An Error-Components Framework for Joint Choice Models of Activity Timing and Duration

Dick Ettema[1]
Fabian Bastin[2]
John Polak[3]
Olu Ashiru[3]

[1] Utrecht University
Faculty of Geosciences
Urban and Regional research centre Utrecht (URU)
PO Box 80.115
3508 TC Utrecht
The Netherlands

[2] CERFACS
42, Avenue Gaspard Coriolis
31057 Toulouse Cedex 01
France

[3] Centre for Transport Studies
Department of Civil and Environmental Engineering
Imperial College London
Exhibition Road
London, SW7 2AZ
United Kingdom

Abstract

This paper develops a model of activity and trip scheduling that combines three elements that have previously been investigated in isolation: the duration of activities, the time-of-day preference for activity participation and the effect of schedule delays on the valuation of activities. The model is in essence a GEV discrete choice model, describing individuals' choice between alternative activity patterns. The utility function is formulated in a flexible way, applying a bell-shaped component to represent time-of-day preferences for activities. The model was tested using a 2001 data set from the Netherlands. The estimation results suggest that time-of-day preferences and schedule delays associated with the work activity are the most important factors influencing the scheduling of the work tour. Error components included in the model suggest that there is considerable unobserved heterogeneity with respect to mode preferences and schedule delay.

1. Introduction

The recent growth of interest in activity-based methods has focused particular attention on travellers' decision making process regarding the timing and duration of their participation in activities. Models of timing and duration choice have direct application to a wide range of demand management policies and are at the core of many recent activity based modelling systems. However, to date in the literature these two dimensions of activity participation have been largely treated separately, despite the compelling observation that in general, the benefit that an individual derives from participating in an activity will depend inter alia both upon the time at which the activity is undertaken and the amount of time devoted to the activity. Moreover, since many of the influences on the timing and duration of activity participation (such as the perceived quality of available travel modes and activity opportunities and the intensity with which an individual undertakes activities) will vary by individual and context and some are inherently difficult to completely characterise via conventional travel or time use data, it is likely that such decisions will be characterised by a significant degree of heterogeneity.

This paper proposes a model for the simultaneous choice of the timing and duration of activities and travel mode. The model is based on earlier work by the authors (Ashiru et al., 2004; Ettema et al., 2004), which formulated the model of activity timing and duration as a discrete choice system. It extends the previous work in two important aspects. First, it accommodates the influence of explicit schedule constraints, in the form of a schedule delay concept, thus providing an important point of connection between the recent activity scheduling literature and the earlier literature on trip re-scheduling. Second, it accounts for both observed and unobserved heterogeneity in activity scheduling behaviour. Observed heterogeneity is accommodated through explicit segmentation of model parameters according to socio-demographic characteristics whereas residual, unobserved inter-personal heterogeneity is accommodated through allowing key parameters to vary randomly across individuals.

The structure of the paper is as follows. The next section of the paper provides a brief review of the existing literature on activity timing and duration choice. The third section introduces the theoretical approach, which assumes that the marginal utility derived from activities encompasses two distinct components; one derived from the duration of activity involvement and the other derived from activity participation at a particular time-of-day, possibly in relation to existing anchor points, such as the work start time. The fourth section discusses the data that were used to test the model empirically. The fifth section describes the estimation methodology. Particular attention is given to the use of advanced optimisation techniques needed to estimate the non-linear utility function expressing individuals' timing and duration preferences. The sixth section discusses the estimation results and the conclusions that can be drawn with respect to individuals' decision making regarding timing and duration of activities. The paper closes with some overall conclusions and a discussion of future research directions.

2. Relevant literature

According to activity based travel theory (Ettema and Timmermans, 1997), trips can be regarded as a necessary means to connect spatially remote activities that will logically precede or follow these activities. This implies that the timing of trips not only depends on trip characteristics that vary by time-of-day (such as travel time and delays) but also on preferences with respect to the *timing* and *duration* of activities. Consequently, when modelling trip-timing decisions, these should be regarded in the context of the activity-scheduling process (e.g. Ettema and Timmermans, 2003).

With respect to modelling the timing and duration of activities, various approaches have been taken within the activity-based framework. Focusing on timing decisions, a first group of models (Bowman and Ben-Akiva, 1998; Arentze and Timmermans, 2005), although applying widely different decision-making mechanisms, have essentially treated the timing of activities as being a choice between a limited number of discrete time intervals. For instance, Bowman and Ben-Akiva (1998) conceptualise the timing of activities as the choice between the morning, afternoon or evening. Models of this type are based on the assumption that for particular activities, certain periods of the day are more attractive, resulting in higher utilities. A limitation of this approach is that only a limited number of broad periods are assumed, and that the exact timing of activities is not directly modelled. In addition, it is assumed that utility is gained from the execution of activities as a whole, thereby ignoring the variation in the utility gained during the execution. This may be a suitable approach for shorter activities, but in case of activities with a long duration, such as work, it is desirable to allow for variation in the attractiveness of the activity during its execution. This is especially the case if one is interested in for instance responses of travellers to congestion which involve retiming of the commute trip, resulting in retiming of the work activity.

In order to obtain a more flexible representation of activity timing, various approaches have described the scheduling of activities in continuous time, rather than in terms of a limited number of discrete time slots. The central idea of these approaches is that for each time-of-day t, there exists a marginal utility (which may vary over time), expressing the utility gained from one time unit of activity participation. One of the first approaches of this type is described by Wang (1996), who assumes that the marginal utility of activity participation at time t equals the observed share of the sample involved in the activity at that time. This is based on the problematic assumption that each activity can be performed at the preferred time, which is unrealistic given the many constraints applying to activity scheduling processes. To overcome this problem, Ettema and Timmermans (2003) propose an alternative model, in which the marginal utility of an activity is a direct function of time-of-day. A similar marginal utility model formulation was also earlier proposed by Polak and Jones (1994).

Although the Polak-Jones-Ettema-Timmermans (henceforth, PJET) models provide a flexible approach to modelling activity-timing decisions, two problems need to be addressed. The first problem is the neglect of the duration component within their marginal utility formulations. Many activities are likely to be subject to fatigue effects, implying that the utility derived from one time unit of activity participation diminishes with increasing duration. The PJET models in contrast, assume that one unit of activity engagement at time-of-day t will always yield the same utility, irrespective of the duration of activity engagement.

A way to account for fatigue effects is offered by time allocation models, which are based on the seminal work of Becker (1965), who treated time as a finite resource, which can be allocated to activities, resulting in a certain level of utility. Time allocation is in this view regarded as an optimisation problem under the restriction of a fixed amount of resources (time). The duration of activities is then determined by the allocation of time to activities such that the overall utility is maximised. Becker’s model was elaborated by, amongst others, Evans (1971) and De Serpa (1971) in order to account for the consumption of goods given consumption rates, prices and the available monetary budget. Other extensions of this approach include the modelling of time allocation on the household level (Zhang et al., 2002) and the specification and testing of advanced time-dependent utility functions (Joh et al., 2003). If the utility derived from an activity is defined as a log-function of the time spent on an activity, the time allocation model can be formulated as a system to be estimated using for instance seemingly unrelated regressions (Kitamura, 1984; Bhat and Misra, 1999). An important property of this type of formulation is that the marginal utility of activities decreases with their duration, representing the onset of activity fatigue. Although the Becker-type models are able to describe how individuals maximise utility by allocating time to activities, they do not take into account the preferences that individuals have with respect to the timing of activities. Recently, Ashiru et al. (2004) and Ettema et al. (2004) have formulated and empirically tested models that combine the time-of-day dependent marginal utility function of the PJET-models and the duration dependent Becker-type models.

A second problem of the JPET formulation is the assumption that the timing of activities is purely based on continuous marginal utility functions associated with activities. This assumption overlooks the fact that the timing of many activities is partially determined by constraints such as work or school arrangements and opening hours of stores and facilities (Hägerstrand, 1970). Such constraints may lead to discontinuities in the marginal utility function relative to anchor points such as work start time or the opening and closing time of facilities.

An approach explicitly accounting for such discontinuities is the schedule delay approach (Small, 1982). This approach focuses especially on the desired start time of activities. It is typically assumed that associated with each activity is a preferred start time. Likewise, the trip to this activity has some preferred arrival time (PAT). Deviations from the preferred arrival time (schedule delays) result in a negative utility. In Small's model trip utility is not only a function of travel time *t*, but also of schedule delay:

$$V_t = at + g_1 SDE + g_2 SDL \qquad (1)$$

where:

SDE is an early schedule delay, defined as $max((PAT - t^a), 0)$;

SDL is a late schedule delay, defined as *max* $((t^a - PAT), 0)$;

t^a is the actual arrival time;

This formulation implies that there is overlap between the schedule delay approach and the JPET models. In particular, the SDL-parameter can be interpreted as the (constant) marginal utility of the activity to start after arrival, since each time unit the activity starts later results in a loss of utility of g_2. Likewise, the SDE-parameter can be interpreted as being associated with the duration of the activity preceding the trip in a similar way, since arriving too early goes at the cost of the activity at the trip origin. On the other hand, there are also differences between the approaches. For instance, the schedule delay approach does not address the valuation of activity involvement when started. For instance, the utility gained from an activity only depends on its start time and not on its duration. In addition, the valuation of an activity that starts before the preferred arrival time is not directly addressed.

From a conceptual point of view, one can argue that the JPET models assume that activities can be scheduled at any time, independent of timing constraints. This scheduling process is then based on some intrinsic time-of-day dependent utility. In reality, however, the timing of most activities is at least to some extent guided by constraints, even if applying to other activities. For instance, the timing of an activity such as "taking a walk" which can principally take place any time, will be determined by obligations to spend time on work and household obligations, limiting the available time window. In empirical studies, it may therefore be difficult to properly disentangle the pure time-of-day preference and the impact of constraints when estimating the marginal utility functions. The schedule delay approach, on the other hand, is much more focused on the direct effect of scheduling constraints on the timing of activities and trips.

Schedule delay models have mostly been applied to model trip departure time choice in the context of single activities, such as work. Occasionally, the schedule delay approach has been applied to model the choice between activity patterns, combining schedule delay of the work trip with duration effects of work and other activities (Hess et al., 2005). A limitation, however, is that the schedule delay approach assumes the existence of some anchor point representing the optimal point in time to arrive or start an activity. This may work well for activities such as work, which are relatively fixed in time, but is more cumbersome for e.g. leisure activities.

From the above, one can conclude that three sources may be identified that influence the timing and duration of activities and thereby the emergence of activity patterns:

- 'intrinsic' time-of-day preferences, represented by continuous marginal utility functions (Ettema et al., 2003; Polak and Jones, 1994);
- fatigue effects, stemming from diminishing marginal returns, as described in the Becker-type time allocation models (Kitamura, 1984; Misra and Bhat, 1999).
- scheduling constraints of the Hägerstrand type as described by the schedule delay approach (Small, 1982);

An important goal of this paper is to test to what extent these three factors influence the timing and duration of activities. To this end, the models recently developed by Ashiru et al. (2004) and Ettema et al. (2004), including both time-of-day dependent components and duration dependent components, are extended with a schedule delay component. The resulting model is tested in the context of the combined choice of work tour and mode under a road-pricing regime. In specifying and testing the model, particular attention is given to incorporating both observed and unobserved heterogeneity, by including socio-demographic factors and error components.

3. Theoretical framework

3.1. Theoretical model

Our theoretical model follows some basic assumptions put forward by a number of other authors, namely that:

1. Individuals derive a certain utility from allocating time to activities (Becker, 1965; Yamamoto and Kitamura, 1996) and this utility depends both on the amount of time allocated and the time of day at which participation in the activity takes place (Ettema et al., 2004);
2. Individuals derive a certain (dis)utility from the time spent travelling (Ben-Akiva and Lerman, 1985);

3. Individuals aim at optimising the utility of their overall activity pattern, being the sum of the individual activity and trip utilities (Becker, 1965; Jara-Diaz; 1998a, 1998b Meloni et al., 2004).

Mathematically speaking, we assume that individuals maximise their utility by solving:

$$\max V = \max\,(V^T + V^A) \tag{2}$$

where V^T is the total utility derived from trips and V^A the total utility derived from activity participation. These utilities are the sums of the utilities of individual trips and activities:

$$V^T = \sum_m V_m^T \tag{3}$$

$$V^A = \sum_n V_n^A \tag{4}$$

Since our study focuses on timing and duration effects associated with activities, the utility of each individual trip *m* is defined as a relatively simple function of travel time ($R_m(s_m)$) and travel cost ($C_m(s_m)$) associated with trip *T* made at start time s_m. In addition, a constant D_m^l is included to represent the constant utility of a trip made by mode *l*:

$$V_m^T(s_m) = D_m^l + nR_m(s_m) + mC_m(s_m) \tag{5}$$

where m and n are the travel time and cost parameter respectively. It is noted that additional trip characteristics can be added without materially changing the approach. Scheduling costs, which represent the disutility of the diversion of some preferred arrival time for the trip, are not included in the utility of trips. Instead, these are represented in the utilities of activities through the implications for activity duration and timing. It is also noted that socio-demographics are included as adjustments to a specific parameter as follows:

$$b^s = \left[b_0 + \sum_{j=1}^{J} b_j d_j \right] \tag{6}$$

where:

b^{s} is the parameter for a person with socio-demographics defined by a *J*-dimensional vector *s*, containing dummy-variables;

b_0 is the base value of the parameter;

b_j is the adjustment for the *j*-th dummy variable;

d_j takes value 1 if the *j*-th element of *s* is unity and zero other wise.

The utility derived from an activity depends, as noted before, on three distinct elements:

- the time-of-day at which an activity is performed. In this respect we assume that there is some intrinsic preference for the time-of-day at which certain activities are undertaken;
- the duration of the activity, assuming that with increasing duration fatigue effects will come into play, resulting in a diminishing utility with increasing duration;
- the start time of the activity, relative to some reference point. In this respect, we assume that specific constraints, such as work hours are store-opening hours directly affect the timing of activities.

With respect to timing and duration, the above implies for instance that the first minute spent on an activity may be valued differently than the 10-th or 50-th minute, but the 10-th minute may be valued differently when engaged in at 7.00 AM or 2.00 PM. In addition, the start time of the activity will have an effect on the utility, which is, however, independent of the duration of the activity. To capture these effects, we define the utility U_n^A, derived from engagement in activity *n* as:

$$V_n^A = f(V_n^H(t_n, s_n), V_n^D(t_n), V_n^S(s_n, s_n^*)) \tag{7}$$

where:

V_n^H is the time-of-day dependent utility, depending on start time s_n and duration t_n;

V_n^D is the duration component, depending on duration t_n;

V_n^S is the schedule delay component depending on the preferred activity start time s_n^* and the actual start time s_n.

Focusing first on the three separate components, the time-of-day component is specified as the baseline utility profile, specifying the user benefit of being involved in an activity at a particular time of day. The time-of-day dependent utility is best understood in terms of the marginal utility $V'^{H}_{n}(t)$ specifying the amount of utility gained from participation during one time unit at time of day *t*. Although alternative specifications are available (see Ettema et al., 2004), the time-of-day component selected is based on a Cauchy distribution:

$$V'^{H}_{n}(t) = \frac{1}{c_n p\left[\left(\frac{t-b_n}{c_n}\right)^2 + 1\right]} * V_{\max,n} \qquad (8)$$

In this function, *b* defines the optimum location, that is where the utility is a maximum, *c* defines the width of the curve (which is symmetrical), which gives the time period in which an acceptable level of utility is gained, and finally V_{max} scales the Cauchy distribution (see Ettema et al., 2004 for examples of the effects of the parameters on the utility shape). The essence of the marginal utility component V'^{H}_{n} is to express that the utility derived from activity engagement intrinsically depends on the time-of-day. In this case we have assumed a bell-shaped curve to represent the time-of-day dependent utility, implying that the high marginal utility is concentrated in one period. Without materially changing the approach, however, alternative functions may be specified, implying for instance multiple periods with a high marginal utility. Socio-demographic variables can principally affect the time-of-day dependent utility through V_{max}, *b* or *c* in the same way as indicated in equation 6.

Given the marginal utility function $V'^{H}_{n}(t)$, the utility gained from activity *n* can be determined by integration if the start time s_n and duration t_n are known:

$$V^{H}_{n}(s_n, t_n) = \frac{V_{\max,n}}{p}\left(\arctan\left(\frac{s_n + t_n - b_n}{c_n}\right) - \arctan\left(\frac{s_n - b_n}{c_n}\right)\right) \qquad (10)$$

With respect to the duration dependent utility V^{D}_{n}, we assume that the utility follows a logarithmic function, as proposed by Yamamoto et al. (2000) and Bhat and Misra (1999):

$$V_n^D(t) = h_n \ln(t_n) \qquad (11)$$

This results in the following marginal utility:

$$V'^D_n(t) = \frac{h_n}{(t)} \qquad (12)$$

An important implication of this function is that marginal utility decreases with increasing duration, representing the fatigue effect, which is intuitively plausible. Socio-demographics can affect the duration dependent utility by modifying the constant h, as specified in equation (6). Finally, we define the schedule delay dependent utility of an activity n as:

$$V_n^S = g_n^e SDE_n + g_n^l SDL_n \qquad (13)$$

where ESD and LSD are the early and late schedule delay respectively, defined as (see Hess et al., 2005).

$$\begin{aligned} ESD_n &= \max(0, (S_n^* - S_n)) \\ LSD_n &= \max(0, (S_n - S_n^*)) \end{aligned} \qquad (14)$$

where S_n^* is the preferred start time of the activity. Having specified the components V_n^D, V_n^H and V_n^S, the total utility derived from an activity, V_n^A, is specified as a function of the respective parts. It is recognised that the components can be combined in different functional specifications (additive, multiplicative, logarithmic or combinations of these). Since this paper constitutes a first exploration of these utility components, we have chosen to use the most straightforward additive function:

$$V_n^A(t_n, s_n, s_n^*) = V_n^H(t_n, s_n) + V_n^D(t_n) + V_n^S(s_n, s_n^*) \qquad (15)$$

Because the components V_n^D and V_n^H are scaled by h and $V_{\max}$ respectively, it is not necessary to add weights to each component. The utility of activity n is thus defined by:

$$V_n^H(S_n, t_n, S_n^*) = \left[\frac{1}{p}\left(\arctan\left(\frac{S_n + t_n - b_n}{c_n}\right) - \arctan\left(\frac{S_n - b_n}{c_n}\right)\right)\right] V_{\max,n} \qquad (16)$$
$$+ h_n \ln(t_n) + g_n^e ESD + g_n^l LSD$$

3.2. Operational model

The operational model is applied to a home-based work tour and is defined in the current study as follows. Following the approach of Polak and Jones (1994), we assume that travellers choose the departure time of trips from home to work and from work back to the home. This effectively divides the day into three periods (pre-work, work, and after-work), which we regard, for simplicity, as single activities of which the utility is defined by equation (16). This implies that the total utility of a commuters' activity pattern *i* can be formulated as:

$$V_i = V_1^T + V_2^T + V_1^A + V_2^A + V_3^A + e_i \qquad (17)$$

with V_n^T as defined in equation 5 and V_n^A as defined in equation 16. Schedule delays were, however, only included for the work activity. In this respect, early and late schedule delays are defined relative to the current work start time, which is assumed to be the preferred start time. In the current study we will assume that an individual chooses between a limited number (say *N*) of feasible activity patterns [P_1,...,P_N] characterised by total utilities [V_1,...,V_N]. It is assumed that the choices made in the SP experiment reflect the preferences for certain time allocation patterns. In particular, the chosen alternative may be considered to be the closest match to an individual's unconstrained allocation outcome. Therefore, the discrete choice data can be used to disentangle the marginal utility functions that guide time allocation on a continuous scale. Assuming a Gumbel distribution for e_i in equation 17 leads to a multinomial logit model as the base specification.

Thus, it is assumed that discrete choice theory provides an adequate framework to model the choice of activity patterns. As noted previously, our model can account for heterogeneity associated with socio-demographic characteristics. However, heterogeneity in preferences may also arise from unobservable sources, such as taste variations. Such variations can be accounted for by more flexible formulations of the traditional discrete choice models (Train, 2003; Hensher and Greene, 2003) in which error components are included in the utility function. In particular, in this study, we assume that the evaluation of a certain utility component may vary from person to person, leading to addition of a random parameter, following a normal distribution with zero mean and unknown standard deviation. A complication in this case is that we work with stated preference data, in which multiple observations per person are available. In order to take account of correlations between choices that are expressed by the same individual, we will assume here that the error component terms only vary from individual to individual. Taking the schedule delay formulation as an example this is expressed as:

$$U^{S}_{n,ir} = (g^{e}_{n} + z^{e}_{n,i})ESD_n + (g^{l}_{n} + z^{l}_{n,i})LSD_n + e \qquad (18)$$

where $U^{S}_{n,ir}$ is the schedule delay utility of activity *n* experienced by individual *i* in replication *r*. In this formulation, $z^{e}_{n,i}$ is an error component with mean zero and some standard deviation $s^{e}_{n,i}$ to estimate. In the current study, error components will mainly be used to test for the heterogeneity in the evaluation of attributes. That is to say error components can in principle be added to any parameter $b_n, c_n, V_{\max,n}, h_n, g^{e}_{n}$ or g^{l}_{n} in equation (16). We will in addition use an error component term associated to sets of modes, in order to take a possible nest structure into account.

4. Stated Preference Data

The model proposed in section 2 was empirically tested using a stated preference data set, collected on various sites in The Netherlands in 2000 as part of a project to assess commuters' potential responses to various road user-charging schemes. Respondents were recruited by means of detailed screening and quota control criteria in which drivers undertaking work, employers business, shopping and social and leisure tours were selected.

The stated preference experiments involved respondents being offered realistic choices between alternative tour patterns. In order to avoid highly unattractive or highly unrealistic SP alternatives, these alternatives were developed based on the characteristics of the individual's current tour, which could include any type of activity.

During the SP experiment respondents were provided with a) re-timing options involving shifts earlier or later relative to the most temporally constrained activity; b) activity duration options; c) total two-way travel time options; and d) total road price charge options. In the survey, a public transport tour, similar to the most attractive existing PT tour, was offered as an alternative for the road pricing options. Thus, each respondent was offered four basic alternatives:

1. a car tour with departure times around the current departure times;
2. a car tour with departure times earlier than the current departure times;
3. a car tour with departure times later than the current departure times;
4. a public transport tour.

The data set provides data regarding the relevant choice dimensions incorporated in the model: activity timing and duration, trip duration and mode choice and is therefore suitable to test the model. To evaluate the model data for respondents who indicated that their current tour was a work trip were selected, as the resulting home-based tour is considered most likely to represent a daily activity pattern. After tests for data consistency and completeness, this resulted in some 1,382 observed choices, delivered by 188 individuals. For each subject, a limited number of socio-demographics were available, along with information regarding their working arrangements.

5. Estimation Procedure

Estimating the model outlined previously involves finding the parameters that maximise the goodness-of-fit of the logit model. Following Ben-Akiva and Lerman (1985) the log-likelihood function is formulated as:

$$LL(q) = \frac{1}{I} \sum_{i=1}^{I} \sum_{n} y_n \log P_n^i \qquad (19)$$

Where:

P_n^i is the probability that individual i chooses activity pattern n,

y_n^i is a dummy variable indicating whether individual i chooses alternative n,

I is the population size.

The parameters q are then computed by solving the program

$$\max_{q} f(q) = LL(q) \qquad (20)$$

The highly non-convex character of the log-likelihood function (19), caused by the non-linear utility specification, leads us to consider nonlinear programming approaches, especially trust-region methods. The main idea of a trust-region algorithm involves the calculation, at iteration k (with current estimate θ_k), of a trial point $q_k + s_k$ by approximately maximizing a model m_k of the objective function inside a trust region defined as

$$B_k = \{q \text{ such that } \|q - q_k\| \leq \Delta_k\}, \qquad (21)$$

where D_k is called the trust-region radius. We can for instance use a quadratic model:

$$m_k(s) = LL(q_k) + s^T \nabla_q LL(q_k) + \frac{1}{2} s^T H_k s, \qquad (22)$$

where H_k is a symmetric approximation of the Hessian $\nabla^2_{qq} LL(q_k)$. The predicted and actual increases in the value of the objective function are then compared by computing the ratio:

$$r_k = \frac{LL(q_k + s_k) - LL(q_k)}{m(q_k + s_k) - m(q_k)}. \qquad (23)$$

If this ratio is greater than a certain threshold, set to 0.01 in our tests, the trial point becomes the new iterate, and the trust-region radius is (possibly) enlarged. More precisely, if r_k is greater than 0.75, we set the trust-region to be the maximum between D_k and $2s_k$, otherwise we set $D_k = 0.5 D_k$. If the ratio is below the bound, the trial point is rejected and the trust region is shrunk by a factor of 2, in order to improve the correspondence of the model with the true objective function. We have followed Conn et al. (2000) in our choice of the parameters.

We additionally constrain the parameters $V_{\max,n}$ end c_n of the marginal utility (8) to be strictly positive, since its integral (10) is discontinuous at c_n equal to zero. Moreover we assume that

V_n^H is positive, so $V_{\max,n}$ and c_n must be of the same sign. Note however that if $V_{\max,n}$ or c_n converges to zero, both corresponding time of day marginal utility and its integral vanish. Therefore, if some of the positiveness constraints are active at the solution, the associated time of components do not add useful information to the model, and can be excluded from it. The resulting model is then unconstrained, and can be estimated using standard nonlinear programming techniques.

5. Empirical results

Using the above algorithm, the model according to equation (17) was estimated. Since an important aim of the study is to disentangle the different utility components of activity timing and duration the following approach was taken. First, a base model was estimated, including trip characteristics and duration dependent components. Thus, no timing or scheduling preferences are included in this model. Next, the time-of-day dependent and schedule delay components were added, first separately and then jointly, to test the effect of both components. Finally, error components were added to test for unobserved heterogeneity. While this slightly differs from the specification given in equation (16), alternative specific constants were added to the early and late car trip alternatives, as well as to the public transport alternative, since this significantly increased the goodness-of-fit of the models. In addition, by estimating a nested logit model, we tested whether the IIA property of the MNL model was violated as a result of correlations in the error terms of the car alternatives. This appeared not to be the case, suggesting that the MNL provides a good starting point for the estimation efforts. Based on the MNL formulation various models were estimated, which are summarised in Table 1.

Based on this consideration, the base model (Model 1) was specified describing the choice of mode and timing of the out and in bound commute trip, displayed in Table 1. Note that this model follows a standard MNL formulation. The base model includes mode and timing specific constants, activity durations and trip characteristics. With respect to the inclusion of socio-demographic variables, it should be noted that various models have been estimated, including different specifications of socio-demographic variables. The models that are presented include socio-demographic variables that were influential in various model specifications. The results suggest that travellers value an earlier or later retiming of their current tour negatively. Since the time-of-day dependent component U^H is not included in this model, this may reflect scheduling constraints stemming from the work organization but may also represent pure time-of-day preference.

Table 1: estimation results of models of activity pattern choice

	Model 1	**Model 2**	**Model 3**	**Model 4**	**Model 5**
$D^{car,early}$	-1.08***	-1.14***	-0.45***	-0.46***	0.20
$D^{car,late}$	-2.44***	-1.83***	-0.95***	-0.74***	-0.43
$D^{public\,transport}$	-2.03***	-2.21***	-2.08***	-2.25***	-5.51***
$h_{pre-work}$	1.99**	5.72***	3.48**	6.70***	6.79*
h_{work}	1.26**	0.45	0.96	0.58	-0.03
$h_{post-work}$	0.81	0.26	-0.34	-0.37	-1.69
$h_{post-work,male}$	0.46	0.71	1.32**	1.40**	2.59***
$h_{pre-work,high-educ}$	1.04*	3.99***	1.78***	3.80**	0.40
ESD_{work}			-0.0067*	-0.0141***	-0.067***
LSD_{work}			-0.0183***	-0.0093**	-0.031***
b_{work}		579.76***		569.05***	580.93***
c_{work}		66.801***		69.696***	83.72***
$V_{max,work}$		4.66***		6.24***	14.47***
$V_{max,work,higheduc}$		2.23*		1.18	-1.95
n_{car}	-0.0076***	-0.0028	-0.0070***	-0.0005	-0.0092*
n_{PT}	-0.0061**	-0.0017	-0.0062**	-0.0003	-0.0123
m_{car}	-0.040***	-0.0607***	-0.0582***	-0.075***	-0.1799***
m_{PT}	0.0002	0.0035	0.0015	0.0050	0.0108
z_{PT}					4.028***
$z_{ESD,work}$					0.042***
$z_{LSD,work}$					0.030***
LL(0)	-1904.77	-1904.77	-1904.77	-1904.77	-1904.77
LL(b)	-1323.16	-1281.5	-1254.72	-1219.09	-952.286

* : significant at α=0.10
** : significant at α=0.05
*** : significant at α=0.01

An additional finding is that they also prefer a car-based tour to a tour by public transport. With respect to the utility of activity durations, the results suggest that activity involvement leads to an increased utility for all activities (although not significant for the post work activity). One way of interpreting this is that activity engagement is preferred over travel. The negative coefficients for travel time by car and public transport support this view. Comparing the activity types, time spent before work appears to be the most important factor in activity scheduling. An important observation is that apparently, the scheduling of the work activity is driven both by duration considerations and by constraints or time-of-day preferences. Looking at the effect of socio-demographics, we find that highly educated individuals value the pre-work time higher. Finally, car cost is valuated negatively, as expected. Public transport cost is not found to significantly affect the choice of activity pattern.

The second model is an extension of the base model in the sense that the time-of-day dependent marginal utility of work is included in the model. Time-of-day dependent components of the other activities did not prove to be significant. The reason for this is probably that the pre-work and the post-work period comprise of multiple activities, making it more difficult to link preferences to particular times of day. Work utility, however, not only depends on the duration of activities, but also on the time at which one engages in the activity. Adding the time-of-day dependent component results in a significant improvement over model 1 (Table 2). The estimation results indicate that including the additional factors affect some of the parameters in the base model 1. The work duration parameter is not significant in this model, suggesting that the time-of-day dependent utility component better captures the scheduling preferences of the work activity. However, the alternative specific constants for early and late work start are still significant. This suggests that both time-of-day preferences and scheduling constraints affect the scheduling of the work tour. Another change is that the travel time parameters are not significant in this model, implying that the travel time implications are accounted for by the time-of-day dependent effect of work participation. The estimated parameters suggest that the highest marginal utility of work is gained around 9.40 AM (b_{work}). The c_{work} estimate implies that the time-of-day dependent utility is considerable between 8.30 AM and 10.30 AM. The $U_{\max, work} * HighEduc$ parameter suggests that highly educated individuals have a higher time-of-day dependent work utility than lower educated individuals, which is probably related to their higher wage rates. The marginal utility derived from activity participation, based on U_n^H and U_n^D is displayed in Figure 1. The figure illustrates clearly that the marginal utility of the pre-work and the post-work period is continuously increasing, since it is only determined by the U_n^D components. The marginal utility of work is jointly determined by the duration and the time-of-day dependent component, resulting in a utility that is more or less stable utility in the period between 8.30 AM and 10.30 AM and than decreases gradually. The figure also illustrates the differences in marginal utility between different socio-demographic groups.

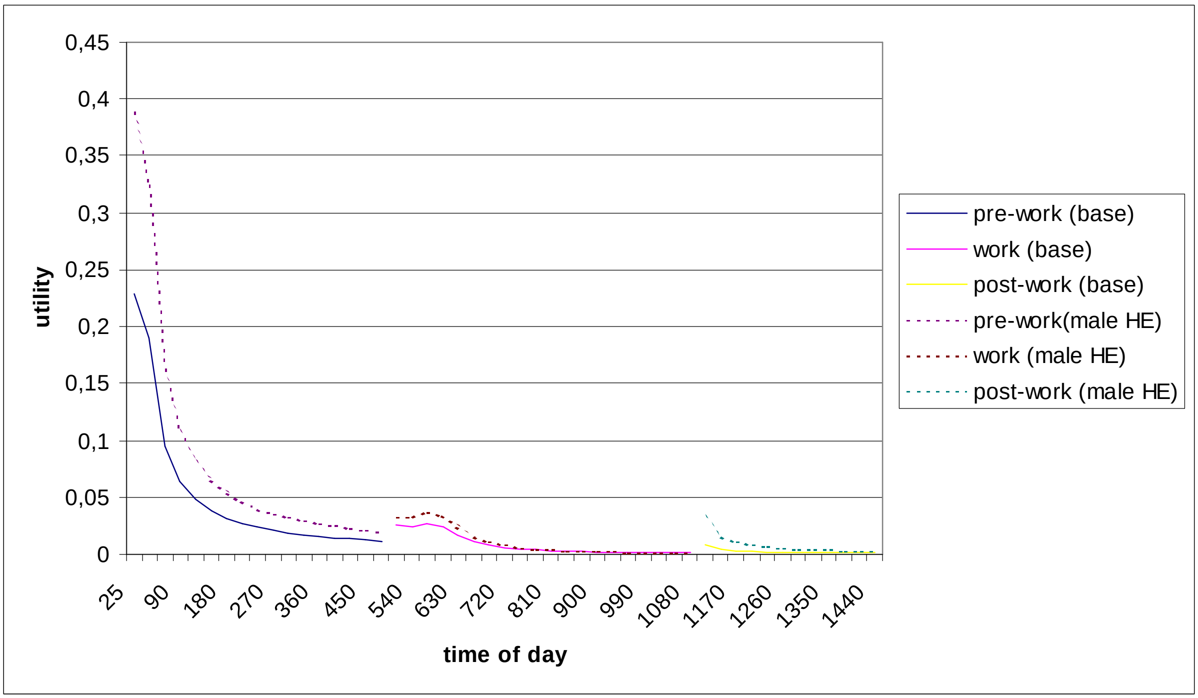

Figure 1: marginal utilities by time of day for different socio-demographic segments

A third model extends the base model 1 by adding schedule delay factors related to the work activity. Also in this case, including the additional factors lead to a significant improvement in the goodness-of-fit (Table 2). It does not change most parameters of the base model significantly. Again the work duration parameter is no longer significant, suggesting that the schedule delay parameter also captures a considerable part of the effect of work duration. As expected, the constants for early and late departure are affected by the inclusion of schedule delay. Both the early and late work start constants diminish in size, suggesting that the constant effect is now captured by the schedule delay parameters. They are, however, still significant, suggesting that there exists an a priori reluctance to deviate from the current work start time. As indicated by the SDE and SDL parameters, both early and late schedule delay are valued negatively, with late schedule delay being about three times as negative as early schedule delay. These findings are in line with the literature in this area (e.g. Small, 1982; De Jong et al., 2001). The negative parameters for travel time by car and public transport (model 1) are not affected by including the schedule delay factors. This suggests that the valuation of travel time implications does not coincide with the valuation of schedule delays, which is in line with the expectations.

The fourth model combines the time-of-day dependent utility effects with the schedule delay effects and the duration effects. Combining these effects leads to a model with a significantly better goodness-of-fit than the previous models (Table 2), suggesting that individuals' choice of activity patterns is affected by all three time related components (duration, schedule delay and time-of-day dependent utility). Comparing this model with model 3, it is seen that the early schedule delay parameter is affected by the inclusion of the time-of-day dependent variables in size and significance. The early schedule delay parameter increases in size and significance. Thus, starting work before the preferred start time produces a positive utility, based on the time-of-day dependent utility component, but also a disutility stemming from the early schedule delay. These are two counterparts of the same situation that become only apparent in this model specification. As in model 2, the inclusion of the time-of-day dependent component leads to an increase in the pre-work duration parameter. This suggests that the involvement in work at particular times is weighted against the time one can spend at home before work. Again, inclusion of the time-of-day dependent utility of work makes the time parameters of car and public transport non-significant, suggesting that the valuation of changes in travel duration coincide with the valuation of changes in involvement in work. The GOF of model 4 is significantly better than models 2 and 3. This suggests that all three time related components capture different components of individuals' timing preferences and that including all three adds to the explanation of activity scheduling behaviour. The estimation results of this and previous models also suggest, however, that there are rather subtle relationships between the various components of the utility of activities and trips, which may partly overlap and correlate. Based on the estimation results we can conclude, however, that scheduling constraints pertaining to the work activity, the time-of-day dependent utility of work, the duration of the pre-work time and the duration of the post-work time (for males) seem to be the most important determinants in deciding how to schedule the work tour. Travel time seems to be important to the extent that it affects the involvement in work in the period between 8:30 and 10:30 when the marginal utility of work is highest.

The fifth and final model extends the fourth model by including various error components. As noted previously, various error components were tested, maintaining only those that are significant. The final model includes error components for $D^{public\ transport}$, ESD_{work} and LSD_{work}, suggesting a high degree of unobserved heterogeneity with respect to the associated parameters and preferences. A first conclusion is that including the error components has a large effect on the goodness-of-fit (Table 2), suggesting that the activity scheduling process that we are modelling is surrounded by much unobserved heterogeneity. For the mode error component z_{PT} (associated with the public transport dummy), the existence of heterogeneity is in line with findings in other studies indicating that travellers may have intrinsic preferences for a particular mode. In addition, the mode error component seems to affect the alternative specific constants. The public transport constant is larger (negative), while the early and late rescheduling constants are not significant in this model. The error components associated with early and late schedule delay, $z_{ESD,work}$ and $z_{LSD,work}$, suggest that significant variation exists in the valuation of early schedule delay. Such variation may be due to differences in the household situation, such as obligations for childcare and serve passenger trips, or differences in the flexibility of work hours. This taste variation may also affect the early and late rescheduling constants. It is also noted that other parameter are affected by the inclusion of the error components. The valuation of pre-work time ($h_{pre-work}$) in general, but also by highly educated travellers ($h_{pre-work,high-educ}$) suggests that this parameter may capture part of heterogeneity, but the classification into education strata is not in itself relevant. Finally, the scaling of travel time and cost parameters seems to be affected by inclusion of the error components.

Table 2: Likelihood ratio tests of nested models

	model 2/ model 1	**model 3/ model 1**	**model 4/ model 2**	**model 4/ model 3**	**model 5/ model 4**
c^2 of likelihood ratio	83.32	136.88	124.82	71.26	533.61
Degrees of freedom	4	2	2	4	3
Significance level	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01

6. Conclusions

This paper has proposed a utility-theoretic framework for timing and duration preferences embedded in a multi-dimensional choice model, which can be formulated in a flexible way as a GEV-model. Doing so, the proposed framework combines a utility-theoretic underpinning in GEV modelling with a very flexible formulation of time and duration preferences, which includes the effect of activity duration, the effect of time-of-day of activity involvement and the effect of start time of the activity relative to some anchor point. It is noted that from a theoretical point of view, the time-of-day dependent utility and the valuation of schedule delay overlap to some extent. It may therefore be difficult to disentangle these effects in empirical studies. The model framework further includes socio-demographic and context variables, that may affect individuals' valuation of timing and duration.

The model was empirically tested using a Dutch data set, accounting for context variables such as gender and education level. Various model specifications were estimated, differing with respect to the timing and duration related variables that are included. Some important conclusions can be drawn from the estimated models. First, the three distinct aspects of utility (duration, time-of-day and schedule delay) all appear to have some effect on activity scheduling preferences. The effect differs, however, between activities. Based on the estimation results we can conclude that scheduling constraints pertaining to the work activity, the time-of-day dependent utility of work, the duration of the pre-work time and the duration of the post-work time (for males) seem to be the most important determinants in deciding how to schedule the work tour. An important finding is that travel times do not appear significant in the model if time-of-day dependent utility is included in the model. Apparently, travel time primarily seems to be important to the extent that it affects the involvement in work in the period between 8:30 and 10:30 when the marginal utility of work is highest. Including error components results in a significant improvement of the model fit, suggesting that the activity scheduling process that we are modelling is surrounded by much unobserved heterogeneity. The unobserved heterogeneity is found in the valuation of early and late schedule delay and preference for public transport.

In addition, the valuation of activity involvement varies between socio-demographics segments. Without testing all socio-demographic impacts exhaustively, the results suggest that highly educated individuals value the pre work time more highly, and that the post work period is appreciated more by males.

The reported work provides a starting point for further research in various ways. First, more extensive estimation efforts have to be made, including a broader range of socio-demographic and context variables.

Second, as the activity patterns used in this study only entail timing, duration and mode as choice dimensions, more elaborate models need to be estimated which include additional choice dimensions such as activity choice and destination choice. These models will then constitute a realistic base for the further development of activity-based models.

References

Ashiru, O, Polak, J.W., Noland, R.B. (2004), The utility of schedules: A theoretical model of departure time choice and activity time allocation with application to individual activity schedules. *Transportation Research Record*. In press.

Arentze and Timmermans (2005), *ALBATROSS version 2.0. A Learning Based Transportation Oriented Simulation System*, EIRASS, Eindhoven.

Becker, G. (1965), A Theory of the Allocation of Time, *Economic Journal*, **75**, 493-517.

Ben-Akiva, M. and J.L. Bowman (1998), Integration of an Activity-Based Model System and a Residential Location Model, *Urban Studies*, **35**, 1131-1153.

Ben-Akiva, M. and S.R. Lerman (1985), *Discrete Choice Analysis: Theory and Application to Travel Demand*, MIT Press, Cambridge, Massachusetts.

Bhat, C.R. and R. Misra (1999), Discretionary Activity Time Allocation of Individuals between In-home and Out-of-home and between Weekdays and Weekends, *Transportation*, **26**, 193-209.

Conn, A.R., Gould, N.M. and Toint, Ph. L. (2000), *Trust-Region methods*, SIAM, Philadelphia, USA.

De Jong, G., A. Daly, M. Pieters, C. Vellay, M. Bradley and F. Hofman (2003), A model for time of day and mode choice using error components logit, *Transportation Research E*, **39**, 245-268.

DeSerpa, A. (1971), A Theory of the Economics of Time, *The Economic Journal*, **81**, 828-846.

Ettema, D. (2005), Latent activities: modelling the relationship between travel times and activity participation. To appear in *Transportation Research Record*.

Ettema, D., O. Ashiru and J. Polak (2004), Modeling Timing and Duration of Activities and Trips in Response to Pricing Policies. *Transportation Research Record*. In press.

Ettema, D.F. and H.J.P. Timmermans (1997), *Activity-Based Approaches to Travel Analysis,* Pergamon, Oxford.

Ettema, D.F. and H.J.P. Timmermans (2003), Modeling departure time choice in the context of activity scheduling behavior, *Transportation Research Record*, **1831**, 39-46.

Evans, A. (1971), On the Theory of the Valuation and Allocation of Time, *Scottish Journal of Political Economy*, **2**, pp.1-17.

Hägerstrand T, 1970, "What about people in regional science?" *Regional Science Association Papers,* **24,** 7-21.

Hensher, D.A. and W.H. Greene (2003), The Mixed Logit model: The state of practice, *Transportation*, **30**, 133 - 176.

Hess, S., Polak, J.W., Daly, A.J. & Hyman, G. (2005), Flexible Substitution Patterns in Models of Mode and Time of Day Choice: New evidence from the UK and the Netherlands, Transportation, forthcoming.

Jara-Diaz, S.R., (1998a), *Time and income in travel choice: Towards a microeconomic activity based theoretical framework*. In: Gärling, T., Laitila, T., and Westin, K. (eds.), *Theoretical foundations of travel choice modelling*, Pergamon, Elsevier, Amsterdam, 51-73.

Jara-Diaz, S.R., (1998b), A general micro-model of users behavior: The basic issues. In: Ortúzar, J.D., Hensher, D.. and Jara-Diaz, S.R. (eds.), *Travel Behavior Research: Updating the state of play*, Pergamon, Oxford.

Joh, C-H, T.A. Arentze and H.J.P. Timmermans (2002), A theory and simulation model of activity-travel rescheduling behavior, submitted for publication in *Transportation Research A*.

Kitamura, R. (1984), A Model of Daily Time Allocation to Discretionary Out-Of-Home Activities and Trips, *Transportation Research B*, **18**, 255-266.

Polak, J. and P. Jones (1994), Travellers'choice of time of travel under road pricing, paper presented at the 73rd Annual Meeting of the Transportation Research Board, Washington D.C.

Small, K.A. (1982), The scheduling of consumer activities:work trips, *The American Economic Review,* **72,** 467-479**.**

Wang, J.J. (1996), Timing utility of daily activities and its impact on travel, *Transportation Research A*, **30**, 189-206.

Yamamoto, T., S. Fujii, R. Kitamura, H. Yoshida (2000), An analysis of time allocation, departure time and route choice behavior under congestion pricing, paper presented at the 79th Annual Meeting of the Transportation Research Board, Washington, D.C.

Zhang, J., Timmermans, H., Borgers, A. (2002), A utility-maximizing model of household time use for independent, shared and allocated activities incorporating group decision mechanisms, *Transportation Research Record* 1807, 1-8.