Estimating the parameters of a dynamic need-based activity generation model

Citation for published version (APA):

Document status and date:
Published: 01/01/2009

Document Version:
Publisher’s PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:
• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher’s website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:
www.tue.nl/taverne

Take down policy
If you believe that this document breaches copyright please contact us at:
openaccess@tue.nl
providing details and we will investigate your claim.

Download date: 01. Apr. 2021
Estimating the Parameters of a Dynamic Need-Based Activity Generation Model

Linda Nijland, Theo Arentze and Harry Timmermans, Eindhoven University of Technology, The Netherlands

Abstract Several activity-based models made the transition to practice over the last decade. However, modeling dynamic activity generation and especially, the mechanisms underlying activity generation are not well incorporated in the current activity-based models. This paper describes a first step in estimating the parameters of a need-based activity generation model. A survey was carried out to collect activity data for a typical week and a specific day among a relatively large sample of individuals. The diary data includes detailed information about activity history and future planning. Furthermore, person-level needs on relevant dimensions were measured using Likert scales. Estimation of the model involves a range of shopping, social, leisure and sports activities, as dependent variables, and socioeconomic, day preference, and need variables, as explanatory variables. The results show that several person, household and dwelling attributes have an influence on activity-episode timing decisions in a longitudinal time frame and, thus, on frequency and day choice of conducting the social, leisure and sports activities.

1. Introduction

There has been considerable progress in development and application of activity-based models over the last decade. Examples of fully operational models are CEMDAP (Bhat and Singh 2000), Famos (Pendyala et al. 2005), TASHA (Roorda et al. 2007), and Albatross (Arentze and Timmermans 2000). Currently, the models are making the transition to practice where they find application as instruments for planning support and policy evaluation. However, there is still ample room for improvement. High on the research agenda are the generation of activities...
based on the needs they satisfy or induce, interactions between activities, scheduling at the household level and activity scheduling for a multi-day period.

Mechanisms underlying activity generation are still poorly understood and not-well represented in current activity-based models (Habib and Miller 2007; Roorda et al. 2007). The notion that daily activities of individuals are driven by basic needs lies at the core of the activity-based approach since the pioneering work of Chapin (1974) and is further emphasized by Miller (2004) and Axhausen (2006). Miller derived some elements of his framework for modeling short- and long-term household-based decision making from Maslow’s hierarchy of needs. Meister et al. (2005) partially implemented needs into their operational model of activity scheduling.

Arentze and Timmermans (2009) developed a theoretical framework based on the assumption that activities are driven by a limited and universal set of subjective needs at person and household level. The needs grow autonomously over time according to a logistic curve with parameters depending on the nature of the need and characteristics of the individual and the household. The model predicts the timing and duration of activities in a longitudinal time frame taking into account time budget constraints, possible interactions between activities, and both household-level and person-level needs. The results of numerical simulations supported the face validity of the suggested framework and modeling approach, demonstrating the possibility of incorporating positive or negative substitution effects between activities and complex dynamic interactions between activities in general. Up till now, however, their approach lacks empirical validation.

The goal of the research project underlying this current paper is to test the suggested approach empirically and to estimate parameters of supposed relationships using data specifically collected for that purpose. The present paper describes the results of a survey, designed to model and predict the timing of activities with respect to underlying needs. The questionnaire focuses on social, leisure and sports activities (as those activities are most likely to be substitutable), a typical week and a specific sampled day. Shopping and some service activities (e.g., going to the library, post office) were included as well, as those activities more or less complete the daily activity agendas. Factors included in the survey consist of socioeconomic and demographic variables, activity history and future planning variables (e.g., time elapsed since last performance), available time for discretionary activities, and scores on (statements of) needs. The survey was held among a sample of approximately 300 individuals through a web-based questionnaire.
The organization of the paper is straightforward. First, we will briefly summarize the need-based concepts and model. This is followed by a description of the survey and the sample. Section 5 describes the results of the parameter estimations. The paper closes with a discussion of the main findings of the study and remaining problems for future research.

2. Need-Based Model

In this section we will briefly outline a model for predicting the timing of activities in a multi-day time frame that is proposed in Arentze et al. (2008, 2009). The model is based on concepts from a more theoretical needs-based model of activity generation, which we cited above, and has parameters that should be identifiable based on activity diary data.

The model predicts a multi-day activity pattern for a given person for a period of arbitrary length. Rather than solving some resource allocation optimization problem, the model assumes that individuals make activity-selection decisions on a daily basis. Although the model is able to take into account interactions between activities and between persons (in a household context), we will consider here a more limited situation where an individual is faced with a decision to conduct an activity \( i \) on a current day \( d \) given that the last time the activity was conducted was on day \( s < d \) (this means that the time elapsed equals \( d - s \) days). To simplify notation, we drop the subscript for individuals in the equations below.

In the basic model, the utility of conducting an activity of type \( i \) on a given day \( d \) is defined as:

\[
U_{sd} = V_{sdi} + V_{di} + \varepsilon_{sdi}
\]

where \( d \) is the current day, \( s \) is the day activity \( i \) was conducted the last time before \( d \), \( V_{sdi} \) is the utility of satisfying the need for activity \( i \) built-up between \( s \) and \( d \), \( V_{di} \) is a (positive or negative) preference for conducting activity \( i \) on day \( d \) and \( \varepsilon_{sdi} \) is an error term.

The utility components can be interpreted as follows. The first term \( (V_{sdi}) \) represents the amount of the need that has been built up across the elapsed time and that will be satisfied if the activity is implemented. The second term \( (V_{di}) \) represents a base utility dependent on preferences for day \( d \). Note that events that are not driven by needs, but rather take place on a certain fixed day, can be modeled as activities with zero need growth.
\((V_{sd} = 0)\) and a relatively high utility for the day \((V_{dd} >> 0)\) when the event is to take place.

Implied by the first term is that a need for an activity grows over elapsed time since day \(s\). There are several functional forms conceivable for a need’s growth curve. The original model assumed a logistic growth function, but also suggests that under normal conditions need growth only moves around the area around the reflection point where the curve is approximately linear. To reduce the number of parameters, we therefore assume a simple linear function here:

\[
V_{sd} = \beta_i t
\]

where \(\beta_i\) is a growth rate and \(t\) is the length of the need growth period between \(s\) and \(d\) \((t = d - s)\).

A decision heuristic that takes into account limited time-budgets states that an activity \(i\) should be conducted on day \(d\) if \(d\) is the earliest moment when the utility of the activity per unit time exceeds a threshold. The utility-of-time threshold imposes a constraint on activity generation and represents an individual’s scarcity of time. The smaller a time budget for activities, the larger the threshold needs to be. When the threshold is well adjusted, the rule leads to fully use of available time (i.e., the budgets are exhausted). At the same time, the rule ensures that every activity generates approximately an equal utility per unit of time when it is conducted. In that sense, the heuristic, even though it is very simple, will lead, as a tendency, to patterns where the utility of activities across a longitudinal period cannot be improved by a revision of activity timing decisions when thresholds are well-adjusted to existing time budgets.

As a first step in estimating the model, we will leave activity duration out of consideration here. Incorporating this choice facet is left for future research. This means that we assume here that the threshold is defined on the level of utility of the activity rather than utility per unit time. The decision rule then becomes, conduct the activity at the earliest moment when the following condition holds:

\[
U_{sd} \geq \mu_d
\]

where \(\mu_d\) represents a threshold for implementing activities on day \(d\), given all existing time demands on that day. Note that defined in this way, the need-growth parameter \(\beta\) for some activity will capture the time needed to overcome the threshold taking into account a (average) duration of that activity. For example, keeping everything else equal, the need-
growth speed will be smaller, i.e. it takes longer to overcome the threshold, if the activity has a longer duration.

Assuming the error term is Gumbel distributed, we can use a logit framework for calculating the choice probabilities as follows ($V_{idt} = V_{di} + V_{sd}$):

\[
P_t(d \mid s) = \frac{\exp(V_{i,d,t} - \mu_d)}{1 + \exp(V_{i,d,t} - \mu_d)} - \max_{k=1}^{t-1} \left[ \frac{\exp(V_{i,d+k,t-k} - \mu_{d+k})}{1 + \exp(V_{i,d+k,t-k} - \mu_{d+k})} \right], \quad t > 1
\]

\[
P_s(d \mid s) = \frac{\exp(V_{i,d,t} - \mu_d)}{1 + \exp(V_{i,d,t} - \mu_d)}, \quad t = 1
\]

(4a)

(4b)

Note that the conditional probabilities sum up to one across days after $s$:

\[
\sum_{d>s} P_t(d \mid s) = 1
\]

(5)

Thus, $P$ defines a choice probability distribution across days after $s$. In other words, the model predicts for a given activity and individual the probability of an interval time ($t = d - s$), thereby taking into account possible day-varying conditions related to day preferences and time budgets, in addition to need build-up rates. In that sense the model is more flexible than existing hazard models for predicting inter-episode durations of activities. Incorporating interactions between activities requires an extension of structural utility terms. We leave this, however, for future research.

3. Design of the Survey

In order to estimate the parameters of the above model, data had to be collected. The questionnaire was administered through the internet to reduce respondent burden and shorten the data entry time. In total, 37 social, sports, leisure and service-related activities were included in the survey. The questionnaire consisted of six different parts. For estimating the parameters we focus on five of them, namely:
- Socio-economic and demographic variables; e.g. gender, age, household composition, income, dwelling type, education level, number of children, age youngest child, living area, car availability, and driver’s license.
- The activity pattern of the day before; the activities the subjects conducted the day before they filled out the questionnaire and some characteristics of those activities (e.g., duration, travel time, planning time horizon, and accompanying persons)
- History: The last time subjects conducted the activities; respondents had two ways to indicate this. First, they could indicate the date, which could be selected with the help of a calendar. Second, they could indicate how many days, weeks or months ago they last performed the activity. A third option was n/a (not applicable) which could be marked if it was longer than 6 months ago or if they never do the activity. The history information was requested for the exhaustive list of 37 activities (not just the activities conducted on the day before).
- Future: If and when the activities were already planned; similar as in the previous part, respondents could indicate the date. If they did not know the date yet, they could indicate in which term they were planning to conduct the activity. Not applicable (n/a) could be marked if the subject did not plan the activity (yet).
- Needs: Scores on statements for six needs; preceding the questionnaire described in this paper, two surveys were carried out to identify and establish the needs underlying activity generation (see Nijland et al. 2010). This resulted in six needs, namely Physical exercise, Social contact, Relaxation, Fresh air / being outdoors, New experiences, and Entertainment. For each of the needs, four statements were included in the current questionnaire as indicators of the need: two of them were positively oriented and the other two negatively. The statements generally started with: “I think it is important to …”, “I like to …” and “I have hardly any need for …”. Using Likert scales, subjects had to indicate to what extent they agreed with the statements (totally disagree, disagree, neutral, agree or totally agree). For each need, sum of scores across items was taken as a measure of the size of the need of the person.

4. Sample

Subjects were selected from a sample of neighborhoods in the Eindhoven region. In the last two weeks of June 2009, 4000 invitation cards were distributed to households in the selected neighborhoods. Furthermore,
individuals who in an earlier survey (Sun et al. 2009) had indicated their willingness to participate again in an Internet survey were approached by e-mail. In this way, approximately 400 individuals were invited additionally to participate in the survey. As an incentive, twenty vouchers of 50 Euros were allocated to respondents through a lottery. In total, 438 individuals started and 290 of them completed the questionnaire.

Table 1 describes the sample and the Dutch national population with regard to some relevant socio-economic variables. The sample is reasonably representative except that above-average educated groups are overrepresented. This bias is typical for surveys in general (Adler et al. 2002; Bricka and Zmud 2003). The elderly (65+ years) and young persons (< 25 years) are somewhat underrepresented and households consisting of two persons (married or living together) are a little overrepresented.

**Table 1. Composition of the sample**

<table>
<thead>
<tr>
<th></th>
<th>Sample (%)</th>
<th>Population (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gender</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>53</td>
<td>50.5</td>
</tr>
<tr>
<td>Male</td>
<td>47</td>
<td>49.5</td>
</tr>
<tr>
<td><strong>Age</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15 -&lt; 25 yr</td>
<td>7</td>
<td>15</td>
</tr>
<tr>
<td>25 -&lt; 45 yr</td>
<td>48</td>
<td>37</td>
</tr>
<tr>
<td>45 -&lt; 65 yr</td>
<td>34</td>
<td>33</td>
</tr>
<tr>
<td>65 -&lt; 85 yr</td>
<td>10</td>
<td>16</td>
</tr>
<tr>
<td><strong>Education</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Below average</td>
<td>14</td>
<td>35</td>
</tr>
<tr>
<td>Average</td>
<td>25</td>
<td>41</td>
</tr>
<tr>
<td>Above average</td>
<td>61</td>
<td>24</td>
</tr>
<tr>
<td><strong>Household</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single, no children</td>
<td>23</td>
<td>35</td>
</tr>
<tr>
<td>Single, children</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>Double, no children</td>
<td>38</td>
<td>29</td>
</tr>
<tr>
<td>Double, children</td>
<td>33</td>
<td>29</td>
</tr>
<tr>
<td>Multiple persons</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

For the analyses in the current paper we used the activity data of which we know the interval time, namely the cases where: 1) the activity was conducted the day before and the respondent indicated the date of (or the time passed since) the last performance of the activity, 2) the activity was conducted the day before and the respondent filled out the date when the activity was planned again, 3) the activity was not conducted the day before, however the date of the last performance and the planned date were known. This resulted in about 1500 cases to be used for the analyses. Note that this constitutes a pseudo random sample of interval times: the
interval times of segments 2 and 3 are random only in as far as no correlation exists between the length of an interval time and knowing the future time it will be conducted again. Since such a correlation likely does exist, we should keep in mind that interval times of this section might be somewhat underestimated.

5. Results

Before starting the parameter estimation for all activities, we first estimated the parameters for each of six activity groups (Table 2), namely: daily shopping, non-daily/fun shopping, social visits, going out, sports and walking/cycling (as an activity). This enabled us to add and delete variables each time after estimation runs, as calculation time of the complete model was relatively long compared to likelihood estimation of ordinary logit models (note: as implied by the second term of Eq. 4a the utility of all days within an observed interval day need to be calculated for determining the probability of each observation). The significant variables per activity group were eventually incorporated in the final estimation of the need-based model (not reported here).

Table 3 shows the variables that were included in the analyses. Most variables were dummy coded, except for the scores on the needs, the number of children in the household and, for some activities, the hours spent on work/education a day. The statistical computing language R (Chambers et al. 2009) was used to program the log-likelihood function for the models and estimate the models using a standard likelihood estimation method. As indicators of \( V_{di} \) we included the days of the week as dummies in the following way (arbitrary choosing Wednesday as a reference):

\[
V_{di} = \alpha_1*\text{Mon}_d + \alpha_2*\text{Tue}_d + \ldots
\]  

(6)

where \( \text{Mon} \) and \( \text{Tue} \) are zero-one variables indicating whether day \( d \) is a Monday, Tuesday, etc. and \( \alpha \) are day-preference parameters. For \( V_{sid} \), a constant and person, household and dwelling attributes shown in Table 3 were included, as follows:

\[
V_{sid} = (\beta_0 + \sum_k \beta_kx_k) * t
\]  

(7)
where $X_k$ are attribute variables and $\beta$ are need-growth parameters. The threshold value ($\mu_d$) could be influenced by the amount of hours spent on work/education a day, e.g.:

$$\mu_{id} = \mu_{i0} + \sum_l \mu_{il} X_{ld}$$

(8)

where $X_{ld}$ are attributes influencing time budgets and $\mu$ are threshold parameters. On this level, work hours, either dummy coded or as a continuous variable depending on the model, was used as explanatory variable.

**Table 2.** Activity groups and their activities included in the estimations (base level in italics and bold)

<table>
<thead>
<tr>
<th>Activity group</th>
<th>Code</th>
<th>Activities included</th>
</tr>
</thead>
<tbody>
<tr>
<td>Daily shopping</td>
<td><em>act_ds</em></td>
<td>Daily shopping</td>
</tr>
<tr>
<td>Non-daily/ Fun shopping</td>
<td><em>act_nond</em></td>
<td>Non-daily shopping</td>
</tr>
<tr>
<td></td>
<td><em>act_funsh</em></td>
<td>Fun shopping</td>
</tr>
<tr>
<td>Social visits</td>
<td><em>act_svisit</em></td>
<td>Visiting relatives/friends</td>
</tr>
<tr>
<td></td>
<td><em>act_recsv</em></td>
<td>Receiving visitors</td>
</tr>
<tr>
<td>Going out</td>
<td><em>act_diner</em></td>
<td>Going out for dinner</td>
</tr>
<tr>
<td></td>
<td><em>act_thea</em></td>
<td>Visiting a theatre</td>
</tr>
<tr>
<td></td>
<td><em>act_conc</em></td>
<td>Attending a concert</td>
</tr>
<tr>
<td></td>
<td><em>act_cafe</em></td>
<td>Visiting a café, bar or discotheque</td>
</tr>
<tr>
<td></td>
<td><em>act_cin</em></td>
<td>Going to the cinema</td>
</tr>
<tr>
<td></td>
<td><em>act_sport</em></td>
<td>Visiting a sports event</td>
</tr>
<tr>
<td></td>
<td><em>act_day</em></td>
<td>A day out (visit a city, recreation park)</td>
</tr>
<tr>
<td></td>
<td><em>act_ebb</em></td>
<td>Play cards, billiards or bingo</td>
</tr>
<tr>
<td></td>
<td><em>act_club</em></td>
<td>Club/union activity (no sports)</td>
</tr>
<tr>
<td>Walking/cycling</td>
<td><em>act_shortw</em></td>
<td>Going for a short walk</td>
</tr>
<tr>
<td></td>
<td><em>act_wpn</em></td>
<td>Walking in a park or nature</td>
</tr>
<tr>
<td></td>
<td><em>act_cycl</em></td>
<td>Touring by bike</td>
</tr>
<tr>
<td>Sports</td>
<td><em>act_outcl</em></td>
<td>Sports outdoors, club/union context</td>
</tr>
<tr>
<td></td>
<td><em>act_outflex</em></td>
<td>Sports outdoors, flexible</td>
</tr>
<tr>
<td></td>
<td><em>act_indcl</em></td>
<td>Sports indoors, club/union context</td>
</tr>
<tr>
<td></td>
<td><em>act_indflex</em></td>
<td>Sports indoors, flexible</td>
</tr>
</tbody>
</table>
Table 3. The explanatory variables considered for the need-based model (base level in italics and bold)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Code</th>
<th>Description / range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Day of the week</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Monday</td>
<td>mon</td>
<td>Monday</td>
</tr>
<tr>
<td>Tuesday</td>
<td>tue</td>
<td>Tuesday</td>
</tr>
<tr>
<td>Wednesday</td>
<td>wed</td>
<td>Wednesday</td>
</tr>
<tr>
<td>Thursday</td>
<td>thu</td>
<td>Thursday</td>
</tr>
<tr>
<td>Friday</td>
<td>fri</td>
<td>Friday</td>
</tr>
<tr>
<td>Saturday</td>
<td>sat</td>
<td>Saturday</td>
</tr>
<tr>
<td>Sunday</td>
<td>sun</td>
<td>Sunday</td>
</tr>
<tr>
<td>Gender</td>
<td>male</td>
<td>Male</td>
</tr>
<tr>
<td></td>
<td>female</td>
<td>Female</td>
</tr>
<tr>
<td>Household composition</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single, no children</td>
<td>hh_s_no</td>
<td>Single, no children</td>
</tr>
<tr>
<td>Double, no children</td>
<td>hh_d_no</td>
<td>Double, no children</td>
</tr>
<tr>
<td>Single or Double, with child(ren)</td>
<td>hh_sd_c</td>
<td>Single or Double, with child(ren)</td>
</tr>
<tr>
<td>Living in at (grand)parents/relatives</td>
<td>hh_par</td>
<td>Living in at (grand)parents/relatives</td>
</tr>
<tr>
<td>Student accommodation, group accommodation</td>
<td>hh_stud</td>
<td>Student accommodation, group accommodation</td>
</tr>
<tr>
<td>Household income</td>
<td></td>
<td></td>
</tr>
<tr>
<td>below average</td>
<td>ibav</td>
<td>below average</td>
</tr>
<tr>
<td>average</td>
<td>ilav</td>
<td>average</td>
</tr>
<tr>
<td>Between 1 &amp; 2 times average</td>
<td>i12av</td>
<td>Between 1 &amp; 2 times average</td>
</tr>
<tr>
<td>about 2 times average</td>
<td>i2av</td>
<td>about 2 times average</td>
</tr>
<tr>
<td>more than 2 times average</td>
<td>i2pav</td>
<td>more than 2 times average</td>
</tr>
<tr>
<td>Number of children</td>
<td>nchild</td>
<td>Continuous</td>
</tr>
<tr>
<td>Age youngest child</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0 – 5 years old</td>
<td>aych06</td>
<td>0 – 5 years old</td>
</tr>
<tr>
<td>6 – 11 years old</td>
<td>aych612</td>
<td>6 – 11 years old</td>
</tr>
<tr>
<td>12 – 17 years old</td>
<td>aych1218</td>
<td>12 – 17 years old</td>
</tr>
<tr>
<td>18 years and older</td>
<td>aych18</td>
<td>18 years and older</td>
</tr>
<tr>
<td>Hours spent work a day-1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Continuous</td>
<td>tswork</td>
<td>Continuous</td>
</tr>
<tr>
<td>Hours spent work a day-2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;3 hours a day</td>
<td>wrk0</td>
<td>&lt;3 hours a day</td>
</tr>
<tr>
<td>3-6 hours a day</td>
<td>wrk36</td>
<td>3-6 hours a day</td>
</tr>
<tr>
<td>&gt;=6 hours a day</td>
<td>wrk8</td>
<td>&gt;=6 hours a day</td>
</tr>
<tr>
<td>Education level</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>edu1</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td>edu2</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td>edu3</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td>edu4</td>
<td>Low</td>
</tr>
<tr>
<td>High</td>
<td>edu5</td>
<td>High</td>
</tr>
<tr>
<td>Living area</td>
<td></td>
<td></td>
</tr>
<tr>
<td>City</td>
<td>city</td>
<td>City</td>
</tr>
<tr>
<td>Village, countryside</td>
<td>village</td>
<td>Village, countryside</td>
</tr>
</tbody>
</table>
Table 3. (continued)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Code</th>
<th>Description / range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dwelling type</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>dwap</td>
<td>Flat, apartment</td>
</tr>
<tr>
<td></td>
<td>dwrow</td>
<td>Terraced house, row house</td>
</tr>
<tr>
<td></td>
<td>dwcorn</td>
<td>Corner house, end house</td>
</tr>
<tr>
<td></td>
<td>dwsemid</td>
<td>Semi-detached house</td>
</tr>
<tr>
<td></td>
<td>dwdet</td>
<td>Detached house</td>
</tr>
<tr>
<td>Car availability</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>carav_yes</td>
<td>Yes, always</td>
</tr>
<tr>
<td></td>
<td>carav_oth</td>
<td>Yes, to be agreed with others</td>
</tr>
<tr>
<td></td>
<td>carav_no</td>
<td>No</td>
</tr>
<tr>
<td>Driver’s license</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>driversl</td>
<td>yes</td>
</tr>
<tr>
<td></td>
<td>nodriversl</td>
<td>no</td>
</tr>
<tr>
<td>Age group</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>age&lt;30y</td>
<td>&lt; 30 years old</td>
</tr>
<tr>
<td></td>
<td>age30-40y</td>
<td>30 - 39 years old</td>
</tr>
<tr>
<td></td>
<td>age40-50y</td>
<td>40 - 49 years old</td>
</tr>
<tr>
<td></td>
<td>age50-60y</td>
<td>50 – 59 years old</td>
</tr>
<tr>
<td></td>
<td>age&gt;60y</td>
<td>60 years and older</td>
</tr>
<tr>
<td>Dominant activity</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>dowork</td>
<td>Paid work</td>
</tr>
<tr>
<td></td>
<td>dostudy</td>
<td>Education/study</td>
</tr>
<tr>
<td></td>
<td>dohh</td>
<td>Household tasks, taking care of child(ren)</td>
</tr>
<tr>
<td></td>
<td>doret</td>
<td>Retired, looking for a job, voluntary work</td>
</tr>
<tr>
<td>Needs (scores on statements)</td>
<td>n_exerc</td>
<td>Physical exercise</td>
</tr>
<tr>
<td></td>
<td>n_fair</td>
<td>Fresh air/being outdoors</td>
</tr>
<tr>
<td></td>
<td>n_newexp</td>
<td>New experiences</td>
</tr>
<tr>
<td></td>
<td>n_social</td>
<td>Social contact</td>
</tr>
<tr>
<td></td>
<td>n_relax</td>
<td>Relaxation</td>
</tr>
<tr>
<td></td>
<td>n_entert</td>
<td>Entertainment</td>
</tr>
</tbody>
</table>

A model was estimated for each activity group separately. Table 4 shows the results of the estimations for each activity group. We included only significant variables on an alpha 5 % level (with a t-value bigger than 1.96 or smaller than -1.96) in the final estimation. However, in some cases the t-values in the final estimations were not significant, but leaving out those variables did not lead to a better result in terms of ability to identify effects of other variables. So in the final estimation results of three of the activity groups, some of the variables are not significant.

The results should be interpreted in the following way: the number of times the beta parameter (including effects of X variables) fits in the
threshold value (including effects of work hours) roughly represents an expected interval time if no specific day preferences would exist. E.g., for Daily shopping and base levels of beta and threshold this is \((1.639/0.705=)\) 2.3 days. Person, household and dwelling attributes influence the value of beta. E.g., if the respondent is male (t-male) it decreases the value of beta for Daily shopping with 0.204. A decrease of beta means an increase of the interval time. So, we find that men go less often to the supermarket or other store for daily shopping than women after having corrected for possible differences in available time (given work hours) and specific day preferences. On the other hand, keeping everything else equal (in particular thresholds), singles (hh_s_no) do grocery shopping more often than persons living in a household consisting of at least two individuals. Furthermore, individuals that own a driver’s license and subjects of which the age of the youngest child is between 12 and 18 years old, have faster build-up times for needs for daily shopping. In case of Social visits, respondents living with their parents, elderly persons and subjects with the dominant activity paid work show a lower need-recover rate for social visits. On the contrary, having a car available causes a higher need recover speed for visiting relatives or friends. If the activity concerns receiving social visits, the interval time decreases as well for a given level of the threshold. The results of Non-daily and Fun shopping indicate that if age increases, persons have an increased need to go shopping. In case of Fun shopping the interval time increases, keeping everything else equal. The activity group Going out shows negative effects for \(\beta\) values when the household income is higher than average and the age of the youngest child is 18 years or older. The activities ‘visiting a café, bar or discotheque’, ‘going out to play cards, billiards or bingo’ and ‘club/union activities (no sports)’, on the other hand, increases the level of the needs which lowers the interval time compared to visiting a theatre. The results of walking/cycling show that the activities walking in a park or nature and touring by bike have a faster need rebuild time compared to going for a short walk. Furthermore, living in a city and a higher personal need for relaxation increases the interval time for a given level of the threshold. Conversely, respondents with a higher personal need for entertainment have a higher need for walking or cycling. Keeping the threshold constant, the frequency of conducting a sports activity increases when the subject is single and decreases if it concerns activity ‘sports indoors in club/union context’, a higher income, respondents between 40 and 50 years old and a moderate education level.

Some variables can also have an impact on the threshold value. For this study we only included the amount of work hours a day as an explanatory attribute so far. In all of the activity groups the amount of time spent on
paid work increases the threshold value which increases the interval time. The subjects that work 6 hours or more on the particular day seem to decrease the frequency of all considered activities significantly, given levels of need-rebuilt times. Only in case of Walking/cycling the continuous variable ‘hours spent on work’ showed a significant influence. If we look at day preferences, we see that individuals tend to have an intrinsic preference for doing grocery shopping on Mondays, Non-daily and Fun shopping on Mondays and Saturdays, and Going out on the weekend.

The Rho squares of the estimations were calculated by using the log-likelihood of the estimations and the log-likelihood of a null-model. A complete null model, where all parameters are set to zero, shows very high Rho squares (between 0.57 and 0.75), but is not a good indicator of the reference goodness-of-fit in that the need-growth and threshold value can impossibly be zero. In order to find an appropriate reference goodness-of-fit we used ‘mean’ values of beta and the threshold to calculate the Log-likelihood of a null-model. For beta we chose 0.5 and for the threshold value 1.25. The Rho-squares calculated on that basis vary between 0.17 for Walking/cycling and 0.49 for Going out. These values indicate a satisfactory performance of the models.

Table 4. Estimation results

<table>
<thead>
<tr>
<th>Daily shopping</th>
<th>Social visits</th>
</tr>
</thead>
<tbody>
<tr>
<td>variable</td>
<td>estimate t-value</td>
</tr>
<tr>
<td>t</td>
<td>0.705 5.517</td>
</tr>
<tr>
<td>thr-base</td>
<td>1.639 9.689</td>
</tr>
<tr>
<td>mon</td>
<td>0.144 2.621</td>
</tr>
<tr>
<td>thu</td>
<td>-0.164 -2.773</td>
</tr>
<tr>
<td>fri</td>
<td>-0.181 -2.414</td>
</tr>
<tr>
<td>t-male</td>
<td>-0.204 -4.483</td>
</tr>
<tr>
<td>t-hh_s_no</td>
<td>0.294 3.543</td>
</tr>
<tr>
<td>t-dwrow</td>
<td>-0.089 -2.177</td>
</tr>
<tr>
<td>t-aych1218</td>
<td>0.285 1.998</td>
</tr>
<tr>
<td>t-driversl</td>
<td>0.220 3.896</td>
</tr>
<tr>
<td>t-n_entert</td>
<td>-0.034 -2.737</td>
</tr>
<tr>
<td>thr-wrk8</td>
<td>0.252 3.479</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>LL0</td>
<td>592.57</td>
</tr>
<tr>
<td>LL model</td>
<td>449.29</td>
</tr>
<tr>
<td>Rho square</td>
<td>0.242</td>
</tr>
<tr>
<td>Nr. of obs.</td>
<td>195</td>
</tr>
</tbody>
</table>
Table 4. (continued)

<table>
<thead>
<tr>
<th>variable</th>
<th>estimate</th>
<th>t-value</th>
<th>variable</th>
<th>estimate</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>t</td>
<td>0.169</td>
<td>8.063</td>
<td>t</td>
<td>0.049</td>
<td>10.880</td>
</tr>
<tr>
<td>thr-base</td>
<td>1.354</td>
<td>7.662</td>
<td>thr-base</td>
<td>1.305</td>
<td>10.164</td>
</tr>
<tr>
<td>mon</td>
<td>0.104</td>
<td>5.973</td>
<td>wnd</td>
<td>0.013</td>
<td>1.746</td>
</tr>
<tr>
<td>fri</td>
<td>-0.076</td>
<td>-2.579</td>
<td>t-act_café</td>
<td>0.083</td>
<td>5.295</td>
</tr>
<tr>
<td>sat</td>
<td>0.234</td>
<td>5.780</td>
<td>t-act_cbb</td>
<td>0.142</td>
<td>4.410</td>
</tr>
<tr>
<td>t-act_funsh</td>
<td>-0.057</td>
<td>-2.710</td>
<td>t-act_club</td>
<td>0.043</td>
<td>4.248</td>
</tr>
<tr>
<td>t-age30</td>
<td>-0.031</td>
<td>-1.872</td>
<td>t-i12av</td>
<td>-0.018</td>
<td>-3.724</td>
</tr>
<tr>
<td>t-age4050</td>
<td>0.063</td>
<td>1.876</td>
<td>t-i2av</td>
<td>-0.025</td>
<td>-4.625</td>
</tr>
<tr>
<td>t-age5060</td>
<td>0.144</td>
<td>3.326</td>
<td>t-ljk18</td>
<td>-0.034</td>
<td>-5.640</td>
</tr>
<tr>
<td>t-age60p</td>
<td>0.150</td>
<td>3.657</td>
<td>t-edu4</td>
<td>-0.009</td>
<td>-2.388</td>
</tr>
<tr>
<td>thr-wrk8</td>
<td>0.233</td>
<td>13.437</td>
<td>thr-wrk8</td>
<td>0.016</td>
<td>2.064</td>
</tr>
<tr>
<td>LL&lt;sub&gt;0&lt;/sub&gt;</td>
<td>527.58</td>
<td></td>
<td>LL&lt;sub&gt;0&lt;/sub&gt;</td>
<td>2172.29</td>
<td></td>
</tr>
<tr>
<td>LL model</td>
<td>386.10</td>
<td></td>
<td>LL model</td>
<td>1099.57</td>
<td></td>
</tr>
<tr>
<td>Rho square</td>
<td>0.268</td>
<td></td>
<td>Rho square</td>
<td>0.494</td>
<td></td>
</tr>
<tr>
<td>Nr. of obs.</td>
<td>119</td>
<td></td>
<td>Nr. of obs.</td>
<td>229</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>variable</th>
<th>estimate</th>
<th>t-value</th>
<th>variable</th>
<th>estimate</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>t</td>
<td>0.937</td>
<td>5.246</td>
<td>t</td>
<td>0.140</td>
<td>8.298</td>
</tr>
<tr>
<td>thr-base</td>
<td>1.134</td>
<td>6.454</td>
<td>thr-base</td>
<td>1.082</td>
<td>6.736</td>
</tr>
<tr>
<td>sat</td>
<td>-0.209</td>
<td>-3.907</td>
<td>t-act_ind</td>
<td>-0.028</td>
<td>-2.035</td>
</tr>
<tr>
<td>t-n_relax</td>
<td>-0.042</td>
<td>-3.055</td>
<td>t-age4050</td>
<td>-0.043</td>
<td>-3.304</td>
</tr>
<tr>
<td>t-n_entert</td>
<td>0.022</td>
<td>2.709</td>
<td>t-hh_s_no</td>
<td>0.044</td>
<td>2.164</td>
</tr>
<tr>
<td>t-act_wpn</td>
<td>-0.248</td>
<td>-3.632</td>
<td>t-i2av</td>
<td>-0.043</td>
<td>-2.451</td>
</tr>
<tr>
<td>t-act_cycl</td>
<td>-0.264</td>
<td>-4.098</td>
<td>t-edu3</td>
<td>-0.079</td>
<td>-4.180</td>
</tr>
<tr>
<td>t-city</td>
<td>-0.118</td>
<td>-2.623</td>
<td>t-edu4</td>
<td>-0.024</td>
<td>-1.494</td>
</tr>
<tr>
<td>thr-tswork</td>
<td>0.040</td>
<td>4.778</td>
<td>t-aych18</td>
<td>-0.026</td>
<td>-0.926</td>
</tr>
<tr>
<td>thr-wrk8</td>
<td>0.037</td>
<td>1.889</td>
<td>thr-wrk8</td>
<td>0.037</td>
<td>1.889</td>
</tr>
<tr>
<td>LL&lt;sub&gt;0&lt;/sub&gt;</td>
<td>438.03</td>
<td></td>
<td>LL&lt;sub&gt;0&lt;/sub&gt;</td>
<td>776.44</td>
<td></td>
</tr>
<tr>
<td>LL model</td>
<td>361.92</td>
<td></td>
<td>LL model</td>
<td>593.63</td>
<td></td>
</tr>
<tr>
<td>Rho square</td>
<td>0.174</td>
<td></td>
<td>Rho square</td>
<td>0.235</td>
<td></td>
</tr>
<tr>
<td>Nr. of obs.</td>
<td>141</td>
<td></td>
<td>Nr. of obs.</td>
<td>139</td>
<td></td>
</tr>
</tbody>
</table>
6. Conclusions and Discussion

This paper described a first attempt in estimating a model of activity generation that is based on notions of dynamic needs. Data used were especially collected for this purpose. The survey included, for a list of 37 activities, the time elapsed since last performance of the activity, if the activity was conducted the day before and if and when the activity was already planned. As indicators of six basic needs for activity generation which were the result of surveys described in an earlier study, four statements for each need were incorporated in the questionnaire.

The results of the estimations of the parameters indicate that several socioeconomic and dwelling variables have an impact on episode interval timing decisions of the shopping, social, leisure and sports activities considered in the present study. As the parameters of the model in this paper are estimated for each of six activity groups, we plan to estimate an overall model in the near future. We also want to add several variables to threshold value, like we did with the hours spent on work a day. An interesting addition may be the scores on the statements concerning the six needs. A possibility to increase the amount of data that can be used for the estimations is to include the data of which the history is known, but lacks the interval time, as they were not conducted on the diary day. What also can be done in the future is to validate the results with the large dataset of the Dutch national travel survey (MON 2004), which consists of activity-travel diaries of about 47,000 respondents. Furthermore, we plan to carry out analyses on the data collected for the purpose of the need-based model to find out to what extent which activities are substitutable.

References


