Reducing interviewer costs in large-scale household travel surveys

John Peachman  
Data Operations Manager, TDC  
NSW Department of Transport

Lita Mendigorin  
Project Officer, TDC  
NSW Department of Transport

Michael Tanner  
GIS Specialist, TDC  
NSW Department of Transport

Abstract:

In 1997 the Transport Data Centre (TDC) of the NSW Department of Transport commenced the Household Travel Survey (HTS), a continuous survey of personal travel behaviour in the TDC Study Area, an area which is broadly equivalent to the Greater Sydney Metropolitan Region. The HTS is conducted using a 24-hour travel diary and a face-to-face interview collection method.

In any survey where the face-to-face interview method is used to collect data, the major cost will usually be for interviewer time. An interviewer must travel to a household and conduct an interview at that household. The relative cost significance of these two processes will depend on the nature of the survey. In a survey with a large geographical scope such as the HTS, interviewer travel time becomes the major influence on costs. This paper discusses a number of options for reducing this cost without significantly affecting the quality of data collected.

Contact author:

John Peachman  
Data Operations Manager  
Transport Data Centre  
NSW Department of Transport  
GPO Box 1620  
SYDNEY NSW 2001

Telephone: (02) 9268 2954  
Fax: (02) 9268 2853  
Email: peachmaj@transport.nsw.gov.au
**Introduction**

In 1997 the Transport Data Centre of the NSW Department of Transport commenced the Household Travel Survey (HTS), a continuous survey of personal travel behaviour in the TDC Study Area, an area which is broadly equivalent to the Greater Sydney Metropolitan Region.

The HTS is conducted using a 24-hour travel diary and a face-to-face interview collection method. The face-to-face interview collection method is used because it maximises response levels, provides a very high level of data quality, and allows for the collection of a wider variety of data items required by users. The high response levels obtained using face-to-face interview minimises the effects of non-response bias, and the presence of an interviewer ensures that questions are more clearly understood by respondents (Richardson, et al., 1995, Ampt, 1996).

There are three major aims when designing a survey like the HTS:

- **Optimise demographic spread** i.e. ensure that the spread of households selected is sufficient to represent the different demographic groupings that have significantly different travel behaviour.

- **Optimise geographic spread**. Areas can have similar demographic characteristics, but still produce significantly different travel patterns, due to transport-related local differences, such as the nature and quality of the road network and the availability and quality of public transport. Therefore, it is important to cover as wide an area as possible to allow for these local differences.

- **Optimise temporal spread** i.e. ensure that the spread of interview dates is sufficient to represent the different time periods throughout a year that produce significantly different patterns in terms of travel behaviour. There are two broad time distributions that need to be allowed for:
  - Weekly distribution: differences in travel behaviour for each day of the week (Monday – Sunday) must be properly represented, and
  - Annual distribution: differences in travel behaviour due to seasonal differences must be properly represented. At the broadest level, holiday and non-holiday periods must be accounted for.

The need to optimise demographic, geographic and temporal spreads has to be balanced against the need to keep costs to a minimum. This paper discusses the implications of these competing needs.

**Interviewer time**

In any survey where the face-to-face interview method is used to collect data, the major cost
Reducing interviewer costs will usually be for interviewer time. An interviewer must travel to a household and then conduct an interview at that household. The relative cost significance of these two processes will depend on the nature of the survey. If, for example, the survey involves a lengthy interview, but this interview can be conducted within a short distance of the interviewer's home, then interview time will be a more important cost factor than travel time. However, in a survey with a large geographical scope such as the HIS, travel time becomes the major influence on costs.

Figure 1  TDC Study Area - Census Collector Districts (CDs) sampled in first quarter of the Household Travel Survey
In the HTS itself, the geographical scope is particularly large, as the IDC Study Area ranges from Shoalhaven in the south to Port Stephens in the north and the Blue Mountains in the west. The annual gross sample is 5,000 households, but even with a sample of this size the need to cover such a wide geographical area means that the distribution of sampled households in some areas can be quite sparse. Figure 1 illustrates a quarterly distribution of sampled households in the HTS (the basic geographical unit sampled in the HTS is the ABS Census Collector District (CD); within each CD, seven households are selected with each household being randomly allocated to one of the seven days of the week).

It can be seen from Figure 1 that because of the large physical size of the IDC Study Area, sampled CDs outside the inner/middle Sydney area are separated by significant distances, and (though the scale of the figure cannot show it) distances between inner/middle Sydney CDs are non-trivial. Since the HTS is conducted using the face-to-face interview method, this means that the cost of interviewer travel time in moving to, from and between sampled CDs is by far the major factor in survey costs. Irrespective of how much overall survey costs can be reduced by rationalising other cost components (e.g. project management or data-entry costs) interviewer travel time will remain the predominant cost factor. The remainder of this paper, therefore, concentrates on examining options to reduce interviewer travel time.

Reducing interviewer travel time

There are two broad options for reducing interviewer travel time:

1. Minimising the time required to travel to each household, and
2. Minimising the number of calls to each household.

Minimising the time required to travel to each household

The time required to travel to a household is largely dependent on the distance between the interviewer’s home and the surveyed household. Minimising this distance is therefore a key requirement if interviewer travel time is to be reduced.

The most obvious way to minimise the distance between interviewers’ homes and surveyed households is by specifically recruiting an interviewer team with this purpose in mind. But when recruiting an interviewer team interviewers should be selected who live as close as possible to the surveyed households for which they are responsible. However, when doing so, a number of practical limitations need to be considered.

Firstly, although the benefits of proximity are clear, an interviewer should not live too close to the households for which they are responsible. For instance, it would be running a risk to allocate a household to an interviewer who lived in the same, or even a nearby, street because of the possibility that household members would know the interviewer either personally or by sight, and as a result might be unwilling to divulge personal information

686
Reducing interviewer costs

to that interviewer. The sample loss through non-response that could occur might be unacceptably high.

Secondly, and more importantly, in order to have a situation where each interviewer lives close to the households for which they are responsible, it may be necessary to have an interviewer team that is too large to be practical. There are three main reasons for this:

1. Workload per interviewer is inversely proportional to the size of the interviewer team. That is, for a fixed number of survey households, the more interviewers there are, the less work there is for each interviewer. Consequently, if the interviewer team becomes too large, the workload per interviewer may become so small that the number of hours of employment provided is not enough to make it worthwhile for the interviewer to continue working on the survey. The threshold where this becomes a problem will vary for individual interviewers, but in general will be related to the availability of alternative interviewing work.

2. The overall quality of the interviewer team can diminish if too much priority is given to selecting interviewers on the basis of their proximity to the surveyed households. A balance needs to be struck between the need for proximity and the need to ensure that the best interviewers possible are selected.

3. The larger the interviewer team, the greater are the management costs involved in training, maintaining and monitoring the team.

Minimising ‘dead’ travel time

Assuming that, within the limitations described above, an interviewer team has been selected to minimise travel time by optimising proximity, there are still travel time inefficiencies that can occur. This is due to constraints that arise from the need to have a temporal spread of households.

Interviewer efficiency is greatest when an interviewer can reduce overall travel time by overlapping a visit to one household with visits to other households i.e. one interviewer trip away from home produces more than one household visit before returning home. The extent to which this aim can be achieved depends on the number of households being surveyed in the same area at the same time. If during a trip to one household there are no other households to be interviewed on that day then clearly no overlapping of visits is possible—the interviewer must travel to and from the area involved for the sake of a single interview.

The inefficiencies arising from non-overlapping visits are increased the more an annual temporal spread (see Introduction) is incorporated into the survey sample design. Theoretically, the annual distribution of travel behaviour should be optimised by ensuring that sampled CDs that are geographically close are interviewed at different times of the year. Unfortunately, this separation of geographical and temporal proximity can lead to interviewer travel time inefficiencies.
A simple example will illustrate the extent to which optimising temporal spread can lead to travel time inefficiencies. For the sake of the example, assume that an interviewer has a workload of two CDs, and the distance (x) between the interviewer's home and each CD is equal, as is the distance between each CD. The CDs are separated in time (interviewed during different weeks).

**Figure 2a. Contiguous CDs separated in time**

As can be seen from Figure 2a, the separation of the two CDs in time means that the interviewer must travel a minimum distance of 4x, since no overlapping of visits is possible.

**Figure 2b. Contiguous CDs paired in time**
Reducing interviewer costs

However, in Figure 2b the CDs are paired in time (interviewed during the same week), which means that an interviewer can now take advantage of trip-chaining i.e. visiting the two CDs within the same trip and thus minimising total travel time. For the example given, the interviewer only needs to travel a distance of 3x, thus saving 25% in total travel time. Thus, by pairing the CDs in time, the interviewer’s ‘dead’ travel time (i.e. time spent travelling back from a household without any possibility of contacting another household while en-route) has been significantly minimised.

While it is clear that pairing CDs in time can reduce the amount of interviewer travel through trip-chaining, estimating the extent of the reduction can be difficult. The simplified model shown in Figures 2a and 2b needs to be replaced by a real-world model that takes into account the actual road network that interviewers travel on. This is best undertaken using a Geographic Information System (GIS), where actual travel times can be accurately estimated using appropriate shortest path algorithms.

An additional restriction on obtaining benefits from CD pairing is that interview times tend to be clustered i.e. households tend to prefer interviews at the same times of the day, usually the times when all members are at home together and there is sufficient free time for an interview. For this reason, by far the most popular interview period is the evening; in the HTS, 50% of interviews are conducted between 4:00pm and 8:00pm.

The clustering of household interview times means that even if it is geographically possible for an interviewer to interview multiple households in one day, it may not be logistically possible, because some of the households may require interviews at the same (or an overlapping) time. The longer the actual interview time, the more likely it is that household interview times may conflict.

In the HTS, the average time for a household interview is approximately 35 minutes. Allowing additional time for a short break and checking of questionnaires, this means that on average only one household can be interviewed per hour. Therefore, during the interviewing peak period of 4:00pm to 8:00pm the theoretical maximum number of interviews is four. In practice, it is most unlikely that this number would be attainable because (i) it would require each household to be available at consecutive hours, with no overlap, and (ii) the households may be in different CDs, which though geographically close in general, will still be far enough apart to require a separate trip to go from one to the other, and hence take up additional time. Taking all these factors into account, it is practical to reduce travel time by clustering two CDs in time, but to cluster more than that leads to logistical difficulties.

Minimising the number of calls to each household

The number of calls an interviewer has to make to a household obviously affects total interviewer travel time. The extent to which calls can be minimised to reduce travel time will depend on the survey design.

There are two stages to the interview process in the HTS: the pre-Travel day interview (household recruitment), and the post-Travel Day interview (interviews with household
Minimising pre-Travel Day calls

Pre-Travel Day calls are made by interviewers to recruit households for the survey. To ensure consistency of approach, thresholds for when interviewers can attempt to contact households must be set. The two thresholds are the 'Earliest First Attempt Date' and the 'Latest First Attempt Date'. The 'Earliest First Attempt Date' is the earliest date that an interviewer can first attempt to contact a household for the recruitment interview. The 'Latest First Attempt Date' is the latest date that an interviewer can first attempt to contact a household for the recruitment interview.

Using Interviewer Workload Charts

As will be discussed in detail below, determining the optimal organisation of interviewer workloads can be a complex process that involves balancing the design constraints of the survey against the practical limitations presented to the interviewer. This process is greatly facilitated through the use of interviewer workload charts that graphically display the workload activity during a particular period.

For this paper, simple spreadsheet charts have been used to simplify presentation, though more complex approaches are possible. A sample chart is shown in Figure 3.

Figure 3. Example interviewer workload chart

This chart represents one interviewer workload for a CD. The bolded day (Monday) is the first travel day for the week, indicated by 'TD1' underneath. The second travel day is Tuesday, indicated by 'TD2' underneath, and so on. The numbers to the left of each travel day (-) represent the number of days before the travel day, and the numbers to the right (+) the number of days after the travel day. It will be seen in the discussion to follow that such charts are an essential tool for analysing the interaction of the various requirements and constraints of the interviewing task.

The Earliest First Attempt Date

As mentioned above, the 'Earliest First Attempt Date' is the earliest date that an interviewer can first attempt to contact a household for the recruitment interview. In determining the Earliest First Attempt Date the main aim is minimise the number of days between the
Reducing interviewer costs

recruitment date and Travel Day. This is to ensure that:

1. **Response is not affected:** if there is too big a gap between the recruitment interview and Travel Day, response may be affected because:
   
   (i) Households may be unwilling/unable to commit themselves to participating in the survey too far in advance of Travel Day, and
   
   (ii) The greater the time between the recruitment date and Travel Day, the greater the possibility that unforeseen events (including respondent change-of-heart) may affect the household’s willingness/ability to participate

2. **The quality of travel data obtained is not affected:** even if households do respond, a big gap between the recruitment interview and Travel Day may mean that the quality of travel data obtained is poor. An important function of the recruitment interview is to explain to the respondent the nature of the upcoming post-Travel Day interviews, and to provide him/her with memory joggers along with instructions as to their usage. The greater the time between the recruitment date and Travel Day, the less likely it is that the respondent will properly remember the detail of these instructions

In the initial stages of the 1997/98 HTS interviewers were instructed that the Earliest First Attempt Date should be five days before Travel Day. This threshold of five days was based on the experience of a previous travel survey, the 1991/92 Household Interview Survey (HIS), where it was found that extending the Earliest First Attempt Date from three to five days significantly reduced costs (Peachman, 1992).

After the first quarter of the 1997/98 HTS the fieldwork operations were reviewed and it was found that even with the Earliest First Attempt Date set at five days, inefficiencies were still arising from the fact that interviewers were unable to maximise their presence in an area by recruiting households that fell outside the five day Earliest First Attempt limit. As a result of this review it was decided to extend the Earliest First Attempt Date to seven days before Travel Day.

Increasing the Earliest First Attempt Date from five to seven days in the 1997/98 HTS had a significant effect on costs as the greater opportunities afforded to interviewers to recruit households when already in the area meant that the number of pre-Travel Day visits to CDs was reduced by around 25%, as shown below:

<table>
<thead>
<tr>
<th>Earliest First Attempt Date</th>
<th>5 days</th>
<th>7 days</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average visits per CD</td>
<td>5.3</td>
<td>3.9</td>
</tr>
</tbody>
</table>

Comparison of the survey results using an Earliest First Attempt Date of seven rather than five days indicated that extending the threshold produced no significant, if any, negative impact on response levels or the quality of travel data reported. Therefore, **the results of the 1997/98 HTS indicate that using seven days before Travel Day as the Earliest First Attempt Date is more cost-efficient than using a lesser number of days.**
It should be noted that extending the Earliest First Attempt Date beyond seven days was not considered for the HTS, because, apart from any other considerations, it can lead to confusion during the recruitment interview. Up to seven days before Travel Day, the interviewer can say to the respondent “your Travel Day is next Monday (Tuesday etc)”. However, if the recruitment date is more than seven days before Travel Day, then Travel Day will be the “Monday after next Monday”, say, which is contrary to the simple wording required at this initial stage of contact with the household.

The Latest First Attempt Date

The 'Latest First Attempt Date' is the latest date that an interviewer can first attempt to contact a household for the recruitment interview, and in the absolute sense this is obviously the day immediately before Travel Day. However, making the day before Travel Day the Latest First Attempt Date is not acceptable because it assumes that every household will be contactable on the day before Travel Day, and clearly that will not always be the case. Some households will require visits on more than one day before contact is made, and the Latest First Attempt Date needs to contain a contingency for this fact. The 1997/98 HTS adopted a Latest First Attempt Date of three days before Travel Day, consistent with the threshold used in the 1991/92 HIS. Reducing this threshold to two days offers potential cost savings but this potential benefit is offset by a significantly higher risk that a household may not be contacted before Travel Day.

Minimising post-Travel Day calls

There is a crucial difference between pre and post-Travel Day interviews. For post-Travel Day interviews there is less capability for conducting multiple interviews on the one day than there is for pre-Travel Day interviews, and hence less capability for reducing the number of times the interviewer needs to visit the CD. There are two main reasons for this:

1. Interview time for the post-Travel Day interview is considerably longer than that for the pre-Travel Day interview, since all members of the household are interviewed in the post-Travel Day interview, and each interview involves a potentially lengthy reporting of trip details

2. As much as possible, post-Travel Day interviews need to be conducted within a day or two of Travel Day, since the post-Travel Day interview requires respondents to recall trips from memory (though, ideally, using a memory jogger), and beyond a day or two the quality of response may decline significantly

The interviewer workload chart shown in Figure 4 illustrates this point. The two shaded sections in the chart show the days when it is logistically possible for an interviewer to optimise their presence in an area by conducting five to seven interviews in the one day (an interviewer can conduct a recruitment interview 1 or 2 days before Travel Day, as long as the first attempt was made on an earlier date). For pre-Travel Day interviews, there is no negative impact in conducting multiple interviews in the one day; the more the better. However, for post-Travel Day interviews a similarly cost-efficient clustering of interviews can only be achieved...
Reducing interviewer costs

The First Travel Day of the week

At first glance, it would seem irrelevant what the first Travel Day of a week is; no matter what the first day is, seven days are sampled consecutively and each day is subject to the same rules for contact. However, costs can, in fact, be affected by the ordering of days. The main reason for this is that the weekend is by far the best time for interviewing households, and the order of Travel Days impacts on the extent to which weekend interviewing is maximised. Weekend interviewing is generally the best time for interviewing because:

- It increases the likelihood that an adult member of the household will be available (Groves, 1989);
- Interviewing need not be confined to evenings (between 4:00 pm and 8 pm);
- There is a greater chance that respondents are more relaxed and more cooperative.

The effect of choosing the wrong First Travel Day is shown in Figure 5a:

Figure 5a Don't make Saturday the First Travel Day!
As can be seen from Figure 5a, making Saturday the First Travel Day is the worst option possible, as there are only two households (Travel Days 1 and 2) that can be recruited on the weekend. What about Friday?

Friday as the First Travel Day is better than Saturday, but not by much as only three households (Travel Days 1, 2 and 3) can be recruited on the weekend.

If we apply this procedure for each day of the week we can ascertain the relative merits of each day as the First Travel Day:

<table>
<thead>
<tr>
<th>First Travel Day</th>
<th>Households that can be interviewed on weekend</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sunday</td>
<td>5</td>
</tr>
<tr>
<td>Monday</td>
<td>6</td>
</tr>
<tr>
<td>Tuesday</td>
<td>6</td>
</tr>
<tr>
<td>Wednesday</td>
<td>5</td>
</tr>
<tr>
<td>Thursday</td>
<td>4</td>
</tr>
<tr>
<td>Friday</td>
<td>3</td>
</tr>
<tr>
<td>Saturday</td>
<td>2</td>
</tr>
</tbody>
</table>

This data shows clearly that Monday and Tuesday are the best options, both days allowing for the possibility of six households being recruited on the preceding weekend, as shown in Figures 5c and 5d:

<table>
<thead>
<tr>
<th>MON</th>
<th>TUE</th>
<th>WED</th>
<th>THU</th>
<th>FRI</th>
<th>SAT</th>
<th>SUN</th>
<th>MON</th>
<th>TUE</th>
<th>WED</th>
<th>THU</th>
<th>FRI</th>
<th>SAT</th>
<th>SUN</th>
</tr>
</thead>
<tbody>
<tr>
<td>-7</td>
<td>-6</td>
<td>-5</td>
<td>-4</td>
<td>-3</td>
<td>-2</td>
<td>-1</td>
<td>TD1</td>
<td>-7</td>
<td>-6</td>
<td>-5</td>
<td>-4</td>
<td>-3</td>
<td>-2</td>
</tr>
<tr>
<td>-7</td>
<td>-6</td>
<td>-5</td>
<td>-4</td>
<td>-3</td>
<td>-2</td>
<td>-1</td>
<td>TD2</td>
<td>-7</td>
<td>-6</td>
<td>-5</td>
<td>-4</td>
<td>-3</td>
<td>-2</td>
</tr>
<tr>
<td>-7</td>
<td>-6</td>
<td>-5</td>
<td>-4</td>
<td>-3</td>
<td>-2</td>
<td>-1</td>
<td>TD3</td>
<td>-7</td>
<td>-6</td>
<td>-5</td>
<td>-4</td>
<td>-3</td>
<td>-2</td>
</tr>
<tr>
<td>-7</td>
<td>-6</td>
<td>-5</td>
<td>-4</td>
<td>-3</td>
<td>-2</td>
<td>-1</td>
<td>TD4</td>
<td>-7</td>
<td>-6</td>
<td>-5</td>
<td>-4</td>
<td>-3</td>
<td>-2</td>
</tr>
<tr>
<td>-7</td>
<td>-6</td>
<td>-5</td>
<td>-4</td>
<td>-3</td>
<td>-2</td>
<td>-1</td>
<td>TD5</td>
<td>-7</td>
<td>-6</td>
<td>-5</td>
<td>-4</td>
<td>-3</td>
<td>-2</td>
</tr>
<tr>
<td>-7</td>
<td>-6</td>
<td>-5</td>
<td>-4</td>
<td>-3</td>
<td>-2</td>
<td>-1</td>
<td>TD6</td>
<td>-7</td>
<td>-6</td>
<td>-5</td>
<td>-4</td>
<td>-3</td>
<td>-2</td>
</tr>
<tr>
<td>-7</td>
<td>-6</td>
<td>-5</td>
<td>-4</td>
<td>-3</td>
<td>-2</td>
<td>-1</td>
<td>TD7</td>
<td>-7</td>
<td>-6</td>
<td>-5</td>
<td>-4</td>
<td>-3</td>
<td>-2</td>
</tr>
</tbody>
</table>

**Figure 5c  Monday as First Travel Day**
Reducing interviewer costs

Are Monday or Tuesday thus equally preferable as the First Travel Day?

No. The charts allow for further analysis here. It can be seen from Figure 5c that if the First Travel Day is Monday then the Latest First Attempt Date for the household with Travel Day One is Friday, which means that even if all the other households are recruited on the weekend the interviewer must make a prior trip to the CD to recruit the household with Travel Day 1, a potentially very inefficient trip. Figure 5d shows that Tuesday does not have this disadvantage. Households with Travel Days 1 - 6 can be recruited on the weekend, leaving only the household with Travel Day 7, which can be recruited on any of the upcoming days Wednesday, Thursday or Friday when the interviewer should be in the area anyway to conduct post-Travel Day interviews i.e. a special trip to the CD for recruitment is not required.

It is clear then that the choice of First Travel Day is by no means irrelevant to costs, and should never be chosen arbitrarily. For the Latest First Attempt Date of three days applied in the 1997/98 HIS, Tuesday is probably the best choice as First Travel Day. If the Latest First Attempt Date were different, an analysis using that threshold would have to be undertaken to determine the exact effect on the First Travel Day. In the current example, if the Latest First Attempt Date had been two rather than three days then Monday would not have had the disadvantage referred to above.

Summary

For a large-scale survey of travel behaviour using the face-to-face interview method by far the most significant cost is for interviewer travel time. There are a number of options for reducing this cost without significantly affecting the quality of data collected.

The major influence on interviewer travel time is simply the distance between the interviewer’s home and the households for which they are responsible. Therefore, the number of interviewers should be as large as possible to ensure the geographical scope of each interviewer’s workload is minimised. However, the number of interviewers must also be consistent with the need to have a quality, stable and manageable team.
Interviewer travel time can also be significantly reduced by ensuring that sampled CDs that are geographically close are interviewed at the same time. Pairing CDs in this way is cost efficient and logistically possible, while still retaining an acceptable temporal spread of sampled households. However, extending the principle to three or more CDs is likely to present logistical problems.

Finally, interviewer workloads must be organised to ensure that interviewers can maximise the number of households visited in each trip to a sampled CD. The ‘Earliest First Attempt Date’, ‘Latest First Attempt Date’ and ‘First Travel Day’ have a significant impact on this number. Analysis of the HIS indicates that an Earliest First Attempt Date of seven days is more cost-efficient than using a lesser number of days, and that with a Latest First Attempt Date of three days the optimal First Travel Day is Tuesday.

References


Groves, Robert M. Survey Errors and Survey Costs, The University of Michigan, 1989
