURRENT PRACTICE

### Response Rates

As in most areas and surveys, response rates in Santiago and Sydney have tended to decline with time in spite of the efforts spent on re-training interviewers and following strict fieldwork control rules. In the Santiago 2001-2002 wave, the response rate was 68% (Ampt and Ortúzar, 2004), but as roughly half of the non response corresponded to no contacts, many of which were probably non-eligible, the real response rate was probably higher. The situation in Sydney is depicted in Figure 1[12]; full response corresponds to the case when household and all trip and person details from all members are obtained. Part response is when all household data is obtained, but full trip and person details are not obtained for one household member.

In Santiago, the final statistics for the 2004-2007 wave have not yet been released, but we have learned that due to cost reasons the strict procedures mentioned above were relaxed as follows. As each interviewer (only 10 to 16 for the whole city during the second wave) had to cover a very large area (Santiago has more than 1.5 million households and the study area is some 2,000 km$^2$), they were given not one, but three sample addresses.

[12] We are grateful to Tim Raimond for having provided us with this information.

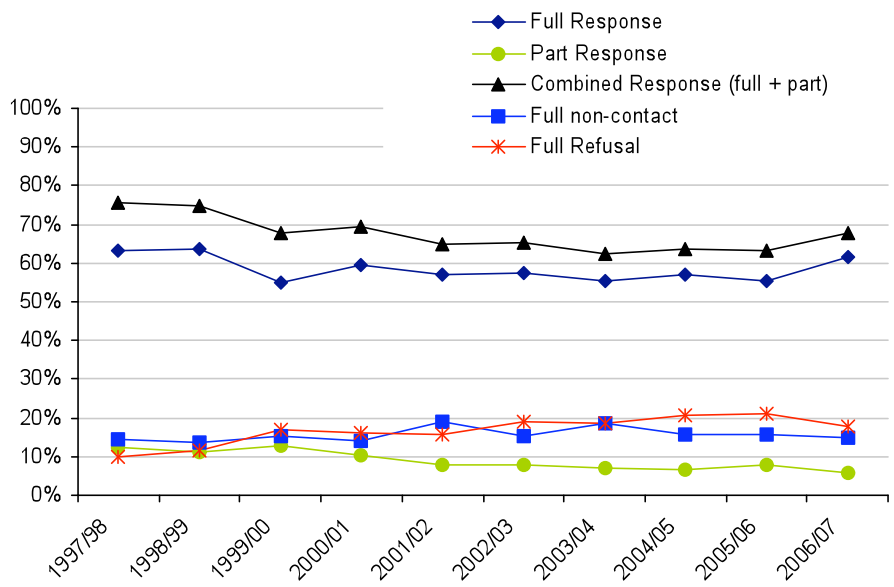


**Figure 1: HTS Response Rates Wave 1 (1997/98) to Wave 10 (20060/07)**

The idea was that they should try the first in the first place, and if after four visits they could not get a response, try the second and so on; however, it is obvious that the interviewers – in general – used the addresses to their advantage (although the 20% validation checks should guarantee that at least they tried a couple of times before changing). The final result was that in the three years the interviewers contacted (or, at least, were provided with) 34,669 addresses; of these, they were rejected or found nobody in 18,316 addresses (furthermore, 795 households had to be replaced because they had false data). So, as the final sample consisted of 15,558 complete households, we may deduce that the response rate dropped to 45%.

In the VATS survey, the response rate started out at around 60% in 1994, but had dropped to about 50% by 1999. This could have been due to a general decline in response rates in all surveys over that period, or to a problem with maintaining ongoing staff motivation over an extended period. After the major design and administration changes in 2000, the VATS response rate was reported to have fallen to about 25%, but lack of complete documentation of the survey in those years precludes a precise observation.

In all of the surveys using the current VISTA design, from the Perth survey in 2002 through to the Sunshine Coast survey in 2007, an average response rate of 55% has been obtained, ranging from a low of 49% to a high of 60%. The smaller cities tended to deliver higher response rates, with smaller response rates in the larger cities.

**Respondent Burden**

Using an on-going survey methodology, with 1-2 days travel, poses relatively low levels of respondent burden. The burden varies with the survey method, and self-completion is considered by some to be more difficult for the respondent for the following reasons:

- It takes longer for each individual to complete the forms because they are not as familiar with the questions as an interviewer would be.
- It is more difficult for people who are not used to completing forms (this can include people for whom English is a second language).

This can lead to under-reporting by certain socio-demographics and by people making certain trips (e.g. those who make many trips). (see discussion on respondent burden by Bricka, S. 2006) Nonetheless, in the various Australian self-completion surveys from VATS through VISTA, the proportion of survey respondents not born in Australia (who might be thought to have more difficulties with self-completion surveys printed in English only) is virtually identical with the proportion of persons not born in Australia as identified in the Census.

On the other hand, in contrast to face-to-face interviews, self-completion forms are able to be completed at any time of day or night meaning that certain types of people find them much easier than a face-to-face approach. In face-to-face interviews, much of the burden of response is taken by the interviewer who reads the questions and encourages the respondent to continue to the end of the survey. Given that the methodology discussed is based on cross-sectional sampling, the considerable burden of panels where people are asked to contribute on a repeating basis is not an issue.

**Weighting and Expansion**

For standard cross-sectional travel surveys there are often four types of weighting (and expansion): household weights, person weights, non-responding household weights, and non-reported trip weights. The household weights may be obtained by comparing the survey results with the Census results on the basis of household size (number of persons) and dwelling type within geographic regions. The person weights can be based on a comparison with the Census on the basis of age and gender of persons within geographic regions.

Non-responding household weights are not used in some surveys (i.e. in VISTA and in the Santiago cases), and for self-completion surveys non-reported trip weights are calculated based on who completed the travel diary (self-reporting vs proxy-reporting) and when the travel diary was filled in (on the Travel Day compared to after the Travel Day). In Santiago, non reported trip weights have been used for cases where there is a suspicion of laxity on the part of the interviewers (i.e. just prior to a full re-training exercise).

As always in the calculation of expansion weights, there is a trade-off between the desire for precision (lots of controlling variables and categories) and the limitations of zero and small counts in cells of the survey data cross-tabulations. Clearly, larger sample sizes can allow greater disaggregating in the expansion variables, and this may be improved as the

sample size increases over many years of an ongoing survey. If Census data is used for the control data set for expansion purposes, the timing of the travel survey and the Census is important. Many transport agencies think it is a good idea to conduct a travel survey in the same year as the Census (years ending in 1 and 6 in Australia; unfortunately the Census is only every decade in Chile and most developing countries). However, this can pose several problems. Firstly, running a travel survey over the same period as the Census can create severe respondent burden problems, with negative impacts on response rates for the travel survey (which, unlike the Census, is not compulsory). Secondly, even if the travel survey is run in the same year as the Census, it is quite normal to experience a significant delay before the release of the Census data (in Australia, where Census data release is relatively quick, the results from the 2006 Census were not released for about 15 months; t he same happens in countries

such as Chile). Therefore, there can be a delay in being able to expand the travel survey data up to the Census data. Thirdly, even though the new Census data will eventually be available for weighting, there are many activities in the design phases of the travel survey (such as selection of stratification regions and drawing of the sample) that rely on Census data. This will necessarily have to be done with data from the previous Census, and this can pose a problem if Census regions change between the old and the new Census. This can mean that the sample is drawn with one set of geographic boundaries, whereas the expansion is then to be done with a different set of boundaries. If the Census is to be used for expansion purposes, then it is probably best, for survey design purposes, to do the travel survey a year or so after the Census, so that the new Census data can be used for both sample design and expansion purposes.

With on-going surveys there are two major questions to be addressed in weighting and expansion. Firstly, how many years of data can be aggregated for expansion purposes, and secondly, which year of Census data should be used for this purpose. The answer to the first question will depend on how rapidly travel patterns are changing in the region, and on the sample size collected in each year. For example, as noted earlier, the Sydney HTS pools three years of data to obtain reliable trip estimates at the zone level. Similar procedures were adopted with VATS, although for some analyses where the data was not expected to change significantly over time, all six years of the database (1994-99) were pooled for analysis.

For a single year, or for a pool of several years of survey data, the question often arises as to which year of Census data should be used for expansion purposes, but this may not matter all that much as the purpose of expansion is to transform the sample data so that it represents the travel behaviour of a much larger population. Therefore, if the underlying travel data does not change much over time (e.g. household trip rates by demographic characteristics), then expanding it to different Census years merely changes the magnitude and distribution of the expanded result. Indeed, one could also expand the data to a

hypothetical distribution of households corresponding to some future development that has not yet occurred.

In the Santiago case, they had to face the problems associated with a rapidly expanding city. The 2001 Census was used to generate the sample, but by 2005 it was obvious that large parts of the study area, particularly to the North of the city, had acquired a significant middle to high income population with much higher car ownership than similar households in the more traditional quarters of the city. This meant that the car ownership figures in the survey were smaller and biased in their spatial distribution, in relation to independent data available; the same happened with other sources of independent data such as number of school children.

Furthermore, drastic changes occurred in relation to travel supply: four large urban highways were built after 2001 (and started operating at the end of 2005) and a large extension to the underground (one and a half lines serving massively populated new quarters) also started operating in 2006. This meant that the "state of travel" in 2006, could only be adequately represented by the data for that year, but the typical sample bias related to household size distribution and socio-demographics were difficult to ascertain as the Census data had become really obsolete.

Correcting the problem of rapid large-scale growth requires, among other things, 'listing' households in large areas where there were no households at the beginning of the project. Usually this is not only a data issue, but also needs additional funding. It also needs devising complex data expansion and re-weighting methods and this requires further research.

## USES OF THE DATA

We have discussed the way data is used in general in the first section. The uses of the data depend to a large extent on the needs of the commissioning organisation. As we noted earlier, these needs are often for modelling of current and forecast travel patterns in cities and urban areas – always for cars, but with increasing frequency also for public transport, cycling and walking (especially in developing countries).

The data uses are not, however, confined to modelling and, to some extent, only lack of imagination can limit the uses of the data. In particular, the data is often used for planning and policy purposes. Examples that we know of include working out the influence of changing the size of free travel zones to school, the impact of all children within 1 km of a school walking or cycling, and estimating differences in car-km between those people who travel to work and those that work at home. There are also non-traditional uses of the data outside the transport field, such as petrol companies using it to optimise the location of their service stations along arterial roads., and outdoor advertising agencies trying to tailor the content of roadside billboards to the demographics of passing motorists.

There are many advantages to encouraging the widespread, imaginative use of the data, since the more it is used, the more likely that continuing funding will be forthcoming. As an example, in the early days of the VATS survey, the main funding source was from users of the data – a good way to encourage innovative uses. We would note, however, that over the longer term this was not a viable way to fund the data collection process.

In the current political/social environment where there is a growing interest in climate change and its effects, it is likely that there will be new uses for the type of data we are discussing. For example, the type of vehicles people drive, on what roads, at what times of day, and potentially at what speeds are likely to be important elements of information in policy and decision making of the future. As an example, this could give information on who/what type of trips/people are contributing most to congestion. And all of these data elements could potentially be included in ongoing travel surveys – particularly if new data collection mechanisms such as GPS devices can be used.

## ONGOING SURVEYS AND CLIMATE CHANGE

In this section we focus on how ongoing surveys are likely to need to change to reflect the 21st Century focus on climate change – both in the data they collect and the way it is collected.

### Collecting Data Related to Climate Change

If, in an era where climate change is likely to be the most significant issue to be addressed in transport planning, the travel survey data is to be used for measuring changes in energy use and greenhouse gas emissions, then new types of data might need to be collected, such as much greater detail in information about household vehicles (i.e. make, model, year of manufacture, engine capacity, and fuel type). Also, coupling the survey with a GPS device carried by respondents can allow information such as speed and route to be added to the already rich data collected; this has been tried in the Sydney HTS. Finally, as can be seen from the following example, some surveys have already included and used typical on-going data for environmental modelling purposes.

A study conducted for Transport Canada by TUTI (Richardson and Seethaler, 2002) addressed the potential for developing an integrated vehicle-use and fuel consumption survey, where previously these had been undertaken as separate surveys in Canada. As part of that project, it was demonstrated how relatively conventional travel survey data (VATS) could be used to estimate fuel consumption and greenhouse emissions across an urban area, Using data from the 1995 VATS survey, the analysis estimated greenhouse emissions from passenger car travel in the Melbourne region, taking account of the following factors:

- Vehicle type used on trips (make, model, engine size).
- Fuel consumption rates by vehicle type.

- Emissions factors (kg CO2 per litre fuel consumed).
- Linking each vehicle to each trip
- Trip length.
- Average trip speed.
- Engine temperature (as a function of time since last trip and length of current trip).
- Ambient temperature (as measured by meteorological records).

The emissions associated with each trip could then be analysed as a function of all the other characteristics recorded in the travel survey (such as trip purpose, trip location, trip timing and driver characteristics), to obtain an understanding of how and why greenhouse emissions were being created as a result of passenger vehicle usage[13].

**The Climate Change Impacts of Data Collection**

A further issue that is likely to be raised in the near future is the actual impact of the survey collection on the environment. There are three types of emissions that must be included when estimating the carbon footprint from an activity:

- Scope 1: emissions created directly by people (e.g. fuel used in a private vehicle).
- Scope 2: indirect emissions from energy consumption (e.g. from fossil fuels burned to generate electricity).
- Scope 3: indirect emissions embodied in products or services consumed by an individual (e.g. paper consumed, flights taken).

Self-completion surveys require the use of large amounts of paper, which have embodied emissions. The emissions created by questionnaire delivery in mail-out surveys (i.e. VATS) are proportional to their volume and weight, as a fraction of total emissions is generated by mail deliveries. But VISTA has given much higher response rates as a result of personal drop-off and collection which usually involves car travel – at least in Australia, the US and Europe.

Face-to-face interviews have less paper involved, but almost always involve a larger amount of vehicle travel, even when samples are clustered. Further clustering could reduce vehicle travel but the need to obtain high response rates is likely to mean successive follow-ups to the same area even if the sample is clustered. Notwithstanding, in countries such as Chile the interviewers travel by public transport in most cases and some also use bicycles.

It could be argued that on-line methods will reduce climate change effects to the greatest extent, but this will need to be considered in the context of computer ownership and use, each of which has significant Scope 2 and 3 emissions. As an example of the magnitude of likely emissions associated with field surveys, and the costs of offsetting these emissions,

[13] This study is similar in methodology to the Fuel Diary in the German Mobility Panel (Appendix 1); it has run since 1994 and covers vehicle use and fuel consumption for passenger transport in the whole of Germany.

the current VISTA survey of 17,250 responding households in metropolitan Melbourne and regional Victoria has been estimated to require a total of approximately 300,000 km of field staff travel by car for the various components of the survey. This will generate about 85 tonnes of $CO_2$.

These emissions can, however, be offset for a total cost of about $1,000 (e.g. www.treesmart.com.au), an amount that has been built into the budget for VISTA. This cost is not excessive, but emphasises the fact that greenhouse emissions can be alleviated at modest cost if everyone, including survey organisations, does their bit to reduce and then offset their own emissions.

## KEY ISSUES FOR REFLECTION AND FOR THE FUTURE

This paper has covered a wide range of issues associated with large-scale on-going mobility surveys. Since it is a resource paper, we conclude by raising some key questions for thought:

- What are the best ways to deal with issues of retaining staff over time, and with keeping their levels of motivation and performance at a desired level?
- What types of questions will be needed in on-going surveys in the future, given the likely emphasis on issues of climate change?
- When do on-going surveys in urban areas give the most benefits?
- How can commissioning organisations expand the number of users of the data they collect?
- When might it be appropriate to use other methods (e.g. internet) as part of on-going data collection?
- Do you need to have equal sample sizes for each year, or are there advantages to taking a large sample every (say) four years and carrying out surveys of smaller samples in the intervening years?
- How can on-going travel surveys resist the over-surveying of the population, with consequent continuing falls in response rates?
- How do you tackle the integration weighting of successive "sample years" when drastic or massive changes in transport supply are experienced in the city?

### ACKNOWLEDGEMENTS

We are grateful to Martin Lee-Gosselin, Tim Raimond, Dirk Zumkeller and Johanna Zmud for their useful comments and support. As usual, any errors remaining in this paper are our responsibility alone. We also wish to thank FONDECYT, through Project 1050672 and the Millennium Institute on Complex Engineering Systems (Project P05-004F) for supporting the research activities of the second author.

## REFERENCES

Ampt, E.S. and J. de D. Ortúzar (2004) On best practice in continuous large-scale mobility surveys. *Transport Reviews* **24**, 337–363.

Battellino, H. and J. Peachman (2003) The joys and tribulations of a continuous survey. In *Transport Survey Quality and Innovation* (P. Stopher and P. Jones, eds.). Pergamon, Amsterdam.

Bratley, P., B. Fox and H. Niederreiter (1992) Implementation and tests of low discrepancy sequences. *ACM Transactions on Modeling and Computer Simulation* **2**, 195-213.

Bricka, S. (2006) Scheduling Considerations in Households Travel Surveys, in *Travel Survey Methods: Quality Survey Methods: Quality Considerations*, Stopher, P and Stecher, C., Elsevier Ltd.

Brög, W. and A.H. Meyburg (1982) Consideration of non-response effects in large-scale mobility surveys. *Transportation Research Record* **80**7, 39-46.

Brög, W, K. Fallast, H. Katteler, G. Sammer and B. Schwertner (1985) Selected results of a standardised survey instrument for large-scale travel surveys in several European countries. In *New Survey Methods in Transport* (E.S. Ampt, A.J. Richardson and W. Brög, eds.). VNU Science Press, Utrecht.

Kan, M.Y. and J.I. Gershuny (2004) British household panel survey calibrated time use data. Institute for Social and Economic Research, University of Essex.

Madre, J.-L, Axhausen, K.W., Brög, W. (2007) Immobility in travel diary surveys*, Transportation,* Springer, **34**, pp 107-28)

Meurs, H. and G. Ridder (1997) Attrition and response effects in the Dutch National Mobility Panel. In *Panels for Transportation Planning* (T.F. Golob, R. Kitamura and L. Long, eds.). Kluwer Academic Publishers, Norwell.

Mistro, R. D., and Behrens, R. (2008) How variable is the variability in traffic? Can TDM succeed? *Proceedings of the 27th South African Transport Conference*, Pretoria, South Africa, 7-11 July.

Murakami, E. and C. Ulberg (1997) The Puget Sound Transportation Panel. In *Panels for Transportation Planning* (T.F. Golob, R. Kitamura and L. Long, eds.). Kluwer Academic Publishers, Norwell.

Pas, E. (1988) Weekly Travel Activity Behavior, *Transportation,* Vol 15. 88-109.

Peachman, J., L. Mendigorin and L. Tanner (1998) Reducing interviewer costs in large-scale household travel surveys, *Proceedings Australasian Transport Research Forum*, Sydney.

Richardson, A.J. (2003) Behavioural mechanisms of non-response in mailback travel surveys. *Transportation Research Record* **1855**, 191-199.

Richardson, A.J. and H. Battellino (1997) Similar, yet different: some emerging trends in travel surveys in Australia. *International Conference on Transport Survey Quality and Innovation*, Grainau, Germany.

Richardson, A.J. and R.K. Seethaler (2002) A study of the potential of developing an integrated vehicle-use and fuel consumption survey. *TUTI Report 9-2002*, The Urban Transport Institute, Taggerty.

Richardson, A.J., Ampt, E.S. and Meyburg, A.H. (1995) *Survey Methods in Transport Planning,* Eucalyptus Press, Melbourne. http://www.transportsurveymethods.com.au/

Tourangeau, R., M. Zimowski and R. Ghadialy (1997) *An Introduction to Panel Surveys in Transportation Studies*. Prepared for FWHA, Washington, D.C.

Wermuth, M.J. (1985) Non-sampling errors due to non-response in written household travel surveys. In *New Survey Methods in Transport* (E.S. Ampt, A.J. Richardson and W. Brög, eds.). VNU Science Press, Utrecht.

## Appendix 1: National Data Sources

| Country | Survey Name | Year- Duration | Reporting Period |
|---|---|---|---|
| Austria | Austrian National Travel Survey | 1995 | Daily mobility: one day<br>Long distance travel: two weeks |
| Belgium | Belgian National Travel Survey (MOBEL) | 1998 – 1999 | Daily mobility: one day<br>Long distance travel: four weeks |
| Denmark | Danish National Travel Survey | 1992 – 2003 (continuous survey), 2006, 2007 | Daily mobility: one day (evenly distributed throughout the year) |
| Finland | Finnish National Person Travel Survey | 2004 – 2005 | Daily mobility: one day<br>Long distance travel: one month (evenly distributed throughout the year) |
| France | French National Travel Survey | 1993 – 1994 | Daily mobility: one weekday, two weekend days<br>Long distance travel: three months<br>Seasonality: eight waves of six weeks distributed throughout the year |
| Germany | Deutsches Mobilitätspanel (MOP) | Continuous since 1994 | One week |
| Germany | Mobilität in Deutschland | 2002 | Daily mobility: one day<br>Long distance travel: three months |
| Germany | INVERMO | 1999 – 2002 | Eight weeks |
| Greece | Athens Metro Development Study | 1996 | Daily mobility: one day |
| Italy | ITALY NHTS | 2004 / 2005 | Daily mobility: one day<br>Long distance travel: seven days + six months for international journeys, reporting periods evenly distributed throughout the seasons |
| Italy | 14. Censimento Generale della Popolazione e delle Abitazioni | 2001 | Daily communing: one day |
| Netherlands | Onderzoek verplaatsingsgedrag (OVG) / Mobiliteitsonderzoek Nederland (MON) | Ongoing since 1978 | Daily mobility and long distance travel: two days evenly distributed throughout the year |
| Norway | Norwegian National Person Travel Survey | 2005 | Daily mobility: one day<br>Long distance travel: one month (evenly distributed throughout the year) |
| Switzerland | Swiss Microcensus on Travel Behaviour | 2005 (conducted every five years since 1974) | Daily mobility: one day<br>Long distance travel: two week excursions; Journeys with overnight stays: four months |
| Sweden | Swedish National Travel Survey (RES) | 2005 / 2006 | Daily mobility: one day<br>Long distance travel: one month for 100 km journeys, two months for 300 km |
| Spain | Spanish MOVILIA Survey | 2000 / 2001 – 2002 | Daily mobility: one weekday, one weekend day (October - December)<br>Long distance travel: 12 months |
| United Kingdom | Great Britain National Travel Survey | January 2002 – January 2005 | Daily mobility: seven days<br>Long distance travel: four weeks (evenly distributed throughout the year) |