Setting up a Continuous Panel for Collecting Travelling Information: Discussion on Methodological Issues

Elke Moons, Geert Wets

Hasselt University

Transportation Research Institute

Wetenschapspark 5

B-3590 Diepenbeek

Belgium

E-mail:{elke.moons, geert.wets}@uhasselt.be

Abstract

Modelling travel behaviour has always been a major area of concern in transportation research. Since 1950, due to the rapid increase in car ownership and car use in Western Europe and in the US; several models of transport mode, route choice and destination have been used by transportation planners. The most recent models clearly indicate that the use of household travel data is a critical component of the travel-demand forecasting process. The data are typically generated through a household-based survey in which a sample of the population records their travel patterns over a given time period. This information is combined with socio-demographic information about the sample to develop relationships between individual/household characteristics and their observed travel patterns.

Traditionally, travel data on households and individuals is collected cross-sectionally every five years in Flanders on two thousand five hundred households that have to report their travel behaviour for two consecutive days. This means that if we want to investigate the changes from one wave to another, we can only speak in terms of moving averages over five years.

It has been established that calendar events (holidays, long weekends, etc.) have a serious impact on travel demand. However, the effect of these regular events has never been combined with different household characteristics, only the impact on traffic intensities has been assessed. This can perfectly be encompassed by a longitudinal survey. Therefore, the idea was suggested of setting up a continuous panel and to ask the respondents to report their travel behaviour (in relation to the other household members).

The aim of this paper is to discuss some of these methodological issues that will arise if one transfers from a multiple cross-sectional method to a more continuous approach of collecting data. Some of the issues that will be addressed are the survey method, the sample rotation and the handling of non-response.

Keywords: Continuous data collection, Travel behaviour data, Survey methods, Sample replacement, Non-response

Introduction

Modelling travel behaviour has always been a major area of concern in transportation research. Since 1950, due to the rapid increase in car ownership and car use in Western Europe and in the US; several transportation planners attempted to model people's choice of transport mode, route choice and destination. These models were necessary to predict travel demand on the long run and to support investment decisions in new road infrastructure that originated from this increased level of car use. In these days, travel was assumed to be the result of four subsequent decisions that were modelled separately. However, in the mean time, one acknowledged that these initial models clearly had some drawbacks (Jovicic, 2001), like eg. the focus on individual trips, where the interrelationships (spatial, temporal, intra-household) between trips and its characteristics are ignored.

The models that were developed in the seventies and eighties accounted for this problem, but still limited insight was offered into the relationship between travel and non-travel aspects. Travel has an isolated existence in these models and the question why people undertake trips is completely neglected. This is where activity-based travel demand models comes into play and these activity-based models have certainly set the standard for the last decade of modelling travel demand. The major idea behind activity-based models is that travel demand is derived from the activities that individuals and households need or wish to perform. Travel is merely seen as just one of the attributes (Jones *et al.*, 1983). Shortly, travel patterns are the manifestation of the implementation of activity programs over time and space. In turn, activity patterns emerge as the interplay between the institutional context, the urban/physical environment, the transportation system and individuals' and households' needs to realise particular goals in life and to pursue activities (Ben-Akiva and Bowman, 1998). This clearly shows why household travel surveys, combined with individual surveys, continue to be an essential component of transport planning and modelling efforts.

In the past, data collection for transportation planning was mainly done on a cross-sectional basis. But, since the main interest now is shifting towards a better understanding of individual behaviour, the means of data collection will have to be adapted too. If one wants to understand how people behave and in which way they are behaving in different situational contexts, data are necessary which are catching the same population in different situations.

Indeed, the variability in behaviour, the flexibility of a person to react and the identification of constraints and regimes can only be detected from a longitudinal perspective. Moreover, it can be argued that a solution to environmental issues is a change in travel behaviour towards a more sustainable mobility. However, the usual cross-section or snapshot-oriented surveys of the behaviour of one day give only poor descriptions of ongoing changes and

hardly allow to distinguish real changes in behaviour from external evolutions caused by specific trends.
Thus data sources are necessary which are enabling the researcher to understand which processes are ongoing and how changes are happening, which results in a need to develop the temporal dimension in our data.

The aim of this paper is to discuss some of the methodological issues that will arise if one transfers from a multiple cross-sectional method to a more continuous approach of collecting data. The next section specifies some of the data requirements. Section 3 addresses the survey method, the sample size and sample rotation and the handling of non-response. A final section gives some ideas for future research.

Data requirements

As stated above, the collected data can serve many goals. On the one hand, the aim is to describe ongoing processes and analyse the change in travel demand, while on the other hand, we also want to investigate intra-person changes and the impact that this might have on a person's mobility. In this paper, we will focus on the first goal, i.e. analysing the impact of regular events on travel behaviour.
It has been established (Weijermans and van Berkum, 2004) that holidays (Christmas, …) and making a long week-end has a serious impact on travel demand. E.g. people undertake their Christmas shopping activities in the week(s) before Christmas, so then there is more traffic towards large shopping centrums. This has been assessed by relating the traffic intensities at several spots along the main roads to calendar days. But not only calendar events have an impact on travel demand, also the environment. It seems only logical that e.g. the choice of a transport mode can rely heavily on the weather or other seasonal components. The effect of the combination of several regular events as well as the impact of these regular events on different types of households has never been examined. This can perfectly be encompassed by a longitudinal survey. However, as stated before, nearly all household travel surveys that are conducted in the past were "snapshots" of travel behaviour in a region. Therefore, the models that are developed up to now can capture cross-sectional variation, i.e. variation among individual respondents, but no changes to individual or household behaviour over time. In order to anticipate future travel demand, there is need for dynamic models based on longitudinal data. At the same time, this type of data provides important information to examine the dynamics behind some general characteristics like employment, household composition, availability of transport modes, evolution of car ownership, etc. that clearly will play an important role in predicting travel demand.

In a typical household travel survey, a sample of the population is asked to record their activity patterns over a given time period. This information is combined with socio-demographic information about the sample to develop relationships between individual/household characteristics and their observed travel patterns. More precise, people are asked to write down for some consecutive days which activities they conducted, where, when, with whom, for how long and which transport mode was used to arrive at the location of the activity. Above this information, some general household information was gathered as well, such as household composition, socio-economic status of the household, availability of transport modes, etc.
Traditionally, travel data on households and individuals is collected about every five years in Flanders on two thousand five hundred households that have to report their travel behaviour for two consecutive days. These 2,500 households are different at each wave of surveying. This means that if we want to investigate the changes from one wave to another, we can only speak in terms of moving averages over five years.

In contrast to the household travel survey, traffic intensity is measured at a traffic control centre every minute of every day at the main roads throughout Flanders. So, one can establish the impact of a specific calendar day on the road, but we cannot make the connection with some important characteristics of a household/individual. It might be the case that certain types of households (e.g. with versus without children) react differently to different types of holidays (e.g. school holidays versus other). Ideally, we would like to be able to make the connection between both type of data and since the household travel survey asks for each travel activity the origin and destination of travel, it seems feasible to set up a micro-simulation model for Flanders for every day of the year.

To achieve this goal, the idea was suggested to set up a continuous panel on a smaller amount of households and to ask the individuals in each household to report their travel behaviour (in relation to the other household members) for some time after which the sample was refreshed.
Several questions arose… What would be the ideal sample size within a wave, what survey instrument is best used, how about sample replacement, how long would each wave take, how many waves, what can we do about non-response and how can we lighten the respondent burden? Some of these questions will be answered in the next section.

Survey method and sampling issues

The problem we are facing in this study is two-fold. On the one hand, one wishes to have a more or less continuous image of travel behaviour in order to determine changes in travel behaviour quicker (compared to a survey conducted once every five years) and, if possible, to be able to point at the moments of change. On the other hand, we want to measure the reaction of the target population to regular events, like a long week-end or a holiday. This means that we need to have data for a sufficient amount of time such that we can determine the population's "normal" travel behaviour in a "normal" week, versus the change in behaviour on a regular event.

Method of data collection

First of all, we have to make clear what is meant by a longitudinal survey and a panel in this particular study. Since the main interest lies in establishing changes in travel behaviour for the whole population and not in determining the intra-person changes due to some specialised events, we can state that a "true" panel survey, where observations are made on different discrete time points on the same subjects is not necessary here. What we actually need are repeated cross-section surveys on independent representative samples of the target population. The continuity lies in the fact that data are not collected once every five years, but that we have observations of the years in between as well.

By using the strict definition of a longitudinal survey, we need to measure a certain observation continuously, i.e. for instance the measurement of travel behaviour over one week or longer. Sometimes the term "short panel" is used for this in transportation literature (Axhausen *et al.*, 2000). In a strict sense, we would not need such a "short panel", since we are not particularly interested in the changes in travel behaviour due to some events for each individual separately, but we need to have a global view of the changes of the population. But, in this study, we opted to use such a "short panel", mainly for administrative reasons. If these people agreed to take part in the survey, we at least have information for some observations points, while taking random samples at each observation point would lead to a greater non-response. Sample rotation will be discussed in the next sub-section.

The cross-sectional study (PHA, 1993), that has been carried out for some years now, is a mail survey with recruitment and follow-up carried out by phone. Applying a mail survey allowed the surveyors to increase the sample size at a relative low cost and at the same time, it provides a sense of anonymity. Moreover, compared to interviews, any bias that might be introduced by interviewers has been removed, such that it reduces measurement error.

However, due to the recent increase in mobile phone technology (and its accompanying privacy issues) the number of people still having a phone at home, beside their mobile phone diminishes every year. Therefore, the recruitment is under review again for the next cross-sectional survey, since now about 75% of the households in Flanders have a phone number and this number goes down every year.

An option that was proposed is to use a multi-method survey in the sense that the people who can be contacted will be given the choice of filling out the questionnaire either via paper and pencil or via the internet. In this way, we can have an initial idea about the changes in travel behaviour much faster. Moreover, we can counter some of the measurement errors by building-in consistency checks in the internet survey. This all seems positive, so why let not everybody fill out the survey on the internet? This might not be a good representation for the people of Flanders, since more than 50% of the people over fifty do not have a computer at home. Although everybody is granted access to the internet (at the library), we may assume that this will cause coverage error. We therefore suggest to select people at random via the National Registry and if they are in the opportunity of choosing between survey methods, they can do so. We have to realise, however, that this might cause coverage and/or non-response error. These sources of error will be discussed further on in this section.

Sample rotation

There was agreed on the use of a short panel for administrative reasons though the sample refreshment, and hence the rotation of the sample is still a point of discussion.

An “ideal” scenario, of course, would be to keep the sample for a whole year and refresh it thereafter, hereby measuring people’s activity and travel behaviour at each of the regular events (summer holidays, religious holidays, yearly returning long week-ends or typical days like mother’s day, father’s day, labour day, etc.). This is very positive in the sense that one can observe the rhythm of the activity patterns and that one has a clear image of the locations where the activities take place. However, it is very unrealistic to keep people motivated to write down everything they have done - and more - for over a whole year. Moreover, since we are not interested in activity patterns for a specific individual in this paper, but more in the travel behaviour of the population as a whole, it is not necessary to have the same individuals in the survey for a whole year. What we do have to pay attention to is that every sample is a representative sample of the whole population.

A first scenario is to ask respondents to register their activities twice every week, perhaps with some more density around a regular event. This would come down to about 120 measurements for each respondent. However, one should be realistic and acknowledge the problem of fatigue after a while.

Judging by Axhausen et al. (2006), six weeks can be a break-point for fatigue in activity-travel surveys. Therefore, a second scenario would be to ask respondents to write down their activities in some kind of travel diary for six weeks, meaning 42 measurements on each person, and the sample would be refreshed every six weeks.
Since, in this latter option, we have a clearer view of people's travel behaviour during an entire week, instead of some snapshots of different weeks, we are inclined to opt for this scenario. Some other sampling issues will be discussed further. And, if a test survey would reveal that fatigue appears sooner than at the six-week breakpoint, an adjustment to the time frame will have to be made.

Sources of error
There are several potential sources of error and each data collection method is susceptible to each kind of error in varying degrees. Although it is impossible to remove the error completely, it can be minimised. The extent to which each error needs to be minimised will depend on the purpose of the survey and the accuracy required. The four major sources of error that are usually regarded are:

- ***Coverage error***: are all groups of interest accurately represented in the frame? Does everyone in the target population have a chance of selection in the sample?
- ***Sampling error***: is the sample large enough to accurately reflect the population?
- ***Measurement error***: are answers to questions being reported and recorded accurately? Do they reflect the respondent's actual views? Have errors been introduced by relationship between the respondent and the people doing the survey? Is the question being asked really capturing the concept you are trying to measure?
- ***Non-response error***: do the people who did not respond to the survey belong to a particular group and is their non-response likely to significantly affect the results of your study?

These are all very important questions that have to be addressed within this context.
Concerning *coverage error*, we can state that this seems not an issue, since we take a random sample of the population.
If the sample size is taken large enough, at least so large such that we would cover the same amount of people after one year as would be covered by the cross-sectional survey, we assume that *sampling error* is not more a problem then in the cross-sectional case (see also next sub-section).
In order to reduce the *measurement error*, a test survey will be conducted first to filter out most of the problems. However, these latter two categories of errors seem to be the hardest to control.

The *non-response error* remains a very important type of error and the only thing one can do to get an initial idea on its magnitude is to try to question people who don't respond by asking them for their reason of non-response or for a reason of drop-out, if some response was recorded. Some more detail will be provided in the next Section.

Sample size determination

It is very hard to find an unambiguous formula for the determination of sample sizes in a longitudinal survey design. In clinical trials or experimental settings where there are some subjects exposed to a certain risk and when you know the exact number of people exposed and those not exposed, it is not so hard to determine the sample size as described in Bryant and Morganstein (1987). However, everybody is exposed to transportation, even those people who do not use any motorised vehicle, are still exposed, cause they can go by bike or on foot. So these formulas can not be used in our study.

Diggle *et al.* (2002) describe the formula for the estimation of the sample size for a comparison of two groups at a single time point (N_S) and they provide an extension to the comparison of two groups across time (longitudinal, N_L). This is done, as shown in the next equation, by multiplying the determined number of subjects at a single time point by a factor, i.e.

$$N_L = N_S \times \frac{1+(n-1)\rho}{n}.$$

As we can see, this factor accounts for the correlation (ρ) between the repeated measures, as well as the time regarded (*n* is the number of time points). We clearly observe that for 1 time point or for a perfect correlation between the repeated measures, we get the same number of subjects as in the design with a single time point. If the correlation equals zero (cross-sectional design), we can observe that for a longitudinal design, we have to divide the number of subjects by the number of time points to get the same result as if we would have used just a single time point. This seems logical.

To be able to use this formula, we need to know the number of time points and the possible correlation between them. The number of time points depends on the scenario that will be chosen, so this can go from 42 (in the six week scenario) up to 365 (in the "ideal" scenario). From a test phase of another study, which is very much alike the one proposed here (i.e. having people fill out an activity-diary for over a week), we try to determine the correlation between the number of activities that were carried out between successive days. The correlation ranged from 0.4582 to 0.8477. If we consider the highest of all correlations, we

certainly have enough subjects. With a confidence of 95% and an accuracy ranging from 0.1 to 0.01, we can see quite some differences in sample sizes, as can be found in Table 1.
For reasons of comparison: the sample sizes (N_S) for the same confidence level and accuracy range at a single time point are respectively 96, 384, 2401 and 9604.

Accuracy	"Ideal" scenario 365 time points	First scenario 120 time points	Second scenario 42 time points
0.1	82	82	82
0.05	326	327	328
0.02	2037	2039	2045
0.01	8146	8154	8176

Table 1: Sample size determination

We can clearly observe an increase in the number of subjects that need to be questioned if the number of time points decreases. However, although there is a large difference in number of time points, the increase in number of subjects needed is rather low, so this would not have a large impact on the choice of the final scenario. Note that in this sample size calculation the non-response was not yet accounted for. This will be discussed in the next subsection.

Non-response

This will prove to be a very important issue in this type of research, since filling in the questionnaire has proved to be quite burdensome in previous cross-sectional studies.

Actually, there are two types of non-response that need to be considered: item non-response and unit non-response. Item non-response is regarded as the "failure to obtain 'full' and complete data from each respondent" (Zmud and Arce, 1997). Part of the item non-response can be captured by using new technologies (Internet, Personal Digital Assistant) in which consistency checks can be built-in (e.g. "you cannot proceed unless you fill out this question"). So the right choice of the survey instrument might resolve part of the problem of item non-response. Another possible solution for item non-response might be to impute the missing responses, though this should be done with great caution! Just doing an analysis on the complete cases or filling in the average on that particular item is often not sufficient.

The second type of non-response is unit non-response. This is the absence of information from some part of the target population of the survey sample (Black and Safir, 2000). This is a significant and growing problem in household travel surveys that needs to be dealt with in

an appropriate way. As stated in Stopher *et al.* (2004), there are two broad categories in unit non-response: refusals (hard and soft refusals and drop-outs) and non-contacts (busy, no reply or answering machines). As stated above, the number of non-contacts is growing if the recruitment of the travel survey is still carried out by phone.

High rates of unit non-response are generally associated with non-response error (see also in the subsection on sources of error), which can be seen as a function of the non-response rate and the difference between respondents and non-respondents to the statistic of interest (Keeter *et al.*, 2000). As can be learned from Groves and Couper (1999) and Stopher *et al.* (2004), characteristics from non-respondents to travel surveys are more likely to be people with a low or a high household income and households with low or high mobility rates. A possible solution for this can be to over-sample these type of households, though this may violate the representativeness of the sample.

Some standards and guidelines are recommended as a means to attempt to reduce the phenomenon of unit non-response. Examples are the use of pre-survey monetary incentives, the use of a pre-notification letter and reminders, reducing the respondent's burden, an increased effort in contacting the households that are difficult to contact and to undertake a non-response survey. The latter might also give an initial idea on the non-response error.

Future research

The most important decisions regarding sample design, sample size and survey methods will be taken after a test survey has been carried out. An initial idea about non-response will help to set the sample size and to determine the appropriate method of data collection.

Future research will mainly focus on the analysis of the data. If we take the individual as basis of analysis, we have to account for the fact that the data are clustered in households as well as correlated over time. However, if, as was suggested in the section on data requirements, only the difference between mobility on "normal" days versus regular events is aimed at for a whole population, then taking care of the time aspect will probably suffice.

References

Armoogum, J., Chlond, B., Madre J.-L., and Zumkeller, D. (2004) Panel surveys. Workshop paper presented at the ISCTSC Conference, Costa Rica.

Axhausen, K.W., Löchl, M., Schlich, R., Buhl, T. and Widmer, P. (2006) Fatigue in long duration surveys. *Transportation*, forthcoming.

Axhausen, K.W., A. Zimmermann, S. Schönfelder, G. Rindsfüser and Haupt, T. (2000) Observing the rhythms of daily life: a six-week travel diary. *Arbeitsberichte Verkehrs- und Raumplanung*, **25**, Institut für Verkehrsplanung und Transporttechnik, ETH, Zürich.

Ben-Akiva, M.E. and Bowman, J.L. (1998) Integration of an activity-based model system and a residential location model. *Urban Studies*, **35(7)**, 1231--1253.

Black, T. and Safir, A. (2000) Assessing non-response bias in the national survey of America's families. *Proceedings of the Survey Research Methods Section of the American Statistical Association*, 816--821.

Bryant, E. and Morganstein, D.R. (1987) Sample size determination for longitudinal surveys. *Proceedings of the Survey Research Methods Section of the American Statistical Association*, 189 – 193.

Diggle P.J., Heagerty, P., Liang K.-Y., and Zeger, S.L. (2002). *Analysis of Longitudinal Data - 2nd edition*. New York: Oxford University Press.

Groves, R.M. (1989) *Survey Errors and Survey Costs*. John Wiley and Sons, Inc.

Groves, R.M. and Couper, M.P.(1998) *Nonresponse in Household Interview Surveys.* John Wiley and Sons, Inc.

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.

Keeter, S., Miller, C., Cohut, A. Groves, R.M. and Presser, S. (2000) Consequences of reducing nonresponse in national telephone survey. *Public Opinion Quarterly*, **64(2)**, 125--148.

Provinciale Hogeschool Architectuurinstituut. (1993) Onderzoek verplaatsingsgedrag in Vlaanderen (Travel behaviour research in Flanders) – Feasability study. Final report.

Stopher, P.R., Wilmot, Ch.G., Stecher, Ch. and Alsnih, R. (2004) Household travel surveys: Proposed standards and guidelines. Paper presented at the ISCTSC Conference, Costa Rica.

Weijermars, W.A.M. and van Berkum, E.C. (2004) Variations in urban road traffic - Analysis of spatial and temporal variations in urban traffic. *Electronic proceedings of the 8th Trail conference World of transport, infrastructure and logistics.*

Zmud, J., and Arce, C. (1997). Item nonreponse in travel surveys: Causes and solutions. *Transport Surveys: Raising the Standard*. Proceedings of an International Conference on Transport Survey Quality and Innovation.