PARAMETER ESTIMATION OF A DYNAMIC NEED-BASED ACTIVITY GENERATION MODEL

Linda Nijland, Theo Arentze and Harry Timmermans, Eindhoven University of Technology, PO Box 513, 5600 MB Eindhoven, The Netherlands, e-mail: e.w.l.nijland@tue.nl; t.a.arentze@tue.nl; h.j.p.timmermans@tue.nl

ABSTRACT

Several activity-based models made the transition to practice over the last decade. However, modelling 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 an adequate sample of individuals. The diary data include detailed information on 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, and person-level needs influence activity-episode timing decisions in a longitudinal time frame and, thus, the frequency and day choice of conducting the social, leisure and sports activities.

Keywords: activity-based modelling, dynamic activity generation, travel-demand modelling, estimation

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 2008; Roorda et al. 2008). 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 modelling 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 modelling approach, demonstrating the possibility of incorporating positive or negative substitution effects between activities and complex dynamic interactions between activities in general. In a follow-up study, the authors developed a RUM model and explored the extent to which the model can be estimated on existing one-day datasets (Arentze et al. 2010). Until now, however, their approach lacks a full empirical validation based on data specifically collected for that purpose.

The goal of the research project underlying this current paper is to test the suggested approach empirically and to estimate parameters of the 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 (constructed scales for) 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 RUM specification of the need-based concepts and model. This is followed by a description of the survey and the sample. Section 6 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.

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. (2009, 2010). 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 agenda 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). The utility of conducting an activity of type *i* on a given day *d* is defined as:

$$U_{nid}(s) = V_{1ni,d-s} + V_{2,nid} + \varepsilon_{1nis} + \varepsilon_{2nid} \qquad (1)$$

where *n* is an index of individual, *d* is the current day, *s* is the day activity *i* was conducted the last time before *d*, $V_{1ni,d-s}$ is the utility of satisfying the need for activity *i* built-up between *s* and *d*, $V_{2,nid}$ is a (positive or negative) preference for conducting activity *i* on day *d* and ε_{1nis} and ε_{2nid} are error terms related to need build-up (ε_1) and day (ε_2).

The utility components can be interpreted as follows. The first term (V_1) 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_2) 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 modelled as activities with zero need growth ($V_1 = 0$) and a relatively high utility for the day ($V_2 >> 0$) 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 inflection point where the curve is approximately linear. To reduce the number of parameters, the RUM model, therefore, assumes a simple linear function here:

$$V_{1nit} = \beta_{ni} t \qquad (2)$$

where β_{ni} 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 multi-day 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, the existing model leaves activity duration out of consideration. This means 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:

$$U_{nid}(s) > u^{o}_{nd} \qquad (3)$$

where u^{o}_{nd} represents a threshold for implementing activities on day *d*, given existing time demands on that day. Note that defined in this way, the need-growth parameter β 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.

The model we estimate here is derived from the assumption that ε_2, is either simulated or zero and the first error term, ε_1, is Gümbel distributed. Given this assumption, an ordered-logit framework of the following form can be derived from decision rule (3) (Arentze et al. 2010):

$$P_{ni}(d \mid s) = \frac{\exp[Z_{nid}(s)]}{1+\exp[Z_{nid}(s)]} - \frac{\exp[\max_{k=s+1}^{d-1}[Z_{nik}(s)]]}{1+\exp[\max_{k=s+1}^{d-1}[Z_{nik}(s)]]} \qquad (4)$$

where

$$Z_{nid}(s) \equiv V_{1ni,d-s} + V_{2nid} + \varepsilon_{2nid} - u^{o}_{nd} \qquad (5)$$

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

$$\sum_{d>s} P_{ni}(d \mid s) = 1 \qquad (6)$$

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.

The model represent dynamics of activity-generation decisions that follow from the fact that needs take time to re-build, and preferences and time budgets for conducting the activity may differ from day to day. A preference for a certain day of the week generates secondary effects on probabilities for other days. Secondary effects emerge because a need for the activity needs time to rebuild after the activity has been conducted. A static model which does not incorporate need build-up time, is not able to represent secondary effects of day-preferences and, hence, would make wrong inferences about intrinsic day preferences.

MODEL ESTIMATION

As expressed in Equation (6), Equation (4) defines a probability distribution across days *d* after *s*. Whether or not this form can be used to determine likelihoods of observations depends on the nature of observations. In the survey conducted to estimate the model (see below) individuals recorded their activity agenda for a given day (*d*) and for an exhaustive list of activities the day the activity was performed the last time (*s*). In case of such observations, we know that the activity has not been conducted in the time between *s* and *d*. According to the model, the probability that the activity has not been conducted in the period from *s*+1 and *d*-1 is defined as:

$$Q_{ni}(d \mid s) = 1 - \frac{\exp[\max_{k=s+1}^{d-1}[Z_{nik}(s)]]}{1 + \exp[\max_{k=s+1}^{d-1}[Z_{nik}(s)]]} \qquad (7)$$

Therefore, the probability of observing *i* in the agenda for day *d* knowing that the activity has not been conducted until that day is given by:

$$L_{ni}(1 \mid d, s) = P_{ni}(d \mid s) / Q_{ni}(d \mid s) \qquad (8)$$

$L_{ni}(1 \mid d, s)$ is the likelihood of observing activity *i* given observation day *d* and recalled last day *s*. This likelihood has the following property:

$$L_{ni}(1 \mid d, s) + L_{ni}(0 \mid d, s) = 1 \qquad (9)$$

The likelihood for a sample of observations can be defined as a function of the model's parameters as follows:

$$L(Y \mid \theta) = \prod_n \prod_i L(y_{ni} \mid \theta) \qquad (10)$$

where *Y* is a sample of individuals, θ is the set of observations included in the model, y_{ni} is a binary variable of observing activity *i* in case of individual *n* and *L*(*y*) is the likelihood defined by Equation (8).

The likelihood function (or loglikelihood function) appears to be non-smooth in the area of the optimum values of beta parameters in particular. Furthermore, due to the dependency relationship between activity probabilities across days, i.e. the secondary effects, convergence of search processes for optimal parameter values in standard loglikelihood methods is very slow. To circumvent these problems, we used a Bayesian method of estimating parameters. Bayesian methods are known to be more robust, as they do not use a function maximization process (Rossi et al. 2005).

The Bayesian method we used for the present estimation task is based on the following equation:

$$K(\theta_i \mid \overline{\theta}_{i-}^{\,n}, \overline{\theta}_{i+}^{\,n-1}, Y_n) = \frac{L(y_n \mid \overline{\theta}_{i-}^{\,n}, \theta_i, \overline{\theta}_{i+}^{\,n-1}) K(\theta_i \mid \overline{\theta}_{i-}^{\,n}, \overline{\theta}_{i+}^{\,n-1}, Y_{n-1})}{\sum_{\theta} L(y_n \mid \overline{\theta}_{i-}^{\,n}, \theta, \overline{\theta}_{i+}^{\,n-1}) K(\theta \mid \overline{\theta}_{i-}^{\,n}, \overline{\theta}_{i+}^{\,n-1}, Y_{n-1})} \qquad (12)$$

where θ_i is the *i*-th parameter of the model, $K(\theta_i)$ is either a posterior (LHS) or prior (RHS) probability distribution across values of parameter θ_i, y_n is the *n*-th observation in the sample, Y_n is the set of observations up to n ($Y_n = y_1 \ldots y_n$), $\overline{\theta}_{i-}$ is a vector of expected values for parameters $\theta_1, \theta_2, \ldots, \theta_{i-1}$, and $\overline{\theta}_{i+}$ is a vector of expected values for parameters $\theta_{i+1}, \theta_{i+2}, \ldots, \theta_m$ (m = number of parameters of the model). Equation (12) describes an incremental Bayesian learning process. Initially, a uniform distribution across some predefined wide-enough range is assumed for each parameter of the model, reflecting the assumption that no prior knowledge about parameter values exists. Observations are processed one at a time in sequence $y_1, y_2, \ldots$. For each observation the posterior distribution is determined one parameter at a time in sequence $\theta_1, \theta_2, \ldots, \theta_m$ using Equation (12), whereby all other parameters are set to their current expected values (denoted as $\overline{\theta}$). The priors in each next case are set to the posteriors obtained from the last case. After all cases have been processed, the posterior distributions represent final estimates. Note that in this method each observation is used only once to update beliefs about the parameters.

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 constructed scales 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.

SAMPLE

Subjects were selected from a sample of neighbourhoods in the Eindhoven region. In the last two weeks of June 2009, 4000 invitation cards were distributed to households in the selected neighbourhoods. 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 (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

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

The activity data used for the analyses in the current paper consists of the cases where the respondent indicated the date of (or the time passed since) the last performance of the activity. The variable 'time passed since last performance' showed the amount of days between the last performance and the day before the respondents filled out the questionnaire. The activity could either be conducted or not be conducted on the latter day. Both of these options were included in the model estimation, Altogether about 4200 cases could be used for the analyses.

RESULTS

The selection and categorization of explanatory variables on individual and household levels to be included in the analysis was based on number of cases available for each (dummy) variable. This number may not be too low in order to get a reliable result. A threshold of 400 cases was used. Table 2 shows the variables that were included in the analyses. Most variables were dummy-coded, except for the scores on the needs and the hours spent on work/education a day. By taking some of the most frequently conducted activities together, six activity groups were created, namely: daily shopping, non-daily/fun shopping, social visits, going out, sports and walking/cycling (as an activity). Table 3 shows which activities were put together. In total those activities contain 2837 cases that can be used for the estimation of the parameters of the need-based model.

The language C was used to program the need-based model and estimate the model using the Bayesian estimation method described above (Eq. (12)).

Table 2 – The explanatory variables considered for the need-based model (base level in italics and bold)

Variable	Code	Description / range
Day of the week	mon	Monday
	tue	Tuesday
	wed	Wednesday
	thu	Thursday
	fri	Friday
	sat	Saturday
	sun	Sunday
Gender	male	Male
	female	Female
Age group	age30-	< 30 years old
	age3040	30 – 39 years old
	age4050	40 – 49 years old
	age5060	50 – 59 years old
	age60+	60 years and older
Household composition	hh_s_no	Single, no children
	hh_sd_c	Single or Double, with child(ren)
	hh_rst	Double, no children, living in at (grand)parents/relatives, student accommodation, group accommodation
Household income	ibav	below average
	i1av	average
	iaav	above average
Age youngest child	aych06	0 – 5 years old
	aych6+	6 years and older
Hours spent work a day	tswork	Continuous
Education level	edul	Low
	edu1av	
	eduh	High
Living area	city	City
	village	Village, countryside
Dwelling type	***dwap***	Flat, apartment
	dwgarden	House
Car availability	carA	Yes, always
	carO	Yes, to be agreed with others
	carN	No
Needs	n_exerc	Physical exercise
(scores on constructed scales)	n_frair	Fresh air/being outdoors
	n_nwexp	New experiences
	n_social	Social contact
	n_relax	Relaxation
	n_entert	Entertainment

As indicators of V_{2di} we included the days of the week as dummies in the following way (arbitrary choosing Wednesday as a reference):

$$V_{2di} = \alpha_{i1} * Mon_d + \alpha_{i2} * Tue_d + ... \qquad (13)$$

where *Mon* and *Tue* are zero-one variables indicating whether day *d* is a Monday, Tuesday, etc. and α are day-preference parameters. For V_{1t}, a constant and person, household, and dwelling attributes shown in Table 2 were included, as follows:

$$V_{1it} = (\beta_{i0} + \Sigma_k \beta_{ik} X_k) * t \qquad (14)$$

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

$$u^o{}_d = \mu_0 + \Sigma_l \mu_l X_d \qquad (15)$$

where X_d are attributes influencing time budgets and μ are threshold parameters. In the current analysis, work hours (as a continuous variable) and car availability (dummy coded) were used as explanatory variables. The threshold value (u^o) was estimated over all activities. Both V_1 and V_2, on the other hand, were estimated for each activity group separately.

Table 3 – Activity groups and their activities included in the estimations

Activity group	Activities included
Daily shopping	Daily shopping
Non-daily/ Fun shopping	Non-daily shopping
	Fun shopping
Social visits	Visiting relatives/friends
	Receiving visitors
	Visiting (e.g., birthday) party
Going out	Going out for dinner
	Visiting a theatre
	Attending a concert
	Visiting a café, bar or discotheque
	Going to the cinema
	Visiting a museum
	A day out (visit a city, recreation park)
Walking/cycling	Going for a short walk
	Walking in a park or nature
	Touring by bike
Sports	Sports outdoors, club/association context
	Sports outdoors, flexible
	Sports indoors, club/association context
	Sports indoors, flexible

Table 4 shows the results of the parameter estimation. The results should be interpreted in the following way. In terms of need-build up, the beta0 parameter represents the intercept when all other beta variables are zero. Person, household and dwelling attributes influence the value of beta. For example, if the respondent is single (hh_s_no) it decreases the value of beta for Walking/cycling (leisure) with 0.133. Keeping everything else equal, a decrease of beta means an increase of the interval time, which is defined as the amount of time between conducting an activity and conducting the same activity again. Thus, we find that singles go less often walking or cycling as an activity than persons living in a household consisting of at least two individuals after having corrected for possible differences in available time (given work hours), car availability, and specific day preferences. In case of Social visits this counts as well for respondents having a higher education level and subjects living in a house with garden. They have longer build-up times for needs for Social visits. On the other hand, keeping every thing else equal (in particular thresholds), elderly people (50+) show a higher need-recover rate for social visits than younger persons. The results of Sports show that individuals living in a house with garden, respondents that are part of a household consisting of one or two adults and children, and higher educated subjects have a longer need rebuild time. In contrast, keeping the threshold constant, the frequency of conducting a sports activity increases when the person considered lives in a city.

Some of the scores on the constructed scales for the needs show significant results as well. A higher need for New experiences causes shorter need build-up times for Social visits and, on the other hand, a higher need for Fresh air/being outdoors results in a lower need-recover rate for this activity group. In case of Walking/cycling (leisure) higher needs for Physical exercise and Fresh air/being outdoors show a decrease of the interval time. Conversely, a higher need for Entertainment increases the interval time of Walking/cycling.

If we look at day preferences, we see that individuals tend to have an intrinsic preference for doing grocery shopping on Thursdays, Fridays, and Saturdays, Non-daily and Fun shopping on Mondays and Saturdays, Social visits on Thursdays, Fridays, and Saturdays, Going out on Saturdays, Sports on Tuesdays and Thursdays, and Walking/cycling on Tuesdays and Sundays. On the other hand, individuals do not prefer Daily shopping on Sundays (or stores are closed on that day), Non-daily or Fun shopping on Tuesdays, Social visits on Mondays, Tuesdays, and Sundays, Going out on Sundays and Walking/cycling on Mondays and Thursdays.

Some variables can also have an impact on the threshold value. For this study we only included the amount of work hours by day of the week and car availability as an explanatory variable. The results show that the amount of time spent on paid work on a day increases the threshold value and, hence, decreases the probability of conducting the activity on that day. In this study, car availability does not have a significant impact on the threshold value.

Table 4 – Estimation results (significant estimates in bold)

	Daily shopping		**Non-daily-/Fun shopping**		**Social visits**	
Variable	estimate	t-value	estimate	t-value	estimate	t-value
beta0	**0.630**	**3.344**	**0.639**	**3.160**	**0.331**	**2.666**
betaMale	0.034	0.276	0.017	0.129	-0.067	-0.672
betaAge30	-0.046	-0.363	-0.064	-0.531	-0.088	-1.021
betaAge4050	-0.046	-0.359	0.007	0.051	-0.055	-0.446
betaAge5060	0.032	0.255	0.044	0.339	**0.122**	**2.346**
betaAge60	0.006	0.046	-0.022	-0.164	0.036	0.275
betaHh_sd_c	-0.017	-0.131	0.068	0.556	-0.094	-0.889
betaHh_s_no	0.004	0.028	-0.047	-0.366	0.130	1.105
betaDwgarden	0.049	0.403	0.053	0.453	**-0.224**	**-7.228**
betaIbav	-0.054	-0.440	-0.061	-0.488	-0.107	-1.058
betaIaav	0.060	0.480	0.014	0.107	-0.047	-0.363
betaEdul	0.020	0.155	-0.036	-0.277	-0.026	-0.200
betaEduh	0.010	0.075	0.031	0.261	**-0.167**	**-2.027**
betaAych06	0.011	0.080	0.012	0.088	-0.104	-0.963
betaCity	0.029	0.237	0.028	0.216	0.010	0.074
beta-n_exerc	0.019	1.348	-0.006	-0.387	-0.003	-0.300
beta-n_frair	0.027	1.631	0.012	0.638	**-0.017**	**-2.200**
beta-n_nwexp	0.018	1.387	-0.014	-0.809	**0.034**	**7.165**
beta-n_social	-0.021	-1.248	-0.003	-0.115	0.020	1.587
beta-n_relax	0.029	1.880	0.001	0.066	0.015	0.800
beta-n_entert	-0.007	-0.309	-0.024	-0.894	0.002	0.099
alphaMon	0.001	0.006	**0.591**	**5.746**	**-0.609**	**-8.329**
alphaTue	-0.109	-0.853	**-0.487**	**-3.128**	**-0.236**	**-2.245**
alphaThu	**0.260**	**5.097**	-0.154	-0.456	**0.103**	**2.494**
alphaFri	**0.405**	**6.567**	0.240	1.201	**0.432**	**3.497**
alphaSat	**0.670**	**24.918**	**0.670**	**13.962**	**0.694**	**8.037**
alphaSun	**-0.659**	**-17.118**	0.029	0.192	**-0.638**	**-4.948**

	Going out		**Sports**		**Walking/cycling (leisure)**	
variable	estimate	t-value	estimate	t-value	estimate	t-value
beta0	0.342	1.502	**0.658**	**2.447**	**0.522**	**5.621**
betaMale	0.026	0.201	-0.009	-0.065	0.147	1.797
betaAge30	-0.004	-0.030	0.080	0.667	-0.068	-0.560
betaAge4050	0.021	0.134	-0.234	-1.780	-0.036	-0.276
betaAge5060	-0.008	-0.059	0.134	1.445	0.095	0.863
betaAge60	0.007	0.049	-0.029	-0.222	0.082	0.704
betaHh_sd_c	-0.038	-0.223	**-0.194**	**-2.024**	0.026	0.203
betaHh_s_no	0.022	0.166	0.001	0.005	**-0.133**	**-2.397**
betaDwgarden	-0.094	-0.843	**-0.259**	**-3.347**	0.064	0.541
betaIbav	-0.023	-0.173	0.099	0.889	-0.004	-0.057
betaIaav	-0.017	-0.129	-0.049	-0.233	0.040	0.314
betaEdul	-0.007	-0.056	0.029	0.221	0.029	0.221
betaEduh	0.049	0.392	**-0.263**	**-3.956**	-0.070	-1.055
betaAych06	-0.011	-0.081	0.030	1.366	-0.005	-0.036
betaCity	-0.006	-0.041	**0.097**	**2.800**	-0.088	-1.249
beta-n_exerc	-0.021	-0.848	0.008	0.466	**0.013**	**2.089**
beta-n_frair	-0.033	-1.366	0.028	0.988	**0.027**	**3.877**
beta-n_nwexp	-0.016	-0.754	-0.018	-0.773	0.000	-0.072
beta-n_social	-0.008	-0.287	0.022	1.289	-0.003	-0.363
beta-n_relax	-0.026	-0.989	-0.042	-0.773	0.011	1.440
beta-n_entert	-0.022	-0.621	-0.047	-1.244	**-0.029**	**-2.549**
alphaMon	0.284	1.297	0.337	1.011	**-0.396**	**-2.170**
alphaTue	0.154	0.580	**0.653**	**4.678**	**0.589**	**5.492**
alphaThu	-0.272	-1.043	**0.649**	**8.120**	**-0.627**	**-6.662**
alphaFri	0.361	0.852	-0.039	-0.228	-0.070	-0.648
alphaSat	**0.393**	**2.224**	0.000	0.000	0.145	1.284
alphaSun	**-0.496**	**-2.372**	0.000	0.000	**0.628**	**10.061**
Thr0	**1.872**	**37.576**			LL model	-1211.500
ThrTswork	**0.068**	**3.756**			LL0	-1735.050
ThrCarA	0.128	1.249			Rho square	0.302
ThrCarO	0.011	0.127	Nr. of obs.	2837	Rho sq (adj.)	0.203

The Rho square of the estimation was calculated by using the log-likelihood of the estimated model and the log-likelihood of a null-model. A complete null model, where all parameters are set to zero is not a good indicator of the reference goodness-of-fit in that the need-growth and threshold value cannot be equal to zero. In order to find an appropriate reference goodness-of-fit we used 'mean' values of the intercepts of beta and a value close to the threshold intercept parameter to calculate the Log-likelihood of a null-model. For all intercept betas we chose 0.5 and for the threshold intercept a value 2. The Rho-square calculated on that basis is 0.302. This indicates a satisfactory performance of the model. However, the adjusted Rho-square is noticeably lower with a value of 0.203, which suggests a lack of data compared to the number of variables incorporated in the model.

CONCLUSIONS AND DISCUSSION

This paper described a first attempt of 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 parameter estimations indicate that several socioeconomic and dwelling variables have an impact on episode interval timing and day choice decisions of the shopping, social, leisure and sports activities considered in the present study. Day preferences and the scores on the statements concerning the six needs show significant effects as well.

The purpose of the present study is to show that it is possible to collect data which can be used to estimate the parameters of a dynamic need-based activity generation model. Although the size of the sample is somewhat limited for the number of variables included in the model, we demonstrated that the developed methodology is feasible. New data should be collected all year round, to capture seasonal influences, and in larger amounts. An interesting avenue is to validate the results with data from a national travel survey, such as for example the Dutch travel survey (called the MON). The Bayesian estimation method used in the present study supports pre- specification of a-priori distributions of parameters that could be set based on other data sources such as the MON. In that approach, data collected specifically for the model would be used for fine tuning rather than estimating parameters from scratch. There are also meaningful ways of extending the model. In the near future, we plan to carry out analyses on the data collected for the purpose of identifying interactions between activities such as to find out to what extent activities are substitutable in the framework of the need-based model.

REFERENCES

Arentze, T.A. and H.J.P. Timmermans (2000). Albatross: A Learning-Based Transportation Oriented Simulation System, EIRASS, Eindhoven.

Arentze, T.A. and H.J.P. Timmermans (2009). A need-based model of multi-day, multi-person activity generation, Transportation Research B, 43, 251-265.

Arentze, T.A., D. Ettema and H.J.P. Timmermans (2009). A longitudinal household based model of activity generation: approach and estimation on one-day activity-travel data, Paper prepared for presentation at the 12th Conference on Travel Behavior Research, Jaipur, India, Dec. 13-18.

Arentze, T.A., D. Ettema and H.J.P. Timmermans (2010). Estimating a model of dynamic activity generation based on one-day observations: method and results. Submitted for publication.

Axhausen, K.W. (Ed.) (2006). Moving Through Nets: The Physical and Social Dimensions of Travel. Elsevier, Oxford.

Bhat, C.R. and S.K. Singh (2000). "A comprehensive daily activity-travel generation model system for workers", Transportation Research Part A, 34, 1-22.

Bricka S, and J. Zmud (2003). Impact of Internet Retrieval for Reducing Nonresponse in a Household Travel Survey. CD-ROM. Proceedings 82nd Annual Meeting of the Transportation Research Board, Washington, D.C. (CD-ROM).

Chapin, F.S. (1974). Human activity patterns in the City. John Wiley and Sons, New York.

Habib, K. M. N. and E. J. Miller (2008). Modelling Daily Activity Program Generation Considering Within-Day and Day-to-Day Dynamics in Activity-Travel Behaviour. Transportation, 35 (4), pp. 467–484.

Meister, K., M. Krick and K.W. Axhausen (2005). A GA-based Household Scheduler. Transportation, 32 (5), 473–494.

Miller, E. (2004). An Integrated Framework for Modelling Short- and Long-run Household Decision-making. In: Paper Presented at the Progress in Activity-Based Analysis Conference, Maastricht, The Netherlands.

Nijland, E.W.L., T.A. Arentze and H.J.P. Timmermans (2010). Eliciting needs underlying activity-travel patterns and their covariance structure: results of multi-method analyses, Transportation Research Record, forthcoming. Proceedings 89th Annual Meeting of the Transportation Research Board, Washington, D.C. (CD-ROM).

Pendyala, R.M., R. Kitamura, A. Kikuchi, T. Yamamoto and S. Fujii (2005). "Famos, The Florida activity mobility simulator", in: Proceedings 84th TRB Annual Meeting, Washington, D.C.

Roorda, M.J., E.J. Miller and K.M.N. Habib (2008). Validation of TASHA: a 24-hour Activity Scheduling Microsimulation Model. Transportation Research Part A: Policy and Practice 42 (2), pp. 360–375.

Rossi, P.E., G.M. Allenby, R. McCulloch (2005) Bayesian Statistics and Marketing, John Wiley & Sons, Sussex, UK.

Sun, Z., T.A. Arentze, and H.J.P. Timmermans (2009). A Heterogeneous Latent Class Model of Activity Rescheduling, Route Choice and Information Acquisition Decisions under Multiple Uncertain Events. Proceedings 88th Annual Meeting of the Transportation Research Board, Washington, D.C. (CD-ROM).