THE OPTIMAL DURATION FOR A TRAVEL SURVEY
– Empirical Observations –

Metin SENBIL
Ph.D., Department of City and Regional Planning
Gazi University
Ankara, Turkey

Ryuichi KITAMURA
(Deceased)
Former Professor of Civil Engineering
Department of Urban Management
Kyoto University
Kyoto, Japan

(Received September 15, 2009)

Empirical observations were carried out using a six-week travel survey, Mobidrive, in order to probe into the optimal duration problem of a single wave travel survey. The regressions on certain travel elements (i.e., number of daily trips, VKM per day, number of home based trip chains, daily travel time expenditure) suggest that a two-week survey duration might serve well to represent personal variability in various dimensions. However further analysis carried out by using nonparametric survival analysis points to a longer period as the elapsed time between activity participation does not seem to converge to true values, at least for rare activities such as recreation.

Key Words: Multi-day travel survey, Optimal survey duration, Mobidrive, Inter-personal variation, Nonparametric survival analysis

1. INTRODUCTION

Relatively little attention has been directed to the duration of multi-day travel surveys. This is particularly the case in Japan where most large-scale household travel surveys have adopted one-day data. Starting from the 1970s, a new interest in travel behavior research has emerged in the area of the variability and dynamics of travel behavior, which was a consequence of the paradigm shift in transportation planning from road expansion to travel demand management. This inevitably demanded more insights into individual or group behavior such as daily activity-travel dynamics, rhythms observed in certain time periods, decision making processes, etc.; accordingly multi-day and/or multi-period surveys have gained increasing popularity. An early example, the Uppsala, Sweden, household travel survey conducted in 1971 had a 35-day duration\(^1\)\(^2\). A 1973 activity survey in Reading, England, had a seven-day survey period\(^3\). Later examples include the Dutch National Mobility Panel survey\(^4\) and the Puget Sound Transportation Panel survey\(^5\). A recent example, the Mobidrive survey in Germany, offered a multi-week data collected in 1999\(^6\).

Based primarily on the studies by Pas and his colleagues\(^3\)\(^7\)\(^-\)\(^10\), recent surveys in the United States have adopted multiple days\(^5\). Pas and colleagues decomposed the total variability of travel behavior into inter-personal variability and intra-personal variability\(^9\). In this new approach, variation between individuals refers to inter-personal variability, variation that one person exhibits over different days refers to intra-personal variability. These variations can further be decomposed into systematic (or explained) and stochastic (or unexplained) parts. Therefore, in addition to variation between individuals, multi-day surveys offer day-to-day variation in travel behavior. Moreover, multi-day data facilitate in-depth investigation of activity-travel behavior such as scheduling behavior.

The analytical framework developed by Pas and colleagues was limited to the identification of trip rates along with survey costs. The desirable number of survey days, however, would vary depending on what one wishes to measure and at what level of accuracy. In other words, the optimal survey duration is a function of the purpose of the survey and the level of precision required of it. Besides, one has to consider the trade-off between survey cost and precision: precision of various observations improves with the survey duration which, in turn, increases the survey cost. Furthermore, precision depends on the sample size too. However, it is obvious that a longer duration is needed to measure characteristics of trips that take place only infrequently. Therefore, for best results, survey duration, sample size and survey cost must be balanced for optimal survey design.

In this study, we offer new insights into optimal survey duration by presenting empirical evidence obtained from a multi-day travel survey, Mobidrive, conducted over six-weeks collected information about time-use by 362 individuals in two cities of Germany, i.e., Karlsruhe.
and Halle, in 1999. As a preliminary stage toward this end, we show the relationship between estimation quality and survey duration in a modeling framework. We use one wave of the survey only and disregard the further multi-period benefits. In the rest of the study, we firstly supply information that form the background to the analyses carried out in the next section, empirical findings. Panel linear regression is used for models that account for several mobility variables in the empirical findings section. Besides, elapsed time between two consecutive activity participations is presented for various activities. Our conclusions are presented in the last section.

2. BACKGROUND

In designing panel surveys, the spacing between survey waves and the duration of a wave represent two of the important issues. The second author of this study elaborated on the spacing between survey waves. In that study, the spacing between waves is related to the probability of drawing true inferences about the phenomenon of interest, with the assumption that the transition from one behavioural state to another is a time-homogenous Markov process. The study highlights the difficulty of obtaining unbiased parameter estimates of discrete-state stochastic processes based on discrete-time observations obtained from survey waves. Another conclusion derived in the study was the very slow improvement in parameter accuracy as the spacing between survey waves decreases. Thus, the study concluded that incorporating recall questions into the surveys would be helpful in detecting behavioural transitions. The study was oriented towards capturing changes in behaviour that materialize over a longer time span, e.g., change in automobile ownership. This present study, on the other hand, focuses on the duration of a survey wave.

Based on the fact that individual and household travel behaviour varies daily, Pas and Koppelman argued that understanding day-to-day variations in travel behaviour is important for designing transportation services. To address this, multi-day data offering insights into intra-personal variability is a prerequisite. Sources of day-to-day variability encompass diverse factors such as unexpected events, weather conditions, as well as expected events recurring at regular intervals, e.g., poker games every Wednesday night. To probe into these factors, multi-day travel surveys must be developed to capture intra-personal variation in addition to inter-personal variation (Fig. 1).

Earlier studies reported findings on intra-personal variation by using different aspects of travel behaviour. For example, using multi-day data obtained from the Reading Activity Diary Survey, Pas and Koppelman observed average intra-personal variances for different population segments by using trip frequency and carried out statistical tests on intra-personal variability. The study offered empirical evidence on intra-personal variability and differences across population groups. Similar studies by Hanson and Huff, Jones and Clarke, Schlich and Axhausen analysed activity-travel patterns to detect intra-person variability.

From another perspective, Mahmassani et al examined daily trip chaining, departure time and route choice decisions of a sample of commuters in Austin, Texas. Two intra-personal variability measures developed in this study were the variation measured by inspecting successive days, and the variation measured by comparing daily measures to “usual” measures, i.e., the median departure time and the most frequently chosen route. In spite of finding significant intra-personal variation, the study concluded with a caveat that intra-personal variability is dependent on the criterion variability that is defined. For example, while a 5-minute criterion for departure time change might capture a substantial variation, a 10-minute criterion might produce a much smaller variation. A study by Mannerings, using data collected from commuters in Seattle, Washington, the other hand, reported very few changes in both departure time and route for commute trips. In the study, a change was defined as any deviation from the usual departure time and route.

Using the Uppsala data, Hanson and Huff investigated systematic components in intra-personal variation in individuals’ daily travel. This would aid in determining “the appropriate time period for collecting data on travel and modelling it” (p. 112). However, they noted the difficulty in deriving an appropriate measure of intra-personal variability, which is “sensitive to the level of detail” (p. 117), e.g., activity type at destination, time of day, mode
used, location, etc. While examining only trip-generation rates gave way to the simplest measure of intra-personal variability, examining the complete daily patterns would generate the most complex one. Hanson and Huff\(^3\) characterized the former as a “pale measure of complexity” (p. 117) since a person could make the same number of trips everyday but for different purposes, at different times, and by different modes. Accordingly, Hanson and Huff\(^3\) proposed three measures for intra-personal variation that are equivalence class, representative day, and core stops. By using these measures, Hanson and Huff assessed whether one-week worth of data are sufficient enough for capturing intra-personal variation. Their findings indicated that although the number of journeys and trips per day were stable over time, types of travel behaviour were rather different; therefore longer periods of data collection were required to collect more detailed information.

Similarly, Jones and Clarke\(^4\) elaborated on measuring variability of travel behaviour over time too. They proposed various measures by using graphical and numerical techniques. Using the same data source as the current study, Schlich and Axhausen\(^15\) carried out a thorough study by examining different variability measures proposed by Pas\(^16,17\), Hanson and Huff\(^2\), and Jones and Clarke\(^4\). Consequently, Schlich and Axhausen\(^15\) suggested that “empirical surveys should cover a period of at least two weeks” (p. 33).

The earliest attempt, known to the authors, to develop an optimum survey duration is the study by Pas\(^3\). In his study, Pas examined statistical efficiency of generalized least squares (GLS) estimators of trip generation models with both single-day and multi-day data, with the dependent variable being the number of daily trips by an individual. The relationship between survey sample size and survey duration was given as:

\[
N_M \text{var}(\beta_M) = N_S \frac{1 + a(T-1)}{T} \text{var}(\beta_S)
\]

(1)

where \(N_M\) and \(N_S\) represent sample sizes of multi-day and single-day data, respectively, \(\beta_M\) and \(\beta_S\) are the coefficient vectors of the regression models estimated using the multi-day sample and the one-day sample, respectively, \(T\) is the number of observation days in the multi-day survey, \(a\) is the “the portion of the total variance that is attributable to unobserved differences across individuals” (p. 75). In Eq. 1, the variance ratio, \(a\) (= \(\sigma_u^2/(\sigma_u^2 + \sigma_w^2)\)), is derived from error components \(u_j [\sim N(0,\sigma_u^2)]\) and \(w_{jt} [\sim N(0,\sigma_w^2)]\) in the linear regression trip generation model. In this, component \(u_j\) is unique to individual \(j\) and invariant over time, while component \(w_{jt}\) is a purely random for individual \(j\) at time \(t\). Thus, the variance ratio, \(a\), is in fact represents the error correlation between two days for individual \(j\).

The equality in Eq. 1 indicates that, with equal sample sizes, estimation derived from multi-day observation is more efficient than the one obtained from the single-day observation given that the weight factor \(a\) is less than one. However improvement in multi-day parameter, \(\beta_M\), efficiency with respect the single-day parameter, \(\beta_S\), converges to the weight factor \(a\) as the survey duration, \(T\) increases given that the sample sizes are equal and the variance of parameter estimated from single-day data is unity (Fig. 1). Also given a smaller value of variance ratio than unity, i.e., \([0<a\leq 1]\), smaller sample size for the multi-day survey can achieve the same level of estimator variance with the single-day survey data (Fig. 2). In both cases, the condition of smaller than unity variance ratio requires the existence of intra-personal variation.

As illustrated above in terms of parameter efficiency, it is also desired that a longer duration is needed to measure certain types of activity-travel behavior too. For example, there are activities that take place regularly at a fixed location and time (e.g., private tennis lesson at a sports club every Thursday) or irregular, infrequent activity engagement (e.g., going to an opera). Another example can be given for travel behavior such as commuter trip departure time or route choices that necessarily call for longer observation periods. However, observation pe-

![Fig. 2 Number of days in the survey (adapted from Pas, 1986, p. 77).](image-url)
period increases with respondent fatigue which decreases response rates\textsuperscript{18,19}.

Studies conducted recently by Stopher and his associates\textsuperscript{20-22} illustrated various advantages of multi-day data over one day data. It was found that means of travel variables such as daily trips are significantly different from one day data\textsuperscript{20}. On the other hand, they highlighted the importance of overcoming respondent fatigue associated with multi-day data collection by passive data recording technologies using GPS recording devices or reading odometers\textsuperscript{21,22}. Technological breakthroughs bring a significant leap towards collecting multi-day data and reduction in sample sizes. However, while these recording technologies might be useful in recording certain aspects of travel behavior, important aspects of activity-travel behavior such as activity pursued, accompanying persons etc. might still require active recording.

3. EMPIRICAL FINDINGS

In this section we present findings using Mobidrive dataset\textsuperscript{6}. Accordingly, we firstly conducted regression analyses on certain mobility measures (i.e., the number of person trips per day, vehicle kilometers traveled (V KM), number of home-based trip chains, and total travel time by a person). We compute intra-personal variability parameter, $\alpha$, given in Eq. 1 by using estimated variances of error components in the regression analyses. This is followed by analyses on elapsed time between two consecutive participations of an activity. Using results of both analyses, we derive information pertaining to minimum duration required for a survey.

The data set is based on six-week travel diaries kept by a total of 317 persons over 6 years of age, from 139 households in Karlsruhe and Halle, two German cities of about 270,000 inhabitants each. The survey duration of six weeks is sufficiently long to investigate dynamic aspects of travel behavior in a wave. The data used in the regression analyses are based on the first wave of travel diaries completed by 91 individuals in Karlsruhe.

In this study, 42 days worth of data sets are divided into shorter multi-day datasets with lengths ranging from 2 to 42 days. Thus, 41 separate datasets are produced from the original data with the number of cases depending on the number of days. A random effect linear regression model\textsuperscript{23} is applied to each of the datasets thus created. The model contains an error component (equivalent to $u_j$ in the previous section) and facilitates the estimation of the variance ratio, $\alpha$. Accordingly, we derived inferences on survey duration based on how estimated parameter, $\alpha$, varies with the observation period.

The dependent variables of the regression models are the number of person trips per day, vehicle kilometers traveled (V KM), number of home-based trip chains, and daily travel time expenditure by a person (see Appendix I). Regression models have shown that the variance ratio, $\alpha$, decreases with survey duration, as expected (Fig. 3).

Evident in all regression results presented in Figure

![Figure 3](image-url)
3 is a sharp decrease in the variance ratio as the observation period increases from two to seven days, suggesting influences of weekly activity cycles on the variance ratio. Also notable are the different patterns of change in variance ratio among the dependent variables. Nonetheless, it appears that one week is a threshold in the consideration of survey duration.

Also evident in Figure 2 is the weekly cyclic pattern in estimated variance ratios after the first seven days of observation. In almost all cases, the variance ratio first increases in the first half of a week and then decreases. However, there are anomalies for the number of daily trips and the number of home-based trip-chains on days 27 and 28, respectively. We were not able to explain this abrupt change in the variance ratio. Based on the results of the regression models, we can conclude that a week seems a logical unit for observing both inter- and intra-personal variations.

Nevertheless, an observation period longer than a week might be desirable in order to delineate a reasonable map of individual activity-travel behavior. Therefore, in addition to panel regressions which located the relationship between parameter efficiency and the observation period, another line of analysis can be adopted by observing activity participations. For this, one may be interested in how often one participates in a given activity, therefore might refer to rhythms in activity-travel behavior. Pertinent measure for this is the mean elapsed time between two consecutive engagements of an activity. Figure 4 presents change in this measure by observation period for ten different activities.

The average elapsed time between two consecutive participations in an activity can be calculated only when there are at least two participations of an activity; if there is only a single or no observation, the case is excluded in the tabulation. According to Figure 4, regular activities like work, school, and return home show stable averages once the observation period exceeds a week, however irregular activities like shopping and recreation reveal average elapsed times increasing with the observation period. Therefore, the optimum survey duration is likely to be longer than one week.

Assuming that optimal survey duration is longer than a week, we investigated how parameter efficiency changes for survival and hazard rates of activity participation. Accordingly, we organized six data sets as multiples of a week, i.e., one, two, …, or six weeks of observation. Using these six data sets, non-parametric analysis of the elapsed time between successive activity participations was performed. Only the irregular, “other” activity was examined here. The intent here was not to estimate the hazard or survival rate of activity participation, but to observe the change (or improvement) in their standard errors (see Appendix II for the method used). Figure 5 presents estimated standard errors.

For infrequent activities grouped as “other” activity, there was a smooth decrease in standard errors for hazard and survival rate with the number of weeks—there was an approximately 50% improvement in the estimated standard errors as the observation increases up to six weeks. Together with the results on the average elapsed time between activity participations, it appears more appropriate to conduct a travel survey longer than a week; therefore, one week may be considered as the minimum duration for a travel survey. Since the analyses of this study are concerned only with the quality of estimates, which improves as the survey duration increases, the study does not offer empirical results concerning the upper limit which depends on other variables such as survey cost, survey and item non-response, respondent fatigue and other problems that worsens with survey duration.

4. CONCLUSIONS

With the intent of observing how survey duration influences estimates of individuals’ mobility indicators, synthetic data sets of various durations were created in this study using the six-week Mobidrive data set. We found that the fraction of the intra-person variability for the mobility measures adopted in this study, and inter-activity elapsed time both decrease as the observation period, i.e., travel survey duration, increases up to one week. Longer than a week observation revealed stabilized values of intra-personal variation, but the efficiency of hazard and survival rates still continue to decrease with the observation period. Our tabulations of average elapsed time between successive activity participations for different types of activities have revealed that, for regular activities such as work and school, the average elapsed time generally stabilizes in a week; but for irregular activities such as shopping or recreation, the average elapsed time increases with the observation period. This suggests that the true values might be obtained by observing travel behavior longer than a week. Standard errors for hazard and survival rate of activity participation support this finding and favor longer survey durations.

However, longer observation periods, especially more than two weeks, might pose serious problems. First of all, all surveys are subject to a financial constraint which sets a strict limitation on the observation period. Even if
Fig. 4 Average elapsed time between two consecutive activity participations
one may overcome this physical barrier, respondent frustration still stands as another important barrier. In order to overcome this barrier, certain technological breakthroughs have been incorporated into the travel surveys. While technology allows us to observe certain mobility variables for longer periods, activity-travel behavior is not fully reducible to passive recording techniques, the need for active participation of individuals into the travel surveys still exist for complex activity-based travel demand models.

REFERENCES


ACKNOWLEDGEMENTS

We kindly thank Prof. Kay Axhausen and his colleagues at The ETH Zürich for providing the dataset. Some parts of this study were completed after Prof. Kitamura’s passing. We would like to thank the anonymous reviewers for their comments.
This appendix provides remaining details of the regression analyses reported in Section 4. In the analyses, seven independent variables other than the constant value are used in the regression analysis. All of them refer to socio-economic and demographic variables of the individual. The data in the regressions are retrieved from the part of the survey gathered from the city of Karlsruhe at Wave 1 covered between September 20, 1999 and October 31, 1999. Data prepared for regression analyses has provided information on 91 individuals; Table A1 provides basic descriptive statistics of pertinent variables used in the analysis.

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>Min.</th>
<th>Max.</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Person trips per day</td>
<td>3422</td>
<td>0</td>
<td>15</td>
<td>3.97</td>
<td>2.03</td>
</tr>
<tr>
<td>VKM per day (meters)</td>
<td>3422</td>
<td>0</td>
<td>79,968</td>
<td>10,937</td>
<td>11,209</td>
</tr>
<tr>
<td>Number of home-based chains</td>
<td>3422</td>
<td>0</td>
<td>8</td>
<td>1.68</td>
<td>0.83</td>
</tr>
<tr>
<td>Total travel time</td>
<td>3422</td>
<td>0</td>
<td>830</td>
<td>49.05</td>
<td>78.23</td>
</tr>
<tr>
<td>Sex (1: male, 0: female)</td>
<td>91</td>
<td>0.51</td>
<td>0.50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>91</td>
<td>7.00</td>
<td>85.00</td>
<td>45.21</td>
<td>18.83</td>
</tr>
<tr>
<td>Household income (in 10 thousand DM)</td>
<td>91</td>
<td>0.00</td>
<td>9.00</td>
<td>3.75</td>
<td>2.44</td>
</tr>
<tr>
<td>Main car user dummy (1: if main user of one of household cars, 0: else)</td>
<td>91</td>
<td>0.44</td>
<td>0.50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employment dummy (1: employed; 0: else)</td>
<td>91</td>
<td>0.00</td>
<td>1.00</td>
<td>0.55</td>
<td>0.50</td>
</tr>
<tr>
<td>Parent dummy (1: parent, 0: else)</td>
<td>91</td>
<td>0.00</td>
<td>1.00</td>
<td>0.26</td>
<td>0.44</td>
</tr>
<tr>
<td>Marriage dummy (1: married, 0: else)</td>
<td>91</td>
<td>0.00</td>
<td>1.00</td>
<td>0.60</td>
<td>0.49</td>
</tr>
<tr>
<td>Weekend dummy</td>
<td>3422</td>
<td>0.27</td>
<td>0.44</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Non-parametric analysis of activity participation has been carried out according to the methodology proposed by Cutler and Ederer\(^2\). Standard error of survival rate \(P_j\) and hazard rate \(\lambda_j\) are calculated as follows:

\[
P_j = (1-q_j)p_{j-1} ; p_1 = 1
\]

\[
se(P_j) = P_j \left( \sum_{k=1}^{j-1} \frac{m_j/r_k}{r_j(1-m_j/r_k)} \right)^{1/2}
\]

\[
\lambda_j = \frac{2m_j/r_j}{h[2-(m_j/r_j)]}
\]

\[
se(\lambda_j) = \lambda_j \left( \frac{1-(h\lambda_j/2)^2}{m_j} \right)^{1/2}
\]

\[n_k\] number of all individuals observed on day \(k\)
\[m_k\] number of activity participants on day \(k\)
\[C_k\] number of observations censored on day \(k\)
Number of observations who can pursue activity \(r_k = n_k - 0.5C_k\)
Activity participation rate \(q_k = m_k / r_k\)
\(h\) is the interval width (that is one day)