# Design Issues for Before and After Surveys of Travel Behaviour Change

**A.J. Richardson, R.K. Seethaler**
**The Urban Transport Institute, Taggerty, Australia**

**and**

**P.L. Harbutt**
**Department of Infrastructure, Victoria, Australia**

## ABSTRACT

Before and After surveys are a common method of measuring the effect of specific policies and projects designed to cause changes in travel behaviour. The purpose of this paper is to consider some issues involved in the design of Before and After surveys required for the evaluation of projects designed to change travel behaviour. The paper covers the following topics:

- What data is to be collected?
- What type of survey will be used?
- From whom is the data to be collected?
- Over what period is the data to be collected?
- What length of time should be allowed between the Before and the After survey?
- How many After surveys should be performed?
- What magnitude of difference is to be detected in the Before and After surveys?
- What is the inherent variability of the parameters to be measured?
- What levels of confidence are required in the results?
- How should changes in the background environment be handled?

Specific comments are made on the inter-temporal Coefficient of Variation of various measures of travel (trips, kilometres, minutes) as a car driver and a public transport user over different periods of time by different travel units (persons and households). This analysis is based on data from the MobiDrive 6-week travel survey in Germany, and adapted to the Melbourne situation. The paper concludes with implications for sample design and sample size for projects such as the TravelSMART travel behaviour change program in Victoria.

*Presented at the 26th Australasian Transport Research Forum,*

*Wellington, New Zealand*

*October 1-3, 2003*

# 1 BACKGROUND

In 2001/2002, the Department of Infrastructure (DoI) began developing methodologies for a Victorian TravelSMART Program, in preparation for implementing pilot projects in 2002. Methodologies were developed for working with schools, the community and workplaces.

The objective of the Victorian TravelSMART Program is:

> *"to reduce the negative impacts of car travel through a reduction in vehicle trips and kilometres travelled, achieved through voluntary changes by individuals, households and organisations towards more sustainable travel choices".*

The DoI also engaged a contractor to provide Evaluation Services to determine the impact of the Victorian TravelSMART Program. Specifically this engagement required the contractor to undertake the following tasks:

- Develop an evaluation methodology;
- Pilot the evaluation methodology and refine as required;
- Conduct an evaluation of the TravelSMART Pilots;
- Analyse the data collected; and
- Report on the evaluation.

The Urban Transport Institute (TUTI) was engaged by the DoI to provide independent advice on a range of issues associated with the evaluation of a TravelSMART program. The object of this paper is to outline recommendations made to the Department of Infrastructure on sampling issues to be considered in the design of an evaluation program and the determination of the required sample size for the evaluation of the impact of a TravelSMART program.

This paper will cover the following topics:

- What data is to be collected?
- What type of survey will be used?
- From whom is the data to be collected?
- Over what period is the data to be collected?
- What length of time should be allowed between the Before and the After survey?
- How many After surveys should be performed?
- What magnitude of difference is to be detected in the Before and After surveys?
- What is the inherent variability of the parameters to be measured?
- What levels of confidence are required in the results?
- How should changes in the background environment be handled?

While the issues raised in this paper are applicable to any Before and After surveys conducted in transport, they are discussed in the context of TravelSMART programs because it is this area where Before and After surveys are currently being used to a much greater extent than in any other area of transport planning, and where guidelines are most urgently needed. However, the issues discussed are also of general relevance to other areas of transport planning.

# 2 ISSUES FOR TravelSMART PROGRAM EVALUATION

The design of a survey for evaluation of a TravelSMART program needs to consider the following issues before a final survey design and sample size can be determined.

## 2.1 WHAT DATA IS TO BE COLLECTED?

Clearly, a TravelSMART program is multi-dimensional with potentially significant economic, environmental and social implications. However, as noted by the Department of Infrastructure in their Request for Tender (T068) for TravelSMART evaluation services (DoI, 2000), "it is unlikely that an effective and efficient evaluation would be able to evaluate the complete range of desired outcomes that are possible and measurable. The contractor will therefore need to prioritise the desired outcomes, and specify performance measures that are practicably measurable". As further noted in the Request for Tender, "As a priority the evaluation will require specific quantified 'before' and 'after' measures of VKT [vehicle kilometres of travel], air quality, GHG [greenhouse gas] emissions, changes in modal split for trips and distances".

While a comprehensive measurement of all the impacts of a TravelSMART program would be desirable, there is an inevitable trade-off between the amount of data collected from each respondent, the sample size of respondents, and the resources available for the evaluation survey. If all the desirable data was collected from a sufficiently large sample of respondents to make meaningful comparisons between before and after conditions, the cost of the evaluation survey could easily outweigh the cost of the TravelSMART program itself. It was therefore recommended that the proposed evaluation concentrate on "quantified 'before' and 'after' measures of VKT, air quality, and GHG emissions". While not diminishing the value of the other Program Performance Measures (such as changes in modal split), it was considered that unless significant and sustained reductions in VKT, air quality and GHG emissions could clearly be demonstrated from the pilot studies, then precise measurement of the other performance measures would be futile. On the other hand, if this evaluation of the pilot studies could demonstrate significant and sustained reductions in VKT, air quality and GHG emissions, then future evaluations may be able to concentrate more on the other program performance measures.

It was also possible that a concentration on other performance measures could be misleading. For example, while mode shift is an important objective of the TravelSMART program, it should not be used as a primary measure of the success of the program. For example, by increasing the awareness of available public transport services, a TravelSMART program may increase the usage of public transport without decreasing the usage of the private car. While this would result in an increase in modal share for public transport, it has done so by increasing total travel rather than by decreasing car usage.

Therefore, while not ignoring other program performance measures, it was recommended that the evaluation methodology and survey program should concentrate on the quantification of VKT, air quality and GHG emissions changes before and after the TravelSMART programs. It was therefore assumed that the evaluation surveys would focus on the measurement of household vehicle use before and after the TravelSMART program implementation. If other parameters were to be measured (e.g. changes in public transport use, walking and cycling), a more complex person-based survey instrument would have been needed and the sample size would have needed to increase to measure changes in the more variable parameters associated with these modes of travel (see later in this paper for more discussion of the variability of public transport usage). Within the constraints of the available budget for an evaluation study, a trade-off had to be made between approximate measurement of a wide range of parameters versus more precise and accurate measurement of a limited number of more critical parameters.

## 2.2 WHAT TYPE OF SURVEY WILL BE USED?

There are two types of survey that might be used in the evaluation study: a repeated cross-sectional survey and a longitudinal panel survey. In identifying changes in behaviour, a longitudinal panel survey is clearly the preferred option (statistically) since the between-sample variance is eliminated. This enables statistically significant changes to be identified with a smaller sample size in the “before” and “after” surveys.

However, a major problem with a longitudinal panel survey, is the reduced response rate, especially in the “after” survey. This is compounded in the TravelSMART programs, where a multiday recording period might be deemed to be necessary to pick up re-allocation of travel within a household between the days of the week. If this “attrition” between the “before” and “after” surveys is a function of the parameters to be measured (e.g. do households who don’t change their behaviour drop out of the “after” survey because they are not interested in the topic), then adjustments must be made for this “attrition bias” before conclusions can be drawn about the success of the TravelSMART program. A panel survey also suffers from a number of other potential biases (in addition to the biases potentially present in any cross-sectional survey). These additional potential biases include:

- Coverage bias
- Refreshment bias
- Demographic shift bias
- Household composition bias
- Panel fatigue bias
- Travel awareness bias
- Seasonal bias
- Procedural change bias

A more complete description of the biases that might occur in a panel survey of this nature is contained in Richardson (2002).

Thus, while panel surveys are statistically superior for before and after surveys, repeated cross-sectional surveys should not be dismissed, especially if sufficient attention is not paid to minimizing the biases that can occur with panel surveys. The reduction in sampling error obtained by the use of a panel survey may well be outweighed by the increase in sampling (and other types of) bias introduced by the use of a poorly designed panel survey.

## 2.3 FROM WHOM IS THE DATA TO BE COLLECTED?

The major decisions to be made here are, firstly, whether data is to be collected about people's travel patterns or about vehicle's travel patterns and, secondly, whether data is to be collected from all people (or vehicles) in a household or from only one person (or vehicle) in a household.

Collection of data about vehicle travel patterns is appropriate when the prime emphasis is on the measurement of VKT and vehicular use. Data on people's travel pattern is more appropriate when the emphasis is on the reasons for travelling and on the use of non-private-vehicle modes of transport.

Restriction of the survey to a single person or vehicle means that re-allocation of activities and travel between members of the household cannot be detected. Since one of the major objectives of TravelSMART is to encourage household members to devise more effective ways of undertaking the activities associated with their particular lifestyles, it might be expected that intra-household re-allocation of activities might be an option that needs to be monitored. Therefore, the travel patterns of the entire household need to be measured. In the context of a vehicle-monitoring survey, this means monitoring the usage of all vehicles in the household. For practical reasons, this means monitoring up to three vehicles per household (which will cover 98% of all households).

In the discussion that follows, it is assumed that people are the focus of the measurement task but that sufficient detail will be obtained about the use of specific vehicles to enable the reconstruction of vehicle usage patterns. However, the sampling issues to be considered are equally appropriate when the focus of attention is on measuring vehicle usage (perhaps through the use of various types of odometer survey).

## 2.4 OVER WHAT PERIOD IS THE DATA TO BE COLLECTED?

The major decision here is whether the survey should take place over one day or over a multi-day period. Statistically, the survey could be restricted to one day. However, because of the larger relative variability in daily travel, compared to say weekly travel, a much larger sample size of households would be needed in order to detect a specified difference in travel behaviour before and after the TravelSMART program implementation. For example, data from the MobiDrive surveys in Germany (a 6-week continuous panel survey of 150 households, to be described in more detail later in this paper) showed that the Coefficient of Variation for daily household vehicle kilometres (within the same household) was two and a half times the Coefficient of Variation for weekly household vehicle kilometres (where the Coefficient of Variation is the Standard Deviation divided by the Mean). Since sample size is proportional to the square of the Coefficient of Variation, this would require about six times as many households doing 1-day travel surveys as would be required for households doing 7-day travel surveys.

There is also a particular reason in TravelSMART why a multiday survey would be more appropriate. Just as there may be re-allocation of activities and travel between household members, there may also be re-allocation of activities and travel across the days of the week in order to achieve a more efficient travel pattern (e.g. saving up several activities in one area and then doing them all on one day on a single trip). For this reason, there is an advantage to undertaking a multiday (preferably 7-day) survey that will capture these re-allocations across days of a complete week.

## 2.5 WHAT LENGTH OF TIME SHOULD BE ALLOWED BETWEEN THE BEFORE AND THE AFTER SURVEY?

While the concept of before and after surveys is relatively straightforward, it is not very clear what is meant by “before” and what is meant by “after”. Clearly, the “before” survey should be performed before the program is implemented and, where the program involves direct contact with participants, the before survey should be sufficiently far in advance so as not to influence the performance of the program. For example, some TravelSMART initiatives involve the participants in the completion of travel diaries, on the basis of which they are advised of potential behavioural changes. If they are previously required to also complete a “before” travel diary, then this could affect their willingness to participate in the TravelSMART program itself.

The timing of the “after” survey is even more problematic. Short-term changes can be captured by conducting the “after” survey shortly after implementation of the TravelSMART program. On the other hand, there is a keen interest in seeing whether any behavioural changes are sustainable in the long run. For this purpose, the “after” survey should be performed some time after the implementation of the program.

While there are advantages in increasing the length of time between the “before” and the “after” surveys (from the perspective or dispersing respondent burden and measuring long-term success of the program), there is a major problem with this course of action. By increasing the time between the two surveys, one is increasing the probability that other changes will also be occurring, in addition to the TravelSMART program. Such changes could include changes in public transport, fares, changes in petrol prices and changes in the infrastructure or transport services provided. One is then faced with the problem of disentangling the effects of the TravelSMART program from the effects of all the other external changes occurring in the background.

## 2.6 HOW MANY AFTER SURVEYS SHOULD BE PERFORMED?

To minimize the problems in having a long period of time between the “before” and ”after” surveys, some have recommended a multiple number of “after” surveys in order to pick up the short, medium and long term effects of the program. The problem with this approach is that the increased number of surveys will increase the burden on the respondents, and could lead to increased bias due to increased rates of attrition.

Multiple “after” surveys are particularly a problem when each survey involves interaction with the respondent and the expenditure of effort by the respondent. However, where the survey can be done with minimal effort, the increased burden due to multiple surveys may be minimal. For example, some types of odometer survey require minimal effort on the part of the respondent, enabling long-term monitoring of VKT. Similarly, the use of GPS monitoring of vehicles enables detailed long-term monitoring of vehicle use with virtually no extra effort on the part of the respondent.

## 2.7 WHAT MAGNITUDE OF DIFFERENCE IS TO BE DETECTED?

Because of the nature of “before” and “after” surveys, it is necessary to specify the size of the difference to be detected between the two surveys. Detection of a small difference will require a larger sample size compared to detection of a large difference. One might therefore be tempted to opt for detection of a large difference, if this can be done with a smaller sample size. However, if such a large difference does not in fact exist, then any smaller differences will not be detected (statistically). On the other hand, the collection of a large sample in order to detect a small difference may not be worthwhile if the effect of the difference detected is immaterial. Therefore, one needs to trade-off these two effects, and specify a difference which could reasonably be expected to occur, and, if it was detected, then the effect of this difference would be material. The client needs to specify a difference in the parameter(s) with which they would be satisfied if it was detected.

## 2.8 WHAT IS THE INHERENT VARIABILITY OF THE PARAMETERS TO BE MEASURED?

In determining the required sample size, it is necessary to have an estimate of the inherent variability of the parameter to be measured. A parameter with high variability will require a larger sample size to detect a difference of a specified magnitude than a parameter with lower variability. In the context of a before and after survey, repeated cross-sectional surveys will have greater variability than a panel survey because they include the variability between households as well as the variability within a household over time. A single person (or vehicle) will have greater variability than a household (or all vehicles in the household). Daily vehicle kilometres will have more relative variability than weekly vehicle kilometres. On this basis, a panel survey of weekly kilometres travelled by all household vehicles would have the lowest required sample size, while repeated cross-sectional surveys of daily kilometres travelled by a single vehicle would have the highest sample size.

A major problem in the current study is obtaining any information on the variability of vehicle kilometres travelled within a household across an extended period of time. Most travel surveys (such as VATS – the Victorian Activity and Travel Survey) record travel for any one household on only a single day. Therefore, it is impossible from such surveys to calculate the variability in travel across several days or weeks. One of the very few surveys to collect data from households across an extended period is the MobiDrive survey conducted recently in Germany (Axhausen et al., 2002). This survey collected data from about 150 households for every day over a period of six weeks. From this data, it is therefore possible to gain an idea of the relative variation in vehicle kilometres travelled on a daily or weekly basis. While the absolute number of vehicle kilometres travelled in Melbourne might be different from the figures obtained for Germany, it is expected that the Coefficients of Variation will be of similar size.

For example, the Coefficients of Variation of vehicle kilometres travelled under different conditions are summarised in Table 1.

**Table 1 Coefficients of Variation of Vehicle Kilometres Travelled**

| | **Duration of Survey Period** | |
|---|---|---|
| **Sampling Unit** | One Day | One Week |
| One person in Household | 148% | 60% |
| All people in Household | 125% | 49% |

The Coefficients of Variation of vehicle kilometres travelled on a single day can be reduced substantially, however, by ensuring that the "before" and "after" surveys for any one household are conducted on the same day of the week, since much of the variation in day-to-day travel occurs because of differences in travel patterns on different days of the week (rather than because of differences from week to week). Table 2 summarises the Coefficients of Variation of vehicle kilometres travelled for different days of the week for person and household travel.

**Table 2 Coefficients of Variation of VKT for Different Days of the Week**

| | Day of Week | | | | | | |
|---|---|---|---|---|---|---|---|
| **Sampling Unit** | Mon | Tues | Wed | Thurs | Fri | Sat | Sun |
| One person in Household | 78% | 76% | 75% | 79% | 89% | 116% | 125% |
| All people in Household | 72% | 66% | 63% | 65% | 74% | 96% | 107% |

It can be seen that the Coefficient of Variation for one person's daily travel is about 23% greater than for a complete households daily travel (after matching day of week in the before and after surveys). While the Coefficient of Variation for daily travel is about 150% greater than for complete weekly travel (without matching day of week in the before and after surveys), this difference reduces to about 55% greater than for complete weekly travel (after matching day of week in the before and after surveys).

Whether daily or weekly travel, or person or household travel, is used, the main point that emerges from the German MobiDrive surveys is that the natural variation in VKT is quite high. Even the least variable measurement (weekly travel by a complete household) shows a 49% Coefficient of Variation. Any before and after surveys that attempt to detect differences in VKT that have been caused by an external intervention (such as the TravelSMART program) must contend with the relatively high level of natural variability in the quantity being measured. The effect of this variability on the required sample sizes will be described in more detail later in this paper.

## 2.9 WHAT LEVELS OF CONFIDENCE ARE REQUIRED IN THE RESULTS?

When testing hypotheses, such as occurs in before and after surveys, there are four possible end-states of the hypothesis testing procedure. Two of these states signify that a correct decision has been made while the other two indicate that an error has been made. The four end-states may be depicted as shown in Table 3.

**Table 3 Possible End-States of Hypothesis Testing**

| | TRUE STATE | |
|---|---|---|
| **DECISION** | **$H_0$** | **$H_1$** |
| **Accept $H_0$** | Correct | Type II Error |
| **Reject $H_0$** | Type I Error | Correct |

Thus if the true state is described by the null hypothesis $H_0$ and we accept $H_0$ as being a description of the true state, then we have made no error. Similarly, we will be equally correct if we reject $H_0$ when the true state is actually described by the alternative hypothesis $H_1$.

A Type I Error will have been committed if we reject the null hypothesis when it is in fact true. For example, we conclude that there has been a decrease in vehicle usage following the implementation of TravelSMART when, in fact, there has been no decrease in vehicle usage (the null hypothesis being that there is no decrease in VKT). A Type II Error will have been committed if we accept the null hypothesis when it is in fact false. For example, we conclude that there has been no decrease in vehicle usage following the implementation of TravelSMART when, in fact, there has been a decrease in VKT.

Obviously in testing hypotheses we would be interested in trying to minimise the chances of making either a Type I or Type II error. Which one we would be most interested in avoiding will depend on the relative costs associated with each type of error. The degree to which we wish to avoid each type of error is expressed in terms of the maximum probability that we will accept for making each type of error. The acceptable probability of committing a Type I error is called the level of significance of the test and is denoted by $\alpha$. The acceptable probability of committing a Type II error is denoted by $\beta$. The value $1-\beta$ is often called the power of the test.

### 2.10 HOW SHOULD CHANGES IN THE BACKGROUND ENVIRONMENT BE HANDLED?

As noted above, we often assume that the difference between the before and after surveys is due to the intervention (such as a TravelSMART program) that we deliberately place between the two surveys. However, especially if there is a considerable period of time between the before and after surveys, there are many other conditions that could also be changing in the background and which may have an effect on the behaviour we are wishing to monitor.

In such circumstances, the usual response is to employ a so-called Control Group in which we try to measure only the effect of the background changes. The Control Group should be subject to all the same background changes as the Participant Group, but should not be subjected to the intervention (i.e. the TravelSMART program). While simple in theory, such conditions are difficult to achieve in practice. For example, in order to have the Control and Intervention Groups exposed to the same background changes, this usually requires that they are close in geographic proximity (i.e. in the same or neighbouring suburbs). However, if the TravelSMART program receives any publicity (which might be desirable for maximum effect of the program itself), then it will be difficult to not contaminate the Control Group, since if they hear about the TravelSMART program they may be influenced to also change their travel behaviour.

A second problem with the use of a Control Group is that the Control Group will also exhibit temporal variability in the parameters being measured (e.g. VKT). Even if there really is no change in VKT over time in the Control Group, there could be the appearance of change due simply to statistical chance. Therefore, the Control Group must be of sufficient size to ensure that any conclusions about changes in background conditions are statistically reliable. This can mean that the Control Group can often be of comparable size to the Participant Group. Under limited budget conditions, this is often seen as a dubious use of limited evaluation study resources.

# 3 TEMPORAL VARIABILITY IN TRAVEL PATTERNS

As noted above, in determining the required sample size, it is necessary to have an estimate of the inherent variability of the parameter to be measured. One of the very few surveys to collect data from households across an extended period, from which longitudinal variability in travel behaviour can be estimated, is the MobiDrive survey conducted recently in Germany (Axhausen et al., 2002). The MobiDrive data contains information on 52,273 trips (from 334 people living in 146 households) over a period of 6 weeks in 1999. For each trip, the data contains (among other things) the date, mode, travel time and travel distance of each trip. Some summary statistics from MobiDrive, and the corresponding figures from VATS 95 are shown in Table 4.

**Table 4 Comparison of Travel Behaviour in Germany and Melbourne**

| **MobiDrive** | Daily Car Driver Travel | | | Daily Public Transport Travel | | |
|---|---|---|---|---|---|---|
| | Trips | Distance | Minutes | Trips | Distance | Minutes |
| Household | 2.7 | 28.4 | 50.9 | 1.2 | 9.7 | 32.3 |
| Person | 1.2 | 12.4 | 22.3 | 0.5 | 4.2 | 14.1 |

| **VATS 95** | Daily Car Driver Travel | | | Daily Public Transport Travel | | |
|---|---|---|---|---|---|---|
| | Trips | Distance | Minutes | Trips | Distance | Minutes |
| Household | 4.5 | 43.1 | 83.4 | 0.7 | 7.4 | 16.0 |
| Person | 1.7 | 16.1 | 31.1 | 0.3 | 2.8 | 6.0 |

It can be seen that, compared to VATS 95, the MobiDrive respondents make less car driver trips (about 70% of VATS 95), but more public transport trips (about 50% more trips and distance). The public transport minutes in MobiDrive are much higher than VATS 95 (over twice as large) because the speed of public transport in Germany (with its heavy reliance on trams and buses) is lower than in Melbourne (with its significant heavy rail system).

The MobiDrive data was extracted from the files to calculate the number of trips and the time and distance covered by each mode by each person on each of their 42 travel days in the reporting period. Thus the data was reduced to a 42 by 334 matrix of person travel per day, a 42 by 146 matrix of household travel per day, a 6 by 334 matrix of person travel per week, and a 6 by 146 matrix of household travel per week. For each person (or household), the average amount of travel (trips, distance or minutes) by car and public transport was calculated across all days (or weeks). The standard deviation of these amounts were also calculated, and thence the Coefficient of Variation. In addition to calculating the values across all 42 days, an additional analysis considered the values segmented by days of the week. Thus, for example, as well as calculating the average and standard deviation of distance travelled as a car driver across all 42 days, the average and standard deviation of distance travelled as a car driver on the 6 Mondays (and Tuesdays etc) was also calculated.

The results are presented below for travel as a car driver and by public transport, using the three measures of travel given by number of trips, distance travelled and minutes spent travelling.

## 3.1 VARIABILITY OF PERSON TRIPS

### 3.1.1 Panel Survey Weekly Trips per Person

As shown in the top-left part of Table 5, the average number of car driver trips per person per week in MobiDrive was 8.37. The average standard deviation of the number of car driver trips per week, across the 6 weeks for one person, was 2.78. The average Coefficient of Variation (CoV) was 29% (note that this is not 2.78/8.37, but rather is the average CoV calculated across each of the respondents).

**Table 5 Variability in Person Car Driver Trips in Germany and Melbourne**

| ***Weekly Trips*** | Panel Survey | | Cross-sectional Survey | |
|---|---|---|---|---|
| ***per Person*** | MobiDrive | Melbourne | MobiDrive | Melbourne |
| Average | 8.37 | --- | 8.37 | 11.86 |
| Standard Deviation | 2.78 | --- | 10.79 | --- |
| Coefficient of Variation | 29% | --- | 130% | --- |
| Adjustment Factor | 84% | --- | 84% | --- |
| Adjusted CoV | --- | 24% | --- | 109% |
| ***Daily Trips*** | Panel Survey | | Cross-sectional Survey | |
| ***per Person*** | MobiDrive | Melbourne | MobiDrive | Melbourne |
| Average | 1.19 | --- | 1.19 | 1.69 |
| Standard Deviation | 1.00 | --- | 1.97 | 2.68 |
| Coefficient of Variation | 75% | --- | 168% | 158% |
| Adjustment Factor | 84% | --- | 84% | --- |
| Adjusted CoV | --- | 63% | --- | 141% |
| ***Daily Trips per Person*** | Panel Survey | | Cross-sectional Survey | |
| ***(stratified by day of week)*** | MobiDrive | Melbourne | MobiDrive | Melbourne |
| Average | 1.19 | --- | --- | --- |
| Standard Deviation | 0.81 | --- | --- | --- |
| Coefficient of Variation | 65% | --- | --- | --- |
| Adjustment Factor | 84% | --- | --- | --- |
| Adjusted CoV | --- | 55% | --- | --- |

However, as shown in Figure 1, the CoV is a function of the average number of trips per person per week. Apart from those who don't drive at all (whose CoV is obviously zero), the CoV is highest for the infrequent driver and falls as the average number of car driver trips increases. The CoV is related to the average number of car driver trips per person per week (T) by the equation: $CoV = 1.26/\sqrt{T}$

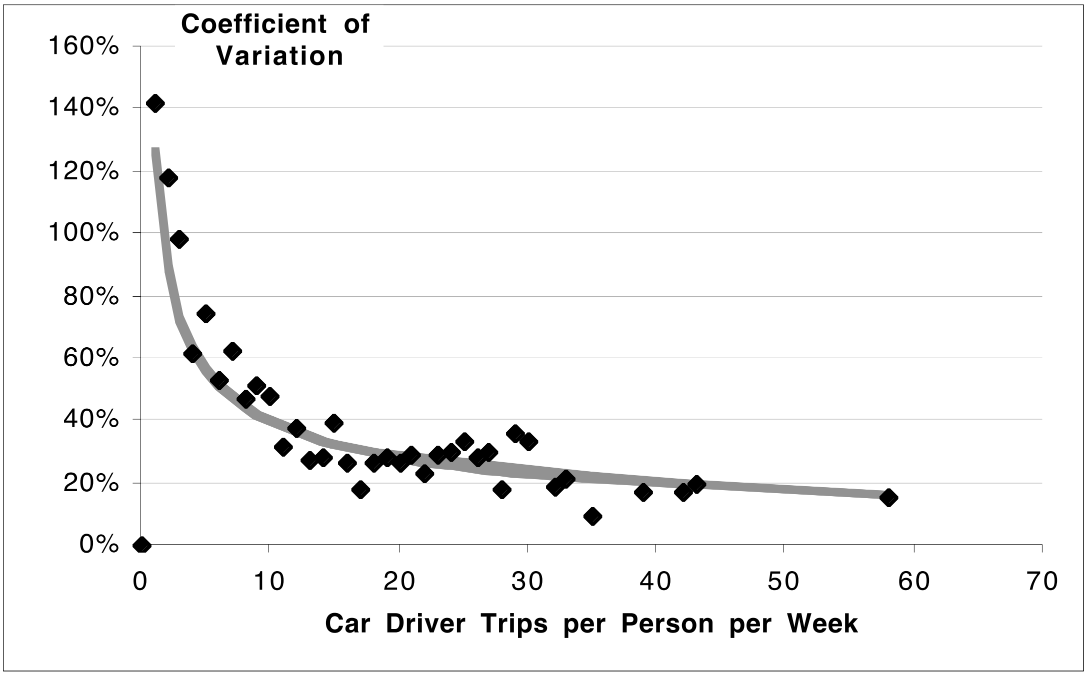


**Figure 1 Coefficient of Variation of Car Driver Trips per Person per Week as a Function of Average Car Driver Trips per Person per Week**

Since people in Melbourne make more trips as a car driver than people in MobiDrive, it would be expected that the CoV of their weekly trip rate would be somewhat lower than in MobiDrive. Using the average weekly car driver trips per person in Melbourne (11.86) and MobiDrive (8.37) and the equation given above, the CoV of weekly car driver trips per person in Melbourne would be expected to be about 24% (i.e. 84% of 29%).

A similar analysis may be performed for public transport trips, as summarized in Table 6. In MobiDrive, the average number of public transport trips per person per week was 3.67. The average standard deviation of the number of public transport trips per week, across the 6 weeks for one person, was 1.67. The average Coefficient of Variation (CoV) was 46%. However, as with car driver trips, the CoV is a function of the average number of trips per person per week. The CoV is related to the average number of public transport trips per person per week (T) by the equation: $CoV = 1.05\sqrt{T}$ . Since people in Melbourne make less trips by public transport than people in MobiDrive, it would be expected that the CoV of their weekly trip rate would be somewhat higher than in MobiDrive. Using the average weekly public transport trips per person in Melbourne (1.89) and MobiDrive (3.67) and the equation given above, the CoV of weekly public transport trips per person in Melbourne would be expected to be about 64% (i.e. 139% of 46%).

**Table 6 Variability in Person PT Trips in Germany and Melbourne**

| ***Weekly Trips*** | Panel Survey | | Cross-sectional Survey | |
|---|---|---|---|---|
| ***per Person*** | MobiDrive | Melbourne | MobiDrive | Melbourne |
| Average | 3.67 | --- | 3.67 | 1.89 |
| Standard Deviation | 1.67 | --- | 5.66 | --- |
| Coefficient of Variation | 46% | --- | 155% | --- |
| Adjustment Factor | 139% | --- | 139% | --- |
| Adjusted CoV | --- | 64% | --- | 216% |
| ***Daily Trips*** | Panel Survey | | Cross-sectional Survey | |
| ***per Person*** | MobiDrive | Melbourne | MobiDrive | Melbourne |
| Average | 0.52 | --- | 0.52 | 0.27 |
| Standard Deviation | 0.61 | --- | 1.05 | 0.89 |
| Coefficient of Variation | 138% | --- | 227% | 331% |
| Adjustment Factor | 138% | --- | 138% | --- |
| Adjusted CoV | --- | 191% | --- | 314% |
| ***Daily Trips per Person*** | Panel Survey | | Cross-sectional Survey | |
| ***(stratified by day of week)*** | MobiDrive | Melbourne | MobiDrive | Melbourne |
| Average | 0.52 | --- | --- | --- |
| Standard Deviation | 0.42 | --- | --- | --- |
| Coefficient of Variation | 119% | --- | --- | --- |
| Adjustment Factor | 139% | --- | --- | --- |
| Adjusted CoV | --- | 165% | --- | --- |

### 3.1.2 Cross-Sectional Survey Weekly Trips per Person

If we treat the MobiDrive data not as a panel survey, but as a series of repeated cross-sectional surveys, we can obtain an estimate of the likely variability in weekly trips per person across the population. The mean, standard deviation and CoV of weekly trips per person was obtained across all people in the survey for each of the six weeks of the survey. As shown in the top-right part of Table 5, the average Coefficient of Variation (CoV) across the six weeks was 130%. This is far higher than the panel survey CoV of 29% for MobiDrive, and reflects the greater variability across people than within the same person over time. Using the same reduction factor to account for the higher trip rates in Melbourne, the cross-sectional CoV in weekly person trips is estimated to be about 109%. Using the same logic for public transport trip rates in Melbourne, as shown in Table 6, the cross-sectional CoV in weekly person public transport trips is estimated to be about 216%.

### 3.1.3 Panel Survey Daily Trips per Person

The two previous sections have considered variability in weekly trip rates. The next two sections consider the variability of daily trip rates. As shown in the middle portion of Table 5, the average number of car driver trips per person per day was 1.19. The average standard deviation in the number of car driver trips per day, across the 42 days for one person, was 1.00. The average Coefficient of Variation (CoV) was 75% (compared to a CoV of 29% for car driver trips per week). However, as with trips per week, the CoV is also an inverse function of the square root of the average number of trips per person per day.

Since people in Melbourne make more trips as a car driver per day than people in MobiDrive, it would be expected that the CoV of their daily trip rate would be somewhat lower than in MobiDrive. Using the average daily car driver trips per person in Melbourne (1.69) and MobiDrive (1.19) and the inverse square root relationship, the CoV of daily car driver trips per person in Melbourne would be expected to be about 63% (i.e. 84% of 75%). A similar analysis for public transport daily trip rates, as shown in the middle portion of Table 6, gives a CoV of daily public transport trips per person in Melbourne of about 191%.

### 3.1.4 Cross-Sectional Survey Daily Trips per Person

By treating the MobiDrive data as a series of repeated cross-sectional surveys, we can obtain an estimate of the variability in daily trips per person across the population. As shown in Table 5, the estimated cross-sectional CoV in daily person car driver trips is about 141% (based on adjustment of the MobiDrive data). This estimate can be compared with a direct estimate using the VATS 95 data, since VATS 95 is indeed a cross-sectional survey of daily trips. The standard deviation of car driver trips per person per day in VATS 95 is 2.68, giving a CoV of 158% (compared to the 141% estimated from the adjusted MobiDrive data). Similarly, for daily public transport trips, the adjusted MobiDrive estimate of CoV is 331%, compared to the direct VATS 95 estimate of 314%. It can therefore be seen that the MobiDrive data, after adjustment for the different levels of mobility, can be used to give reasonably good estimates of CoV for the Melbourne situation.

### 3.1.5 Panel Survey Daily Trips per Person (stratified by day of the week)

The preceding analysis of daily trip rates in a panel survey has made no distinction between the days of the week, i.e. it has simply calculated the variability across all 42 days making no distinction, for example, between weekdays and weekends. It is well-known however that there are significant variations in travel across the days of the week. In a panel survey, this variation can be removed from the design by ensuring that households are approached on the same day of the week in each wave of the panel, thereby ensuring that differences observed in the waves are not simply due to a change in day of week between the waves for that household. The variability in trip rates on the same day of the week across the 6 weeks of the MobiDrive data were therefore investigated.

The average number of car driver trips per person per day is still 1.19 (as observed earlier when all 42 days were considered together). However, the average standard deviation in the number of car driver trips per day, when each day of the week has been considered as a separate strata, is (as shown in the bottom-left of Table 5) reduced to 0.81 (compared to 1.00 when all 42 days are considered together). The average Coefficient of Variation (CoV) is therefore 56% (compared to a CoV of 75% when all 42 days are considered together). However, as with trips per day across all 42 days, the CoV is also an inverse square-root function of the average number of trips per person per day.

Using the average daily car driver trips per person in Melbourne (1.69) and MobiDrive (1.19) and the inverse square-root function, the CoV of daily car driver trips per person in Melbourne, after ensuring that the same day of the week is used in each wave of the panel survey, would be expected to be about 57% (i.e. 84% of 68%) (note that these calculations are not relevant to repeated cross-sectional surveys, since each household is only surveyed once in a repeated cross-sectional survey). A similar analysis for daily public transport trips gives a CoV of 85% (which is much less than the 191% when no control is exerted over the day of the week in each wave of the survey, because public transport usage is even more variable across the days of the week, particularly between weekdays and weekends).

## 3.2 VARIABILITY OF HOUSEHOLD TRIPS

The preceding section has considered the variability in the number of trips undertaken by a person. However, the Before & After surveys may be conducted on the basis of an entire household's travel patterns before and after TravelSMART implementation, in which case information is required about the variability in trip rates on a household basis. This section therefore repeats the previous analysis, but uses the household as the unit of analysis. Since the commentary would be very similar for this section as in the previous section, only the main results are presented in tabular format in Tables 7 and 8 for car driver and public transport trips, respectively.

**Table 7 Variability in Household Car Driver Trips in Germany and Melbourne**

| ***Weekly Trips*** | Panel Survey | | Cross-sectional Survey | |
|---|---|---|---|---|
| ***per Household*** | MobiDrive | Melbourne | MobiDrive | Melbourne |
| Average | 19.1 | --- | 19.1 | 31.7 |
| Standard Deviation | 5.0 | --- | 16.3 | --- |
| Coefficient of Variation | 30% | --- | 85% | --- |
| Adjustment Factor | 78% | --- | 78% | --- |
| Adjusted CoV | --- | 23% | | 66% |
| ***Daily Trips*** | Panel Survey | | Cross-sectional Survey | |
| ***per Household*** | MobiDrive | Melbourne | MobiDrive | Melbourne |
| Average | 2.72 | --- | 2.72 | 4.53 |
| Standard Deviation | 1.79 | --- | 2.92 | 4.66 |
| Coefficient of Variation | 77% | --- | 109% | 103% |
| Adjustment Factor | 77% | --- | 77% | --- |
| Adjusted CoV | --- | 60% | | 85% |
| ***Daily Trips per Household*** | Panel Survey | | Cross-sectional Survey | |
| ***(stratified by day of week)*** | MobiDrive | Melbourne | MobiDrive | Melbourne |
| Average | 2.72 | 4.53 | --- | --- |
| Standard Deviation | 1.48 | --- | --- | --- |
| Coefficient of Variation | 66% | --- | --- | --- |
| Adjustment Factor | 78% | --- | --- | --- |
| Adjusted CoV | --- | 51% | --- | --- |

**Table 8 Variability in Household PT Trips in Germany and Melbourne**

| ***Weekly Trips per Household*** | Panel Survey | | Cross-sectional Survey | |
|---|---|---|---|---|
| | MobiDrive | Melbourne | MobiDrive | Melbourne |
| Average | 8.42 | --- | 8.42 | 5.14 |
| Standard Deviation | 2.87 | --- | 11.09 | --- |
| Coefficient of Variation | 49% | --- | 133% | --- |
| Adjustment Factor | 128% | --- | 128% | --- |
| Adjusted CoV | --- | 62% | | 170% |
| ***Daily Trips per Household*** | Panel Survey | | Cross-sectional Survey | |
| | MobiDrive | Melbourne | MobiDrive | Melbourne |
| Average | 1.19 | --- | 1.19 | 0.73 |
| Standard Deviation | 1.09 | --- | 1.95 | 1.79 |
| Coefficient of Variation | 142% | --- | 184% | 244% |
| Adjustment Factor | 127% | --- | 127% | --- |
| Adjusted CoV | --- | 181% | | 235% |
| ***Daily Trips per Household (stratified by day of week)*** | Panel Survey | | Cross-sectional Survey | |
| | MobiDrive | Melbourne | MobiDrive | Melbourne |
| Average | 1.19 | 0.73 | --- | --- |
| Standard Deviation | 0.78 | --- | --- | --- |
| Coefficient of Variation | 106% | --- | --- | --- |
| Adjustment Factor | 127% | --- | --- | --- |
| Adjusted CoV | --- | 135% | --- | --- |

## 3.3 VARIABILITY OF PERSON DISTANCE TRAVELLED

The preceding sections have used trips as one measure of travel. This section will use the distance travelled as a car driver and by public transport as another measure of travel. As before, the analysis will be performed for persons and households, daily and weekly, and for panel and cross-sectional surveys. Once again, since the commentary would be very similar for this section as in the original sections, only the main results are presented in tabular format in Tables 9 and 10 for car driver and public transport trips by persons, and in Tables 11 and 12 for car driver and public transport trips by households, respectively.

**Table 9 Variability in Person Car Driver Distance**

| ***Weekly Kilometres per Person*** | Panel Survey | | Cross-sectional Survey | |
|---|---|---|---|---|
| | MobiDrive | Melbourne | MobiDrive | Melbourne |
| Average | 86.8 | --- | 86.8 | 112.9 |
| Adjusted CoV | --- | 40% | --- | 144% |
| ***Daily Kilometres per Person*** | Panel Survey | | Cross-sectional Survey | |
| | MobiDrive | Melbourne | MobiDrive | Melbourne |
| Average | 12.4 | --- | 12.4 | 16.1 |
| Adjusted CoV | --- | 103% | --- | 196% |
| ***Daily Kilometres per Person (stratified by day of week)*** | Panel Survey | | Cross-sectional Survey | |
| | MobiDrive | Melbourne | MobiDrive | Melbourne |
| Adjusted CoV | --- | 67% | --- | --- |

**Table 10 Variability in Person Public Transport Distance**

| ***Weekly Kilometres*** | Panel Survey | | Cross-sectional Survey | |
|---|---|---|---|---|
| ***per Person*** | MobiDrive | Melbourne | MobiDrive | Melbourne |
| Average | 29.6 | --- | 29.6 | 19.1 |
| Adjusted CoV | --- | 113% | --- | 234% |
| ***Daily Kilometres*** | Panel Survey | | Cross-sectional Survey | |
| ***per Person*** | MobiDrive | Melbourne | MobiDrive | Melbourne |
| Average | 4.22 | --- | 0.5 | 2.73 |
| Adjusted CoV | --- | 294% | --- | 395% |
| ***Daily Kilometres per Person*** | Panel Survey | | Cross-sectional Survey | |
| ***(stratified by day of week)*** | MobiDrive | Melbourne | MobiDrive | Melbourne |
| Adjusted CoV | --- | 155% | --- | --- |

**Table 11 Variability in Household Car Driver Distance**

| ***Weekly Kilometres*** | Panel Survey | | Cross-sectional Survey | |
|---|---|---|---|---|
| ***per Household*** | MobiDrive | Melbourne | MobiDrive | Melbourne |
| Average | 199 | --- | 199 | 302 |
| Adjusted CoV | --- | 37% | | 84% |
| ***Daily Kilometres*** | Panel Survey | | Cross-sectional Survey | |
| ***per Household*** | MobiDrive | Melbourne | MobiDrive | Melbourne |
| Average | 28.4 | --- | 28.4 | 43.1 |
| Adjusted CoV | --- | 81% | | 117% |
| ***Daily Kilometres per Household*** | Panel Survey | | Cross-sectional Survey | |
| ***(stratified by day of week)*** | MobiDrive | Melbourne | MobiDrive | Melbourne |
| Adjusted CoV | --- | 67% | --- | --- |

**Table 12 Variability in Household Public Transport Distance**

| ***Weekly Kilometres*** | Panel Survey | | Cross-sectional Survey | |
|---|---|---|---|---|
| ***per Household*** | MobiDrive | Melbourne | MobiDrive | Melbourne |
| Average | 68 | --- | 68 | 52 |
| Adjusted CoV | --- | 86% | | 177% |
| ***Daily Kilometres*** | Panel Survey | | Cross-sectional Survey | |
| ***per Household*** | MobiDrive | Melbourne | MobiDrive | Melbourne |
| Average | 9.7 | --- | 9.7 | 7.4 |
| Adjusted CoV | --- | 182% | | 280% |
| ***Daily Kilometres per Household*** | Panel Survey | | Cross-sectional Survey | |
| ***(stratified by day of week)*** | MobiDrive | Melbourne | MobiDrive | Melbourne |
| Adjusted CoV | --- | 133% | --- | --- |

## 3.4 VARIABILITY OF PERSON TRAVEL TIME

The preceding sections have used trips and distances travelled as measures of travel. This section will use the time spent travelling as another measure of travel as a car driver and by public transport. As before, the analysis is performed for persons and households, daily and weekly, and for panel and cross-sectional surveys. Once again, since the commentary would be very similar for this section as in the original section, only the main results are presented in tabular format in Tables 13 and 14 for car driver and public transport trips by persons, and in Tables 15 and 16 for car driver and public transport trips by households, respectively.

**Table 13 Variability in Person Car Driver Travel Time**

| ***Weekly Minutes per Person*** | Panel Survey | | Cross-sectional Survey | |
|---|---|---|---|---|
| | MobiDrive | Melbourne | MobiDrive | Melbourne |
| Average | 156 | --- | 156 | 219 |
| Adjusted CoV | --- | 36% | --- | 116% |
| ***Daily Minutes per Person*** | Panel Survey | | Cross-sectional Survey | |
| | MobiDrive | Melbourne | MobiDrive | Melbourne |
| Average | 22.3 | --- | 22.3 | 31.2 |
| Adjusted CoV | --- | 90% | --- | 155% |
| ***Daily Minutes per Person (stratified by day of week)*** | Panel Survey | | Cross-sectional Survey | |
| | MobiDrive | Melbourne | MobiDrive | Melbourne |
| Adjusted CoV | --- | 59% | --- | --- |

**Table 14 Variability in Person Public Transport Travel Time**

| ***Weekly Minutes per Person*** | Panel Survey | | Cross-sectional Survey | |
|---|---|---|---|---|
| | MobiDrive | Melbourne | MobiDrive | Melbourne |
| Average | 111 | --- | 111 | 41 |
| Adjusted CoV | --- | 145% | --- | 263% |
| ***Daily Minutes per Person*** | Panel Survey | | Cross-sectional Survey | |
| | MobiDrive | Melbourne | MobiDrive | Melbourne |
| Average | 15.8 | --- | 15.8 | 5.9 |
| Adjusted CoV | --- | 377% | --- | 408% |
| ***Daily Minutes per Person (stratified by day of week)*** | Panel Survey | | Cross-sectional Survey | |
| | MobiDrive | Melbourne | MobiDrive | Melbourne |
| Adjusted CoV | --- | 199% | --- | --- |

**Table 15 Variability in Household Car Driver Travel Time**

| ***Weekly Minutes per Household*** | Panel Survey | | Cross-sectional Survey | |
|---|---|---|---|---|
| | MobiDrive | Melbourne | MobiDrive | Melbourne |
| Average | 357 | --- | 357 | 584 |
| Adjusted CoV | --- | 30% | | 68% |
| ***Daily Minutes per Household*** | Panel Survey | | Cross-sectional Survey | |
| | MobiDrive | Melbourne | MobiDrive | Melbourne |
| Average | 50.9 | --- | 50.9 | 83.4 |
| Adjusted CoV | --- | 73% | | 92% |
| ***Daily Minutes per Household (stratified by day of week)*** | Panel Survey | | Cross-sectional Survey | |
| | MobiDrive | Melbourne | MobiDrive | Melbourne |
| Adjusted CoV | --- | 56% | --- | --- |

**Table 16 Variability in Household Public Transport Travel Time**

| ***Weekly Minutes per Household*** | Panel Survey | | Cross-sectional Survey | |
|---|---|---|---|---|
| | MobiDrive | Melbourne | MobiDrive | Melbourne |
| Average | 253 | --- | 253 | 112 |
| Adjusted CoV | --- | 107% | | 210% |
| ***Daily Minutes per Household*** | Panel Survey | | Cross-sectional Survey | |
| | MobiDrive | Melbourne | MobiDrive | Melbourne |
| Average | 36.2 | --- | 36.2 | 16.0 |
| Adjusted CoV | --- | 276% | | 300% |
| ***Daily Minutes per Household (stratified by day of week)*** | Panel Survey | | Cross-sectional Survey | |
| | MobiDrive | Melbourne | MobiDrive | Melbourne |
| Adjusted CoV | --- | 166% | --- | --- |

# 4 CALCULATION OF REQUIRED SAMPLE SIZE

Given the estimates of variability described above, this section considers the required sample size for a survey that measures kilometres of travel as a car driver and by public transport before and after the implementation of TravelSMART. Assume, for the moment, that the survey will be a longitudinal panel survey with the primary objective of measuring total kilometres of travel undertaken by all people in a household in a week. Assume that the intention of the Before and After surveys is to test whether there has been a reduction of total kilometres of car driver travel and an increase in public transport kilometres after implementation of the TravelSMART program. Assume that the changes required to be detected are a reduction of 10% in vehicle kilometres, and an increase of 10% in public transport kilometres. Further assume, for the moment, that there is not expected to be any changes in background factors and that no Control Group survey is being conducted (all of these assumptions can later be relaxed).

In order to calculate a sample size, it is necessary to estimate the variability of the parameter to be measured. Assuming a longitudinal panel survey, it is therefore necessary to estimate the variability of vehicle and public transport kilometres within a household from week to week. From the analysis of the MobiDrive survey data described above, the CoV of weekly household kilometres has been estimated (for Melbourne) as 37% for car drivers trips (Table 11) and 86% for public transport trips (Table 12).

In estimating the required sample sizes for a panel survey, a number of different approaches can be considered, as depicted in Table 17. In the most general situation, there may exist a Participation Group (who will participate in the TravelSMART program) and a Control Group (who will not participate, but will be monitored over the evaluation period). Each member of these groups will participate in a Before survey and an After survey (assume for the moment no problems with panel attrition). For each respondent, we will have a measurement of the parameter of interest (e.g. kilometres of travel) in the before and after survey. Since we are working with a panel, where each respondent provides two measurements (one before and one after), we could "pair" these observations and calculate a difference in kilometres travelled for each respondent. Thus, for each group we have three distributions of parameters (the before, the after and the paired differences), each described in terms of a mean and a standard deviation. Thus for the Participation Group, we have a "before" mean ($pm_b$) and standard deviation ($ps_b$), an "after" mean and standard deviation ($pm_a$ and $ps_a$), and a paired difference mean and standard deviation ($pm_d$ and $ps_d$), with similar values for the Control Group.

**Table 17 Options for Calculating Panel Survey Sample Sizes**

| | Before Survey | After Survey | Paired Observations |
|---|---|---|---|
| Participation Group | $pm_b$<br>$ps_b$ | $pm_a$<br>$ps_a$ | $pm_d$<br>$ps_d$ |
| Control Group | $cm_b$<br>$cs_b$ | $cm_a$<br>$cs_a$ | $cm_d$<br>$cs_d$ |

If we don't have a Control Group survey, then we are essentially assuming that $cm_b=cm_a$, or $cm_d=0$. In such a situation, we can use one of two hypothesis tests on the Participation Group data. Either we can work with the distributions of results from the before and after surveys and test the null hypothesis that $pm_b=pm_a$, or we can work with the paired observations distribution and test the null hypothesis that $pm_d=0$.

It should be realized that even if $pm_b=pm_a$ and $ps_b=ps_a$ (the before and after distributions are statistically identical), then while $pm_d=0$ it is not essential that $ps_d=0$, since it is not guaranteed that there will be perfect correlation between the before and after distributions. Thus, someone who travels more in the Before survey than they usually do will not necessarily also travel more than they usually do in the After survey. While the mean difference in kilometers travelled will be zero, it is not required that all the individual differences in kilometers travelled also be zero. Because of this fact, two different sample size equations can be used, depending on whether one works with the individual before and after distributions or with the distribution or paired observations.

If one compares the before and after distributions of kilometers travelled per household per week, the required sample size for hypothesis testing in both the before and after surveys (Richardson, Ampt and Meyburg, 1995, pg 122) is given by:

$$n = \frac{2\left(z_{\alpha} + z_{\beta}\right)^2\left(\sigma^2\right)}{\delta^2}$$

where n = required sample size
α = the probability of making a Type I error
β = the probability of making a Type II error
σ = the standard deviation of kilometres per week per household
δ = the required difference in kilometres per week per household

If one tests whether the mean of the paired observations distribution of kilometres travelled is significantly different from zero, the required sample size is given by:

$$n = \frac{\left(z_{\alpha} + z_{\beta}\right)^2\left(\sigma^2\right)}{\delta^2}$$

where σ = the standard deviation of the difference in kilometres travelled per week per household

Using the first method and assuming that α=β=5%, σ=37% (of the mean) and δ=10% (of the mean), then the required sample size to detect the reduction in kilometres of car driver travel is about 290 households. Assuming that σ=86% (of the mean) and δ=10% (of the mean), then the required sample size to detect the increase in kilometres of public transport travel is about 1570.This calculation assumes that the sample is being drawn from an infinite, or at least very large, population. However, in the TravelSMART project the size of the population of households in each of the study areas is relatively small (about 1500 households in each of three study areas). With such a small population (N), it is necessary to multiply the estimated sample size (n) by a Finite Population Correction Factor (FPCF), where:

$$\text{FPCF} = \frac{1}{1 + n/N}$$

With a population of only 1500 households, the required sample size for detecting the difference in car travel is reduced to about 240 households, while the required sample size for detecting the increase in public transport travel is reduced to about 770 households. That is, in order to measure a statistically significant reduction of

10% in weekly household vehicle kilometres in the after survey (when the inherent variability of weekly household vehicle kilometres is 37% of the mean), a sample size of 240 households would be required in both the before and after surveys, while in order to measure a statistically significant increase of 10% in weekly household public transport kilometres in the after survey (when the inherent variability of weekly household public transport kilometres is 86% of the mean), a sample size of 770 households would be required in both the before and after surveys. The larger sample required to detect the change in public transport usage is because of the inherently greater variability in public transport usage over time.

The above calculations has been based on a number of specific assumptions, namely:

Type of Survey: Panel

Unit of Measurement: Households

Period of Measurement: One week

Variable being Measured: Kilometres of travel (car driver and PT)

Population Size: 1500 households

Coefficient of Variation of Parameter: 37% (car) and 86% (PT)

Detected Difference: 10% of mean

Probability of making a Type I error ($\alpha$): 5%

Probability of making a Type II error ($\beta$): 5%

By varying some of these parameters, we can see that, for a specific set of conditions, detecting a 10% change in travel with a confidence level of 95%, from a population of 1500 households in each study area, the sample sizes shown in Table 18 would be required as a function of the type of survey (panel or cross-sectional survey), the unit of measurement (person or household), the quantity being measured (trips, kilometres or minutes), the mode of transport (car driver and public transport), and the time period of the survey (week, day or matched day-of-week).

**Table 18 Sample Sizes Required for Various Before & After Survey Designs**

| | Car | | | Public Transport | | |
|---|---|---|---|---|---|---|
| | Trips | Kilometres | Minutes | Trips | Kilometres | Minutes |
| Panel Survey | | | | | | |
| Person | | | | | | |
| Week | 119 | 279 | 234 | 556 | 972 | 1129 |
| Day | 547 | 909 | 810 | 1261 | 1389 | 1430 |
| Matched DOW | 453 | 590 | 504 | 1196 | 1164 | 1278 |
| Household | | | | | | |
| Week | 110 | 242 | 168 | 539 | 778 | 931 |
| Day | 510 | 726 | 648 | 1237 | 1239 | 1375 |
| Matched DOW | 410 | 586 | 471 | 1085 | 1077 | 1200 |
| Repeated Cross-sectional Survey | | | | | | |
| Person | | | | | | |
| Week | 947 | 1123 | 990 | 1306 | 1332 | 1364 |
| Day | 1112 | 1270 | 1165 | 1401 | 1436 | 1440 |
| Household | | | | | | |
| Week | 582 | 758 | 597 | 1210 | 1227 | 1297 |
| Day | 762 | 998 | 825 | 1332 | 1378 | 1393 |

Several features emerge from this comparison. Firstly, larger sample sizes are generally required to detect changes in either distance travelled or travel time than in trips undertaken. Secondly, larger sample sizes are required to detect changes from repeated cross-sectional surveys than from a panel survey. Thirdly, larger sample sizes are required to detect changes when using a daily travel diary compared to using a weekly travel diary (although this difference can be substantially reduced in a panel survey by maintaining the same day of the week for each household in later waves of the panel). Fourthly, larger sample sizes are required to detect changes from person travel data than from household travel data. Finally, larger sample sizes are required to detect a 10% change in public transport usage than a 10% change in car usage.

Traded off against these sample size differences, however, is the fact that some of the parameters enabling smaller sample sizes also give rise to survey designs which are more difficult to undertake. For example, panel survey data is more difficult to obtain (with full control of other biases) than repeated cross-sectional data. Weekly travel diaries are more burdensome than daily travel diaries. Getting travel data from all household members is more difficult than getting data from one member of the household.

# 4 CONCLUSIONS

The purpose of this paper was to consider some issues involved in the design of Before & After surveys required for the evaluation of the impact of programs such as the TravelSMART program in Victoria, Australia. To do this, it was important to obtain a quantitative understanding of the underlying variability of the parameters to be measured (in particular, the variation over time in travel by car). This was obtained by a detailed analysis of the MobiDrive data from Germany, and the estimation of Coefficients of Variation in key travel parameters for the Melbourne situation.

Following this analysis, the paper estimated the required sample size for a survey that measures trips, vehicle-kilometres and travel time for car and public transport travel before and after the implementation of TravelSMART. Sample sizes were calculated for different Types of Survey, Units of Measurement, Periods of Measurement, Coefficients of Variation of the Parameters of Interest, the desirable Detectable Difference in the before and after surveys, the Probability of making a Type I error (a) and the Probability of making a Type II error (b).

Several features emerged from this analysis. Firstly, larger sample sizes are generally required to detect changes in either distance travelled or travel time than in trips undertaken. Secondly, larger sample sizes are required to detect changes from repeated cross-sectional surveys than from a panel survey. Thirdly, larger sample sizes are required to detect changes when using a daily travel diary compared to using a weekly travel diary (although this difference can be substantially reduced in a panel survey by maintaining the same day of the week for each household in later waves of the panel). Fourthly, larger sample sizes are required to detect changes from person travel data than from household travel data. Finally, larger sample sizes are required to detect a 10% change in public transport usage than a 10% change in car usage.

Traded off against these sample size differences, however, is the fact that some of the parameters enabling smaller sample sizes also give rise to survey designs which are more difficult to undertake. For example, panel survey data is more difficult to obtain (with full control of other biases) than repeated cross-sectional data. Weekly travel diaries are more burdensome than daily travel diaries. Getting travel data from all household members is more difficult than getting data from one member of the household.

As always, the actual sample size chosen will depend on discussions between the client and the consultant, taking account of the available budget for the survey and the required quantity and quality of data collectable within that budget.

# 5 REFERENCES

Richardson, A.J., Ampt, E.S. and Meyburg, A.H. (1995). Survey Methods for Transport Planning. Eucalyptus Press, Melbourne.

Richardson, A.J. (2002). Panel Survey Issues in the Evaluation of the TravelSMART Program. TUTI Report 11-2002, The Urban Transport Institute, Taggerty, April (available at www.tuti.com.au).

Axhausen, K.W., Zimmermann, A., Schönfelder, S., Rindsfüser, G. and Haupt, T. (2002). "Observing the rhythms of daily life: A six-week travel diary". Transportation. 29(2), pp 95-124.