

# Tackling Non Response in Household Travel Surveys - A Case Study

Elke Moons

*Transportation Research Institute, Hasselt University*

*Science Park 5/6*

*B-3590 Diepenbeek, Belgium*

*E-mail: elke.moons@uhasselt.be*

Geert Wets

*Transportation Research Institute, Hasselt University*

*Science Park 5/6*

*B-3590 Diepenbeek, Belgium*

*E-mail: geert.wets@uhasselt.be*

## 1. Introduction

Ever since car ownership and car use started to increase in Western Europe and the US, transportation planners attempted to model people's travel behaviour. A series of models was developed, each successor ameliorating the previous one (Jovicic, 2001), however, the question why people undertake trips was completely neglected. This is where the activity-based models came into play, which have set the standard for transportation and travel behaviour modelling during the last decade (McNally, 2000). The main idea behind these activity-based models is that travel demand is derived from the activities that individuals and households need or wish to perform (Jones, *et al.*, 1983).

In order to receive the necessary information to build these models, data need to be collected. Methodological research in the time use and transportation community is accumulating, offering a good basis to decide on the most appropriate format for data collecting. A questionnaire, asking people for their typical travel behaviour on an average day has long been the dominant form of data collection in transportation research. It has been argued, however, that such a questionnaire may result in under-reporting of trips, especially of off-peak, non home-based trips of short duration (Koppelman, 1981; Dijst, 1993). Stimulated by these findings, Stopher (1992) and Clarke and co-authors (1981) reported that an activity diary outperforms a travel survey with regard to under-reporting of trips, even by 13-16%. Similar figures have been reported in time use literature as well (Nieme, 1993).

Nevertheless, diaries are not perfect either. Collecting diary data is quite demanding for the respondents and therefore, although this is not unique for diary data, there may be evidence of lower response rates and of differential non response by certain socio-demographic groups. For example, Sen *et al.* (1995) reported that larger households had lower response rates when compared to single member households.

Regardless of the potential source of bias introduced in this way, an activity-based model requires detailed information, so a diary seems a good instrument to collect it. In this paper, a possible solution will be presented on how non response in household travel surveys can be tackled. Section 2 outlines the data collection process and tries to provide a solution on how to go about non response, while Section 3 provides the results and gives profiles of the responders of the study. The paper concludes with a discussion of the major findings and possible avenues for future research.

## 2. Sample Design

In January 2006, the Transportation Research Institute started with the collection of activity-diary data among households in Flanders (the Dutch speaking part of Belgium) in order to set up an activity-based model of travel behaviour (Wets, 2005). In the design of the sampling scheme of the

presented case study, both the coverage of the people in Flanders and the logistic feasibility of the fieldwork are important concerns. Even when an exhaustive list (such as the National Registry) is available, direct selection from this list would be too expensive, because the geographical spread would be too wide. Cost savings may allow the investigators to use a larger sample size than they could use for a simple random sample of the same cost. Therefore, a stratified clustered design is opted as a preferable solution in this case study.

The *stratified* sample has as an advantage that it produces smaller bounds on the errors of estimation than would be produced by a simple random sample of the same size. This case study has two stratification levels: one at the provincial level where within each province a proportional representation in the base sample is sought, and one at municipal level. They are preferred to provinces, because the latter are too large to handle and municipalities are stable, established administrative units. The great variation in their size is controlled for by systematically sampling within a province with a selection chance proportional to the municipality's size.

Whereas the stratification effects and the systematic sampling according to municipalities have the effect of increasing the precision, the *clustering* effect might slightly reduce precision. It is well-known that units (individuals) within a cluster (household) will resemble each other more than in a simple random sample. However, only a slight decrease in overall efficiency is to be expected, hence negative effects will probably not outweigh the advantages. People can be sampled from abbreviated listings and the survey's field worker's travel distance will be reduced significantly.

A feasibility study, taking into account accuracy, budget constraints and available logistic means set the total number of successful survey results for the sample to 2500 households in Flanders. Former studies taught us that the response rate in this type of studies is rather low. Accounting for this by estimating a response rate of about 60%, 4200 households were used as a sample. However, this would be overly optimistic: it is possible that people can not be reached, others might not be willing to cooperate and still others might not end the inquiry successfully, since filling in the questionnaire is quite burdensome. In order to 'solve' this problem of non-responders, it was decided upon to replace these households by reserve households. These reserve households have four characteristics in common with the non-participating household: they live in the same municipality as the refusing household, the age of the reference person falls in the same age category as that of the reference person of the initially chosen household, their genders must match and the household composition (number of adults and number of children) needs to be similar. This matching of reserve and reference households is to ensure that people show more or less the same mobility characteristics, since a household without children will surely show different mobility behaviour when compared to a household with three children. A reserve household is chosen when a reference household could not be reached or when it refused to cooperate in the survey and then the same applies to this household. This procedure continues until the list of 5 households has run out. In general, one could state that in this way an overall response rate of about 12% is aimed at.

### **3. Results**

#### **3.1 Non Response**

The survey process was started in January 2006 and it is still ongoing. This subsection provides the initial results of the case study. The first table combines the results of the reference and the reserve households in one figure, which means that 4200 is the total sample size to be regarded. Of these 4200 household groups, 4167 could be reached and an introduction letter was sent to them. Of these group, 3090 households were willing to participate after the first contact, while 1077 refused to do so. The main reasons for refusal to participate were: 'lack of interest', 'lack of time' and perceiving

themselves as ‘too old’ to participate. The households that were willing to participate were sent the questionnaires and an appointment was made to call them back on the second day of the inquiry (the duration of the survey is a week). 2902 households confirmed their participation after receiving the questionnaires, while 188 households no longer wanted to participate (and the same was true for their respective reserve households). Table 1 summarizes the results.

**Table 1: Response rates for the overall survey (for the reference households)**

	January 2006 - April 2007			
Introduction letters sent	4167	100%		
Willing to participate after first contact	3090 (1467)	74% (35%)	100% (100%)	
Willing to participate after receiving the questionnaires	2902 (1388)	70% (33%)	94% (95%)	100% (100%)
Questionnaires sent back	621 (317)	15% (8%)	20% (22%)	21% (23%)
<i>Still in survey process</i>	2281 (1071)	55% (26%)	74% (73%)	79% (77%)

Suppose now that the setting of the reserve households would not have been opted for. Then the results would be quite different (see Table 1, results in parentheses). Of the 4167 households that were reached, only 1467 were willing to participated after the first contact, and after the questionnaires were sent, another 79 households refused to participate. The final number of households that had sent their questionnaires back is only 317, which would mean that the response rate so far equals about 8%, this is much lower than the 60% that was aimed at.

Working with reserve households that can replace the reference household when required proves to be a good solution to cover part of the non response. In the next subsection, we will take a closer look at the respondents’ profiles.

### 3.2 Profiling respondents

Literature learns us that there might be differential non response by certain socio-demographic groups. Therefore, in this subsection, we will focus on the households who filled out complete questionnaires alone, and we will try to profile them by comparing their the number of responders to the number that they should take up based on the sampling frame. In this way, separate response rates will be determined based on five important variables: age and gender of the reference person in the household, number of children, number of people in the household (HH) and the province. The response rates can be found in Table 2. In general, compared to the overall response rate of 15%, one can observe that people between 35 and 64 years of age are more willing to participate than older or younger people. Middle-sized households are clearly more cooperative compared to singles or to large households consisting of seven people or more. Apparently, the more children in a household, the more it is willing to take part in the survey. These latter two results are counterintuitive and in contrast with previous results by e.g. Sen *et al.* (1995). One might expect that households with more children or more people in the household in general would have less time on their hands to fill out a rather burdensome travel survey. Though, this does not seem to be true. Two provinces seem to outperform the other three, but this might be a biased result, as these are the two provinces with the highest number of complete surveys. In the other three provinces, the survey is still ongoing.

*Table 2: Overall response rates on five profiling variables*

Variable	Categories				
Age	<25 12%	25-34 11%	35-44 18%	45-64 19%	65+ 10%
Gender	Male 17%	Female 7%			
HH composition	1 6%	2 14%	3 18%	4 24%	5 21%
	6 20%	7 7%	8+ 10%		
Nr. of children	0 12%	1 15%	2 21%	3 23%	4+ 21%
Province	Limburg 31%	Flemish-Brabant 22%	Antwerp 10%	Eastern Flanders 8%	Western Flanders 7%

#### 4. Discussion and final remarks

Travel surveys collect the critical information that is crucial for modern transportation planning and policy development and diaries are an appropriate instrument for obtaining the required data. However, filling out an activity-diary can be quite burdensome, hence leading to low response rates.

This paper investigates how part of the problem of non response can be overcome. Instead of just increasing the sample size, for each base sample household, four reserves are provided, matched according to the household type. When a previous household could not be reached or when it refused to cooperate in the survey, it will be replaced by a reserve household.

Results show that working with reserve households almost doubles the response rate. In profiling responders, we observe that singles and very large households are hard to reach, whereas households with children and households with a reference person between 35 and 64 years are more cooperative. Further investigation of differential non response can be carried out through a cluster analysis. Together with an exploration of item non response, this forms a basis for future research.

#### REFERENCES

- Clarke, M., Dix, M. and Jones, P. (1981) Error and Uncertainty in Travel Surveys. *Transportation*, **10**, 105–126.
- Dijst, M. (1993) *Gewijzigde Opzet van Dagboekonderzoek Succesvol (in Dutch)*. OSPA, TU-Delft, Delft.
- Jones, P.M., Dix, M.C., Clarke, M.I. and Heggie, I.G. (1983) *Understanding Travel Behavior*. Gower, Aldershot.
- Jovicic, G. (2001) *Activity Based Travel Demand Modelling*. Danmarks TransportForskning.
- Koppelman, F.S. (1981) Non-Linear Utility Functions in Models of Travel Choice Behaviour. *Transportation*, **10**, 127–146.
- McNally, M.G. (2000) *The Activity-Based Approach*. Center for Activity Systems Analysis. Paper UCI-ITS-AS-WP-00-4.
- Nieme, I. (1993) Systematic Error in Behavioural Measurement: Comparing Results from Interview and Time Budget Studies. *Social Indicators Research*, **30**, 229–244.
- Sen, A., Sööt, S., Yang, L. and Christopher, E. (1995) Household Travel Survey Nonresponse Estimates: The Chicago Experience. *Transportation Research Record*, **1493**, 170–178.
- Stopher, P.R. (1992) Use of an Activity-based Diary to Collect Household Travel Data. *Transportation*, **19**, 159–176.
- Wets, G. (2005) *An Activity-Based Approach for Surveying and Modelling Travel Behaviour*. Strategic Basic Research funded by the Institute for the Promotion of Innovation by Science and Technology in Flanders.