

Design and Methodology of 1991/92 Household Interview Survey

John Peachman
Survey Project Officer
Transport Study Group NSW

Abstract:

In 1991/92 the Transport Study Group NSW conducted a household interview survey of 20,000 households in the Sydney, Blue Mountains, Newcastle, Central Coast and Wollongong areas over a 12 month period. The survey parallels a similar survey held in 1981 for the Sydney region alone over a six month period.

This paper details the issues involved in developing the design and sampling methodology for the survey, and discusses the reasons for adopting the final approach. The implications of using a personal interview collection methodology will also be discussed.

The paper also addresses issues relating to the greater geographical and temporal spread of the 1991/92 survey, and details the inclusion of new questions designed to reflect transport-related social and demographic trends in Australia.

Contact Author:

John Peachman
Transport Study Group NSW
Locked Bag 1
HAYMARKET NSW 2000

Telephone: (02) 218 6638 Fax: (02) 218 6625

Introduction

In 1981 the Transport Study Group NSW (TSG), then known as the State Transport Study Group of NSW, conducted the 1981 Sydney Travel Survey. This survey was designed to provide a comprehensive picture of travel behaviour in the Sydney region, and to create a database sufficiently rich to satisfy the transport information needs of a broad spectrum of users, from the general public to transport modellers. The 1991/92 Household Interview Survey (HIS) was undertaken by the TSG largely as a continuation and updating of the 1981 survey.

The two surveys are broadly comparable in their aims and methodology, but the 1991/92 survey differs in having a wider geographical and temporal coverage than the 1981 survey. The 1991/92 survey covers a complete twelve months as opposed to six months, and includes the Newcastle and Wollongong areas as well as Sydney.

There are also some important differences in detail between the two surveys. While some reference will be made in this paper to the similarities between the two surveys, the paper's emphasis will be on describing the areas of difference, and in highlighting features that could impact on planning for future travel surveys. It should also be borne in mind that the paper was compiled in the middle of the 1991/92 survey, so that only preliminary unexpanded data was available at the time of writing.

Administration of the survey

In 1981 the administration of the survey was entirely by the Transport Study Group. In 1991/92, the fieldwork and data processing component of the survey was contracted out to AGB Australia. In general, the contracting out of the day-to-day survey administration to AGB worked very well. While largely chosen for their proven expertise and experience in large scale interviewer-management, AGB also provided extremely valuable feedback and advice at the planning stage, particularly with questionnaire design.

Sample frame

There were two basic approaches to creating a sample frame considered for the 1991/92 survey:

1. Administrative by-product listing
2. Area listing

An administrative by-product listing is a listing compiled by another organisation for its own purposes, but which may be suitable for a sample frame. An area listing is an actual listing of dwellings in the field by trained listers.

In the 1981 survey, electricity authority records were used as the sample frame. For the 1991/92 survey, usage of electricity authority records was considered, but was complicated by the fact that because of the extended geographical coverage, use of such records would have meant accessing data files from a number of electricity authorities, with an attendant increase in file incompatibility problems and increased administrative burden. Electricity authority records were finally rejected as a frame on this account. Australian Electoral Commission (AEC) records as a possible frame were also rejected after an investigation established that there was a significant and systematic underenumeration of dwellings occupied by migrant or 'ethnic' groups.

Because of these problems with administrative by-product listings, an area listing approach was chosen for the 1991/92 survey. The major advantage of an area listing, of course, is that - done properly - it provides the most up-to-date listing of dwellings possible. The main disadvantage, normally, is that being a labour-intensive method it can be expensive.

The pilot test for the 1991/92 survey was conducted using an area listing, and the listing proved to be of high quality. It is worth noting that even if an area listing is not possible for a main survey, it can be an invaluable exercise to conduct such a listing at the pilot stage, as the data obtained can be used as a check of any alternative frames proposed. As an example, the AEC analysis referred to above could only be done because there was an up-to-date area listing against which the AEC files could be checked.

Sample size

In the early stages of planning for the 1991/92 survey it was intended that the sample size for the survey would be 32,000 households, a figure calculated by applying the 1981 rationale of choosing a 2% sample of the population to the greater population catchment area of the 1991/92 survey. However, once detailed cost estimates were available from the pilot test, it became clear that having such a large sample would exceed budget limitations.

It was decided, therefore, to depart from the simple approach of selecting a fixed percentage of the population, and instead to quantify the trade-off between costs and statistical accuracy as a basis for the determination of sample size. The key determinants were:

- (i) The target result was to be a quantifiable level of statistical accuracy at a fixed confidence level for each Statistical Local Area (SLA) in the survey.
- (ii) The confidence level was to be 95%.
- (iii) The data item on which calculations would be based was average number of trips per SLA.

The table below shows the relationship between costs and statistical accuracy calculated on this basis.

<u>Cost(*)</u>	<u>Sample Size</u>	<u>Relative Standard Error</u>
100%	32000	10.6%
95%	30400	10.9%
90%	28800	11.2%
85%	27200	11.5%
80%	25600	11.8%
75%	24000	12.2%
70%	22400	12.6%
65%	20800	13.1%

* As percentage of cost for 32000 households.

As can be seen from this table, reducing the projected sample size from 32000 to 20800 leads to significant cost savings without a commensurate loss in accuracy. For this reason, the latter figure was chosen for the 1991/92 survey.

Sampling methodology

The sampling methodology for the 1991/92 survey was a two-stage cluster sample design with Census Collector Districts (CDs) as the units selected at the first stage of sampling ("primary sampling units"). The number of CDs per SLA was chosen using systematic sampling with probability proportional to size of SLA ('size' refers here to number of dwellings estimated from census data). For each selected CD, blocks were formed and a single block selected at random. Within the chosen block, a cluster of seven dwellings was selected, each of which was allocated a different day of the week as Travel Day.

Cluster sampling was chosen purely for cost reasons. It helps to reduce the distance between survey dwellings, with a consequent reduction in travel time costs, both at the area listing and interview stages. It is worth noting, however, that in the 1991/92 survey the most significant travel time cost savings probably arose not from clustering, but from an extension of the 'First Attempt Date', ie the maximum number of days before Travel Day that an interviewer could attempt to recruit the household.

In the pilot test, interviewers were not allowed to attempt initial contact with households more than three days before Travel Day. The rationale for keeping the contact period as small as possible was that it would reduce the elapsed time from recruitment to completion, and limit the possibility of respondents rethinking their position on co-operation. However, analysis of the pilot, and feedback from interviewers, established that this recruitment threshold was too low. Significant amounts of travel time were being wasted because the following scenario was occurring: "I had just interviewed Household X, and was driving past Household Y.

I could have recruited Household Y right then, but their Travel Day was four days later, so I couldn't. I had to make a special trip back the next day to recruit them."

In response to this problem, the First Attempt Date was extended from three to five days for the main survey. As was expected, this significantly reduced the amount of wasted travel time. At the time of writing of this paper, it was not possible to adequately quantify the actual savings on travel costs, but the overall reduction in travel costs per interview from the pilot to the main survey was around 50%, and it is believed that the First Attempt Date extension was probably the main contributor to this reduction.

In general, this aspect of the survey highlighted the fact that in planning for a survey with a wide geographical scope, it is important to consider interviewer travel as a two-phase process:

1. *Inter-CD travel:* Travel to and between survey CDs.
2. *Intra-CD travel:* Travel within a CD.

For a wide survey area such as the 1991/92 survey, the cost of Inter-CD travel will greatly outweigh Intra-CD travel, since a large amount of time can be required simply to get to a survey location, which in itself may require a comparatively small amount of time to traverse for recruitment and interviewing (even without clustering). It follows, then, that in planning for any similar survey, special attention should be paid to the comparative cost weights of the two phases of travel. If, for example, it is decided that the statistical disbenefits of clustering are particularly undesirable, it may still be possible to offset the cost increases from a non-clustered methodology by a judicious minimisation of the Inter-CD costs

Scope and coverage

As mentioned earlier, the 1991/92 survey differed from the 1981 survey in having a wider geographical and temporal coverage. Having a twelve month's survey allows, of course, for a greater elimination of seasonal effects, while the inclusion of other major urban areas provides information on inter-urban travel behaviour, as well as local information for these other areas. Extending the geographical and temporal coverage of the 1991/92 survey also highlighted the following points:

(i) It made even more critical the need to employ the most experienced and reliable organisation possible to handle the logistics of such a massive exercise. Maintaining interviewer stability and momentum over twelve months, over a wide geographical area, is not for the faint-hearted!

(ii) A greater geographical spread can have a significant impact on the mechanics of selecting an appropriate sample frame. The reference earlier to the rejection of electricity authority data as a sample frame illustrates the fact that the wider the area to be covered, the more likely it is that a number of different

Peachman

administrative by-product collections will need to be utilised to compile a comprehensive sample frame.

(iii) Including less densely urban areas in the study area can also have an impact on sample loss. A notable feature of the 1991/92 sample was the high proportion of vacant dwellings listed in some areas (for example, Port Stephens). The reason for this was that the areas involved had a high incidence of dwellings used as holiday accommodation, which were difficult to establish as such at the listing stage, only being identified as vacant once they were included in the survey.

(iv) The unit cost per interview can be increased due to greater travel costs in less urbanised areas.

In addition to the wider geographical and temporal coverage of the 1991/92 survey, the coverage was further extended as follows:

Visitors: In 1991/92, data was collected for visitors as well as residents of the household. A visitor was defined as any person whose normal residence was outside the study area and who was staying with the household at the start of the Travel Day. In 1981 all visitors were excluded from the survey.

Children Under Four: In 1991/92, trip data was collected for all people within scope of the survey irrespective of age. In 1981 trip data was not collected for children under four years of age.

Collection methodology

As for the 1981 survey, a personal interview methodology was adopted for the 1991/92 survey. The main reasons for this were:

(i) It was established in the 1981 survey (Ampt (1982) Section 8.3.2) that the reportage of trip behaviour is likely to be significantly understated unless a personal interview methodology is adopted.

(ii) A personal interview methodology maximises response rate, as households are more likely to respond to the 'personal touch' that only an interviewer can provide. As a corollary, non-response bias is reduced.

(iii) A personal interview allows for maximum flexibility in the type of information collected, as an interviewer can (non-directively) assist the respondent in understanding the more difficult questions.

The personal interview methodology is not without its problems, however. For one thing, it presupposes that personal contact can actually be made, and this is not always the case. The 1991/92 survey highlighted one area that was a particular problem in this regard, and which is likely to become an even greater problem for future surveys.

It was found during the survey that obtaining response from people living in security blocks of flats was especially difficult to obtain, because it was far too easy for the householder to refuse to co-operate when the initial contact was through an intercom system rather than face-to-face (the whole point of a personal interview methodology, of course, is to maximise face-to-face contact). Inner-city flat dwellers, in particular, can be especially loathe to co-operate, as they tend to get a lot of junk-mail and 'junk-callers'.

There appears to be very little that can be done to alleviate this problem, and, unfortunately, it is a situation which will almost certainly worsen rather than improve. Such blocks are likely to become more prevalent, and the security systems installed more sophisticated. The extent of the difficulties involved can be illustrated by the fact that there are already blocks that don't even have intercoms for notifying residents of a desire to enter - entry can only be effected by using a PIN number).

The significance of this problem should not be under-estimated. The following table illustrates the relationship between dwelling structure and non-contact rates:

Non-contact Rates by Dwelling Structure

Dwelling Structure	Non-Contact Rate	
	1981	1991 (Preliminary)
Separate House	2.6%	6.7%
Semi-detached etc.	3.4%	3.6%
Flats	8.7%	16.5%

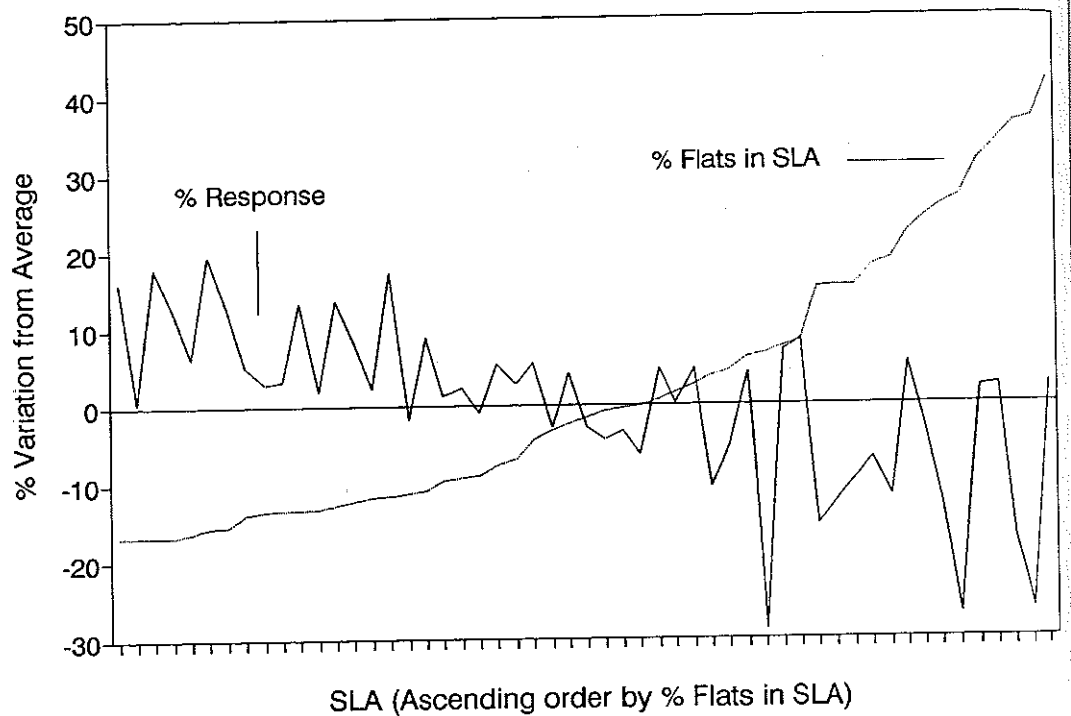
It is clear from this that what was already a significant problem in 1981, was exacerbated for the 1991/92 survey. The 1991/92 survey also showed that there was a clear correlation between the level of non-contact and block density, as follows:

Dwelling Structure	Non-Contact Rate
Flats:1-2 storeys	12.2%
Flats:3 storeys	15.4%
Flats:4+ storeys	18.8%

Peachman

It follows from the above that survey response levels will vary significantly on an area basis depending on the housing density within each area. The strong inverse relationship between the number of flats surveyed within an SLA and the response level in that SLA is illustrated in **Figure 1**.

Survey Response and Housing Density 1991/92 HIS (Preliminary)



The main inference from this is that in designing a similar survey to the 1991/92 survey it is important to take into account the differential response rates that will arise in different areas. The experience of the 1991/92 survey is that it is not cost-effective to attempt to rectify response problems such as these after the event. Even with the dedication of full-time supervisors to 'turning around' initial refusals, only about 25% success rate was achieved. The problems need to be anticipated at the planning stage and addressed then. Over-sampling of likely

problem areas (this would still leave bias problems, of course), or the employment of area-specific collection strategies, for example, might be considered. Even then, however, the possibility remains that the problem might be essentially intractable for a personal interview methodology.

Questionnaire design and content

The 1991/92 survey was designed to provide a broad spectrum of data that reflected current transport-related issues and demographic and socio-economic trends. It included questions which were either (i) entirely new, or (ii) had appeared in some previous surveys individually, but not necessarily together in the one survey, or (iii) had appeared in previous surveys, but in a less detailed form. The general design of the questionnaire was based on that used in the 1981 survey, and similar later surveys (Adelaide 1986 and New Zealand 1989/90).

One important difference between the 1981 and 1991/92 questionnaire designs was the decision in 1991/92 not to include detailed mode choice questions for work, school, shopping and social/recreational trips. The main reason for this decision was that collection of such data is very time-consuming for both interviewer and respondent, and there has been a general acceptance since 1981 that data for choice models is better collected through a panel methodology in longitudinal studies, or an interactive methodology such as HATS (Household Activity Transport Simulator) (Taylor pp.47-50). Nevertheless, because of strong user needs for such data, a compromise was made by including a simple mode choice question for Home to Work trips only, at minimal extra cost.

A detailed list of changes and additions to questionnaire design and content in the 1991/92 survey would be inappropriate for this paper, but some of the more significant new features of the 1991/92 survey questionnaires are described below.

Childcare

The most significant addition to the 1991/92 survey questionnaire was a series of questions relating to childcare arrangements, which parallel those normally asked about employment arrangements viz.

- (i) Name and address of childcare facility
- (ii) Type of childcare
- (iii) Time-period of childcare
- (iv) Attendance in last seven days

These questions were included in an attempt to provide data for a perceived increase in community awareness of the importance of childcare in determining

family travel behaviour, a factor heightened by the increase of female participation in the workforce. The influence of childcare needs and arrangements on travel patterns is greater than a simple summation of the trips generated specifically for movement between childcare establishments. As for employment and education needs, it can be a critical factor in limiting the geographical and mode choices available to people.

To our knowledge, the childcare questions in the 1991/92 are the first time such questions have been asked. It may be of interest, then, to elaborate on some of the issues that were addressed before a final decision was made on their inclusion.

A prerequisite for meaningful collection of this data was the independent decision to extend the scope of the 1991/92 survey to include all residents of the household irrespective of age (ie. children under four years of age, excluded from the 1981 survey, were included in 1991/92). This original decision was made to facilitate capture of travel behaviour driven by the needs of young children. Once this decision was made, it was clear that the same sort of location and activity data that was collected for employment was needed for childcare.

A number of special considerations arose. One concern was that the information being asked would be regarded by respondents as too sensitive; that even respondents who freely gave information regarding their own daily trip patterns might be loathe to do so for their children. This was a particular concern since there were actually a number of child abduction cases in the headlines around this time. There was also concern that there might be an adverse reaction to male interviewers asking such questions.

Given these concerns, it was a pleasant surprise to find that the pilot test showed that there was no significant respondent resistance to answering the childcare questions, nor were there any apparent differences in response related to the sex of the interviewer.

A side-effect of the greater detail required for children in the 1991/92 survey was that it increased the size and complexity of the person details questionnaire. Because of this it was decided to break with the methodology of having one questionnaire for every person (used in the most recent household surveys, Adelaide and New Zealand), and instead return to the approach used in the 1981 survey of having specific questionnaires for adults and children. This turned out to have a number of benefits:

(i) Specialised questionnaires for adults and children meant that each could be streamlined to remove superfluous questions and reduce the sequencing that would have been necessary using one questionnaire for both. For example, the children's questionnaire did not need to have any employment-related questions, and had a greatly simplified set of trip questions. As a by-product of questionnaire specialisation, there was a reduction in printing costs. Data for children could be collected using a 19-page questionnaire rather than the 40 pages or more questionnaire that would have been necessary using a one-questionnaire approach (the adult questionnaire was 39 pages).

(ii) It is well established that the use of proxy interviews leads to significant under-reporting of trips (Ampt (1986) p.4). Because of this, the 1991/92

survey maintained the strict rules applied in the 1981 survey regarding proxies i.e. for adults no proxy interviews were accepted for trip information, though proxies were accepted for children.

Having separate adult and child questionnaires helped to reinforce to interviewers the different rules that applied to each regarding the acceptance of proxy interviews. In fact, by having separate questionnaires, it was possible to specifically disallow proxies through the question sequencing of the adult questionnaire.

This was not an insignificant contribution. In the pilot test, many of the interviewers had problems coming to grips with the rules as to when proxies were or were not acceptable. Their natural inclination was to 'get what they could'.

(iii) The reduced question set for the child's questionnaire had a favourable effect on response. Feedback from interviewers indicated that respondents were less likely to baulk at the completion of a questionnaire for a child when it became obvious that it was less onerous than the adult questionnaire.

Overall, interviewer and public response to the two questionnaire methodology was very favourable. It was generally seen as a useful streamlining of the interview process.

Education

In addition to the new childcare questions a complementary set of questions was asked about education arrangements, both for adults and children i.e.

- (i) Name and address of educational institution
- (ii) Type of educational institution
- (iii) Attendance in last seven days

(In the 1981 survey the name and address of the school attended by a child was asked, but no further information was obtained. No information was obtained for adult education).

In the 1991/92 survey, therefore, data for the three major activity attractors - employment, education and childcare - was collected in an equal and consistent fashion.

Time Choices

The 1991/92 survey questionnaire was also designed to maximise the amount of information available on the time choices available to people for the journey to

Peachman

work. The 1981 survey had collected data on people's working hour arrangements (Fixed hours, Flexitime etc), but had made no attempt to quantify the connection between time choices available and actual time chosen. The 1991/92 survey included questions in which respondents gave the actual work-band hours under which they operate, thus opening up the possibility of analysis of the relationship between people's time options and their actual behaviour.

In an attempt to further widen the analytical possibilities in this area, a question that (to our knowledge) had never previously been asked in travel surveys was also included: 'Reasons for departure time for work'. We believe that this combination of related time-choice questions should provide a rich vein of information for research.

Travel Costs

The actual costs incurred in travelling are, of course, a crucial factor in the generation of trips, and the trip choices (mode, route, time) made for the trip. The 1991/92 survey had a greater emphasis than the 1981 survey on direct quantification of these factors. Respondents were asked to report the fare paid for public transport usage, and the amount paid for parking for private vehicle usage.

Conclusion

There is little doubt that up till now the data obtained from large-scale travel surveys has to some extent been under-utilised. There are various reasons for this, but certainly a major factor in the past was the relative inaccessibility of the data due to the then necessary reliance on cumbersome and expensive mainframe systems (Taylor (1989)). In the 1990s this barrier to accessibility will be largely overcome, through the use of (comparatively) inexpensive hardware and increasingly user-friendly software.

The conjunction of the broad content of the 1991/92 Travel Survey with increased accessibility will provide an ideal opportunity for both the general public and specialist analysts to maximise the information potential of a large scale travel survey. The 1991/92 Travel Survey should provide a firm base for transport decision making in the future.

References

Ampt, E S (1986) Adelaide Household Travel Survey - Report on the Pilot Survey (Director-General of Transport South Australia)

Ampt, E S (1982) The 1981 Sydney Travel Survey - Some Recent Advances In Large Scale Travel Surveys, in Richardson, A J (1982) Transport Survey Methods

Richardson, A J (1982) Transport Survey Methods (Department of Civil Engineering: Monash University)

Taylor, M A P (1989) et al Overview of Travel Surveys Report No. 1 Melbourne Travel Survey Research Project