

Effects of task complexity and time pressure on activity-travel choices: heteroscedastic logit model and activity-travel simulator experiment

Chao Chen · Caspar Chorus · Eric Molin · Bert van Wee

Published online: 17 February 2015

© The Author(s) 2015. This article is published with open access at Springerlink.com

Abstract This paper derives, estimates and applies a discrete choice model of activity-travel behaviour that accommodates potential effects of task complexity and time pressure on decision-making. To the best of our knowledge, this is the first time that both factors (task complexity and time pressure) are jointly captured in a discrete choice model. More specifically, our heteroscedastic logit model captures potential impacts of task complexity and time pressure through the scale of the utility of activity-travel options. We collect data using a novel activity-travel simulator experiment that has been specifically designed with the aim of testing our model. Results are in line with expectations, in that higher levels of task complexity and time pressure are found to result in a smaller scale of utility. In other words, higher levels of task complexity and time pressure lead to more random choice behaviour and as a consequence to less pronounced differences in choice probabilities between alternatives. An empirical illustration suggests that choice probability-differences between models that do and those that do not capture these effects, can be very substantial; this in turn suggests that failing to capture the effects of task complexity and time pressure in discrete choice models of activity travel decision-making might lead to serious bias in forecasts of the effects of transport policies.

Keywords Task complexity · Time pressure · Activity-travel choice · Heteroscedastic logit · Activity-travel simulator experiment

Introduction

Daily activity-travel choices are often highly complex, in the sense that there are many alternatives to choose from and that these alternatives are multi-dimensional (i.e., consist

C. Chen · C. Chorus (✉) · E. Molin · B. van Wee
Transport and Logistics Group, Delft University of Technology, Jaffalaan 5, 2628 BX Delft,
The Netherlands
e-mail: c.g.chorus@tudelft.nl

of combinations of activities and associated travels) and sometimes difficult to compare (i.e., resemble one another in terms of attractiveness). Although the complexity of daily activity-travel decision making has been widely acknowledged since decades (e.g., Recker et al. 1986; Kitamura 1988; Arentze and Timmermans 2004), a limited amount of attention has been directed towards incorporating task complexity in discrete activity-travel choice models—but see Arentze et al. (2003) and (Caussade 2005) for notable contributions. In fields adjacent to transportation, more examples can be found where task complexity is explicitly captured in discrete choice models (e.g., Swait and Adamowicz 2001; DeShazo and Fermo 2002; Dellaert et al. 2011).

A similar but even more salient argument holds for the notion of time pressure. It is obvious that in many activity-travel contexts, pre-, but especially en-route, time pressure¹ is a potentially important factor influencing decision-making processes: consuming too much decision-time comes with the risk of late arrival at the activity location due to, for example, missing a bus or a highway-exit. However, we know of no attempts to explicitly capture, in a discrete choice model, the influence of time pressure on the making of operational activity-travel choices. Also in fields adjacent to transportation, we were unable to find studies that aim to capture time pressure in a discrete choice model.

This paper contributes to the travel behaviour modelling literature (i) by being the first to explicitly and jointly model task complexity and time pressure in a discrete choice model of activity-travel choices, and (ii) by estimating and testing the resulting model using data from a novel activity-travel simulator experiment that has been specifically designed for that purpose. Our modelling approach is inspired by previous studies (e.g., DeShazo and Fermo 2002; Caussade 2005; Dellaert et al. 2011) that have built and estimated heteroscedastic logit models whose scale is conceived to be a function of task complexity (and, in our study, time pressure). Our data collection approach builds on previous studies that used dynamic, interactive simulator-experiments to study activity-travel behaviour (e.g., Bonsall and Palmer 2004; Chorus et al. 2007; Sun et al. 2012).

“[Model development](#)” section presents our discrete choice model of activity-travel behaviour, incorporating task complexity and time pressure. “[The activity based travel simulator](#)” section presents the activity-travel simulator used for data collection. “[Model estimation](#)” section discusses model estimation and parameter interpretation, and “[Results](#)” section presents conclusions and potentially fruitful directions for further research.

Model development

Random utility maximization or RUM (McFadden 1973) is the dominant theoretical paradigm underlying the modelling of discrete choices. The RUM decision rule assumes that decision-makers choose that alternative from a set of choice alternatives from which they derive the highest utility. The utility consists of a systematic part and a random part reflecting the idiosyncrasies of the choice process and possibly unobserved attributes: $U_i = V_i + \varepsilon_i$. Here, U_i is the utility of alternative i ; and V_i is the systematic component of the utility; ε_i is a random component. Although other specifications are possible, the structural utility V_i is typically modelled as linear in the parameters function. In case of K distinctive attributes, V_i has the following form: $V_i = \sum_{k=1}^K \beta_k \cdot x_{ik}$, where β_k is a

¹ This paper does not focus on time pressure of the type that is experienced by someone who aims to fit an overly ambitious activity-travel program in a day’s 24 h. Rather, we focus on the time pressure that is experienced when there is less time than needed or wanted, to make an operational activity-travel decision.

parameter which is to be estimated and which expresses the weight (taste) regarding attribute k ; and x_{ik} is the value of attribute k of alternative i . Depending on the assumptions regarding the distribution of the random utility component, different choice probability formulations arise. If it is assumed that the random component ε_i is independently identically distributed (IID) Extreme value type I, the well-known multinomial logit model (MNL) model arises. In this model, utility is related to choice probability as follows:

$$P(i) = \frac{e^{\mu V_i}}{\sum_{j \in C} e^{\mu V_j}}$$

Parameter μ is the scale parameter, which is inversely related to the variance of the error component: $\text{var}(\varepsilon) = \pi^2/(6\mu)$. The scale parameter is not identifiable jointly with taste parameters, and is therefore normalized. A typical normalization is to set the scale to 1, implying that the variance of the random error equals $\pi^2/6$. The IID assumption underlying the MNL model involves that the random error for alternative j is independent from that of alternative i , and that the error term distributions of all alternatives have the same variance. This latter assumption is called homoscedasticity. One approach that has been successfully used in previous studies (see references in the introduction) is to handle the impacts of task complexity on choice behaviour is to allow for the variance of the random component in the utility function to be a function of task complexity, which is equivalent to the notion that the scale of the utility is a function of task complexity.²

As each choice task may be associated with a different level of task complexity, the scale is no longer identical for all choice tasks, which gives rise to a more flexible model, called heteroscedastic logit (HL). The core feature of HL models is that the random component is no longer identically (i.e., with equal variance) distributed across alternatives [e.g. (Daganzo 1979), Bhat (1995)]. DeShazo and Fermo (2002) utilized a HL model to evaluate the impacts of the complexity of choice sets on choice consistency. Arentze et al. (2003) took a similar approach to demonstrate that the variance of the random component rises with the increase of task complexity. Caussade (2005) and Dellaert et al. (2011) developed HL models with the scale parameter being specified as a function of task complexity.

In this paper, we propose to model the impact of time pressure on a traveller's choices in a similar fashion as the impact of task complexity, and hence, we incorporate time pressure in a heteroscedastic model. Thus, in the resulting HL model³ the scale parameter μ is no longer considered to be a constant but it is parameterized as a function of task complexity and time pressure of the choice task s . This function takes the following form to ensure non-negativity: $\mu_s = \exp(a(D_s, T_s, Int(D_s, T_s)))$, where $a()$ is a linear function of its arguments and associated parameters; D_s is the measurement of task complexity in choice situation s ; T_s is the measurement of time pressure in choice situation s ; $Int(D_s, T_s)$ is the measurement of possible interaction effects involving both task complexity and time pressure.

² As a referee noted, another way to understand and model traveller response to increasing levels of task complexity and time pressure would be to hypothesize that they switch from a linear-additive utility-maximization decision rule, to a presumably less computationally demanding rule such as—for example—elimination by aspects (Tversky 1972) or Satisficing (Simon 1955). We leave the exploration of such alternative responses for further research. Note that we did check if choices under high levels of task complexity and time pressure became Lexicographic (e.g. an individual always choosing the fastest mode, irrespective of the performance on other attributes). We found that none of the individuals in our sample exhibited Lexicographic behavior regarding either the time or cost attribute, which can be considered as some support for our approach to model behavior using the linear-additive utility-maximization rule.

³ More specifically, to accommodate for task complexity and time pressure while allowing for unobserved (i.e., random) heterogeneity in tastes, we use a Mixed Logit version of the HL model.

We expect that if task complexity and/or time pressure increases, decision-makers will have more trouble choosing the alternative from which they derive the highest utility. Hence, task complexity and time pressure are expected to increase the randomness in the choices made, resulting in a larger variance of the error component. Consequently, we hypothesize that if the complexity of the choice task increases and/or time pressure increases, the scale parameter μ_s will become smaller.

Measuring task complexity

We choose to measure task complexity in terms of the time used by the individual to make a decision, under the condition that no time constraint is present. The idea is that if someone takes a long time to reach a decision, this can be considered a proxy for, or a signal of, the complexity of the choice task (see also Diederich 2003). Of course, many other ways exist in which task complexity can be and has been operationalized, such as counting the number of alternatives and attributes in a choice task (Arentze et al. 2003), or computing the entropy of a choice task (Swait and Adamowicz 2001). One advantage of using decision time as a measure of task complexity, is that it generates a substantial amount of intra- and interpersonal variation; as such it allows for the efficient statistical inference of potential relations with dependent variables (such as, in our case, the scale of utility).

However, it should be noted that a longer decision time may—in addition to task complexity effects—reflect other effects, such as for example relating to the similarity of the alternatives on offer, or (the absence of) dominant alternatives. But note that choice set composition was not varied systematically in our experiment, whereas task complexity in terms of the number of alternatives was; as a consequence, we may assume that a substantial part of the systematic variation in decision time will be due to variation in task complexity (if such an effect exists in reality). Furthermore, it is worth noting at this point that decision time may also be related to the presence or absence of habitual behavior: if someone (in real life) always chooses the car-option for her daily commute, this decision will at some point not take a lot of time anymore. Note that our experiment—although designed to be to some extent a realistic account of real life—will still differ substantially from actual day-to-day travel choices made respondents. As such, we assume that habitual behavior (and its effect on decision time) in real life will only transfer to a limited extent to our experimental setting. Finally, it may be noted that we also estimated models where task complexity was measured directly in terms of the number of alternatives on offer and other observable characteristics of the choice situation, to avoid any potential confounding of the decision time proxy with other factors such as choice set composition and habit (see directly above). The fit of these alternative models with the data however, was substantially worse than that of the models which used the decision time as a proxy for task complexity; this made us decide to use the latter in the remainder of our analyses.

Measuring time pressure

Previous studies into time pressure have conceived time pressure in terms of how much time a decision-maker is allowed when making his decision, i.e. the decision time budget. This time-budget is usually a priori constrained and systematically varied by researchers (e.g. Nowlis 1995; Ordóñez and Benson Iii 1997; Dhar and Nowlis 1999). It is subsequently hypothesized that the less decision time budget decision-makers have, the more time pressure they experience, *ceteris paribus*. However, such a measure does not take into

account the actual time that is used by a decision maker to make the decision. For example, if the budget is 60 s, it may be argued that a decision-maker that took 59 s to make a decision felt more pressured than a person that took only 30 s, as the former has almost used up all his or her decision time budget. Therefore we propose a measure of time pressure, which relates the actual decision time consumed to the decision time budget. It is formulated as follows: $DS_s = DT_s/DTB_s$, where DS_s is the time pressure measure for choice situation s ; DT_s is the actual decision time for choice situation s ; and DTB_s is the decision time budget for choice situation s . For the same example above, the value of this new measure DS_s equals 0.98 for the first decision-maker and 0.50 for the second decision-maker.

Note that since in our HL model an increase of time pressure is assumed to be associated with an increase in the randomness of choice, one may be compelled to expect that the scale would become smaller as DS_s increases, and hence, that the scale monotonically decreases as a function of DS_s . However, when the value of DS_s is close to 0, this indicates that the decision maker only used a very small fraction of the available time budget, which may also be interpreted as a signal that he or she did not care about choosing the best alternative. In other words, a low value of DS_s may indicate absence of engagement with the choice task. Following this argument, (very) low values of DS_s are expected to lead to relatively low scale values. When this is indeed the case, an inverted U-shaped curve, rather than a monotonic decreasing relation between time pressure and scale, is to be expected. Of course, which of the two possibilities is correct is an empirical question. To allow for both a time pressure and an engagement effect, our HL-models employ linear as well as quadratic terms for the DS_s variable. To the extent that an inverted U-shape is found in the process of model estimation, DS_s should be considered an engagement/time pressure index rather than a time pressure index alone.

The activity based travel simulator

In order to estimate the model developed in the previous section, choices need to be observed for different contexts of task complexity and time pressure. In order to be able to control for especially time pressure, we rely on stated preference methods. Compared with conventional SP methods, travel simulators typically stimulate respondents to be more actively involved in the experiment, provide illustrative and interactive user interfaces, and most importantly for our study, provide the researcher with more control about experimental conditions (e.g., Chen and Mahmassani 1993; Mahmassani and Jou 2000; Bonsall and Palmer 2004; Chorus et al. 2007; Prendinger et al. 2011). In the following, we describe the activity-based travel simulator (ATS), a travel simulator which we specifically developed for this study.

The starting point for the development of ATS is the notion that a traveller needs to conduct some activities in a normal workday, such as working, grocery shopping, meeting friends, etc., the so-called activity program. In order to limit the number of combinations, only a few typical activities from different activity categories are selected in ATS, that is: (1) *Primary activities* work; (2) *Maintenance activities* grocery shopping and fitness; (3) *Leisure* leisure shopping and meeting friends. In order to execute all activities, people have to travel between respective locations. Conducting a daily activity program usually consists of several trips, for which typically various modes are available to choose from. Moreover, the timely order to execute activities (defined as activity sequence) may differ. To reflect these activity-travel options, an alternative in a choice set in our ATS describes the

execution of a complete activity program for a given day, which contains the following basic elements: (i) a timely ordered sequence of the activities; (ii) geographic locations of the activities; and (iii) a trip between the activity locations, including trip modes, their respective travel time and travel cost.

Participants were asked to assume that they recently moved to the hypothetical travel environment created in the ATS. This environment, displayed in Fig. 1, basically involves two cities where activities can be conducted. The supposed home is located in Stad A (city A), where one can conduct most activities, while the work location is located in Stad B (city B). All the daily activity program alternatives are graphically displayed in the main screen in a sequential animation with arrows indicating the sequence of the activities and required travel. The arrows are accompanied with the icons denoting the travel mode for each trip, accompanied by its related travel time and costs. The mode is depicted as an icon. Some icons additionally illustrate the additional travel burden associated with traveling by that mode for a particular activity, e.g. a bicycle with a shopping bag in the basket in case of a grocery trip by bicycle. In addition, the complete activity schedule including all trips is shown in one string at the bottom of the screen, denoting the same information in a different format. It is important to note that at any moment in time, the specific information of *only a single* alternative can be shown on the interface, so that it is impossible to see all the alternatives in a choice set on a single screen. However, the respondent can easily switch between the different screens to become acquainted with all the alternatives of the choice task.

Travel times and costs that correspond to the travel mode for a trip between two activities, are randomly drawn from a predefined range of values for each choice task and for each participant, as presented in Table 1.



Fig. 1 An example of the ATS computer screen

Table 1 Ranges of attribute-values for the travel trips

	Travel time (min)	Travel cost (€)
<i>Within neighbourhood</i>		
Cycling	6–10	0
Car	3–5	1–2
<i>Between neighbourhood and train station A</i>		
Cycling	12–16	0
Car	6–10	2–3
<i>Between train station A and integrated facility</i>		
Walking	1–3	0
Cycling	1–3	0
<i>Between train station A and train station B</i>		
Train	35–45	4–6
<i>Between train station B and office</i>		
Walking	3–5	0
<i>Between neighbourhood and office</i>		
Car	46–56	7–10

Variations in task complexity and time pressure

Choice task complexity levels are varied by varying: (i) the number of alternatives in the choice set, which varies between 2 and 4, and (ii) by varying the number of activities within each choice alternative, which varies between 1 and 4. Table 2 shows the six task complexity combinations used for this study.

In the first part of the experiment, there is no predefined time budget for making choices (i.e., participants could take as long as they want to make a choice). Every participant is first presented with a series of six choice tasks, one for each of the choice task numbers as presented in Table 2, and in order of increasing complexity. The time it takes the participant to make a decision is recorded. After a break, another series of six choice tasks is presented to each participant, again each representing any of the six choice task numbers of Table 2. However, this time the participant has to make the choices under the condition of a time constraint.

The decision budget for this second series of choices needs to be set in such a way that a participant is neither stressed out by a very small time budget nor overly relaxed due to a

Table 2 Variations in task complexity (CTnr = choice task number)

Nr. of activities in an activity program	Nr. of travel alternatives in choice set		
	2	3	4
1	<i>CTnr 1:</i> work		
2	<i>CTnr 2:</i> work, grocery shopping	<i>CTnr 3:</i> work, fitness	
3		<i>CTnr 4:</i> work, fitness, grocery shopping	<i>CTnr 5:</i> work, meeting friends, fitness
4			<i>CTnr 6:</i> work, leisure shopping, fitness, meeting friends

Table 3 Time factor per choice task number

Choice task Nr.	1	2	3	4	5	6
Value of time factor	0.7	0.7	0.7	1.1	0.9	1.1

very large time budget. The time budget is therefore personalized and depends on the time the participant took to make the first series of choices, multiplied by a time factor. To set the values of the time factor, a small-scale pilot experiment was carried out. 20 people were recruited for this pilot run, who were randomly divided into three groups. For the first group of seven persons, the time factors all equal one. A brief interview was conducted afterwards, asking their opinions about the extent to which they felt time-pressured to make their decisions for each of the tasks. Based on these results, the time factor values were adjusted, and the procedure is repeated for the second and the third group. This finally led to time factors applied in the experiment, which are shown in Table 3. A time factor larger than 1 means that participants are given more time to make a decision than it took them in the first series, however, the resulting assigned time budget may still be perceived as a time pressure as participants may not be aware of the time it took them to make a decision in the first series of choices; because there was no time constraint in the first series, any constraint may add to the notion of feeling pressured.

The time budget assigned to make a choice is presented very clearly in the upper-right corner of the computer screen: a countdown clock showed how many seconds were left for choice making (“Aftellen” in Fig. 1). If the participant fails to reach a decision within the given time budget, ATS would inform him or her that because of this, a choice is randomly and automatically made by ATS instead. Respondents were informed before starting this part of the experiment, that if they would not choose within the allowed timeframe, a random choice would be made for them.

The sample

In May and June 2012, participants were recruited to participate in this experiment by IntoMart, a marketing research company. The company approached persons from its existing panel database, who own a car, work at least 2 days a week and who commute to towns or cities other than their own place of residence. Participants were offered €20 incentive and €10 of travel cost to join. The experiment was executed in a controlled computer room in a sequence of sessions, with a maximum of 40 persons each to ensure that every participant could be closely monitored by an experiment supervisor. Table 4 presents the main characteristics of the participants. In total, 113 persons participated, who almost all have a paid job. 74 % of those with paid jobs commute to work at least 4 days a week. For the rest of the background characteristics, except that half of the participants belong to the category of WO/HBO⁴ in education, the sample is fairly heterogeneous.

Participants’ experiences with ATS

Before presenting the results we briefly discuss the suitability of the simulator. More specifically, in this section, we report on the participants’ experiences with the activity-based simulator. After completing the experiment, participants rated five statements, each

⁴ WO/HBO stands for university education, MAVO/VMBO intermediate vocational education, HAVO higher general continued education, and WO pre-university education in the Dutch vocational system.

Table 4 Characteristics of the experiment participants

Characteristics	Category	Distribution (n = 113) (%)
Job	Paid job	97.3
	Volunteer	0.8
	Others	1.8
Commuting to work (days per week)	≥4 days	74.3
	4> and ≥2 days	16.8
	<2 days	8.9
Age	20–30	14.2
	30–40	11.5
	40–50	25.7
	50–60	39.8
	60–70	8.8
Gender	Male	69.9
	Female	30.1
Education	Lower education	5.3
	MAVO/VMBO	27.4
	HAVO	14.2
	VWO	2.7
	WO/HBO	50.4
Marriage	Single	31.9
	Married	48.7
	Partner/living together	19.4

on a five-point scale ranging from “completely disagree” denoted as 1 to “completely agree” denoted as 5, regarding their evaluations of the experiment, as listed in Table 5.

The table suggests that a large majority of the participants were able to remain focused during the experiment process, felt the information shown in ATS was illustrative, understood the experiment well, and enjoyed the experiment as a whole. A small proportion of the participants felt that activity programs presented to them were not sufficiently realistic to their real life situation. Overall, this feedback suggests rather positive evaluations from the participants. Note that the average ratings obtained here are quite comparable to those attained in another travel simulator (Chorus et al. 2007). In Chorus et al.’s experiment, similar evaluations on the participants’ feedback on their simulator are obtained, with the following four statements: (1) I found it difficult to remain concentrated during the experiment; (2) I found it difficult to identify with the different travel situations; (3) I found the travel simulator easy to understand; (4) I enjoyed participating in the experiment. They found that the average ratings of the four statements are 2.24, 1.94, 4.19 and 4.47 respectively on a scale from 1 to 5 (the small values of the first two are due to the negative formulations in the answers of the two statements). If the first two statements would be reformulated by replacing the word ‘difficult’ with ‘easy’, the ratings of the two might be transposed to 3.76 and 4.06 respectively. It may be argued that the choice task in the travel simulator in this study is complex, but very concrete, while the choice task applied in Chorus’ et al.’s travel simulator was less complex, but more abstract. That comparable results are found for both simulators, indicates that indeed travel simulators succeed in engaging participants in complex choice tasks.

Table 5 Self-reported feedbacks on the experiment

Variable	Frequency	Percentage
<i>It was easy to understand the travel simulator</i>		
1 very much disagree	1	0.9
2 disagree	1	0.9
3 neutral	10	8.8
4 agree	48	42.5
5 very much agree	53	46.9
Average (4.34)		
<i>It was easy to remain focused during the experiment</i>		
1 very much disagree	1	0.9
2 disagree	2	1.8
3 neutral	2	1.8
4 agree	60	53.1
5 very much agree	48	42.4
Average (4.36)		
<i>The information shown in the abstract map was illustrative</i>		
1 very much disagree	1	0.9
2 disagree	2	1.8
3 neutral	1	0.9
4 agree	67	59.3
5 very much agree	42	37.1
Average (4.30)		
<i>The daily activity programs presented in the experiment look realistic for my situation</i>		
1 very much disagree	5	4.4
2 disagree	12	10.6
3 neutral	28	24.8
4 agree	50	44.2
5 very much agree	17	15.0
Average (3.52)		
<i>It was enjoyable to participate in the experiment</i>		
1 very much disagree	1	0.9
2 disagree	1	0.9
3 neutral	1	0.9
4 agree	41	36.3
5 very much agree	69	61.0
Average (4.56)		

Model estimation

The systematic component of the utility function

In addition to the usual travel time and cost attributes, we include the number of interchanges (relevant only for public transportation and multi-modal travel options) as well as a dummy variable for travel options that only included car (to capture intrinsic preferences for the car option beyond time, cost, and interchanges). For the sake of readability of the

equations, the utility function of a choice alternative is formulated from a single representative person's perspective. Therefore, the subscript representing a particular person is suppressed from the equations. As such, the systematic component of the utility function can be formulated as the following linear-in-parameter formulation: $V_i = \beta_{TT} \cdot TT_i + \beta_{TC} \cdot TC_i + \beta_{TI} \cdot TI_i + \beta_{car} \cdot Car_i$, where TT_i denotes total travel time of alternative i ; TC_i denotes total travel cost of alternative i ; and TI_i denotes the number of travel interchanges in alternative i . Dummy Car_i equals 1 when alternative i only employs car as travel mode, and 0 when it does not.

Specification of the scale

The specification of the scale parameter for inclusion in the heteroscedastic model follows from the previous discussion and is written as follows: $\mu_s = e^{(\lambda_{DT} \cdot DT'_s + (\delta_T \cdot DS_s + \theta_T \cdot DS_s^2) + \omega \cdot DT'_s \cdot DS_s)}$, where DT'_s is the indicator of the complexity of the task of complexity task number s and based on the measured decision time under the condition of no time constraint; λ_{DT} denotes the parameter for task complexity, which is expected to have a negative sign, as higher levels of task complexity are expected to decrease the scale; DS_s denotes the engagement/time pressure index; δ_T denotes the parameter for the linear component of the engagement/time pressure index; θ_T denotes the parameter for the quadratic component of the engagement/time pressure index.⁵ In line with the hypothesized non-monotonic relation between the engagement/time pressure index and the scale of utility (see “[Model development](#)” section), the linear parameter δ_T is expected to be positive as very low values of the index are associated with a small scale value and this may increase as the index value increases; the quadratic parameter is expected to be negative, because the index is hypothesized to have an optimum value, somewhere between a very low and a very high value of the index. Finally, ω gives the strength of the interaction effect of task complexity and time pressure, which is expected to have a negative sign as the simultaneous combination of both conditions (i.e., high levels of task complexity and of time pressure) are expected to result in an additional negative effect on the scale.

Estimating the model

The developed HL model is estimated using Python biogeme (Bierlaire 2008). In addition, three other models are estimated: (i) a basic MNL model, (ii) a Mixed Logit (ML) model that takes taste heterogeneity into account, but not the abovementioned scale effects; and (iii) a heteroscedastic mixed logit (HML) model that takes both taste heterogeneity and the abovementioned scale effects into account. As the parameters for time, cost and number of interchanges are expected to be negative, triangular distributions are assumed with the additional constraint that the sum of the mean and the spread takes a negative sign, to ensure that the whole distribution lies within the negative-sign range (Hensher and Greene 2003). A normal distribution is assumed for the car dummy variable. Halton draws were used to simulate the integrals for ML and HML models, and the number of the draws was

⁵ Models with random scale components were also tried out, but these unfortunately did not converge. Note that we also estimated a parameter in the scale function to capture potential learning or fatigue effects; it might be expected that respondents—due to such effects—became quicker or slower in terms of responding to choice tasks, during the course of the experiment. This learning/fatigue parameter which was coupled to the number of the choice task turned out to be statistically insignificant at any reasonable level of significance.

gradually increased to 3000 where stability of the estimated parameters of both the ML and the HML models was achieved. The models are estimated on 1356 observed choices (12 each) made by 113 participants.

Results

The impacts of task complexity and time pressure

As shown in Table 6, each less constrained model performs better in terms of both adjusted rho square and the associated likelihood ratio test, than its more constrained predecessor (presented on its left). This suggests that model fit is gradually enhanced by increased model sophistication in both allowing for random parameters and for heteroscedasticity. It is clear that taking taste heterogeneity into account improves fit much more than taking heteroscedasticity into account. Nevertheless, comparing the HL with the MNL model, and the HML with the ML model, makes clear that adding the impacts of task complexity and engagement/time pressure modestly increases model performance, irrespective of whether unobserved taste heterogeneity is accounted for or not.

Note that the parameter of the mean of travel time (β_{TT}) is statistically insignificant in the HL model that does not take random taste heterogeneity into account, and that its estimate is drastically different from those produced by the other models. Hence, embedding the impacts of task complexity and time pressure into the conventional MNL model without taking random taste heterogeneity into account has led to a bias in the

Table 6 Estimation results

	MNL		HL		ML		HML	
Adjusted rho-square	0.150		0.171		0.271		0.282	
Initial log-likelihood	−1528.12		−1528.12		−1528.12		−1528.12	
Final log-likelihood	−1295.780		−1260.534		−1108.436		−1088.864	
Likelihood ratio test	464.68		535.172		839.368		878.512	
Nr. of draws					3000		3000	
Nr. of parameters	3		6		5		8	
Parameters	Value	t-stat.	Value	t-stat.	Value	t-stat.	Value	t-stat.
Mean (β_{TC})	−0.0415	−2.74	−0.0738	−2.66	−0.115	−4.75	−0.159	−3.88
Spread (β_{TC}) ^a					(0.115)		(0.159)	
Mean (β_{TT})	−0.0125	−2.16	−0.00669	−0.63	−0.0476	−5.10	−0.0636	−3.67
Spread (β_{TT})					(0.476)		(0.0636)	
Mean (β_T)	−0.373	−14.41	−0.690	−3.86	−0.590	−9.72	−0.938	−5.20
Spread (β_T)					0.406	4.21	0.590	2.63
Mean(β_{Car}) ^b	0		0		0		0	
SD (β_{Car})					1.70	6.07	2.92	4.12
λ_{DT}			−0.0101	−6.05			−0.00745	−3.16
δ_T			2.03	3.95			2.56	4.44
θ_T			−3.36	−4.58			−3.96	−5.01

estimate of the taste for travel time. Nevertheless, the estimates of the HL model as far as the *scale* of utility is concerned, are quite comparable to the ones produced by the HML model.

As shown in Table 6, the estimate of task complexity (λ_{DT}) is significant in both the HL and the HML model and has the expected negative sign. This means that, as hypothesized, the more complex a choice task is, the smaller the scale, or in other words, the larger the error variance. The interaction effect between task complexity and time pressure (ω) was found to be not statistically significant in both the HL and the HML models and is therefore dropped from both models and not reported here. This means that, in contrast with our expectations, time pressure and task complexity in our experiment do not reinforce each other in their impact on scale. Possibly this is caused by the fact that more complex tasks in our experiment are given relatively more time in the time pressure condition, as indicated by the time factors presented in Table 3. From hindsight, this correlation between experimental conditions should have been avoided.

With respect to the parameters estimated for time pressure, both the linear (δ_T) and quadratic (θ_T) components are statistically significant and have signs that support the hypothesized non-monotonicity of the relation with scale. More specifically, δ_T and θ_T equal to 2.56 and -3.96 respectively, which results in the relation between engagement/time pressure and scale as plotted in Fig. 2. Note that in this Figure, the dotted line corresponds to the HL model, the continued line corresponds to the HML model. The value of the vertical axis represents the value of the scale divided by $\exp(D_s)$. The Figure clearly shows that if the DS_s index values increases, the scale first increases until an optimal level is reached (at $DS_s = 0.321$), after which it decreases. Thus, as hypothesized, for very low values of DS_s (i.e., in situations where the participant took very little time to make a decision) the scale of the utility is smaller than for intermediate values. We believe that this signals the absence of engagement with the choice task.

Hence, this index indeed has a dual interpretation and is consequently labelled as ‘engagement/time pressure index’ as opposed to a time pressure index. Moreover, also a hypothesized, the results indicate that the choices tend to become more random when the time is running out, hence, when time pressure is high.

Table 6 indicates that the estimates for the total travel cost, the total travel time and the total number of interchanges in the MNL, the ML, and the HML model are all significant

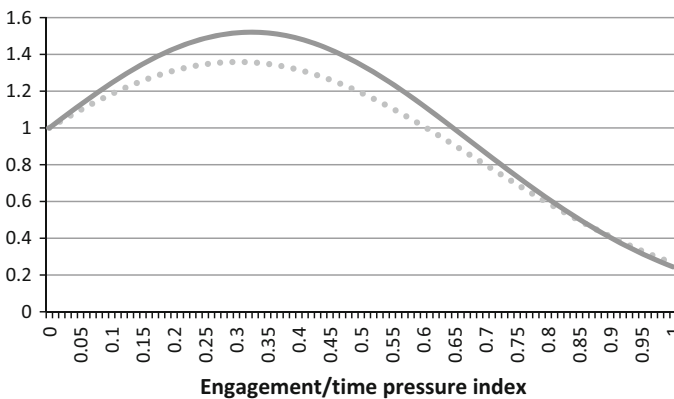


Fig. 2 Scale (y-axis) of the systematic component of the utility function as a function of DS_s (HL model: dotted line, HML model: solid line)

Table 7 Choice task used for illustration

Alternative	Travel cost (euro)	Travel time (min)	Nr. of travel interchanges	Car as dominant travel mode
1	10	118	7	No
2	15	117	9	No
3	12	110	8	No
4	19	121	5	Yes

and have the expected negative sign. The mean of the car dummy is not statistically significant and therefore fixed at zero, but the Car dummy has a relatively large standard deviation, which suggests that both strongly positive and negative basic preferences for car travel exist.

Beyond the inspection of estimation results, an important question relates to the potential differences in choice probability predictions implied by the estimated heteroscedastic models (which capture task complexity and time pressure) and their homoscedastic counterparts (which do not). As will be seen in the following illustrations, this difference—and hence the bias resulting from not accommodating for task complexity and time pressure—can be substantial. As a first illustration, we compute and compare elasticities for travel time and cost, for the Mixed Logit and the Heteroscedastic Mixed Logit models. Note that we ignored the HL model and its comparison with the MNL model, as we distrust the former given its insignificant travel time parameter. Travel time elasticities for the ML and HML models equalled -0.737 and -0.646 respectively. Travel cost elasticities for the ML and HML models equalled -3.118 and -2.758 respectively. The difference, in terms of elasticities, between the two models implies that on the average, travelers in our dataset were less sensitive to changes in travel time and cost than one would conclude based on a model which ignores the impact of task-complexity and time-pressure.

This implication is in line with conclusions that we draw from our second illustration: we selected a choice task that was considered as relatively complex by participants, in the sense that the average decision time (in the condition where no time constraints were present) was higher than those of other tasks. Recall that the task complexity indicator is individual specific, hence even though the average decision time of the selected choice task is relatively high (87 s), there is still much heterogeneity in perceived complexity among the participants. As shown in Table 7, the selected task involved a choice between four alternatives, each containing a relatively large number of travel interchanges.

For this choice task, we predicted choice probabilities for each of the four alternatives using the heteroscedastic Mixed Logit⁶ model and its homoscedastic counterpart. We distinguished between four (two \times two) conditions: first, low task complexity for which we took the average decision time of 87 s; and high task complexity for which we took the highest recorded decision time for this task, being 227 s. Second, time pressure, which was varied in a low value, for which we took the value(s) of the engagement/time pressure variable that corresponds to the highest scale—see Fig. 2), and a relatively high value for

⁶ Given the partly unreliable results obtained for the heteroscedastic logit model (see discussion further above) we choose to focus on the mixed (Heteroscedastic) logit models. Each choice probability was simulated using 1,000,000 multidimensional Halton draws.

Table 8 Illustration of the modelled impact of task complexity and time pressure on choice probabilities

	Mixed logit (homoscedastic)	Heteroscedastic mixed logit			
		Low task complexity		High task complexity	
		Low time pressure	High time pressure	Low time pressure	High time pressure
P (alt1)	0.23	0.22	0.25	0.25	0.25
P (alt2)	0.06	0.05	0.19	0.14	0.23
P (alt3)	0.30	0.27	0.26	0.27	0.26
P (alt4)	0.41	0.47	0.30	0.35	0.27
P (alt4)–P(alt2)	0.35	0.42	0.11	0.21	0.04

which we took the value of 1 for the engagement/time pressure variable). Table 8 reports the simulation results.

The Table reports choice probabilities for the four alternatives as implied by the homoscedastic mixed logit model, as well as by the heteroscedastic mixed logit model (under the four different conditions); in addition, the choice probability difference between the most and least popular alternatives is reported. A first result is that for the condition of both low task complexity and low time pressure levels, the heteroscedastic mixed logit model predicts more profound differences in choice probabilities than its homoscedastic counterpart. When time pressure increases to its maximum level (i.e., right before the time runs out), and keeping task complexity fixed, the heteroscedastic mixed logit model predicts much less profound differences in choice probabilities than its homoscedastic counterpart. For respondents that consider the task to be highly complex (the two columns on the right hand side), the heteroscedastic mixed logit model predicts less profound differences in choice probabilities than its homoscedastic counterpart, and especially so when much time pressure is present. In this latter situation, i.e., involving high levels of both task complexity and time pressure, the difference between the homo- and heteroscedastic models is particularly striking: while the homoscedastic model predicts that the most popular alternative is more than seven times as popular as the least popular alternative, the heteroscedastic model predicts that the two are almost equally popular.

These results are of course fully in line with expectations (and with theory) in the sense that higher levels of task complexity and time pressure were expected to lead to more random choice behaviour. This dependency of choice behaviour on task complexity and time pressure conditions is captured by the heteroscedastic model, but ignored by its homoscedastic counterpart. To the extent that the heteroscedastic model fits the data statistically better than its homoscedastic counterpart (as is the case on our data), these results suggest that failing to incorporate task complexity and time pressure in activity-travel models may lead to non-trivial biases in forecasting.

Conclusion and discussion

This paper presents a discrete choice model of activity-travel behaviour that incorporates the effects of task complexity and time pressure on the scale of the utility. The model is subsequently estimated on data from a novel activity-travel simulator experiment that was specifically designed for the purpose of testing our model. Our main results are as follows:

firstly, high levels of time pressure and task complexity lead to a smaller scale of utility and hence to more random choice behaviour. Secondly, very short decision times also lead to more random behaviour, although in that case there is no evidence of time pressure. We interpret this phenomenon in terms of a lack of engagement with the choice task among those who make a choice within a matter of one or 2 s after being presented with the choice task. Thirdly, contrary to expectations, we find no evidence for an interaction effect between task complexity and time pressure. In other words, the impact of task complexity on choice behaviour in the context of our data does not become more pronounced when there is a high level of time pressure (and neither vice versa). Fourthly, on our data, heteroscedastic models that incorporate the effects of time pressure and task complexity achieve higher levels of model fit than corresponding homoscedastic models that do not accommodate these effects. Fifthly, and more importantly than these differences in model fit, we find that choice probability predictions differ substantially between estimated homo- and heteroscedastic models: the former predict much more (less) pronounced differences in choice probabilities between alternatives than the latter, when there are relatively high (low) levels of task complexity and time pressure. In other words, under these conditions, heteroscedastic models predict a much more (less) even distribution of choice probabilities across choice alternatives, than their homoscedastic counterparts. Our findings are intuitive and suggest that incorporating task complexity and time pressure pays off, in terms of achieving a better model fit and—more importantly—in terms of presumably achieving a more realistic account of activity-travel behaviour and attaining more accurate choice probability forecasts.

This latter dimension (i.e., a potentially substantial improvement in forecasts) implies that capturing the impacts of time pressure and task complexity in discrete choice models of activity-travel behaviour is also important from a practical or policy-viewpoint; this holds even more in light of the fact that in real life, many activity-travel choices are made under conditions of considerable task complexity and time pressure. Our results suggest that by ignoring in choice models the effects of task complexity and time pressure on activity-travel behaviour, policy makers are likely to overestimate traveller sensitivity to changes in the attributes of travel options, when some choices are made under conditions of high levels of task complexity and time pressure, and others are not. Our heteroscedastic models suggest that under high levels of task complexity and time pressure, choice behaviour is governed to a large extent by randomness, implying a limited sensitivity to changes in the availability and characteristics of travel options.

Of course, before our results can be generalized, it is important that they are verified based on other datasets. Although the impact of task complexity has by now been well established, this is not the case for the impact of time pressure (nor for the presence or absence of interaction effects between the two). Whereas we used Stated Preference data collected in a simulator experiment, it would be particularly interesting to see if our results also hold in the context of revealed preference data. Some readers might even argue that what we measured in our experiments is perhaps even more about the impact of task complexity and time pressure in choice experiments, than about their impact on (real life) travel behavior. Although we went through a lot of effort to design a simulator which gives a realistic account of a travel behavior context, it goes without saying that we only partly succeed therein. As a consequence, our manipulation of task complexity and time pressure can only be considered proxies of the variation in task complexity and time pressure that travelers may experience in real life. This in turn implies that our results should be interpreted with the utmost care. In our view, they are only but a first step towards a proper understanding of real life behaviors under varying levels of task complexity and time

pressure. Before any stronger conclusions and generalizations can be drawn, revealed preference data clearly are a necessity.

Needless to say, it is a challenge to collect RP data in a way that they allow the researcher to accurately measure task complexity and time pressure (this was in fact the main reason why we used a carefully controlled SP experiment). However, new technologies (including mobile phone applications) might make it possible to collect reliable RP-data in the not so far away future. The models and analyses presented in this paper provide a stepping stone for these and other possible follow up research efforts.

Acknowledgments We gratefully acknowledge funding by the Netherlands Organisation for Scientific Research (NWO): the research project that has led to this paper, titled ‘Use of supernetworks’, was part of an NWO-funded research program called ‘Sustainable Accessibility of the Randstad’. We thank three anonymous referees for making a range of interesting suggestions based on a previous version of this paper.

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

References

- Arentze, T., Borgers, A., Timmermans, H., DelMistro, R.: Transport stated choice responses: effects of task complexity, presentation format and literacy. *Transp. Res. Part E* **39**(3), 229–244 (2003)
- Arentze, T.A., Timmermans, H.J.: A learning-based transportation oriented simulation system. *Transp. Res. Part B* **38**(7), 613–633 (2004)
- Bhat, C.R.: A heteroscedastic extreme value model of intercity travel mode choice. *Transp. Res. Part B* **29**(6), 471–483 (1995)
- Bierlaire, M.: Estimation of Discrete Choice Models with BIOGEME 1.8. Transport and Mobility Laboratory, EPFL, Lausanne (2008)
- Bonsall, P., Palmer, I.: Modelling drivers’ car parking behaviour using data from a travel choice simulator. *Transp. Res. Part C* **12**(5), 321–347 (2004)
- Caussade, S., Ortúzar, J.D.D., Rizzi, L.I., Hensher, D.A.: Assessing the influence of design dimensions on stated choice experiment estimates. *Transp. Res. Part B* **39**(7), 621–640 (2005)
- Chen, P.S., Mahmassani, H.S.: Dynamic interactive simulator for studying commuter behavior under real-time traffic information supply strategies. *Transp. Res. Rec.* **1413**, 12–21 (1993)
- Chorus, C.G., Molin, E.J.E., Arentze, T.A., Hoogendoorn, S.P., Timmermans, H.J.P., Van Wee, B.: Validation of a multimodal travel simulator with travel information provision. *Transp. Res. Part C* **15**(3), 191–207 (2007)
- Daganzo, C.: Multinomial Probit: The Theory and its Application to Demand Forecasting. Academic Press, New York (1979)
- Dellaert, B.G.C., Donkers, B., Soest, A.V.: Complexity effects in choice experiment-based models. *J. Mark. Res.* **49**(3), 424–434 (2011)
- DeShazo, J.R., Fermo, G.: Designing choice sets for stated preference methods: the effects of complexity on choice consistency. *J. Environ. Econ. Manag.* **44**(1), 123–143 (2002)
- Dhar, R., Nowlis, S.M.: The effect of time pressure on consumer choice deferral. *J. Consum. Res.* **25**(4), 369–384 (1999)
- Diederich, A.: Decision making under conflict: decision time as a measure of conflict strength. *Psychon. Bull. Rev.* **10**(1), 167–176 (2003)
- Hensher, D.A., Greene, W.H.: The mixed logit model: the state of practice. *Transportation* **30**(2), 133–176 (2003)
- Kitamura, R.: An evaluation of activity-based travel analysis. *Transportation* **15**(1–2), 9–34 (1988)
- Mahmassani, H.S., Jou, R.C.: Transferring insights into commuter behavior dynamics from laboratory experiments to field surveys. *Transp. Res. Part A* **34**(4), 243–260 (2000)
- McFadden, D.: Conditional logit analysis of qualitative choice behaviour. In: Zarembka, P. (ed.) *Frontiers in Econometrics*, pp. 105–142. Academic Press, New York (1973)
- Nowlis, S.M.: The effect of time pressure on the choice between brands that differ in quality, price, and product features. *Mark. Lett.* **6**(4), 287–295 (1995)

- Ordóñez, L., Benson Iii, L.: Decisions under time pressure: how time constraint affects risky decision making. *Organ. Behav. Human Decis. Process.* **71**(2), 121–140 (1997)
- Prendinger, H., Nakasone, A., Miska, M., Kuwarhara, M.: Open energy sim: conducting behavioral studies in virtual worlds for sustainable transportation. *Integrated and Sustainable Transportation System (FISTS)*, 2011 IEEE Forum on, IEEE (2011)
- Recker, W.W., McNally, M.G., Root, G.S.: A model of complex travel behavior: part II—an operational model. *Transp. Res. Part A* **20**(4), 319–330 (1986)
- Simon, H.A.: A behavioral model of rational choice. *Q. J. Econ* **69**, 99–118 (1955)
- Sun, Z., Arentze, T., Timmermans, H.: A heterogeneous latent class model of activity rescheduling, route choice and information acquisition decisions under multiple uncertain events. *Transp. Res. Part C* **25**, 46–60 (2012)
- Swait, J., Adamowicz, W.: The influence of task complexity on consumer choice: a latent class model of decision strategy switching. *J. Consum. Res.* **28**(1), 135–148 (2001)
- Tversky, A.: Elimination by aspects: a theory of choice. *Psychol. Rev.* **79**(4), 281 (1972)

Chao Chen received his PhD from Delft University of Technology. In 2014 he joined SAIC Motor, an automotive manufacturing company in China, as a marketing and product manager. His current work is primarily concerned with consumer behaviour research in China's automotive industry.

Caspar Chorus is Professor of choice behavior modeling, and head of the Transport and Logistics group, at TU Delft. His main research aim is to increase the behavioral realism of (travel) choice models, and by implication to provide a stronger foundation for (transport) policy development.

Eric Molin is an Associate Professor of Transport Policy at TU Delft. His main field of interest is behavioral modeling in transport, with a special interest in advanced stated choice experiments, multi-modal transport systems, and new transport technology.

Bert van Wee is Professor of Transport Policy at Delft University of Technology, The Netherlands and Scientific Director of TRAIL Research School.