

A Panel Approach to Evaluating TravelSmart Initiatives in the Short Term – South Australia Pilot Survey

Peter Stopher¹, Stephen Greaves¹, Min Xu², Camden Fitzgerald¹, Natalie Lauer¹

¹ Institute of Transport and Logistics Studies, The University of Sydney, Sydney, NSW, Australia

² NSW DIPNR, Sydney, NSW, Australia

1 Introduction

The Department for Transport, Energy and Infrastructure (DTEI – formerly the Department of Transport and Urban Planning) of the government of South Australia initiated a TravelSmart intervention in early 2005, called “TravelSmart Households in the West”. This intervention is unusual, in that it is designed to be implemented over a period of slightly longer than two years, and that it is being initiated with contact through various community groups, which will subsequently be followed by a “door-knocking” approach. It is also somewhat unusual, in that an independent evaluation has been commissioned of this intervention, which will take place over a period of nearly three years. It is the pilot surveys of this independent evaluation that are the subject of this paper.

As part of the “TravelSmart Households in the West” project, the Institute of Transport and Logistics Studies is undertaking the independent evaluation of the voluntary travel behaviour change initiative. As discussed elsewhere, ITLS is also designing a long-range evaluation procedure for the National Travel Behaviour Change Program (Stopher et al., 2005a). In undertaking the evaluation for the short-term period of the “TravelSmart Households in the West” project, ITLS is using elements of the proposed long-range monitoring program, proposed to the National Travel Behaviour Change Program. The purpose of this paper is to describe the procedures that are being implemented to monitor the introduction of TravelSmart in Adelaide in the “Households in the West” project and to provide some preliminary information on the initial waves of the monitoring effort.

2 The Design of the Short-Term Monitoring Program

The proposed short-term monitoring strategy for South Australia consists of two components. The first component is an annual GPS household panel, that is planned to be surveyed in each of 2005, 2006, and 2007, with the surveys taking place approximately in the May-June period of each year. The second component is a panel of households participating in an odometer survey, every four months from May 2005 to September of 2007. Without discussing at this point the various reasons for selecting these particular surveys, it is important to realise that the issue here is to conduct a series of surveys, with the first survey being conducted prior to any TravelSmart intervention work, thereby providing the “before” case, and then for a series of surveys to be conducted during the time that the intervention takes place (two years), and following completion of it, to track the changes in behaviour that households implement as a result of the TravelSmart interventions. It might be suggested, for example, that surveys should be conducted at least annually for the period from before the interventions begin, until after they conclude. In the case of South Australia, it is desired that the final survey be completed in sufficient time to be able to deliver a report on the results of the intervention by the end of calendar 2007.

2.1 Objectives of the Monitoring

The objectives of the evaluation are to determine whether or not the travel behaviour of those who accept TravelSmart tools changes in directions that are expected and anticipated for this policy instrument, and to determine whether or not the changes that may be implemented by households and people are sustained through the period of this evaluation. Specifically, for each household in the panel, the following items are to be measured, and changes looked for:

- Vehicle kilometres of travel (VKT) for vehicles in personal use by members of the household;
- Number of trips undertaken per day;
- Number of trips by mode of travel;
- Number of trips by purpose; and
- Trip lengths of trips by mode, purpose, and time of day.

These values are to be tracked through the three-year period of this project for households that are participating in TravelSmart and also for households that are not participating, but which are located within the suburbs where the TravelSmart program is introduced.

2.2 Control Groups

There is always a need for some type of control group in evaluations of this nature. The function of the control group is to determine if there are likely to be background changes that may have affected the behaviour of the target households, either in trending towards increased trip making, or towards less. One of the requirements for a valid control group is that it should consist of households who are as similar as possible to the target households, so that other factors, stemming from differences between the control and target groups, do not interfere with the assessment of external changes occurring to the target group.

As more and more of the population is exposed to voluntary travel behaviour change, finding a control group will become more and more difficult. In this particular case, it was decided that there simply was no good candidate locality for a control group and that the only possibility was to use households that did not participate in TravelSmart from within the targeted suburbs form the control group. The advantage of this is that such households will be much more likely to be similar to targeted households, in terms of sociodemographic status and travel options and services. The disadvantage is that there is no way to be certain that there has been no diffusion of the TravelSmart program to these households, and that they are, therefore, good predictors of what would have happened to TravelSmart households if there had been no intervention. Because of the staged roll-out of TravelSmart in the “Households in the West project”, it was decided that this was still the best option. In addition, when any household leaves either one of the panels, they will be asked if they have heard of TravelSmart and, if so, if and how they estimate it has impacted their travel.

2.3 Panels

There is still debate within the transport community on the value of panels. Many transport professionals do not agree on the merits of panels. However, in a situation in which the main goal of a series of surveys is to measure change, there should be little debate on the fact that a panel is clearly the best strategy. The merits and problems of panels are discussed here as a precursor to describing the application of these in South Australia.

There are, in general, two options for conducting repeated surveys of a population. On the one hand, each sample for each survey in a series can be drawn randomly and

independently of any other sample. Consider the case at hand. In South Australia, there are a total of approximately 60,000 households in the region to be covered by this TravelSmart project. One could design an initial “before” survey to take place in the period between April and June 2005, before interventions are offered to households, with subsequent surveys, consisting of the “after” surveys, conducted periodically through the following 27 months. Such periodic surveys should show both the progressive take up of TravelSmart tools, and the short-term sustainability of travel behaviour changes that households receiving tools might undertake. In the independent sampling approach, the samples for these surveys would each be drawn independently of each other, and, unless the sample sizes were very large in proportion to the population of 60,000 households, it would be unlikely that the same households would appear in any two surveys.

The second option is to recruit a panel of households. In this case, the panel of households are surveyed repeatedly throughout the period of the survey. Thus, in the case of South Australia, the panel would be recruited for the before survey, and then the same households would be asked to respond to each subsequent survey. There are several types of panels and the interested reader can find more details about each type in the referenced literature (Kish, 1965; Richardson et al., 2003; Armoogum et al., 2004; Stopher and Greaves, 2004). The first is an overlapping sample. In this case, some households do not repeat the survey on a particular wave, and new households are added each time to the sample, to replace those that fail to repeat the survey. The sample size of the panel remains unchanged, however. It should be noted that households that only participate once in this design provide no information on change. Households must participate on at least two occasions to provide change data. However, the reality in any real-world panel is that there will always be some drop out at each wave of a panel. The second type of panel is called a subsample panel and occurs when those who fail to repeat the survey are not replaced, so that the sample size for the panel drops throughout the period of the panel. The final wave of this panel will be a subsample of the households originally recruited, as will each of the succeeding waves after the first. This panel requires oversampling at the outset, so that the final sample size in the last wave of the panel reaches the desired panel size. The third type of panel is a true panel in which all panel members remain in the panel for the duration of all waves of the panel. In practice with human populations, this is not an achievable panel, because it arises only when there is no loss of sample from any wave of the panel.

Implementation of a panel in the South Australia case would need to be either a subsample panel or an overlapping sample. We have chosen the overlapping sample design, because this design tends to be more efficient, especially when there are three or more waves for the panel. There is, in fact, little difference between a subsample panel and an overlapping panel if there are only two waves.

2.3.1 Advantages of a panel

The advantages of a panel are primarily threefold. First, especially when measuring change, a panel will require a significantly smaller sample than will repeated cross-sectional samples. This results (Stopher and Greaves, 2004) from the effects of the covariance between the panel measurements on two successive occasions, which reduces the overall sampling error substantially. Thus, for a given desired level of accuracy for measuring a change, a panel will require a smaller sample size than two independent cross-sectional samples. As Stopher and Greaves (2004) have pointed out, this reduction in sample size can be on the order of a factor of 2 to 3, even when the correlation between the two waves of the panel is modest. It is also worth noting here that the correlation between pairs of waves of a panel is likely to be significantly higher if more days of measurement are included on each occasion. This will decrease the variances, and increase the correlation, so that the sample size will also fall with increasing number of days of measurement.

Second, a panel provides information on the dynamics of change, which cannot be seen in repeated cross-sectional surveys. In the case of a voluntary travel behaviour change

intervention, for example, if two successive cross-sectional surveys showed that there was a decrease in VKT for household cars, and an increase in public transport riding, the causes of such changes would be indeterminate. It is quite possible that these observed changes are seen because of other factors, such as changes in the economy, changes in bus routes or fares, or other factors. The changes may also not be changes at all, but it could be that the first sample was of households that had a higher average VKT than those in the second sample. Thus, it would not even be clear that a *change* had actually occurred, unless there were ways of establishing that the households in each sample had originally had the same average VKT. On the other hand, with a panel, it is possible to attribute cause and effect, through comparing the same households on each occasion. There also is no possibility of stating that the households sampled on one occasion had higher VKT than those sampled on the second occasion.

The third advantage of panels is that they are generally much less burdensome for the respondents in the longer run. Whereas, in repeated cross-sectional sampling, sample members will need to complete full surveys, including all social and demographic data, panel members, on succeeding occasions, should only need to update their information from the previous wave. This can often be done with relatively little effort. Similarly, if the remainder of the survey requires repetition of the same task, such as filling out a diary, or entering odometer readings onto a card, the households will become used to this task, and will also be able to complete it with relatively little effort.

These three advantages of panels represent the principal reasons for considering panels. One might also add, however, that there is another potential advantage. This is that once having recruited households to take part in the survey, it is often easier to have them participate a second, or even a third time, compared to having to recruit a completely new set of households. In a small pilot panel in Canberra (Stopher and Alsnih, 2005), more than 70 percent of households who completed the before survey also completed the after survey. In this survey, 84 households were initially recruited for four panels (two each of control and target, and two undertaking GPS surveys, with two undertaking a diary survey). Of these 84 households, 51 actually completed the before task (60.7 percent). In the after survey, 37 of the 51 households were able to be re-contacted (72.5 percent), of which 36 (97 percent) completed the after survey. (In using an overlapping design, a further 12 households were recruited for the after survey, bringing the total for the after survey to 49 households. Of those 12 households, 10 completed the survey task, so that the final after sample was 46 households.) To achieve the same sample sizes with repeated cross sections would presumably have required about 80 households to be recruited for the after survey, instead of attempting to re-contact 51, and freshly recruiting a further 12 households.

2.3.2 Disadvantages of a panel

The above illustration section also indicates one of the biggest problems of a panel – attrition. Attrition is the loss of sample resulting from loss of interest by recruited panel members, moving away from the study area, death, dissolution of households, etc. These causes are normally expected to deplete the panel by about 30 percent in the first year, although by decreasing percentages in subsequent years. For example, in the German Mobility Panel (Armoogum et al., 2004), the average attrition between the first and second year has been 30.1 percent, while it has been 25.1 percent between the second and third years (17.5 percent of the first wave sample). In the Puget Sound Transportation Panel, after 12 waves, there are still about 200 households of the original 1,200 that are in the panel. If attrition continued at an average of 27.5 percent per year, then there should only have been 35 households in the twelfth wave that were in the first wave. Attrition can be reduced somewhat by maintaining regular contact with panel households, such as by mailing out information relating to the study results, and undertaking small personal contacts, such as sending out greeting cards at appropriate times. Nevertheless, one should expect somewhere in the vicinity of 30 percent attrition in the first year of a panel, with decreasing levels from then on.

The second disadvantage of a panel is conditioning. This has to do with two occurrences. The first is that panel members may change their behaviours because of their participation in the panel. This may arise because the subject of the panel survey makes panel members more aware of certain issues or behaviours, and creates in them a desire to make changes in their own behaviours. The second aspect of conditioning arises from panel members becoming more skilled at completing the survey task, and either completing the task with increasing accuracy, or becoming more skilled at answering questions in a way that they feel to be more politically acceptable, as opposed to being truthful. Generally, conditioning arises as an issue in panels that run for a long period of time, but is not usually evident in two or three years of participation.

A third disadvantage of panels, especially for subsample panels, is that the panel members may become increasingly less representative of the population. This can be corrected to some degree, if desired, through replacement for attrition in overlapping sample designs. However, the extent to which this can be done, while preserving the panel itself, is often limited, especially if the panel is continuing for some time, and attrition decreases in each successive wave.

A fourth disadvantage of panels is the effort needed to maintain the panel. Doing successive cross-sectional sampling is clearly less involved and requires less effort, in that there is no need to send out interim documents to members of the sample to maintain interest, and there is no issue of updating and correcting address information. Thus, there is much more work involved in maintaining a panel than for doing repeated cross-sectional samples.

Finally, there is a greater likelihood that panels will require incentives, to induce them to continue to participate. The greater likelihood that incentives will be required arises because of the repetition of the tasks required of households in the panel sample. Also, because these households will be more aware of what is expected of them in subsequent panel waves, they may feel that they should receive some compensation for continuing to undertake this task, whereas repeated cross-sectional samples will have no such concerns.

Nevertheless, the advantages of panels clearly outweigh the disadvantages, especially for measuring changes in behaviour over time, in the opinion of the authors of this paper. Therefore, for the project of concern in this paper, panel surveys have been selected as the method for measuring change resulting from voluntary travel behaviour change interventions.

2.4 Survey Methodology

2.4.1 GPS Survey

Travel diaries have proved to be somewhat unreliable for collecting travel behaviour data, as has been clearly demonstrated in several recent studies (Pearson, 2005; Wolf et al., 2003). Our own experience with diary surveys has shown, in three separate instances, that diaries are not as reliable as we would desire. In Canberra (Stopher and Alsnih, 2005), for example, the diaries showed results that were completely contrary to GPS measurements, and only after considerable effort was expended on diary repairs. Similarly, in New South Wales (Stopher et al., 2005b), we have again had to expend considerable effort to repair diary data. The most common problem encountered is people omitting to report trips back to home, especially at the end of the day. Any omitted trips will be especially problematic when one is attempting to measure a relatively small change in travel behaviour. Third, in checking the Sydney Household Travel Survey, we found that the face-to-face interviews underreported the number of trips that people make by about 7 percent (Stopher, et al., 2005c).

Based on these experiences, we decided that a diary approach to evaluation would not be successful, especially because of the relatively small changes we expect to measure from

TravelSmart initiatives. In addition, it is desirable to have as long a period of reporting as possible for the survey, and diaries of more than two days' duration are not really feasible. Indeed, while there is one instance of diaries being completed for a period of six weeks (Axhausen et al., 2000), and there have been experiments with one-week diaries (need references), the general experience with diaries of even two days duration is that there is a drop off in reporting on the second day (Purvis, 1997), while diaries for a longer period suffer from continuing drop off. In a Dutch one-week diary, respondents were called every three days, to encourage response, and the result was a clear decline in reporting every second and third day, with a pick up following each reminder call. As noted earlier, if it is possible to measure multiple days, then this will be expected to reduce the variability in daily travel, thus either increasing the accuracy of measurement, or decreasing the sample size requirements. Therefore, another point against the use of diaries is the fact that multiple days of measurement are not really feasible.

ITLS has been working for several years on the idea of using Global Positioning System (GPS) receivers to collect data on personal travel. The specific approach is to use a passive GPS device, where the person being surveyed has to do as little as possible, other than take the device with them as they travel. The initial devices that were used were restricted to in-vehicle use, being powered from the car's electrical system, and including an antenna/receiver that could be mounted on the roof of the car, or used inside the car, and including a logging box, where the data are processed and stored. An early wearable version, that included a battery pack for power, and required the logging box and battery to be carried in a bag, with the antenna/receiver on the shoulder strap, was used by ITLS in studies in both Canberra (Stopher and Alsnih, 2004) and Sydney (Stopher et al. 2005c). However, we have been working with an Australian firm for the design and development of a new wearable device, which has been described elsewhere (Stopher, Greaves, and FitzGerald, 2005). This latest device is compact and lightweight and is easily carried by respondents. It can be clipped to the belt, carried in a bag or pocket, be used as a key ring, etc. The device shows the same advantages and limitations of all GPS devices to date (see also Stopher, Jiang, and FitzGerald, 2005).

There are several important advantages to a GPS approach for evaluating a TravelSmart initiative. First, the focus of this evaluation is on the vehicle kilometres of travel (VKT) by car, the person kilometres of travel (PKT) by public transport, and the number of trips made in total and by different modes (car driver, car passenger, walking, bicycling, and public transport). Mode of travel can largely be determined from the GPS records by using such things as speed, acceleration, stopping, and also the disappearance of valid measurements for a distance and time that are consistent with use on a public transport vehicle. Trip purpose and travel costs are of relatively little importance to this evaluation, although the former can be measured by GPS, using land-use layers in GIS applications. Diaries and other similar surveys cannot provide accurate measurements of distance travelled, because people are known to be poor at estimating travel distances (Stopher et al., 2005c). Therefore, VKT and PKT cannot be measured accurately through diaries. Also, people omit to report trips, as has been shown in a number of studies (Wolf et al., 2004, Pearson, 2005, Stopher et al. 2005c), so that the number of trips is not known accurately from diary surveys. Second, GPS has the advantage that it measures distance (and time) very accurately, and also can provide very good measurement of the number of trips (usually to within ± 5 percent, whereas people report trip times and distances to within ± 10 to ± 25 percent). Third, because GPS measurement imposes very little burden on the people being surveyed, it can easily be used for much longer periods of time than diary surveys. Most transport planners would agree that diary surveys are limited to about 2 days in most cases, with exceptional cases permitting three days of diary data. However, GPS devices, of the type developed by ITLS, can be used for periods of a week or longer. One of the major advantages of this is that the variability of VKT, PKT, and trips measured over a week is much less than the variability of a one-day or even two-day measurement of these values. For example, in our work in Canberra, we found that the variance of daily VKT from one week of GPS measurement was about two-thirds of the variance from a two-day diary. Richardson, Seethaler, and Harbutt

(2003) have shown that two-day variability is also lower than one-day variability by a factor of about 70 percent. Therefore, this is a distinct advantage of GPS measurement.

On the other hand, this is largely an unproven technology in this application. We do not have reliable information about variances in the key measures from GPS measurement, and therefore cannot determine the sample size accurately. Second, we do not know how the public will accept the new GPS device. We have had some experience with the earlier GPS devices which suggest that GPS is not accepted as well as a face-to-face interview, but is accepted better than postal surveys. Recruitment rates of around 60 percent have been achieved for GPS, with 60 to 70 percent of the recruited households completing the GPS task. This gives an overall response rate of about 40 percent, which is similar to telephone recruitment for CATI surveys, is higher than most postal surveys, and lower than face-to-face surveys. This, however, related to an earlier and much bulkier device. It is not known what effect the new, small and lightweight device will have on these statistics. Third, there is little information available on how much of a full week of measurement people will provide. In a pilot test in Canberra, respondents were asked to take GPS devices for one week, and use them each day. Between the before and after surveys in that pilot survey (which was conducted with a panel), we obtained a total of 245 person-days of travel, out of a potential maximum of 315 days. Given that we could expect that an average of 10 percent of person days, individuals will not leave home, the maximum number of days that should have been expected was about 283. The response obtained represents about 87 percent of the maximum expected number of person days that might have been measured.

A further issue relates to the use of the panel. We do not have reliable information on the attrition rates for a panel using GPS. Again, the Canberra pilot survey suggested that figures quoted elsewhere of about a 30 percent sample loss in a year probably held in that study, although those panel members that were still in the panel at the second wave all completed the GPS task. For the South Australia project, we decided to recruit a panel of 200 households for the GPS task, and to survey them once each year, obtaining from each household one week of GPS data. GPS devices would be provided to all household members over the age of 14 years.

2.4.2 Odometer Survey

Given the unknowns for the GPS survey, it was decided to undertake an odometer survey as both a back-up and a verification tool for the GPS survey. As discussed elsewhere (Stopher et al., 2005a), past experience with odometer surveys has concentrated on asking people to report in one survey contact two odometer readings, one at the beginning of a diary period and one at the end. Generally, experience with this has been that people will remember one of these two reporting activities but often forget the second one. Also, the dates on which people remember to do the task are often not the precise dates that were asked for in the diary survey, such as at the beginning of the first diary day and at the end of the second diary day. Based on these experiences, we decided on a different approach. In this study, households are first recruited to the odometer panel. They must be pre-qualified as owning at least one car, since without a car, there is no odometer to be read. In the recruitment process, details are collected about the household, its members, and its vehicles. Following this, the household is asked to write down the odometer readings of each vehicle available to the household, together with the date on which the reading is taken. It is not necessary for the household to record all odometer readings on the same day. Following this, at approximately four-monthly intervals, the household is asked to record odometer readings again, along with the date on which the recording is made. At that time, we also check whether any characteristics of the household may have changed. The specific date on which the recording is made is of little importance, because we can determine the number of days that have elapsed since the last recording was made, and estimate an average daily VKT for that period. We can do the same for any period for which odometer readings have been obtained for the same vehicles of the same household. Thus, we can compute an average

from approximately eight months, and an average from approximately twelve months, and for any longer period.

It should be noted that we do not ask for the time at which the odometer reading is noted, to reduce respondent burden. However, it is not likely that the time will make any significant difference to the results. The assumption that is integral to not requesting the time is that respondents record their odometer readings at the same time in each wave. However, the worst cases would be presented by a respondent who records odometer reading at the beginning of one day and the end of the day for the next wave. This could effectively add one day extra to the travel period. Similarly, another respondent might do the reverse, thereby subtracting one day from the travel period. However, with a total travel period of approximately 120 days, the level of error provided by this slight uncertainty was considered too small to be of concern.

Because the task is rather simple, because households are only asked to update household, person, and vehicle characteristics from the previous time, rather than completing a new survey, and because households will be contacted every four months, we anticipate that attrition from the odometer panel will be less than that of an annual panel survey. It has been shown in panels that more frequent activity in the panel (every four months compared to once a year) reduces attrition. Also, because the task is small, we anticipate some decrease in the number who do not wish to continue. However, this has yet to be determined in practice. Again, as with the GPS survey, this type of survey has not been undertaken before. Therefore, again, we do not know what the variability in daily VKT will be from such a survey, and cannot easily estimate a sample size for the panel. Without knowledge of the variance in odometer readings, it was decided to recruit a panel of 1,000 households. This panel is distinct from the GPS panel. It is expected that, because the task of recording one's odometer reading is technical in nature, the problem of conditioning the panel will be largely avoided.

3 The Pilot Survey

In keeping with good survey practice, the first step in this project was to design and execute pilot surveys of each of the GPS panel and the odometer panel. Because we are setting up panels that will be measured repeatedly, it was decided that the pilot survey samples in each case would be recruited to remain in their respective panels. The goal was to recruit 40 households to undertake the GPS pilot survey and 60 households to undertake the odometer pilot survey.

3.1 GPS Pilot Survey

The GPS pilot survey was designed so that households would be recruited to the panel, and asked to undertake a one-week GPS survey, followed by a second one-week survey one month later. Panel members would then continue to undertake the GPS survey at one-year intervals in 2006 and 2007. The decision to ask panel members to undertake two one-week surveys in a short space of time was partly to test the procedure of gaining updated information on the households, and also to have the second round of GPS measurements coincide with the timing of the rest of the main panel. Based on previous experience with GPS, it was also decided to offer half of the pilot sample households an incentive to undertake the GPS survey. This incentive was set at \$5 per GPS device. Because of an assumed higher response rate to the incentive, it was decided to offer the incentive to 40 households and to offer no incentive to a further 80 households. The initial expectation was that about 25 percent of contacted households would agree to be part of the panel and actually complete the tasks, and that an incentive might double that to 50 percent. Hence, the initial recruitment was set up with these numbers.

Because it was known that a substantial proportion of households in the study area were either without a land line telephone, or had a “silent number” (not listed in the directory), it was decided to undertake recruitment by initially selecting household addresses, and mailing a recruitment package to the households in the sample. This package consisted of an introductory letter from the Government of South Australia, a consent form required by the University of Sydney, a respondent information sheet, explaining what each person was expected to do, and the forms to complete about household, person, and vehicle information. Households were offered two alternative means to respond to the survey – a stamped, addressed envelope was provided for posting the information back or households could elect to provide the information by telephone. The consent form was to be posted back for either the postal or telephone completion, and households were asked to indicate their preferred telephone number, if they desired to provide the information by phone.

Addresses for posting out the recruitment packages were obtained from the parcel-level GIS of the study area. This identified those parcels in use for residential purposes, and provided addresses for each dwelling on each parcel, as well as indicating how many units there were in multi-family dwellings. This information was used to sample addresses for recruitment. We also attempted to match the sampled addresses to listed telephone numbers and found that generally less than 60 percent of addresses could be matched to a telephone number.

After sending out the recruitment packages, we provided telephone reminders to those households that had not returned these packages and for which we were able to gain a match of a telephone number. We lost a substantial number of households through address problems, where either there was no longer a dwelling at the parcel, the home was empty, or the unit number that we assigned was not recognisable and was returned as undeliverable. In addition, there were ineligible households (due to language problems), and invalid phone numbers from the matching procedure. For households that we were unable to match a telephone number, and for which we had no reason to believe that the address was invalid, we sent postal reminders. After providing reminders, the results from the initial 120 addresses sampled were as shown in Table 1.

Table 1: Sample Disposition for the GPS Pilot Survey – Initial Sample

	AU\$5 Incentive/device	No Incentive	Total
Total Sample	40	80	120
Known Ineligible (Invalid numbers, known ineligibles, return-to-sender)	5	17	22
Known Eligible Households	25	36	61
Eligibility Rate [#]	83%	68%	73%
Unknown disposition	10	27	37
Potentially Eligible Sample^{##}	33	54	88
Refusals	18 (55%)*	19 (35%)*	37 (42%)*
Recruits	7 (21%)*	16 (30%)*	23 (26%)*

[#] Known Eligible/ (Known Eligible + Known Ineligible)

^{##} Known Eligible + Unknown*Eligibility Rate

* Percentage of potentially eligible sample

As can be seen, the total known eligible contacts were only 61 from 120, although 37 addresses remained of unknown eligibility, because no successful contact was made with them. From the 61 eligible contacts, 23 were successfully recruited. However, the incentive was not effective, and the recruitment rate with the incentive was considerably less than the rate without an incentive. Had no incentives been offered, it would appear likely that the response from this initial sample would have been closer to 26 households.

Given that we had not succeeded in recruiting the number of households desired for the pilot, we changed the method of recruitment for the balance of the sample. We drew

additional addresses and matched these to telephone numbers. For those addresses for which we were able to match telephone numbers, a pre-notification letter (only) was sent to the households. Following receipt of the pre-notification letter, the households were phoned and asked to participate in the study. No incentive was offered. This recruitment step included determining the number of people in the household and the number over 14 years of age, and also confirming the household's address, to which the GPS devices should be sent. Following recruitment, a package was couriered to each household, consisting of one GPS device for each person over 14 years of age, a subject information statement, instructions for how to use the GPS device, a charger for each device, and the household, person, and vehicle forms to be completed for the household. From a sample of 300 addresses drawn, we were able to telephone match 160, and sent pre-notification letters to all 160. The results of this recruitment procedure are shown in Table 2.

Table 2: Sample Disposition for the GPS Pilot Survey – Second Sample

Total Sample	Total 160
Un-used sample	18
Known Ineligible (Invalid numbers, known ineligible, return-to-sender)	42
Known Eligible Households	93
Eligibility Rate [#]	69 %
Unknown disposition	7
Potentially Eligible Sample^{##}	98
Refusals	61 (62%)*
Recruits	32 (33%)*

[#] Known Eligible/ (Known Eligible + Known Ineligible)

^{##} Known Eligible + Unknown*Eligibility Rate

* Percentage of potentially eligible sample

At the time of writing, 42 of the households in the pilot group have returned GPS devices, with 94 persons using the GPS devices (2.24 per household). From a preliminary analysis, these 94 persons provided an average of 6.4 days of data. (We would expect 6.3 approximately, assuming that there is an underlying non-mobility rate of 10 percent.) On the average, we have retrieved data on 20.4 trips per person, which gives an average of 3.2 trips per person per day. The average distance of recorded trips is 5.4 kms and the average duration is 11.6 minutes. This covers trips by car (as driver or passenger), by public transport, walk, and bicycle. The numbers of trips are, at this stage, a mix of linked and unlinked trips, because any trip that involved a wait at a location to board a vehicle during the trip will have been recorded as separate trips (broken by the wait), whilst any trips in which no wait occurred would be recorded as a single trip. In some instances, if the person waited in a location where the signal was lost, but did not wait for more than a few minutes, the segments of the trip would not have been separated.

3.2 Odometer Pilot Survey

As for the GPS pilot survey, the odometer pilot survey was designed so that those households recruited would become a part of the main panel. The pilot survey households were asked to provide odometer readings shortly after recruitment, and then again, one month later. At that point, they are also synchronised to the main panel, and would continue to provide odometer readings every four months thereafter, until about August 2007. Initially, we followed a similar procedure to that for the GPS survey, selecting addresses, and sending out the full package of materials, consisting of the introductory letter, the consent form, the subject information sheet, the household, person, and vehicle form, and a sample of the odometer form. Households were again asked to post back the signed consent form, and indicate the method by which they would like to provide the household, person, and vehicle information. As with the GPS survey they could choose to return their information via

post or telephone, and, in addition, they could submit their information using an internet version of the Odometer survey. Each household was provided with a unique identification number and password to log on to the survey website. This same method would also be used to return the household odometer readings. Households that did not respond to the initial mailing, and for whom we could find a telephone number were called to ask them to complete the survey. Table 3 shows the final outcomes of the recruitment process for wave 1 of the pilot odometer survey.

Table 3: Sample Disposition for the Odometer Pilot Survey

	Count
Total Sample	230
Known Ineligible no vehicles & Return-to-sender	29
Known Eligible Contacts	76
Eligibility Rate [#]	72%
Unknown disposition	125
Potentially Eligible Sample^{##}	166
Refusals	41* (25%)**
Recruits	35 (15%)**

[#] Known Eligible/ (Known Eligible+Known Ineligible)

^{##} Known Eligible + Unknown*Eligibility Rate

*Includes 10 households who promised to return forms upon telephone contact, but who failed to do so.

**Percentage of potentially eligible sample

Of the eligible households we were able to contact, 46 percent were recruited to the survey. While the total number of recruits (35) falls short of the target (60), this can largely be attributed to the low productivity of non-contactable households.

After receiving consent forms and the completed household, person, and vehicle information, households were then sent an odometer form for each car indicated in the vehicle information. They were asked to fill in the data and return the information by their chosen method. Households that were late in returning their odometer readings were sent a reminder postcard after a week and subsequently called to remind them to return materials. Table 4 indicates the distribution of the return methods chosen and the rates of return of odometer readings by households.

Table 4: Returned Odometer Readings by Retrieval Method - Pilot Wave 1

Retrieval Method	Post	Telephone	Internet	Total
Recruited	27	3	5	35
Returned Odometer Readings	25 (93%)	2(67%)	5 (100%)	32 (89%)

Wave 2 of the odometer pilot has also been completed. Households received a telephone call to confirm they are still resident at the same address, to confirm vehicle ownership and to remind them that the next package was about to be sent. Thirty two of the 35 pilot households were contacted and thirty confirmed they would continue. Continuing households were sent a household, person and vehicle information form merged with the data they provided during wave 1 and odometer postcards for each vehicle they currently have. If vehicles were sold or acquired between waves, respondents were asked to provide the odometer reading at the point of sale or acquisition in addition to the current readings. Twenty eight of the households completed odometer readings for the second wave. All of the losses of sample in the second wave are from households that originally chose post as their means to return data to us. Those using telephone and internet complied with the request for the second wave, although one telephone household did not provide odometer readings in the end.

At the time of writing, the main survey is underway, with a goal of adding another 970 households to the thirty pilot households, for a total panel of 1,000 households. For the main

survey, we initially used the same method as for the pilot of a mailing to households for recruitment, but then switched to using the second method of contact, which is to phone match households that are drawn from an address-based random sampling, then recruit households by phone, and send out materials to those households that agree in the phone contact to undertake the survey. From the first method, we recruited 672 households of which 619 have completed odometer readings. This was from an initial mailing of 4,000 households. For the second method, we have so far recruited 446 households from a total of 1763 households that were sent pre-notification letters, and have received odometer readings to date from 212 of those. This part of the survey is still underway at this time. However, we have achieved 831 of the goal of 970 households at this time.

4 Conclusions

At this point, with refinement of the recruiting methods, it appears that both the odometer survey and the GPS survey hold considerable promise as an effective method to evaluate travel behaviour change policies. The response rates to both the odometer and GPS surveys are in the same range as the response rates to more conventional diary surveys, where recruiting is done by post or phone. Completion of the odometer survey is significantly better than for diary surveys, and the GPS also appears to be successful in collecting a full week's worth of data from those who were asked to undertake this task. While further analysis of the results, including the second waves of both surveys, are needed for a full determination of the effectiveness of the methods, indications, at this point, are that the surveys are effective.

In addition, while recruitment for a panel has clearly had a slight depressing effect on response rates, the overall response rates for recruitment to a three-year panel are encouraging. Given the gains in accuracy in measuring change that attend the use of a panel, we anticipate that the longer application of these methods will be successful.

5 References

- Armoogum, J., B. Chlond, J. Madre and D. Zumkeller (2004 in press). "Panel Surveys" – A Resource Paper for the 7th ISCTSC, Costa Rica, August 2004.
- Axhausen, K., Zimmerman, A., Schonfelder, S., Rindsfuser, G. and Haupt, T. (2000). "Observing the Rhythms of Daily Life: A Six Week Travel Diary" *Arbeitsbericht Verkehrs- und Raumplanung* 25, March 2000. Retrieved on 21st June, 2005, from: <http://www.ivt.ethz.ch/vpl/publications/reports/ab25.pdf>
- Kish, L. (1965). *Survey Sampling*, John Wiley and Sons, Inc., New York.
- Pearson, D. (2005). "Household Survey Trip Under Reporting: A Case Study in Laredo, Texas", paper presented at the 84th Annual Meeting of the Transportation Research Board, Washington, DC, January.
- Purvis, C. (1997). "The San Fransisco Bay area household panel survey: A response to clean air and mobility initiatives", in *Panels for Transportation Planning: Methods and Applications*, T.F. Golob, R. Kitamura, and L. Long (editors), Kluwer Academic Publishers, Boston, pp193-203.
- Richardson, A. J., Seethaler, R. and Harbutt, P. (2003). "Design Issues for Before and After Surveys of Travel Behaviour Change", *ATRF(26)*, CD-Rom.
- Stopher, P. and R. Alsnih (2005). "Evaluating Voluntary Travel Behaviour Change Strategies – A Case Study", paper presented to the 84th Annual Meeting of the Transportation Research Board, Washington, DC, January.

Stopher, P. and Greaves, S. (2004) “Sample Size Requirements for Measuring a Change in Behaviour”, paper presented at the 27th Australasian Transport Research Forum, Adelaide, September.

Stopher, P R, Greaves, S P, Xu, M, and Lauer, N (2005a) Designing a Procedure to Undertake Long Term Evaluation of the Effects of TravelSmart Interventions, Paper to be presented to the 28th Australasian Transport Research Forum, Sydney NSW, September.

Stopher, P., Xu, M. and Lauer, N. (2005b). Results of a Pilot Application of TravelSmart in New South Wales – A Panel Evaluation, Paper to be presented to the 28th Australasian Transport Research Forum, Sydney NSW, September.

Stopher, P., Xu, M., and FitzGerald, C., (2005c). Assessing the Accuracy of the Sydney Household Travel Survey with GPS, Paper to be presented to the 28th Australasian Transport Research Forum, Sydney NSW, September.

Stopher, P., S.P. Greaves, and C. FitzGerald (2005) “Developing and Deploying a New Wearable GPS Device for Transport Applications”, paper to be presented to the Second International Colloquium on the Behavioural Foundations of Integrated Land-Use and Transportation Models: Frameworks, Models, and Applications, Toronto, June.

Stopher, P., Q. Jiang, and C. FitzGerald (2005) “Processing GPS Data from Travel Surveys”, paper to be presented to the Second International Colloquium on the Behavioural Foundations of Integrated Land-Use and Transportation Models: Frameworks, Models, and Applications, Toronto, June.

Wolf, J., M. Loechl, M. Thompson and C. Arce (2003). “Trip Rate Analysis in GPS-Enhanced Personal Travel Surveys”, in *Transport Survey Quality and Innovation*, P. Stopher and P. Jones (editors), Pergamon Press, Oxford, pp. 483-498.

6 Acknowledgements

The authors would like to acknowledge the funding provided for this work by the Government of South Australia, Department of Transport, Environment, and Infrastructure. Any errors in this paper are solely the responsibility of the authors and do not reflect a policy position of the South Australian Government.