A peek at the future: capturing anticipation effects in discrete choice models

Charisma Choudhury
Institute for Transport Studies & Choice Modelling Centre, University of Leeds
34-40 University Road, University of Leeds
Leeds LS2 9JT (UK)
Phone: (+44) (0)113 343 2659
C.F.Choudhury@leeds.ac.uk

Stephane Hess
Institute for Transport Studies & Choice Modelling Centre, University of Leeds
Lifton Villas, 1-3 Lifton Place
University of Leeds
Leeds LS2 9JT (UK)
Phone: (+44)(0)113 343 6611
s.hess@its.leeds.ac.uk

Abu Toasin Oakil
Independent researcher
The Netherlands
Phone: (+31) (0) 641 726 996
abutoasin@gmail.com

Word count:
5,210 words text + 9 tables/figures x 250 words (each) = 7,460 words

Submission Date:
1 August 2017

ABSTRACT

Many of the choices we make involve 'planning ahead' based on future values of the attributes of the alternatives (increase/decrease in price) and/or anticipation (or apprehension) associated with personal circumstances (e.g. getting married/divorced, having a child, changing a job, etc.). While long term panel datasets allow us to know about future life events and their timeline, there are several challenges in incorporating them in the choice framework: the heterogeneity of the planning horizon, data issues like discrepancies between anticipation and actual future observations and interdependencies among the choices to name a few. In this paper, we investigate modelling methodologies to realistically capture such anticipation effects and demonstrate these using two case studies. The first case study focuses on lane choice in an urban arterial (a short term choice) where we investigate the impact of anticipated downstream traffic conditions on lane choice. The second case study focuses on car-ownership decisions (a medium term choice) where we investigate the potential impacts of anticipated future life events (e.g. change in job, birth of a child, marriage/divorce, etc.), on present choices. Estimation results confirm that for both cases, anticipation effects are significant indeed and there is substantial heterogeneity in planning horizons.

Keywords: Anticipation, Car-ownership, Life-event, Lane choice

1. BACKGROUND

Many of the choices we make involve ‘planning ahead’ based on anticipation and apprehension. For instance, investment decisions may be based on anticipated trends of the market; residential location/re-location choices may be based on anticipated life-events; route choices may be based on anticipated travel times, etc. From a modelling perspective, this has been partly captured in dynamic discrete choice models where a decision maker chooses the alternative in the current period that maximizes his/her expected utility over the current and future periods (1, 2). However, due to data limitations, such studies are often operationalized using stated preference (SP) datasets (e.g. 3, 4), where it might be difficult to really experience anticipation. The studies that are based on field data are based on perfect information of the market and/or assume/estimate a fixed value for the discount factor while calculating the aggregate asset return (e.g. 5). However, there has been evidence that the rate at which decision makers discount the future can vary tremendously across individuals and can differ substantially from the economy-wide asset returns (6,6).

Furthermore, these models have primarily focused on anticipation or apprehension involved with the attributes of the alternatives and less on anticipation and apprehension involving personal circumstances (e.g. getting married/divorced, having a child, changing a job, etc.). Availability of disaggregate life-course panel datasets, where the evolution of personal circumstances is available alongside the evolution of choices and attributes of alternatives, provides us with the opportunity to incorporate these effects in the modelling framework. For example, future demographic events such as household formation and childbirth (happening 1 or 2 years later) has been found to have a significant impact on decrease change in the number of cars in the current year (8). Recent research has also focused on joint modelling of life events and travel choices (e.g.9, 10) where the correlation between the error terms has been accounted for within the Multivariate Probit framework. Both models account for state-dependence, but ignore the planning ahead aspect of decisions.

While long panel datasets allow us to know the future values of attributes as well as life events and their timeline, there are several challenges in incorporating them in the choice framework. *Firstly,* the planning horizon of the decision maker (which can vary significantly across the population) is unknown to the modeller. For example, someone may purchase a property with the long term increase in future price/ distant life events in mind, while another decision maker may have a shorter planning horizon and ignore these considerations. This makes it difficult to ascertain from when the anticipation effect starts to impact the decisions. *Secondly,* although panel data provides us with the opportunity to observe the future attribute values and life events, the attributes and events observed in later time periods may not be the same as the anticipated values. For instance, someone may purchase a property based on anticipation that the price will increase by a certain extent in future, but there may be a change in the market leading to changes in the actual value. It is this actual value that modellers observe in the future data, and this may be different from the anticipated value. Similarly, life events can be unexpected or there may be discrepancies with the planned time and the actual occurrence. Like other plans, those regarding life events may also change dynamically, leading to further modelling complexities. These issues are not investigated in detail in the existing literature and form the motivation for this research.

In this paper, we investigate modelling methodologies to realistically capture these anticipation effects and demonstrate them using two case studies. The first case study focuses on lane choice in an urban arterial (a short term choice) where we investigate the impact of anticipated downstream traffic conditions on choice of target lane. The second case study focuses on car-ownership decisions (a medium term choice) where we investigate the potential impacts of anticipated future life events (e.g. change in job, birth of a child, marriage/divorce, etc.), on present choices. The particular focus is on 1) Capturing the heterogeneity in planning and anticipation 2) Addressing the discrepancy between the anticipated and the observed attributes/events.

The rest of the paper is organized as follows. The generic model formulation is presented first followed by the two case studies. For each case study, the datasets, model formulations and estimation results are presented. The findings are summarized in the concluding section.

2. MODEL FRAMEWORK

2.1 Basic structure

The Random Utility Modelling (RUM) principles are used as the basis of the modelling framework. The utility of an alternative is influenced by the attributes of the alternatives (X_{in}^{time}) and the characteristics of the decision maker (Z_{in}^{time}): where, 'time' can refer to variables corresponding to present (t) and past (1:t-1). It is also influenced by anticipated (or apprehended) future values of the attributes (X_{in}^{t*}) and anticipated life events/demographics(Z_n^{t*}) (cf. Figure 1).

$$U_{in}^t = U(X_{in}^t, X_{in}^{1:t-1}, X_{in}^{t*}, Z_n^t, Z_n^{1:t-1}, Z_n^{t*}) \quad (1)$$

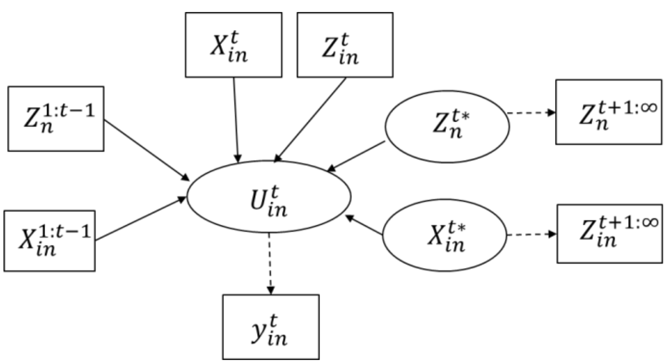

FIGURE 1 Model Framework.

While in previous research, we had focused on the effect of the past (state-dependence) (13, 14), in this research, we investigate the anticipation component in detail and ignore state-dependence. This simplifying assumption leads to the following equation:

$$U_{in}^t = U(X_{in}^t, X_{in}^{t*}, Z_n^t, Z_n^{t*}) \quad (2)$$

Since anticipation (or apprehension), X_{in}^{t*} , involves only the expected value of an attribute and/or an expected life event leading to changes in characteristics of the decision maker Z_n^{t*} , as opposed to the actual values of the corresponding variables in future, these are latent/unobserved. However, the observed future values serve as indicators of these latent variables, and these can be defined as follows:

$$X_{in}^{t+1} = I(X_{in}^t, X_{in}^{t*}, \lambda, \vartheta_n), \quad (3)$$

where, $\lambda$is a vector of unknown parameters, $\vartheta_n \sim D(0, \Sigma_\vartheta)$ is a vector of error terms with a distribution function D which has a zero mean and a variance-covariance matrix denoted as Σ_ϑ and I(.) is a function linear in parameters with additive error terms (though it can be non-linear as well). It is assumed that ϑ_n is not correlated with the error term of the utility function.

The joint probability of the choice and the indicators can be expressed as follows:

$$P(y_n, I_n | X_n, \beta, \alpha, \lambda, \Sigma_\varepsilon, \Sigma_\omega, \Sigma_\vartheta) = \int_{X^*} P(y_n | X_n, X_n^*, \beta, \Sigma_\varepsilon) g\,(I_n | X_n, X_n^*, \lambda, \Sigma_\vartheta) f(X_n^* | X_n, \alpha, \Sigma_\omega) \quad (4)$$

Where $g(I_n | X_n, X_n^*, \lambda, \Sigma_\vartheta)$denotes the joint density function of the indicators conditional on the latent variables and $f(X_n^* | X_n, \alpha, \Sigma_\omega)$ is the joint probability density function of the latent variable X_n^*. This is a standard hybrid choice model (18) which we exploit here not for the typical treatment of attitudes and perceptions but for anticipation of future events.

2.2. Accounting for heterogeneity

To account for the heterogeneity in planning horizon, we assume a latent class framework where the time period of anticipation can vary among the classes. The number of classes can be two in the simplest case (people who plan ahead and people who do not), though the number of classes can be more and can differ depending on the choice context (e.g. long, medium and short term choices; risky vs. non risky choices, etc.). Under the latent class assumption, the joint probability can be extended as follows:

$$P(y_n, I_n | X_n, \beta, \alpha, \lambda, \Sigma_\varepsilon, \Sigma_\omega, \Sigma_\vartheta) = \sum_{C \in G_n} P(y_n, I_n | X_n, \beta, \alpha, \lambda, \Sigma_\varepsilon, \Sigma_\omega, \Sigma_\vartheta, C) P(C), \quad (5)$$

where G_n is the total number of groups, $P(y_n, I_n | X_n, \beta, \alpha, \lambda, \Sigma_\varepsilon, \Sigma_\omega, \Sigma_\vartheta, C)$ is the equivalent of Equation (4) in group C, and $P(C)$ is the probability of belonging to group C. The difference across classes arises in the use of a different time horizon.

3. CASESTUDY 1: ANTICIPATION IN LANE CHOICE MODELS

The short-term case study deals with a lane-choice scenario at an intersection where the 'future path-plan' and 'anticipated delay in queue dissipation' can affect the choice of lanes alongside neighborhood variables (e.g. speed of the vehicle in the front, density in different lanes, etc.).

3.1. Data

The data used in the study is the trajectory data collected from Lankershim Boulevard in Los Angeles, California (19). Figure 2 shows a schematic of the arterial segment. Lane numbering is assigned starting from the left most lane (e.g. in a four lane section, the rightmost lane index is 4). The study site is approximately 1600 feet (488 m) in length. It consists of four signalized intersections and three to four through lanes in each direction in each section.

The last observation of each vehicle in the side street (before entering the intersections) is regarded as the decision point and the first observation of each vehicles after entering the main arterial section indicated the observed choice.

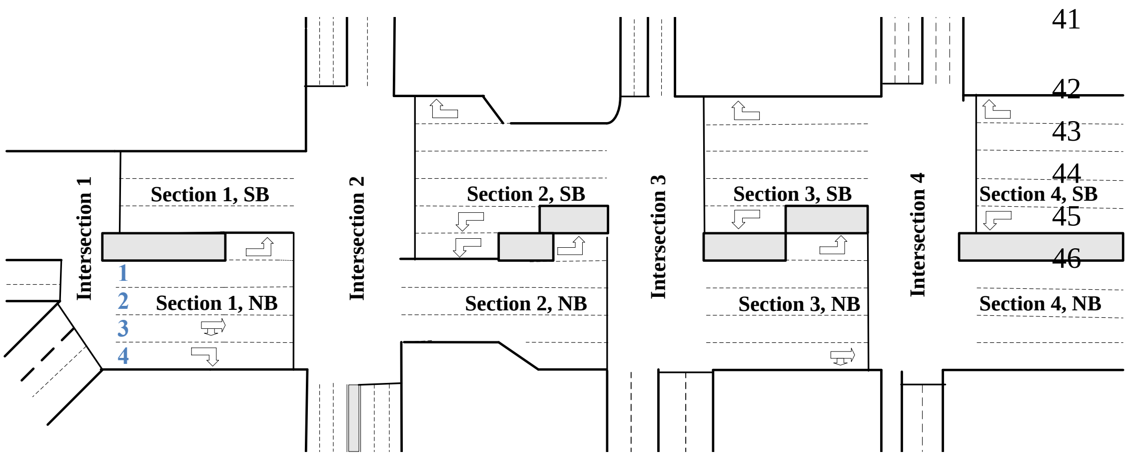

FIGURE 2 Data collection site.

The dataset used for estimating the intersection lane choice model includes 703 observations (1 observation per vehicle). 629 of them are northbound and 74 of them are southbound. The vehicles are mostly passenger cars with only a small percentage (3.5%) of trucks and buses present. The majority of entering vehicles are observed for more than one section (80.2%) with more than half (55.9%) vehicles observed for more than two sections. However, most of the vehicles (90.2%) observed more than two sections do not make any turns within the observed data collection area (i.e. go straight). A detailed analysis of the data is available in the NGSIM Lankershim Data Analysis Report (19).

3.2.Model Structure

The model structure, which is based on the latent plan framework (20, 21) is presented in Figure 3. The driver is assumed to have a two level decision: selection of target lane and selection of the immediate lane conditional on the target lane. At the first level, the driver chooses the most desirable lane as the target lane. The target lane choice set consists of all the available lanes the driver is eligible to move to. The target lane utilities can be affected by a wide range of factors. These include factors related to path-plan considerations, such as the distance to a point where the driver needs to be in a specific lane and the number of lane changes required from the target lane to the correct lane. However, the effects of path-plan in the target lane choice also depend on the planning capability of the driver and his/her familiarity with the network.

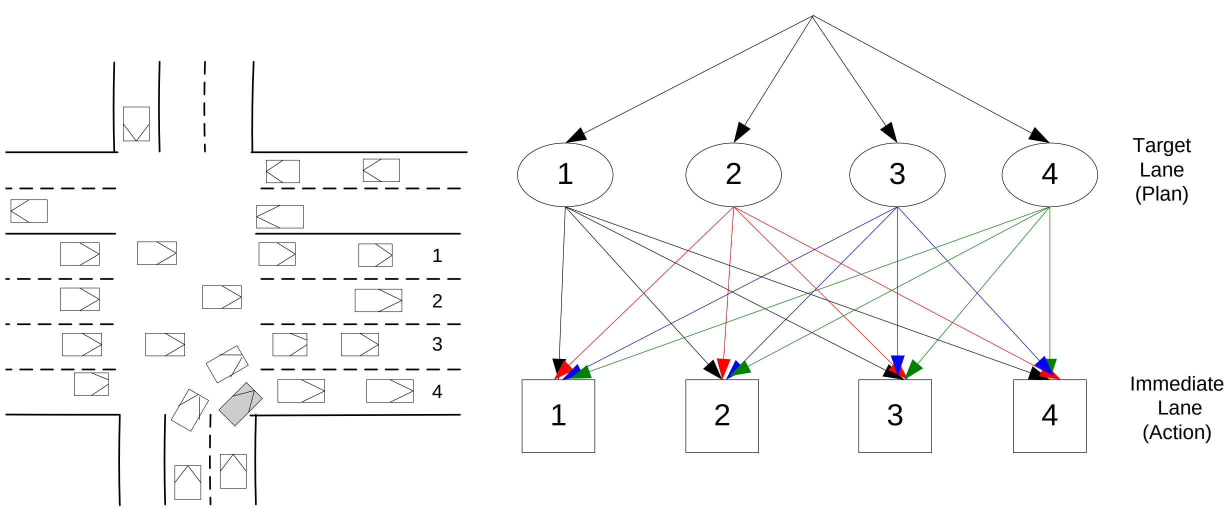

FIGURE 3 Structure of the intersection lane choice model.

Drivers are therefore assumed to belong to either of two classes:

- Myopic drivers: drivers who consider the path-plan and influencing variables only in their immediate subsequent section while making the lane selections

- Drivers who plan-ahead: these drivers consider path-plan and anticipated delay beyond their immediate subsequent section while making lane selections.

The utility of target lane can be expressed as follows:

$$U_{ln}^{C} = V_{ln}^{C} + \varepsilon_{ln}^{C}, l \in L_n, C \in 1,2 \quad (6)$$

Where, V_{ln}^{C} denotes the systematic utility for target lane *l* of driver class *C*.

The systematic utilities for the two groups of drivers are as follows:

- Class 1 (myopic): $V_{ln}^{1} = \beta_{ln} X_{ln} + \beta_{ln}^{1} q_{ln}^{1}$, utilities do not include anticipated delay and path plan beyond immediate following section
- Class 2 (drivers who plan-ahead): $V_{ln}^{2} = \beta_{ln} X_{ln} + \beta_{ln}^{2} q_{ln}^{1,2..k}$, utilities include anticipated delay and path plan for all downstream sections

The immediate lane choice is affected by the maneuverability considerations and is conditional on the target lane selection.

The probability that driver *n* selects lane *j* is the joint probability of selecting lane *j* given target lane *l* and the probability of choosing target lane *l* and can be expressed as follows:

$$P(j) = \sum_{C\in 1,2} \int_{\vartheta} \sum_{l\in L_n} P(j|l,\vartheta_n) P(l|\vartheta_n)\, P(C) f(\vartheta) d\vartheta \quad (7)$$

Where, ϑ_n is the individual specific error term (assumed to represent aggressiveness and capture the correlations among different actions and observations of the same driver) and P(C) is the class membership probability. The probability that the driver belongs to Class 1 (π_1) or Class 2 $(1-\pi_1)$ is estimated from the data along with other parameters.

Assuming that the observations from different drivers are independent, the log-likelihood function for all observed drivers is given by:

$$\mathcal{L} = \sum_{n} \ln(P(j)) \quad (8)$$

The parameters of the target lane choice components and immediate lane choice components are jointly estimated by maximizing this function. Different functional forms of variables, as well as interactions between multiple variables, have been tested during estimation and the functions resulting in best goodness-of-fit are selected.

3.3. Results

The estimation results are presented in Table 1 and discussed in the following subsections.

The Target Lane Choice Model

The target lane choice model describes drivers' choice of lane they would prefer to be in. The target lane choice of the driver is affected by the path-plan, the lane attributes and driver characteristics. Path-plan variables include number of lanes a driver has to cross (if any) in order to take a turn or exit while following the path. Lane attributes include average speed and anticipated delay in each lane. The driver characteristics are however unobserved in the data and represented by the individual specific error terms and the class membership probabilities. The anticipation effects are thus captured by the anticipated delay and path-plan variables and found to vary depending on the estimated class membership.

Estimation results indicate that the probability that a driver belongs to Class 2 (driver who plan ahead) is 18.3 %. The taste heterogeneity between the two groups of drivers has been systematically tested and found to be statistically significant only for the path-plan variables.

The magnitudes of the lane-specific constants indicate that all else being equal, the drivers prefer lanes on the right (the rightmost lane being the most preferred lane). The anticipated delay reduces the utility of the target lane although the effect is not found to be significant. Furthermore, although the value of the anticipated delay variable was different for the two driver classes, the sensitivity to the delay was not found to be significantly different between the two groups.

TABLE 1 Estimation results of the lane selection model

Final Log-likelihood	-2115.8
Number of Parameters	20
Number of Observations	703

Variable		Parameter value	t-Stat
Target Lane			
Lane 2 constant		-0.837	-3.64
Lane 3 constant		1.30	7.62
Lane 4 constant		3.25	8.16
Anticipated delay (second)		-0.477	-0.56
Lanes away from turning lane (myopic)	coefficient-myopic drivers	-0.0240	-0.63
	constant-myopic drivers	1.43	0.83
	heterogenity coefficient -myopic drivers	1.53	0.75
Lanes away from turning lane (with plan-ahead)	coefficient-drivers who plan-ahead	-4.08	-1.98
	constant-drivers who plan-ahead	2.05	3.01
	heterogenity coefficient -drivers who plan-ahead	0.466	0.74
Expected maximum utility from immediate lane*		0.915	7.22
Immediate Lane			
Lanes away from connecting lane:	coefficient	-1.01	-1.19
	constant	0.691	1.94
	heterogenity coefficient	1.96	3.48
Target lane dummy		3.16	4.54
Lanes away from target lane:	coefficient	-4.42	-3.00
	constant	2.12	2.14
	heterogenity coefficient	0.0904	0.36
Conflict dummy		-1.76	-9.63
Driver Class			
Driver population with >1 section lookahead (%)		18.3	2.07

The path-plan of the driver has an important role in the target lane selection. The two classes of drivers are found to have different sensitivities to path-plan considerations, which in this case has been modeled as an interaction between the number of lanes away from the correct lane and the driver specific random term (aggressiveness). The functional form best fitting the data is found to be as follows:

$$\left[\frac{\theta_1}{\psi_1+\alpha_1\upsilon_n}(e_n^{1l})(1-\delta_n)+\frac{\theta_2}{\psi_2+\alpha_2\upsilon_n}(e_n^{2l})\delta_n\right] \quad (9)$$

where,

$\delta_n = 1$ *if the driver plans-ahead beyond the immediate section*

e_n^{1l} = *lanes away from turning lane for myopic drivers*

e_n^{2l} = *lanes away from turning lane for drivers who plan-ahead*

$\theta_i, \psi_i, \alpha_i$ = *coefficients of vehicle class i*

As seen from the estimates, for both classes of drivers, the utility of lanes reduces if they are away from the lane that the driver needs to be in to follow his path. This disutility is, however, less for aggressive drivers, since they are more prone to make aggressive lane changes later if needed (inertia effect is dominant). The disutility was found to be larger and more significant for drivers who plan ahead (Class 2).

The Immediate Lane Choice Model

As seen in Table 2, immediate lane choices were found to be influenced by maneuverability considerations and inertia to continue to the naturally connecting lane. Inertia effects are captured by variables like current lane inertia and number of lanes away from the connecting lane. The inertia effect was greater for aggressive drivers. Aggressive drivers tend to stay in their current lane as long as possible and then make aggressive changes if a lane change is warranted by the path-plan. Drivers were also found to have a strong preference to reach their target lane and lanes closer to their target lanes. Maneuver to a given lane may not be possible due to conflicts with neighboring vehicles. In the case of such obstructions or conflicts, the driver can choose an immediately available lane or can wait until the neighboring vehicle moves and there are no obstructions to maneuver to the intended target lane. As a result, if there are conflicting vehicles in the direction of a lane, the driver was found to have a lower preference for that lane.

The improvement in the goodness-of-fit of the new model was statistically compared with a 'reduced form' model estimated with the same data which does not have anticipated delay, latent class and the associated heterogeneity in the path plan.

4. CASESTUDY 2: ANTICIPAITON IN CAR OWNERSHIP MODELS

The mid-term case study deals with the decision to increase the number of cars in the household. Previous research has revealed that important life course events have a critical role alongside household demographics and a significant anticipation effect is observed for some of the life/ demographic events (9). However, as mentioned, there are potential methodological limitations in the previous research which motivates the current research.

4.1. Data

The retrospective data used in this study was collected in some selected urban and rural areas of the Utrecht province in the Netherlands (see (15) for details). The data was collected in 2010 containing 20-year histories of randomly selected households. A total of 1,200 questionnaires were distributed of which 475 were returned. The main features of the dataset are: 1) information of household members, 2) retrospective information regarding the households' work and income situation, household composition, residential, work and travel situation, 3) prospective information about households intentions regarding the same aspects, 4) information regarding mobility and residential decisions of households' social networks, and 5) households' perceptions of the housing market, job opportunities, and travel costs. An example of the survey is given in Figure 4.

Please indicate in the calendar about important household issues/events –

	1990	1991	1992		2008	2009	2010
Household status							
5. Mention number of person in household at 1990 and only when it is different from earlier.							
6. Please indicate in the calendar by "X" or "√" when you had experienced one or more of the following events.							
(1) You left parent home							
(2) Marriage/Cohabitation							
(3) Birth of children							
(4) Child's home leaving							
(7) Divorce/Separation							
(6) Death of household member (please mention the relationship with the person died)							

Please indicate appropriate response for you in the calendar depicting last few years of your life and also mention the initial status at 1990 for each question –

	1990	1991	1992		2008	2009	2010
Travel and transport							
25. Please mention the number of cars available in your Household							

FIGURE 4 Snapshot of the retrospective survey.

This study is based on a subset of this dataset focusing on an increase in the number of cars in relation to anticipated changes in several aspects of life. This analysis is based on person-year-observations, i.e. every case is an observation from an individual for a particular year. The period covered here is reduced from 21 (20-years history + current year) to 17 years because two years lag and lead effects of events were considered. Therefore, it is possible to have a total of 8,075 (475 individuals*17 years) person-year-observations. However, consideration of missing values leads to 3,656 person-year-observations from 312 households.

The majority of respondents (around 75% of respondents) have a high education level, with university level education or higher vocational education (HBO). Age varies from 20 to 90 years with almost equal gender proportions. About 30% are 30-40 years old and around 15% are older than 60 years for person-year observations. More than 75% of the respondents were living with a partner and dual worker families accounted for about 60% of the person-year observations. The overrepresentation of highly educated persons has some implications for the conclusions we can draw and therefore cannot be generalized for the Netherlands. It is to be noted that except for education and gender, age of the respondent, household income and work-status, household composition and number of car availability are specific to a particular calendar year as reported by the respondents. Therefore, the representativeness of these variables is difficult to assess. In terms of age bias, it should be noted that since we use a 21-year history of all individuals, a bias toward older aged respondents is at least partly offset. Table 3 shows the frequency of occurrence of various events for the aforementioned samples, based on the total number of person-year-observation. As seen in the table, the data includes a total of 157 car ownership changes (additions and reductions) out of 3,656 person-year observations (i.e. 3.2% cases).

Table 3: Descriptive statistics of event history

Number of households	312	
Number of observations	3656	
Events	**Count**	**%**
Start of living together	50	1.4
Birth of the first Child	62	1.7
Home-leaving of the last Child	42	1.1
Separation or divorce	11	0.3
Residential move	273	7.5
Employer change (both or either)	423*	11.6
Employer change for respondent	305	8.3
Employer change for partner	140	3.8
Retirement event (both or either)	61*	1.7
Respondent Took retirement	37	1.0
Partner Took retirement	30	0.8
Increase in number of cars	114	3.2

*if respondent & partner both took retirement/change employer in same year then count is 1.

4.2. Model Structure

The proposed model is an extension of the study by Oakil et al. (8) where a binary logit model was developed using data collected using a retrospective survey to investigate effects of past and future changes in family composition, residential location and changes in job or job status. Results indicated that important life course events have a critical role alongside household demographics and a significant anticipation effect is observed in cases of childbirth and changes in residential location. However, the researchers used the observed life events in future years as direct variables and ignored the heterogeneity in planning ahead. In the proposed extension, we acknowledge the potential measurement error in the anticipated life events and account for the underlying heterogeneity by means of latent classes.

The decision makers are assumed to belong to either of the two classes:

- Myopic: people who do not anticipate effects of life events occurring in the next year on their current decisions
- People who plan-ahead: people who take into account effects of life events occurring in future year(s) in their current decisions.

Further, as proposed in Section 2.1, the observed life events in future years are used as indicators as opposed to direct variables.

The utility of increasing the number of cars in year *t* for an individual *n* belonging to Class 1 (the myopic group) can be expressed as follows:

$$U_{in1}^{t} = \beta_{in1}^{t} X_{in}^{t} + \beta_{n1}^{t} Z_{n}^{t} + \upsilon_n + \varepsilon_{in}^{t} \quad (10)$$

Where, υ_n is the individual specific error term, $\upsilon_n \sim N(0, \sigma_n^2)$

The corresponding utility for an individual *n* belonging to Class 2 (the plan-ahead group) can be expressed as follows:

$$U_{in2}^{t} = \beta_{in2}^{t} X_{in}^{t} + \beta_{n2}^{t}(Z_n^{t} + Z_n^{t*}) \quad + \upsilon_n \; + \varepsilon_{in}^{t} \tag{11a}$$

$$z_{nk}^{t+1} = \delta(z_{nk}^{t*}) + \upsilon_n \; + \delta_k \;, \forall k \tag{11b}$$

Where, z_{nk}^{t*} is the anticipated life event k at time t (latent), z_{nk}^{t+1} is the observation of the life event k at time t+1, δ_k the measurement error corresponding to the life event, $\delta_k \sim N(0, \sigma_k^2)$

The class membership probabilities are assumed to be $f(Z_n^t)$ and the likelihood function is same as Equation (5).

4.3. Results

Different functional forms of variables, as well as interactions between multiple variables, have been tested during estimation and the functions resulting best goodness-of-fit are selected. Effects of life events ranging from new births, deaths, formation of family (marriage/ moving together), residential relocation, and change in job status and/or job location have been tested but only the effect of a new birth was found to be intuitive and statistically significant.

For the class membership model, the impact of different socio-demographic characteristics (e.g. age, income, gender and education levels) was tested but only a 'higher education dummy' (which is 1 if the respondent is a graduate from university or applied universities) was found to be statistically significant.

The estimation results are summarized in Table 4.

As seen in the table, there is significant inertia towards the current levels of car ownership (as expected). Respondents over 60 are found to be less likely to increase the number of cars. The probability of an increase in the number of cars increases with income up to 7,500 Euros/month. People belonging to the highest income group after which it reduces slightly (though the coefficient is not significant at 95% level of confidence) and likely to be due to saturation effect. The utility for getting a new car reduces if the respondent already has 1 car (compared to no cars) and reduces further if he/she already has 2 cars. They are found to be more likely to get a new car if there is a new birth in the same year or the following year (only for the plan-ahead class). Starting to live together with the partner, changing employment and changing residential location in the same year is found to have a significant impact. When the variants of these variables are tested for the anticipation component, no intuitive or significant effect is however found.

The coefficient of the measurement term is found to be statistically significant but the standard deviation of the measurement error term and the individual specific error term are not found to be significant.

TABLE 4 Estimation results of the car-ownership model with anticipation

	Coefficient	Robust t-stat
Choice component		
Inertia term	3.83	7.88
Age >60	-2.26	-2.20
Income 1501-3000 Euros/month	0.78	1.46
Income 3001-4500 Euros/month	1.34	2.45
Income 4501-6000 Euros/month	1.65	3.00
Income 6001-7500 Euros/month	1.83	2.58
Income > 7500 Euros/month	1.35	1.86
Currently have 1 car	-1.10	-4.08
Currently have >1 car	-2.62	-5.48
New birth in the family in same year	0.99	2.37
New birth in the family following year (only for plan-ahead class)	1.34	1.58
Started to live together with partner in same year	1.56	3.79
Changed employment in same year	0.50	1.76
Changed resiential location in same year	0.96	3.10
Measurement component		
Measurement coefficient	0.81	4.14
Std deviation of measurement error term	1.18	0.56
Std deviation of individual specific error term	0.87	0.76
Class membership		
Constant for plan-ahead class	6.81	9.38
Coefficient of higher education in plan-ahead class	0.54	0.91

The goodness-of-fit values of the base model were compared with the proposed model with anticipation. The results indicate improvements due to the proposed anticipation formulation (Table 5).

5. CONCLUDING REMARKS

The paper presents a framework for incorporating anticipation effects in choice models and demonstrates it using a short-term and a medium-term decision context. Results indicate that adding anticipation does improve model fit but the magnitude of improvement can be context specific. Given that the paper had looked at two case studies which are inherently very different, it is premature to reach any conclusions about this and needs further investigation. In particular, comparison of model parameters of the same group of respondents for long, medium and short term decisions can provide further insights on this. Of course, improvements in model fit alone are not the main interest, especially given concerns in the hybrid choice literature (REF) about misattribution of benefits. The key interest in this work is behavioural and our work seems to suggest that anticipation plays a role and can be accommodated in choice models.

Among the two novel components (acknowledging the heterogeneity in anticipation and accounting for measurement errors), the former is found to have a more significant effect. Although the use of indicators results in marginal improvement, it needs further investigation. Also, in our current research, we have focused only on anticipation effect in isolation but a combination of state-dependence and anticipation is behaviourally more realistic and will be addressed in future research.

We also acknowledge that causality may be difficult to establish. For example, does a car-free decision maker buy a first vehicle in a given year because he or she expects to move house the following year, or does the ownership of the new vehicle influence his/her later decision on residential location?

An important area for future research is to establish whether the impact of anticipation and apprehension changes not just as a function of the choice made and its time horizon, but also of the consequences and the time horizon thereof. For example, while lane changing behaviour is a short term decision, failing to be in the correct lane for a motorway exit can have varying levels of impact, depending on the urgency of arriving on time at the destination but also the length of the detour required. Our work potentially also has important applications outside of a pure transport context, for example in a health setting, where discounting of future health consequences may vary extensively across individuals.

REFERENCES

1. Train, K.E., 2009. Discrete choice methods with simulation. Cambridge university press.
2. Cirillo, C. and Xu, R., 2011. Dynamic discrete choice models for transportation. Transport Reviews, 31(4), pp.473-494.
3. Viscusi, W.K., Huber, J. and Bell, J., 2008. Estimating discount rates for environmental quality from utility-based choice experiments. Journal of risk and Uncertainty, 37(2-3), pp.199-220.
4. Dubé, J.P., Hitsch, G.J. and Jindal, P., 2014. The joint identification of utility and discount functions from stated choice data: An application to durable goods adoption. Quantitative Marketing and Economics, 12(4), pp.331-377.
5. Rysman, M. and Gowrisankaran, G., 2009. Dynamics of Consumer Demand for New Durable Goods. National Bureau of Economic Research.
6. Samwick, A.A., 1998. Discount rate heterogeneity and social security reform. Journal of Development Economics, 57(1), pp.117-146.
7. Frederick S., Loewenstein G. and O'Donoghue T. (2002). Time discounting and time preference: A critical review. Journal of economic literature, 40(2), pp.351-401.
8. El Zarwi F., Vij A., Walker J. (2015) Integrated Hidden Markov and Discrete Choice Models: Developing a forecasting framework for the transition matrix model, 14th International Conference on Travel Behaviour Research, Windsor, UK.
9. Oakil, A.T.M., Ettema, D., Arentze, T. and Timmermans, H., 2014. Changing household car ownership level and life cycle events: an action in anticipation or an action on occurrence. Transportation, 41(4), pp.889-904.
10. Crastes dit Sourd, R., Calastri, C., Hess, S., Scheiner, J., and Holz-Rau, C. (2016) Modelling anticipation and interrelations between long-term mobility decisions over the life course: a recursive multivariate probit approach, 5th Annual Conference of the European Association of Research in Transportation, the Netherlands

11. Ehreke, I., Crastes dit Sourd, R., Beck, M., Hess, S., Axhausen, K.W., Holz-Rau, C. and Scheiner, J., (2016) A Dynamic Approach to Long Term Mobility Decisions in the Life Course, Transportation Research Board 95th Annual Meeting, USA.
12. Dalal P. and K.G. Goulias (2013) Preparation and Adaptation Processes in Travel Behavior Dynamics: A Latent Growth Modeling Approach. Geotrans Report 2013-01-01. Santa Barbara, CA.
13. Choudhury, C., Ben-Akiva, M.E., Toledo, T., Rao, A. and Lee, G., 2007. State dependence in lane changing models. In Proc. 17th Int. Symp. Transp. Traffic Theory (pp. 711-733).
14. Choudhury, C.F., Ben-Akiva, M. and Abou-Zeid, M., 2010. Dynamic latent plan models. Journal of Choice Modelling, 3(2), pp.50-70.
15. Oakil, A.T., 2013. Temporal dependence in life trajectories and mobility decisions (Doctoral dissertation, Utrecht University).
16. Verhoeven, M., Arentze, T.A., Timmermans, H.J.P. and Van der Waerden, P.J.H.J., 2008, January. Retrospective surveys: some experiences in the context of measuring lifecycle events. In The 87th Annual Meeting of the Transportation Research Board. CD-ROM. Washington, DC.
17. Daly, A., Hess, S., Patruni, B., Potoglou, D. and Rohr, C., 2012. Using ordered attitudinal indicators in a latent variable choice model: a study of the impact of security on rail travel behaviour. Transportation, 39(2), pp.267-297.
18. Abou-Zeid, M. and Ben-Akiva, M., 2014. 17 Hybrid choice models. Handbook of choice modelling, p.383.
19. Cambridge Systematics, Inc. (2006) NGSIM: Lankershim Data Analysis. http://www.ngsim.fhwa.dot.gov
20. Choudhury, C. and Ben-Akiva, M., 2008. Lane selection model for urban intersections. Transportation Research Record: Journal of the Transportation Research Board, (2088), pp.167-176.
21. Choudhury, C.F. and Ben-Akiva, M.E., 2013. Modelling driving decisions: a latent plan approach. Transportmetrica A: Transport Science, 9(6), pp.546-566.
22. Vij, A. and Walker, J.L., 2016. How, when and why integrated choice and latent variable models are latently useful. Transportation Research Part B: Methodological, 90, pp.192-217.