A 6-Wave Odometer Panel for the Evaluation of Voluntary Travel Behaviour Change Programs

Peter. Stopher[1] and Natalie. Swann[1]
[1]University of Sydney, Sydney, NSW, Australia

1 Introduction

This paper reviews a 6-wave panel survey of household vehicle use in Adelaide, South Australia, and Melbourne, Victoria, which used odometer readings as the key form of measurement. The study was conducted by the Institute of Transport and Logistics Studies (ITLS) for the National Travel Behaviour Change Project (NTBCP). The study was conducted to explore the feasibility of using such a method to evaluate the effectiveness of Voluntary Travel Behaviour Change (VTBC) projects implemented at the household level. This paper reports the significant contributions of this pilot study to the travel research community, including:

- Reviewing the rationale for selecting this particular form of measurement (Section 2);
- Describing the benefits and challenges of implementing this methodology (Section 3);
- Evaluating the willingness of households to voluntarily participate in the study (Section 4.1);
- Investigating the demographic characteristics of responding households and concomitant potential for bias (Section 4.2); and
- Analysing the ability of the method to measure real changes in household vehicle use (Section 5).

Odometer Surveys are limited in their usefulness – they only provide data about vehicle kilometres travelled (VKT) in household vehicles and fail to provide information on implications of changing VKT such as mode-switching or trip-chaining. We demonstrate in this paper, however, that for the purpose of tracking household vehicle use over time Odometer Panel Surveys can be low-burden, highly accurate, and representative of the population.

2 Background to Implementing a Panel Survey of Odometer Readings

The survey described in this paper is designed to evaluate the impact of a VTBC Program. VTBC is a term used to cover a variety of travel demand management approaches which "seek to find the means for individuals and households to change their travel behaviour – adopting approaches where individuals choose their own method of changing travel behaviour rather than simply acting in response to external policies or pressures" (Taylor and Ampt, 2003, p165). That is, the public is provided with tools to equip them to change their travel behaviour voluntarily. In Australia, there has been considerable interest in VTBC (Red3, 2005).

The present study was conducted for the NTBCP, a joint initiative of the Commonwealth and participating jurisdictions aimed at reducing the impact of car use on greenhouse gas emissions (AGO 2007), which "fills a gap in transport emission abatement efforts for urban areas, where increasing passenger vehicle use is driving strong greenhouse emission growth" (Kemp and Anderson 2003). The study is a pilot to test survey instruments and methodological approach for a large-scale national evaluation of VTBC implementations over a period of five years starting in 2008. Because the policy issue is that of reducing greenhouse gas emissions, the primary measure required from the evaluation is aggregate change in VKT, which is then related by an agreed formula to greenhouse gas emissions. The odometer survey described in this paper has been simultaneously conducted in the

western suburbs of Adelaide – in the Local Government Areas (LGAs) of Charles Sturt, Port Adelaide and Holdfast Bay – and in the inner Northern Melbourne locality of Darebin. These areas were chosen because they were being provided currently with VTBC tools (Adelaide) or because a VTBC program had already been conducted recently in the area (Darebin).

This survey's primary aim is to quantify the change in the total VKT by household in the wake of the roll-out of a VTBC program to measure the sustainability of the program's outcomes. Evaluation of VTBC initiatives has consistently been identified as somewhat problematic (Ker, 2002; Taylor and Ampt, 2003; Ampt, 2001). The challenge for evaluators is to identify the occurrence of travel behaviour change, quantify it and describe its character. Odometer surveys conducted with a longitudinal panel have been recommended (Stopher et al., 2005) as a potentially valuable tool for fulfilling at least part of these requirements. They offer the opportunity to measure the trends in car use among households participating in VTBC initiatives, compared to a control of those not participating.

The Odometer Survey methodology described below was developed for the following reasons:

1. It involves very little burden for the respondent;

2. It collects only the information needed to measure changes in VKT; and

3. It is not subject to error in trip reporting and also provides little opportunity for respondents to give a "politically correct" response as opposed to a true response.

The methodology then, provides a basic measure for evaluating whether VTBC implementations have had a measurable impact on VKT at the household level. At the time this research was initiated, very little was known about the variance within and between households in VKT. As such, it was impossible to estimate the required sample size to measure change in household VKT accurately (for a more detailed discussion of this topic see Stopher *et al.,* 2005). While the cost of conducting the survey is very low on a per household basis, it was necessary to determine the sample size needed to assess whether there was significant change.

3 Methodology: Benefits and Challenges

The two areas chosen for study were selected for their exposure to VTBC initiatives. The South Australian sample was randomly selected by household address from all known household addresses in the area. The Victorian sample, however, was taken from a list of all residential addresses in the area that had not previously participated in the short-term VTBC initiative in the year previous. While this was practically necessary to avoid repeatedly contacting the same households, it led to some demographic bias in the sample.

Sampled household addresses were processed to match listed telephone numbers. Households with known telephone numbers were then posted a preliminary notification letter about the study from the sponsoring government agency. Shortly after, households were phoned by a market research company on behalf of Sydney University. Necessarily for an odometer study, households had to own at least one motorised vehicle to be included - households with no vehicles were screened out at recruitment. If recruitment was successful, details of the cars belonging to the household were collected over the phone so that the survey package could be customised for the household. Recruited households were given the choice of which method they wished to use to return their odometer readings: post, telephone, or Internet.

Because the respondent is required to leave the telephone to collect their odometer readings, and particularly because not all vehicles are necessarily at the residence at the time of the call, the survey task could not be completed in the recruitment phone-call. All

households, irrespective of the way they wished to return the data, were posted hard copy forms. This provided the tools to collect the odometer readings from each vehicle and ensured that both telephone and internet responders knew the full scope of the questions they would be asked prior to the retrieval call or logging on to the website. That is, respondents were not surprised by any additional data requested on-line that they were not already aware they would be asked to complete prior to logging onto the website.

In the first wave, households were asked to complete a household and vehicle information form as well as customised cards for each vehicle's odometer reading. If households were slow to return their data, they were reminded at least once by telephone and once by post.

Three months later, households were recontacted with a telephone call to check that their contact details were still the same, to confirm the vehicles that they owned, and to ensure that they were willing to participate again. In subsequent waves, the forms were delivered to households with the details of the previous wave pre-printed so that they only needed to correct errors or update changes. In each subsequent wave some additional random replacement recruitment has been conducted to make up for households that drop out or that cannot be contacted. The process of completion and reminders was the same as for the initial wave. Table 1 illustrates the timing of the survey waves and highlights some of the obstacles to response.

Table 1: The timing of the survey waves

	2005				2006											
	Sept	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec
South Australia	W1			W2 - Xmas			W3			W4			W5			W6 - Xmas
Victoria	W1			W2 - Xmas			W3 - C'wealth Games			W4			W5 - School Holidays			W6 - Xmas

This method had particular benefits and challenges that are briefly introduced below.

3.1 Benefits

Using a panel design is one of the key strengths of this methodology. In each wave of the survey we only collect a single set of readings from each household. By collecting readings over multiple waves we are thereby able to calculate the average daily VKT over a considerable period (3 months in this case). By calculating average daily VKT over such a long period much of the daily and weekly variation is accounted for and, because the variation is reduced, we can more readily identify small, long-term trends. If a household misses a wave – e.g. because they were on holidays – but continues in the survey after their absence, the panel design means we are still able to calculate average daily VKT between more distant waves. Furthermore, because we are repeatedly measuring the same households, we are able to eliminate the variance of introducing new households all the time, and to take advantage of the fact that the measurements in successive waves are correlated and reduce the variance still further.

The task is very easy; it is easy to describe and understand, and the burden placed on respondents is therefore low. We believe that this has influenced the excellent continuation rates and the very clean data set. The task of recording all the odometer readings from household vehicles is very difficult to manipulate – it is difficult to report to the researcher what you think they 'want' to hear. It has been suggested that the medium (e.g., Internet versus pen and paper surveys) can affect the content of the response made by participants

(e.g., Hawthorne 2003; Huang 2006). However, because the odometer data is rather technical, we believe that response mode (post, telephone or internet) can have little effect on the nature of the response. Instead, being able to offer multiple response modes has increased the opportunity for households to manage their own participation.

The simplicity of the task also means that it is easier to quantify VKT at a household level even if only one household member is willing to participate. Collecting the odometer readings from other householders' vehicles is much less burdensome than debriefing reluctant household members on a whole day's travel for a diary, or even than ensuring reluctant household members use a Global Positioning System (GPS) device.

3.2 Challenges

One of the challenges posed for all public household surveys in the current political and cultural setting is obtaining a sample that is unaffected by coverage error; that is, the exclusion of distinct population subsets as a result of sample frame design (Groves 1987). For example, using the telephone to conduct recruitment systematically excludes those without landlines or listed numbers. Unless the population of non-landline owning households are demographically randomly distributed, this is likely to lead to the systematic exclusion of particular types of households. However, sampling by post, rather than phone proved to be prohibitively expensive and unproductive in a parallel survey of 1000 households being conducted for the South Australian State government's short term evaluation. Perhaps, with the activation of the national Do Not Call register, government agencies may be able to obtain more recent listings of addresses matched to telephone listings. The fact that this survey was conducted under the auspices of a government agency by a University may have contributed to its respectability and encouraged households to participate.

While the task is easy, it is repeated very often and could easily become a burden. We have sought to avoid this potential problem by developing Visual Basic programs that build customised versions of the forms used to collect demographic and odometer data. When a household is recruited, they are asked to complete the household and vehicle information forms in full. In subsequent waves, however, their forms are provided pre-printed with the information they have previously supplied and with specific prompts to collect data on new persons or vehicles, fill in missing data and identify changes to persons between waves (e.g. acquisition of a drivers licence). Households are asked in subsequent waves to review the forms and confirm that everything is the same or note down any changes since the last time they participated. This reduces the burden dramatically and, by continuing to feed the information we have on file back to respondents, can actually contribute to cleaning the data set over progressive waves. Because households are shown the data the researchers recorded in previous waves they are able to both complete items they had previously forgotten and identify any data entry errors.

Any longitudinal panel survey will face the problem of attrition, whereby persons or households are no longer available to participate. This has been the case in the odometer survey as people choose not to continue, sell their vehicles, move out of the evaluation zone or simply cannot be reached in subsequent waves. Conducting replacement recruitment supplemented the sample in early waves. Frequent contact with respondents has also enabled us to capture many changes of address within the evaluation zone.

A final and non-trivial challenge is capturing the final and initial readings of vehicles that are sold or acquired between waves. Households are universally provided with a specific form for retrieving these readings when they are recruited. In each subsequent wave, when we call to confirm their continued participation we check to see if there has been any vehicle transition and specifically provide extra forms to these households. If households report a sale or acquisition of vehicles but fail to report their reading, they are followed up either by

phone at the time of data collection, or with another customised form in the next data collection wave. We ask householders to specify whether they are providing exact readings or estimates of the reading. Of course, some household members simply do not know the reading or date of vehicle sale or acquisition, but we have had a high success rate (approx. 80%) in retrieving data about vehicle turn over.

4 Response and Demography: Results and Implications

This section outlines the results of the pilot odometer panel survey in terms of response and demography. The response rates and levels of continuation indicate that this method is feasible in both a medium and a large urban region of Australia. The demographic profiles suggest that even if the survey suffers from coverage error, it succeeds in maintaining a stable demographic profile over time.

4.1 Response Rates

This pilot survey aimed to obtain data from 200 households in each of the South Australian and Victorian evaluation zones. This section describes the recruitment required to obtain this level of returns. In addition, it describes the rates of continuation into later waves and the success of replacement recruitment conducted to maintain the 200 household sample.

In the first wave, 319 households in South Australia and 380 households in Victoria were recruited (see Table 2). Table 2 shows that the initial recruitment was about 25 percent in South Australia and just below 18 percent in Victoria, which is not surprising given the nature of the sample drawn in Victoria as explained at the beginning of section 3. Of those households recruited in Wave 1, 64% of South Australian households returned their forms and 62% of Victorian households returned data. The Victorian sample contained a large number of invalid numbers and unknown disposition households. We have not had the means to investigate why it was not possible to contact such a large number of households in the Victorian sample, although we are aware that the current telephone listings available for market research are long out of date. For those households were no contact was made – i.e. where the eligibility status was unknown – we have estimated the proportion of those numbers that would have been eligible based on the proportions of ineligibility in the households of known disposition.

The process of recruitment is described in Section 2, but it is important to remember that households were notified at the outset of the length and frequency of the survey. The burden is low in any one wave, but the commitment to continue in the survey for nearly two years is a deterrent to participation. The response rate, whilst low, is comparable to the rate that is achieved by most diary surveys today (e.g., Stopher *et al.,* 2004), other than those conducted by face-to-face interviewing. It would definitely be desirable to increase this response rate, but if the sample is representative, this level of recruitment may be acceptable. We consider the demography of the sample in section 4.2.

Table 3 summarises recruitment and continuation across all six waves. In subsequent waves, the recruitment rate for new households varied between 13 and almost 20 percent, but was almost certainly affected by the season. Compare, for example the response rates with the wave schedule in Table 1 – the waves affected by major events such as Christmas or the Commonwealth games appear to dip slightly. On the other hand, the retention rate of continuing households was generally close to 90 percent. The continuation rate describes all those households that had previously participated in the survey in each wave. In each subsequent wave, there are households that specify they no longer wish to participate, households whose contact details are no longer valid and households that we simply fail to reach after 10 call attempts. Those households that still have valid contact details but which we do not reach are still included in the eligible sample for the next wave.

Table 2: Recruitment Rate for the Odometer Panels in Wave 1

	South Australia	*Victoria*
Sample	1242	2270
Ineligible*	431	825
Known Refused	490	766
No Contact	2	299
Recruited	319	380
*Recruitment Rate***	*39.4%*	*28.8%*
Returned Data	205	238
*Return Rate****	*64.3%*	*62.1%*
Response Rate#	**25.3%**	**17.6%**

* Ineligible households are those with no vehicle, with a mismatched/invalid phone number
** Recruitment Rate is Recruited/(Known Eligible + Estimated Eligible). Known Eligible are those numbers contacted and determined to be eligible, while estimated eligible are those numbers where no contact is made, and an eligibility rate is applied to estimate the number that could be assumed to be eligible. We used the same rate as in the known sample – 65% in SA and 58% in Victoria - which makes our estimates more conservative than is probably necessary.
*** Return Rate = Returned Data/Recruited
# Response Rate = Returned/(Known Eligible + Estimated Eligible)

In parallel to these quarterly pilot surveys, we are running a tri-annual Odometer Study of 1000 households in the same area of Adelaide which is continuing through 2007. When we made contact with the South Australian households in Wave 6 we asked them whether they would be willing to voluntarily continue on into the parallel study and report their odometer readings on three more occasions. Of the remaining eligible households, an extraordinary 187 of them (81%) agreed to this request.

Table 3: Percentage Response Rate for the Odometer Panels for All Waves

		Returned/(Known Eligible + Estimated Eligible)					
		Wave 1	*Wave 2*	*Wave 3*	*Wave 4*	*Wave 5*	*Wave 6*
SA	New Recruit Households	25.3	15.5	18.9	NA	NA	NA
	Continuing Households	NA	89.8	88.1	90.6	90.1	89.8
Vic	New Recruit Households	17.6	19.5	NA	12.7	NA	NA
	Continuing Households	NA	88.3	87.4	88.5	88.3	89.3

4.2 Demography of the Participating Households

Due to the high levels of continuation in the survey the demographics of the sample are remarkably stable across each of the waves. Tables 4 & 5 illustrate the stability of the sample on variables of key significance to travel surveys. The mean displayed is calculated across those households that actually returned data in each of the given waves. While the magnitude of the mean is different in the two different areas, the changes between waves within each State are extremely small. Historically, travel surveys struggle to recruit either very large or very small households (e.g., Ampt and Stopher 2005) but, as Tables 4 and 5 demonstrate, on most variables the sample distribution of households is similar to that of the census.

Table 4: Comparison of the Demographics for the Six Odometer Waves in South Australia with 2001 Census Data for All Households*

Demographic (per household)	Value	South Australia Statistics						
		2001 Census All Households	Wave 1	Wave 2	Wave 3	Wave 4	Wave 5	Wave 6
Number of Persons	1	33.70%	21.60%	20.50%	22.60%	23.30%	20.40%	21.00%
	2	34.20%	42.20%	41.90%	41.50%	41.00%	44.40%	43.40%
	3	14.00%	15.20%	15.30%	14.10%	13.20%	12.50%	11.70%
	4	12.10%	15.70%	15.30%	16.70%	16.30%	16.70%	17.60%
	5+	6.10%	5.40%	7.00%	5.10%	6.20%	6.10%	6.4%
Number of Vehicles	0	15.10%	0%	0%	0%	0%	0%	0%
	1	44.10%	50.70%	50.70%	48.50%	47.80%	46.80%	49.30%
	2	30.50%	38.00%	37.30%	38.70%	41.20%	42.10%	39.50%
	3+	10.20%	11.20%	12.00%	12.80%	10.90%	11.10%	11.20%
Average Number of Adults		1.9	1.95	1.97	2.00	1.97	2.00	1.98
Proportion of Population Adults		80.30%	80.24%	79.40%	82.51%	81.13%	81.32%	80.04%
Average Number of Children		0.47	0.48	0.51	0.42	0.46	0.46	0.49
Proportion of Population Children		19.70%	19.76%	20.60%	17.49%	18.87%	18.68%	19.96%
Average Number of Males		1.15	1.15	1.20	1.20	1.19	1.21	1.20
Average Number of Females		1.22	1.28	1.29	1.22	1.24	1.24	1.27
Average Number of Full-Time Workers		0.62	0.79	0.69	0.71	0.68	0.68	0.66
Average Number of Full-Time Students		0.4	0.45	0.45	0.42	0.46	0.43	0.46
Average Number of University Educated Persons		0.41	0.47	0.45	0.45	0.45	0.46	0.46

* The South Australia census statistics are obtained by aggregating the Western Adelaide Statistical Subdivision (SSD 40510) with the Statistical Local Areas of Holdfast Bay North (SLA 405202601) and Holdfast Bay South (SLA 405202604) to approximate the evaluation zone.

Note that the incidence of 1 person households is lower in the sample than in the census; this is due to the ineligibility of households without a vehicle for an odometer survey. Single person households are more likely to have no cars than larger households. Therefore direct comparisons between the demography of the sample and that of the census are somewhat problematic. It is hoped that future comparisons can be made using unit record files so that demographic analyses can be limited to households owning vehicles.

Questions relating to income are often the most offensive to respondents and therefore most threatening to recruitment and continuation. Furthermore, it is usually a very poorly completed item on surveys, with much ambiguity about drawing distinctions between gross and net income and income from investments. It was deemed to be unnecessary to the evaluation of VKT to ask specifically about income. Instead, proxies such as highest level of education attained by each householder were asked and the average number of University educated persons per household also shows stability and is close to the census average. While there may still be error associated with overall representation in the sample of high or low income earners, it does not appear that either group is particularly predisposed to terminating their involvement in the survey.

Previously, we have shown that the demographic profiles of households choosing different response methods (post, telephone or Internet) are different (Swann and Stopher 2006). In particular, households choosing the Internet were shown to be more likely to be large households with children, higher rates of full-time employment and multiple vehicles – i.e., those households that are typically difficult to recruit to household travel surveys. It may be

possible that offering choice in response mode has enabled people with multiple responsibilities to manage their own participation.

Table 5: Comparison of the Demographics for the Six Odometer Waves in Victoria with 2001 Census Data for All Households*

Demographic (per household)	Value	Victoria Statistics						
		2001 Census All Households	Wave 1	Wave 2	Wave 3	Wave 4	Wave 5	Wave 6
Number of Persons	1	33.40%	18.9%	18.7%	18.5%	17.2%	19.5%	19.9%
	2	32.80%	28.6%	27.8%	29.3%	31.6%	29.3%	27.4%
	3	16.30%	15.1%	16.3%	14.1%	14.4%	13.7%	12.9%
	4	11.70%	24.8%	25.4%	24.4%	23.7%	23.4%	26.4%
	5+	5.90%	12.6%	12.0%	13.6%	13.0%	14.1%	13.4%
Number of Vehicles	0	17.90%	0.0%	0.0%	0.0%	0.0%	0%	0%
	1	45.80%	53.3%	55.2%	56.3%	56.9%	56.3%	57.2%
	2	28.40%	38.8%	38.1%	35.9%	34.7%	33.5%	32.8%
	3+	7.90%	7.9%	6.7%	7.5%	8.3%	10.2%	10.0%
Average Number of Adults		1.97	2.00	1.99	2.04	2.02	2.01	2.01
Proportion of Population Adults		82.60%	70.10%	69.28%	70.90%	70.45%	70.31%	69.66%
Average Number of Children		0.41	0.85	0.88	0.84	0.85	0.85	0.88
Proportion of Population Children		17.40%	29.90%	30.72%	29.10%	29.55%	29.69%	30.34%
Average Number of Males		1.15	1.30	1.33	1.35	1.36	1.32	1.35
Average Number of Females		1.23	1.55	1.54	1.54	1.50	1.54	1.54
Average Number of Full-Time Workers		0.77	0.97	1.01	0.99	1.00	1.00	0.97
Average Number of Full-Time Students		0.39	0.70	0.69	0.72	0.74	0.74	0.73
Average number of University Educated Persons		0.83	1.06	1.08	1.13	1.08	1.14	1.13

* Victorian statistics have been calculated from the Darebin-Northcote Statistical Local Area (SLA 205301891)

4.3 Implications

This section has demonstrated that it is possible to recruit urban Australians into a panel study of household vehicle use. While the response rates are not exceptional when compared to other travel surveys, in the context of this particular methodology, in which people are made aware of the commitment at the outset, they are better than expected. Furthermore, the continuation rates are excellent. We believe this is a reflection of the efforts made to reduce burden – once households realise how easy the survey is to complete they are quite happy to continue.

The sample may have some coverage error. When the 2006 census is released, it is hoped that more detailed statistical analysis can be conducted. Particularly, it will be of interest to limit the census statistics to car owning households for any such comparison. However, the demography of the sample across waves is stable, which indicates that there is no systematic loss of particular subsets of the population. This is good news, because it means that if we do discover coverage error, we could either develop weights that can be applied to the data set during analysis or implement a stratified sampling strategy to ensure we capture the desired sample profile without great risk of losing particular households over the life of the survey.

5 Analysis of Vehicle Use: Results and Implications

Prior to going into the field with this odometer survey, there was concern over whether the public would successfully differentiate the odometer from the trip meter. As a result we carefully make note on each odometer card: “Please make sure you record the Odometer reading and NOT The Trip Meter Reading. The Odometer shows the total number of kilometres that the car has travelled since new”. While we have not calculated the percentage, most respondents completed the task accurately; the readings from the vast majority of vehicles are in sequence across the data collection waves. This bodes well for future implementations of similar surveys. The following section details the analysis of VKT conducted using the odometer readings across all six waves of the study.

5.1 Measuring Change in Vehicle Kilometres Travelled (VKT)

The essential data collected in this survey consist of the Odometer reading and date of reading from household vehicles. The principle measure of interest is average daily VKT, calculated at the vehicle level. (We have used the vehicle level for this analysis, because this is easiest to understand, rather than using household averages that encompass varying numbers of vehicles per household. The data can be aggregated subsequently to the household level, to determine changes at the household level.) As such, it does not matter if the readings from the vehicles in the sample – or even within a household – were taken on different days, so long as the readings are in fact dated when they are noted down by the respondent. Because only one set of readings is collected in each wave, it requires two waves to obtain a value of kilometres driven for a period. Therefore, from six waves, we can estimate an average number of kilometres driven between each pair of waves, i.e., waves 1 and 2, 2 and 3, 3 and 4, 4 and 5, and 5 and 6 and also waves 1 and 3, 1 and 4, 1 and 5, 1 and 6, waves 2 and 4, 2 and 5, 2 and 6, waves 3 and 5, and 3 and 6, and waves 4 and 6. For the purpose of this paper it is enough to report on the pairs of consecutive waves, but this highlights the usefulness of the panel design. The sample sizes for each of these will vary, because only those who provided dated odometer readings in both waves 1 and 2 can be used to provide an estimate for the average between wave 1 and wave 2, and similarly for each of the other values. It is useful first to determine the average number of days that elapsed between each of the waves.

The schedule for conducting the survey indicates that these households would be surveyed every three months. In reality, there was some variation in the length between waves. The first wave of these two panels was conducted over a longer period of time than had originally been intended. It lasted from late July through October 2005 (elsewhere, we have referred to this simply as September, for simplicity); due to the discovery of such a high proportion of ineligible telephone numbers in Victoria, additional recruitment had to be conducted and extended the overall length of the wave. Because of these start-up delays, an average of 113 days elapsed between the first two odometer readings in South Australia and an average of 100 days elapsed for Victoria. Waves 2- 5 rolled out roughly on schedule but in order to complete the survey before running into Christmas 2006, there were 68 days between waves 5 and 6 for South Australia and 66 days for Victoria. In addition to the reduced timeframe, we also found households became better trained at completing the survey promptly in later waves and the speed with which households responded in the sixth wave was shorter than previously experienced. These statistics are summarised in Figure 1, which demonstrates the fluctuation of the mean number of elapsed days around the 90 day mark.

Figure 1 also illustrates that the standard deviation in the number of days is generally quite small, except for wave 1 to 2, with values around 6 to 12 for successive waves. This shows that it generally took about two to four weeks to get all returns back for each subsequent wave. The missing/refused rate for South Australia is substantially better than that for

Victoria, with the former running from 20 to 35 percent while the latter run from 26 to 42 percent.

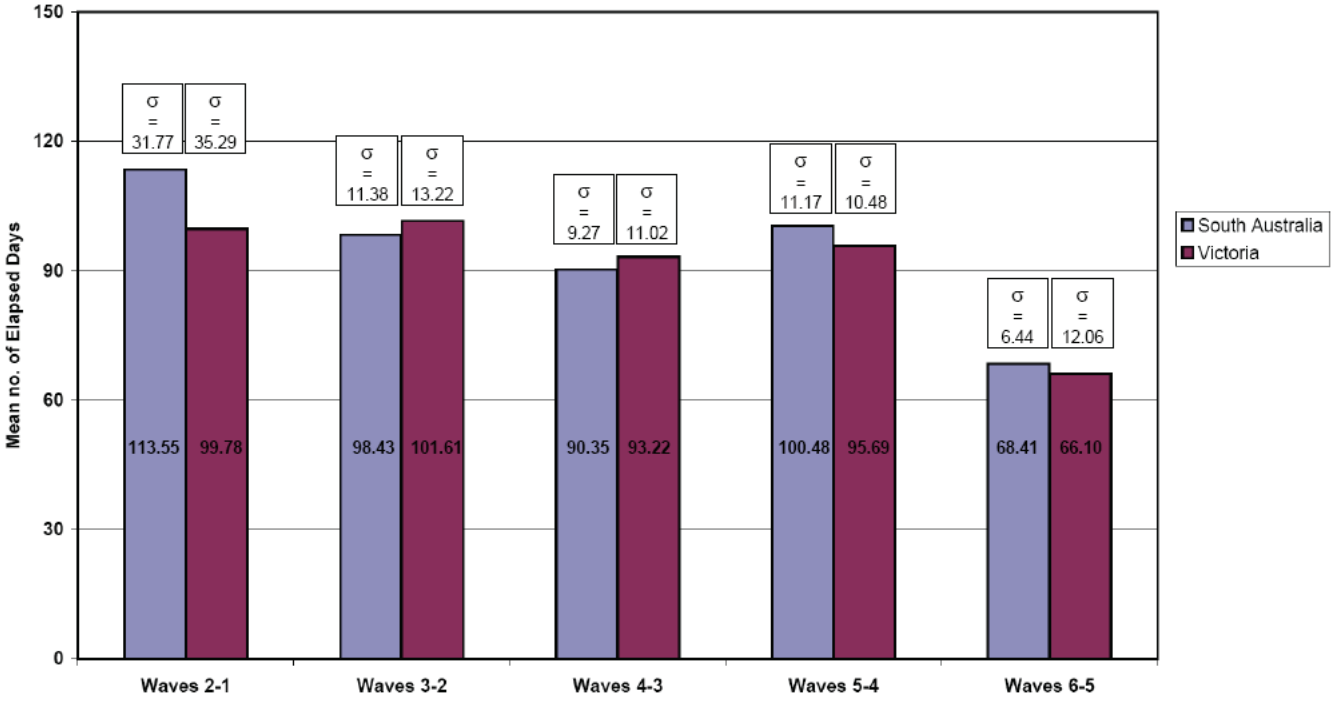

Figure 1: Mean number of elapsed days between waves

The next thing that is important to look at, given the purposes of the odometer survey, is the average number of vehicle kilometres per day. Figure 2 summarises the results for South Australia and Victoria, respectively. It can be seen that both panels produce very similar results, with the average VKT per vehicle being around 30 per day.

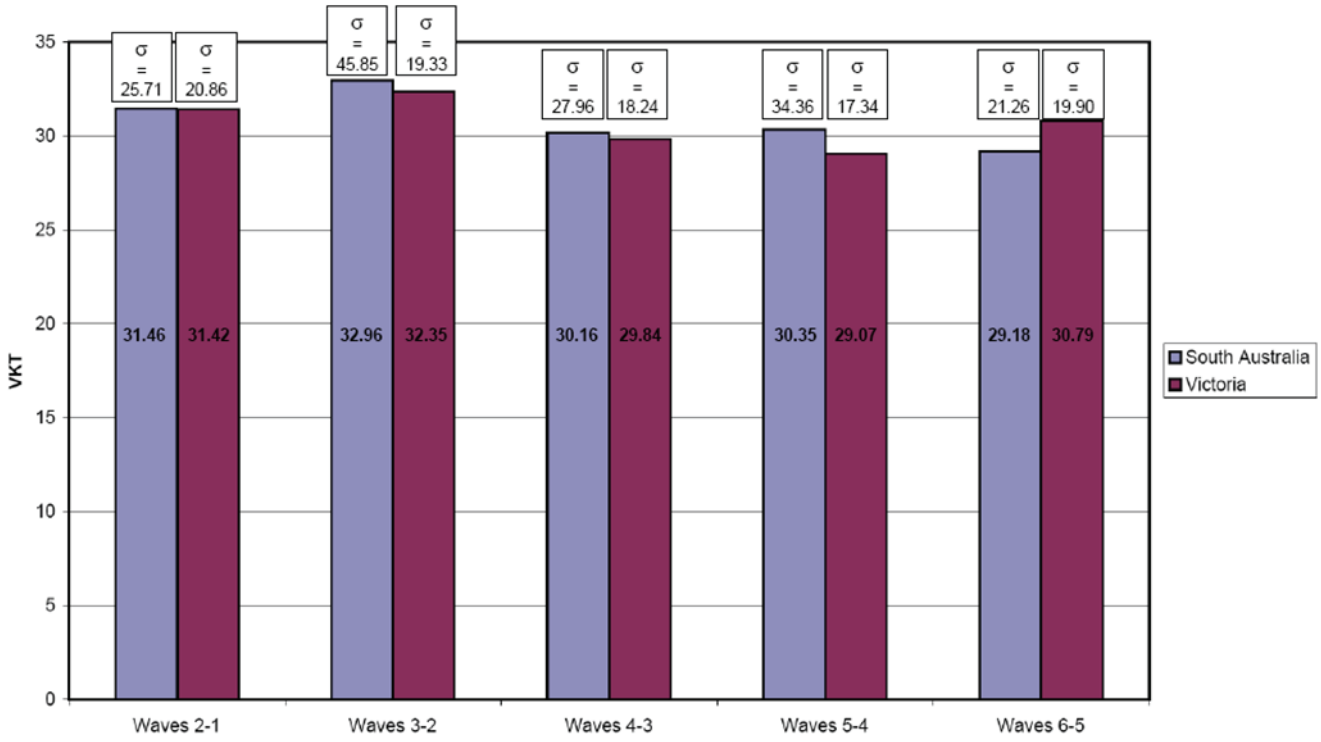

Figure 2: Average daily VKT for each pair of consecutive waves

It is of most importance to note that the difference between waves 1 and 2 represents average VKT per day for the period from September to December in 2005, whilst the difference between waves 5 and 6 represents the same period in 2006. These two pairs of readings are therefore of most interest, because they should be free of seasonal differences. We can note that between waves 1 and 2 in South Australia, the average VKT was 31.5, whilst between waves 5 and 6, it had dropped to 29.2. In Victoria, there is also a drop over the year, with waves 1 and 2 having a value of 31.4 and waves 5 and 6 having a value of 30.8, but the decrease is much less marked. The standard deviations are higher in South Australia than Victoria, with the former generally being around 26, and the latter around 17. In Victoria, as mentioned earlier, none of the sample are participants in a VTBC program. In the South Australia sample, however, no households were participating in wave 1, 10 were by wave 2, while in wave 5 there were 91 participants and in wave 6 there were 113. It could be surmised that the drop in average VKT from 31.5 to 29.2 in South Australia is evidence of the effect of VTBC program participation, while the drop in Victoria from 31.4 to 30.8 is perhaps due to other outside factors, such as fuel prices. What is of most importance here is to understand whether the drop of 0.6 kilometres per day in Victoria is significantly different from zero, and also whether it is significantly different from the drop of 2.3 kilometres per day in South Australia. This is explored further below.

We have also estimated the covariances between all pairs of waves. These are shown in Tables 6 and 7. In these tables, the diagonal is the variance, and the off-diagonal elements are the covariances between pairs of waves. The tables are symmetrical, so only the above-diagonal elements are shown, to make the tables easier to comprehend.

Table 6: Variance-Covariance Matrix for the South Australia Panel Average Daily VKT

Wave Pair	*1 and 2*	*2 and 3*	*3 and 4*	*4 and 5*	*5 and 6*
1 and 2	661	898	590	313	325
2 and 3		2102	1230	577	301
3 and 4			782	571	325
4 and 5				1181	378
5 and 6					452

Table 7: Variance-Covariance Matrix for the Victoria Panel Average Daily VKT

Wave Pair	*1 and 2*	*2 and 3*	*3 and 4*	*4 and 5*	*5 and 6*
1 and 2	435	307	301	227	329
2 and 3		374	261	243	289
3 and 4			333	233	293
4 and 5				301	244
5 and 6					396

From these tables, we can estimate the sampling error for the difference in VKT per vehicle between any two pairs of waves. For our purposes, however, the most important result to look at is the result from waves 1 and 2 versus waves 5 and 6. In South Australia, the variance of waves 2-1 is 661, for waves 6-5, it is 452, and the covariance between these two wave pairs is 325. Table 8 shows the results of sample size calculations according to the difference in VKT between two pairs of waves that would result if these variances and this covariance were to be typical of what might be found in a larger study. However, looking at Table 6, we see a worst case that is shown by comparing waves 1 to 2 against 4 to 5. In this case, the two variances are 661 and 1181, and the covariance is 313. These results (with the variances and covariances rounded) are also shown in Table 8. It can be seen that these

higher variances and lower covariance produce sample sizes that are about double those of the previous results.

Based on these estimates of variance and covariance, then a panel of 1,600 and 3,800 vehicles would be needed to have 95 percent confidence that an actual difference in daily VKT per vehicle of ±1 kilometre was significantly different from zero. (It should be recalled that the 95% confidence limits are 1.96 times the sampling error.) Any difference of less than this amount would not be able to be distinguished from no change at a 95 percent confidence level. Thus, if the samples for these pilot panels had in fact been around 1,600 to 3,800 vehicles, then the drop of 0.6 kilometres in Victoria would not be considered significantly different from zero with 95 percent confidence, whilst the drop of 2.3 kilometres in South Australia would be considered significant.

Table 8: Estimated Sample Sizes for a Quarterly Odometer Panel of Households

Max. Sampling Error (kms)	*Scenario 1 – Low Estimates of Variance*				*Scenario 2 – High Estimates of Variance*			
	Variance 1	*Variance 2*	*Covariance*	*Sample Size*	*Variance 1*	*Variance 2*	*Covariance*	*Sample Size*
±0.05	700	450	325	157,143	700	1200	300	371,429
±0.1	700	450	325	39,286	700	1200	300	92,857
±0.25	700	450	325	6,286	700	1200	300	14,857
±0.5	700	450	325	1,571	700	1200	300	3,714
±0.75	700	450	325	698	700	1200	300	1,651
±1.0	700	450	325	393	700	1200	300	929
±1.25	700	450	325	251	700	1200	300	594
±1.5	700	450	325	175	700	1200	300	413
±1.75	700	450	325	128	700	1200	300	303
±2.0	700	450	325	98	700	1200	300	232
±2.25	700	450	325	78	700	1200	300	183
±2.5	700	450	325	63	700	1200	300	149

To convert the sample sizes to households, we find that the average car ownership in South Australia among car-owning households is about 1.6 cars per household, so that the sample sizes in households would need to range from about 1,000 to 2,400. It should also be remembered that, on average, about 25 percent of households failed to provide an odometer reading in each pair of waves, so that the actual sample that would need to be drawn would be 1,250 to 3,000, so as to achieve an actual sample of 1,000 to 2,400 households providing valid measurements.

It is important to be clear, however, that in using the South Australia figures, we have a sample that is made up of approximately 50 percent participants and 50 percent non-participants. In this case, the comparison of what was found in Victoria with what was found in South Australia is representative of what might be found in a national study that sampled equal sample sizes between areas where VTBC has not been introduced and areas where it has, using random samples in both cases. An implication of this is that, to achieve a control sample that will allow detection of whether participating households are different from non-participating, the sample sizes discussed in this section need to be doubled. The reason for this is that a sample of, say, 1,000 households drawn randomly from an area where VTBC has been introduced will produce about 500 VTBC participants. If the desire is to look specifically at the behaviour of the participants versus the non-participants, then each of the participants and non-participants will have sampling errors based on a 500 sample size, i.e., between about ±1 and ±1.5 kms, with corresponding 95 percent confidence bounds of between ±2 and ±3 kms.

To follow this line of reasoning, we undertook a further more detailed analysis of the South Australia sample. In this analysis, we separated the VTBC participants from the non-

participants, and estimated the means, variances, and covariances separately for each of these two groups. This would be analogous to the situation where sampling was undertaken at random from communities where VTBC has been introduced, and where it is intended that the non-participating households represent the control group. In doing this, the results shown in Table 9 were found.

Table 9: Means, Standard Deviations and Covariances for Participants versus Non-Participants – South Australia Sample

Measure	*Mean*	*Standard Deviation*	*Sample Size*	*Covariance*	*Covariance Sample Size*
Participants					
VKT/vehicle /day (waves 1-2)	31.96	26.46	85	105.83	76
VKT/vehicle/day (waves 5-6)	28.40	14.87	99		
VKT/household/day (waves 1-2)	53.00	42.87	79	988.63	70
VKT/household/day (waves 5-6)	48.34	35.69	98		
Non-Participants					
VKT/vehicle /day (waves 1-2)	29.34	15.44	78	178.77	61
VKT/vehicle/day (waves 5-6)	29.21	16.50	90		
VKT/household/day (waves 1-2)	47.42	34.85	76	1015.92	59
VKT/household/day (waves 5-6)	47.95	32.38	90		

The figures presented here show that the average VKT at the household and the vehicle level decrease among households participating in VTBC but not among non-participating households. However, it is critical to determine whether this is significant or simply the result of chance. To do this, we need to make careful consideration of the amount of variation that could naturally be expected to occur.

The figures of standard deviation and covariance presented in Table 9, can be used to calculate the sampling error: for the non-participants the sampling error for household VKT is ±3.18 kms, whilst for the participants it is ±4.29 kms. This means that the *increase* of 0.53 kms in the non-participants' average VKT per household per day has a 95 percent confidence limit of ±6.23 kms, while the *decrease* in the participants' average VKT per household per day of 4.66 kms has a 95 percent confidence limit of ±8.40 kms. In other words, with these sample sizes, we cannot be statistically sure that there has in fact been any change at all in either group.

We then proceeded to estimate the sample sizes we *would have* needed to use to be 95 percent confident that the 4.66 kms *decrease* in the participants' VKT per household was significantly different from zero. We found that this would require that, instead of having 70 households that provided odometer readings in waves 1, 2, 5, and 6, we would have had to have 225. Because we found that, in a random sample of households from the area where VTBC was introduced in Adelaide, about 63 percent of the original sample of households for wave 1 provided odometer readings in each of the four critical waves (waves 1, 2, 5, and 6), and because we ended up with about 55 percent of households in this group that were VTBC participants, this sample size of 225 for the covariance estimation means that the original wave 1 sample would have had to be about 650 households. Such a sample would have produced about 185 non-participants that provided all four critical odometer readings, which would produce a sampling error for the non-participants of about ±2.5 kms, or 95 percent confidence bounds of ±4.9 kms. We would still not be sure that the *increase* in the VKT of the non-participating households was anything other than random variation in the data.

5.2 Implications

The odometer readings collected by this method provide a useful measure of household vehicle use. This section has demonstrated that, due to the variability of VKT, quite large samples are necessary to be able to report small changes with statistical significance. The panel design shows promise for confidently measuring change, but it would require large sample sizes, particularly if it is necessary to disaggregate the sample by household characteristics such as VTBC participation.

Odometer readings, however, will only ever provide limited information about household travel – even if we confidently demonstrate a positive effect from VTBC programs, an odometer survey will provide no information about the implications of the reduction in VKT for other modes or trip chaining behaviour.

6 Conclusion

This paper has demonstrated that Odometer Surveys are a practically feasible, publicly acceptable, and reliable method for measuring change in household vehicle use, where change in aggregate VKT per household per day is defined as the principal policy measure of interest. The collection of odometer readings was chosen because it was a direct way to quantify distance travelled in private vehicles that was difficult to manipulate or misreport. The use of a panel design was useful for minimising variation and despite the frequency of the burden, by keeping the task simple we have shown it is possible to maintain good continuation rates throughout the life of the survey. While the sample obtained may have been subject to coverage error, we have also shown that the demography of the sample is stable over time, which indicates that there is no selective drop out from particular population segments.

The odometer results show that household VKT in both Adelaide and Melbourne are similar at about 30-35 kms per day, although there is more variation in the South Australian sample than the Victorian sample. The means and measures of variation obtained from this pilot suggest that sample sizes upward of 1000 households would be necessary to conclude with 95 percent confidence that an actual difference in daily VKT per vehicle of ±1 kilometre was significantly different from zero. While the data here show that the average daily VKT per household and per vehicle in South Australia stayed stable for VTBC non-participants and dropped in participating VTBC households, the small sample size means we are unable to state with statistical confidence whether this was significant.

Acknowledgments

The authors would like to acknowledge the funding provided for this project by the Program Partners for the National Travel Behaviour Change Project, namely the ACT Planning and Land Authority, the Australian Government through its Department of the Environment and Water Resources, Queensland Transport, South Australian Department of Transport, Energy, and Infrastructure, and the Victorian Department of Infrastructure. However, the opinions and conclusions in this paper are those of the authors and should not be construed as representing the views of the National Travel Behaviour Change Project partners. Any errors in this paper are solely those of the authors.

References

Ampt, E. (2001) "The Evaluation of Travel Behaviour Change Methods – A Significant Challenge" Paper presented at the 24th Australasian Research Forum, Hobart, April 2001. Accessed on: 8/07/2004 from: http://www.patrec.org/atrf/index.php

Ampt, E. and Stopher, P. (2005) "Mixed Mode Data Collection in Travel Surveys: Challenges and Opportunities". Paper presented at the 28th Australasian Research Forum, Sydney, September 2005. Accessed on: 29/11/2006 from: http://www.patrec.org/atrf/index.php

Australian Greenhouse Office (2007) National Travel Behaviour Change (NTBC) Project Details. Accessed from: http://www.greenhouse.gov.au/ggap/ntbc.html on 13/08/2007.

Groves, R. (1987) "Research on Survey Data Quality", *Public Opinion Quarterly 51(Part 2 Supplement):156-172.*

Hawthorne G (2003) "The effect of different methods of collecting data: Mail, telephone and filter data collection issues in utility measurement "*Quality of Life Research* 12: 1081–1088.

Huang H-M (2006) "Do print and Web surveys provide the same results?" *Computers in Human Behavior* 22: 334–350.

Kemp, Hon. D and Anderson, Hon. J (2003) *$6.5 Million Funding for National Travel Behaviour Change,* Press Release 15th August, 2003. Accessed from: http://www.environment.gov.au/minister/env/2003/mr15aug03.html on 13/08/07.

Ker, I. (2002) "Can evaluating be too prescriptive? Appraisal in the age of the Triple Bottom Line" Paper presented at the Australasian Evaluation Society International Conference, October/November 2002, Wollongong, Australia. Accessed on 2/12/2004 from: http://www.aes.asn.au

Red3 (2005), *Evaluation of Australian TravelSmart Projects in the ACT, South Australia, Queensland, Victoria and Western Australia: 2001–2005*, Report to the Department of Environment and Heritage and State TravelSmart Program Managers. Accessed on 2/11/2006 from: http://www.travelsmart.gov.au/publications/evaluation-2005.html

Stopher, P., C. Wilmot, C. Stecher, and R. Alsnih (2004). "Household Travel Surveys: Proposed Standards and Guidelines", Keynote paper presented to the 7th International Conference on Travel Survey Methods, Playa Herradura, Costa Rica, August.

Stopher, P., Greaves, S., Xu, M. and Lauer, N. (2005) *Stages 1.2 &1.3 Development and Scoping of Options for Long-Term Monitoring of the NTBCP*, report prepared by the Institute of Transport and Logistics Studies for the National Travel Behaviour Change Partners.

Swann N. and Stopher P. (2006) "The impact of offering choice in response mode: A comparison of post, telephone and Internet respondents in a multi-mode longitudinal travel survey" paper presented at the ACSPRI Social Science Methodology Conference, Sydney, Dec 2006.

Taylor, M. and Ampt, E. (2003) "Travelling smarter down under: policies for voluntary travel behaviour change in Australia", Transport Policy 10:165-177.