
We recommend you cite the published version.
The publisher’s URL is: 
http://dx.doi.org/10.1016/j.tra.2010.01.007

Refereed: Yes

(no note)

Disclaimer

UWE has obtained warranties from all depositors as to their title in the material deposited and as to their right to deposit such material.

UWE makes no representation or warranties of commercial utility, title, or fitness for a particular purpose or any other warranty, express or implied in respect of any material deposited.

UWE makes no representation that the use of the materials will not infringe any patent, copyright, trademark or other property or proprietary rights.

UWE accepts no liability for any infringement of intellectual property rights in any material deposited but will remove such material from public view pending investigation in the event of an allegation of any such infringement.

PLEASE SCROLL DOWN FOR TEXT.
Which road do I take? A learning-based model of route-choice behavior with real-time information

Eran Ben-Elia¹, & Yoram Shiftan²

¹ Urban and Regional Research Centre, Faculty of Geosciences, Utrecht University, P.O. Box 80115, 3508 TC, Utrecht, The Netherlands. e.benelia@geo.uu.nl

² Faculty of Civil & Environmental Engineering, Technion - Israel Institute of Technology, Technion City, Haifa, 32000, Israel. shiftan@technion.ac.il

(* correspondence author)

Submitted to Journal “Transportation Research A” on 27/4/2009
Revised on 9/2/2012, Accepted 29.3.2011

Key Words: Experience, information, panel data, Prospect Theory, random utility models, reinforced learning, route-choice, travel time variability, uncertainty.

Abstract

This paper presents a learning-based model of route-choice behavior when information is provided in real time. In a laboratory controlled experiment, participants made a long series of binary route-choice trials relying on real-time information and learning from their personal experience reinforced through feedback. A discrete choice model with a Mixed Logit specification, accounting for panel effects, was estimated based on the experiment's data. It was found that information and experience have a combined effect on drivers' route-choice behavior. Information expedited learning and in the short run increased risk-seeking behavior. Informed participants tended to base their decisions on memorization relating to previous outcomes compared to non-informed participants who tended to rely mostly on recent outcomes. Informed participants were also more prone to risk-seeking and had greater sensitivity to travel time variability. In comparison, non-informed participants appeared to be more risk-averse and less sensitive to changes in levels of variability. These results have important policy implications on the design and implementation of ATIS initiatives. The advantage of incorporating insights from behavioral research to improve estimations obtained from discrete choice models is also discussed.
“Would you tell me, please, which way I ought to go from here?”
That depends a good deal on where you want to get to, said the Cat. I don’t much care where – said Alice. Then it doesn’t matter which way you go, said the Cat. – so long as I get SOMEWHERE, Alice added as an explanation. Oh, you’re sure to do that, said the Cat, if you only walk long enough. ”.

Lewis Carrol, Alice’s Adventure in Wonderland

1. Introduction

Like Alice in the well known two road predicament, many drivers find them selves in frequent dilemmas over their route choice decisions, particularly when real-time information is provided describing the possible travel alternatives that is not congruent with their previous driving experiences. Modeling the impacts of information provision on drivers' route-choice decision making is an essential step in improving the design of an efficient advanced traveler information system (ATIS) and assessing its benefits. ATIS includes all those systems that use information technology to inform, monitor control and even charge travelers for using the roads, usually implemented as part of intelligent transport systems (ITS) initiatives (Bonsall, 2000). Additional information together with advanced technologies like GPS-based path-finders is regarded as likely to contribute to reducing of travel time uncertainty, enabling travelers to choose efficiently among the available routes, save travel time and reduce congestion (European Commission, 2008). However, the exact impact of ATIS is likely to be sensitive to drivers' behavioral and cognitive response to information which is much less understood. In addition effects such as experience and learning seem also to play an important role in the decision making process.

The following hypothetical example (and see Table 1) demonstrates the complexity of correctly predicting drivers' route-choice under uncertainty on the basis of a simple but nonetheless non-trivial predicament:

Suppose Alice (who is trying to get back home from Wonderland) is faced with information, provided in real time, regarding the choice between two routes: a faster route (F) and a slower route (S). The average travel times on the faster route - as predicted by ATIS - is 25 minutes and on the slower route - 30 minutes. Alice has no specific obligation to arrive home at a specific time, but obviously would like to do it in the least amount of time. ATIS provides real-time information regarding the travel time ranges (the deviation around the mean value). The possible ranges are ±5 or ±15 minutes for each route. Consequently Alice is
faced with three possible travel time scenarios (as shown in Table 1) which we name “Fast & Safe”, “Fast & Risky” and “Low Risk”.

Under these assumptions which is the most probable choice Alice will make in each scenario? Answering this question requires formulating a choice model that explains her choices. However, estimating the impact of information within the framework of a choice model is hardly a simple task. Since information is assumed to change the level of uncertainty in the choice situation, the main challenge is to model travelers’ response to information. Thus, the assumption of perfect information generally embedded in random utility-based route-choice models, i.e. without information effects, is discarded. These models assume drivers’ cognition is governed by rational choice. They do not explicitly abstract behavior under uncertainty; rather they apply different sets of assumptions on the distributions of the unobservable factors (Watling & van Vuren, 1993). In this context, flexible error terms as applied in mixed discrete choice models such as Mixed Logit (Srinivasan & Mahmassani, 1999) and Mixed Probit (Mahmassani & Liu, 1999) have proved of added value. Using dynamic specifications is another approach to reduce uncertainty by updating the level of utility over time (Horowitz, 1984). It is also possible to assume that travelers’ cognition reflects bounded rationality (Simon, 1982) whereby choices are made based on simple decision rules or heuristics in the form of thresholds of accepting possible outcomes. Srinivasan & Mahamassani, (2003) applied this approach by using a Mixed Logit specification. Although these approaches have provided valuable insights, to a large degree, there is still a considerable lack of understanding of the cognitive and behavioral aspects of drivers’ decision-making processes under uncertain travel times as well as the behavioral impacts of real-time information. Obviously, this is also a key factor in ATIS design and its efficient deployment.

This paper extends some initial insights from behavioral decision research gained by the authors in a previous paper (Ben-Elia et al., 2008) – mainly that descriptive information, on one hand, and experience from learning, on the other hand, have significant impacts on the route-choice decision process. In this paper we focus on adopting this approach within the framework of econometric choice models. As shown in the next section, different behavioral generalizations can provide distinct and even contradictory descriptions of human behavior under uncertainty in general and particularly in the context of the seemingly trivial route-choice example depicted above. This fact further complicates the task of increasing the behavioral realism of travel behavior models.

The rest of the paper is organized as follows. Section 2 contains a review of the state-of-the-art literature on behavioral decision research and the difficulties in adapting it to study route choice behavior under uncertainty as reflected in the hypothetical problem presented
above. Next in section 3, a route-choice experiment, based on this example, is designed to capture drivers' behavior under uncertainty, followed by a description of data collection and longitudinal analysis. Section 4 details the mixed modeling approach applied in the behavioral analysis. The presentation of results of the model estimation (Sec. 5) is followed by a discussion (Sec. 6) and conclusions addressing policy implications and future research directions (Sec. 7).

2. Literature Review

Traditionally, travel behavior modeling has been based on the axioms of expected utility (Bernoulli, 1738; Von-Neumann & Morgenstern, 1944; Luce & Raiffa, 1957). Random utility based discrete-choice models or RUM provide an econometric interpretation of expected utility theory. This approach is still regarded as the official workhorse for most travel related behavioral modeling. RUM have been developed considerably in the past three decades and specifically for route-choice modeling. Chronologically speaking, the classic Multinomial Logit Model (Daganzo & Sheffi, 1977) was the first to be applied followed by the C-logit model (Cascetta et al., 1996), the IAP logit (Cascetta & Papola, 1998) and Path-Size Logit (Ben-Akiva & Bierlaire, 1999). Following the presentation of the General Extreme Value theorem (McFadden, 1978) more flexible modeling structures were developed including: Nested Logit (Ben Akiva & Lerman, 1985); Cross-Nested Logit (Vovsha, 1997); General-Nested Logit (Wen & Koppelman, 2001) and Paired-Combinatorial Logit (Chu, 1989; Koppelman & Wen, 2000). These GEV-based models were applied to route-choice models with an adaptation of GNL and PCL (Prashker & Bekhor, 1998; Gliebe et al., 1999; Bekhor & Prashker, 2001) and a GNL based Link Nested Logit (Vovsha & Bekhor, 1998). In the last decade, a major breakthrough in modeling capabilities was achieved by the introduction of the Mixed Logit or Logit Kernel model (Ben Akiva & Bolduc, 1996; Bhat, 1998; Bhat, 2000; McFadden & Train, 2000). Several studies have recently adapted Mixed Logit in the context of route-choice modeling: (Bekhor et al., 2002; Srinivasan & Mahamassani, 2003; Jou et al., 2007). A detailed review of such RUM-based route-choice models is provided by (Prashker & Bekhor, 2004).

Behavioral decision research (Kahneman & Tversky, 1984; Tversky & Kahneman, 1992) has revealed empirically systematic violations of some of the assumptions of expected utility theory (EUT) shared by RUM. Some researchers have even raised concern over the validity of RUM in forecasting travel behavior (Gärling & Young, 2001). Three other possible alternative approaches, raised in the cognitive decision research literature, for describing behavior under conditions of uncertainty are the 'Hot Stove', Prospect Theory and Reinforced Learning.
When the feedback available to the decision makers is limited to their obtained payoff, experience reduces the tendency to select risky or highly variable alternatives. Based on the observation that most people refrain from touching for a second time a *hot stove*’s door, Denrell & March, (2001) and Denrell, (2007) demonstrate that this pattern is a natural consequence of the inherent asymmetry between the effect of good and bad experiences. In line with the ‘classic’ law of effect (Thorndike, 1898), good outcomes are expected to increase the probability that a choice will be repeated and facilitate the collection of information concerning the alternative that has yielded the good outcome. Bad outcomes are expected to reduce the probability that the choice will be repeated, and impair the collection of information concerning the alternative that has yielded the bad outcome. As a result, the effect of bad outcomes lasts longer than the effect of good outcomes, and decreases the tendency to choose the option associated with the high variability.

*Prospect Theory* deals with situations involving one-shot decisions, based on descriptional information of the available alternatives. It asserts that decision makers frame possible outcomes as gains or losses based on a point of reference and not according to final-states as the classic interpretation of EUT suggests (Kahneman & Tversky, 1979). In addition, in EUT, decision makers are usually risk-averse. Conversely, according to the predictions of Prospect Theory, decision makers will usually reveal risk-averse behavior in the case of gains, as well as, risk-seeking behavior in the case of losses. For example, most people prefer the gamble, or prospect “there is an 80% chance of losing 4000$, over "loose 3000$ with certainty (100%)". In addition the presence of loss aversion implies that losses loom larger than equivalent gains. Another issue that Prospect Theory addresses is non-linearity in the perception of probabilities.

Unlike one-shot choices, with repeated decisions, decision makers can learn from their own experience. Studies based on *Reinforced Learning* assert that increasing the level of uncertainty in the environment moves behavior towards the direction of random choice (i.e. indifference). This ‘payoff variability effect’ occurs when the decision maker receives no specific information describing the possible outcomes of choice and has to rely on feedback from past experience (Busemeyer & Townsend, 1993; Mayeres et al., 1996; Erev & Barron, 2005). Experience leads to adaptive learning but, at the same time, it is also a function of sampling available information on the basis of past experience which at times can lead to biased estimates (Tversky & Kahneman, 1974).

We demonstrate the significance of these three behavioral generalizations to predicting route-choice within the framework of the hypothetical example presented in Section 1. Let the term *maximization* refer to the proportion of choices of the Fast route. A generalization of the ‘hot stove’ effect implies that repeated experience without information is
likely to lead Alice to deviate from maximization in the direction of risk-aversion. Thus, with information and under scenario Fast & Risky maximization is expected to increase, whereas in the other two scenarios the difference will be negligible. A partial support to a risk-averse generalization, in the case of route-choice, was found by (Abdel-Aty et al., 1997) and (Abdel-Aty & Abdallah, 2004). Avineri & Prashker, (2006) also assert risk aversion when drivers face static pre-trip travel time information.

Since it is reasonable to assume that in the case of commuting, travel time is usually perceived as a loss, a generalization of the common assumptions of Prospect Theory implies that the availability of information - complete description of the travel time distribution - would lead Alice in the direction of risk-seeking behavior. Thus, information is expected to increase the rate of maximization in scenario Fast & Risky and impair it in scenario Fast & Safe. A partial support to a risk-seeking generalization in the case of route-choice was found by Katsikopulos et al.(2002).

Finally, in the case of Reinforced Learning, the payoff variability effect is reduced by precise information, which implies that available information will increase Alice’s maximization in all the scenarios considered here. Moreover, it implies that the effect will be particularly large under scenarios Fast & Risky and Fast & Safe. Avineri & Prashker, (2003) demonstrate the importance of this observation by showing that increasing travel time variability of the route with a higher expected travel time (like in Fast & Safe) increases its attractiveness.

The three described behavioral approaches provide clear evidence for contradictory predictions of individual behavior when both information and experience are involved in route-choice. This is a major drawback if travel behavior models are to be improved by adapting more behavioral realism. Furthermore, previous research endeavors mentioned above, provide key insights for further scientific investigation. First, as the three theories show, different decisions result when choices are based on information compared to experience. For example, when facing unknown delays Japanese drivers tended to rely on broadcasted traffic information and less on their previous experience in the short run (Fujii & Kitamura, 2000). Selten et al.(2007) provide evidence to suggest that both effects play a role in the long run. However they did not specifically study descriptional information, rather their focus was on feedback and foregone payoffs. Second, both the type and content of information could well affect drivers' response. For example, Avineri & Prashker (2006) who studied the effect of pre-trip static information (prior description of routes' mean travel times) with subsequent gaining of experience through feedback, found that information increases the tendency of risk aversion. In contrast, providing information about the range (the minimum and maximum travel times) as depicted by Katsikopulos et al.( 2002), oddly
enough, reveals tendency for risk-seeking behavior. However, since their study was one-shot and learning was unaccounted for, its implications, for understanding behavior in the long run, remain unclear.

Given this incomplete picture of the cognitive and behavioral aspects of route-choice decision making, more research is certainly justified. The most complete approach to study route-choice behavior and the impacts of real-time information would be one based on field studies, i.e. revealed preference (e.g. Kraan et al., 1999). However, as both time, money and the availability of operational ATIS are always an issue of concern, as well as the limited ability to control for unknown factors, the practical solution is the use of stated preference through careful experiment design and proper control (e.g. Wardman & Bonsall, 1997). Moreover, should quality field data of route-choice be available, it may well be possible to combine the two types of data as suggested for example by Ben-Akiva & Morikawa, (1990).

3. Experiment design

The route-choice experiment was designed based on the following principles: First recognition of the importance of both descriptive information and learning from experience. Second, information reflects the level of uncertainty en-route by providing a description of travel time variability through the use of the range as a visual cue. Third, rather than pre-trip and static, information is treated as en-route and dynamic – i.e. real-time information. Since in this study we focus on individual behavior, interaction between drivers and the effects of driver behavior on congestion are not accounted for in the design, unlike Selten et al., (2007). However, in Section 7, this is discussed as a possible important future research direction.

3.1 Setup and design

The experiment was designed in accordance with the hypothetical example in Table 1. A detailed account of the experiment design and a part of the longitudinal analysis are reported by Ben-Elia et al.(2008). To keep the discussion clear, and as this is an important background to the rest of the presentation, we provide here a brief description of the design.

49 participants (undergraduate Technion students) were divided randomly between two groups: informed and non-informed (the control). Upon recruitment, participants filled out a computer-based survey regarding their socio-demographic characteristics and habitual travel behavior. Subsequently, they were requested to conduct a series of route-choice experiments. In both groups participants were faced with all the three traffic scenarios (treatments) of the hypothetical example as presented in Table 1. Each scenario included 100 random choices so in total each participant completed 300 trials. Participants were
randomly assigned to 1 out of the 6 \((=3!\)](n=6\text{!})\) possible orders of running through the scenarios. After each scenario was completed an announcement was given warning of the beginning of the next scenario. However, participants were neither informed in advance how many runs they were expected to complete within each scenario nor how many scenarios were they expected to complete. For each choice situation the informed group \((N=24)\) received real-time information about the travel time range (the minimum and maximum travel times) for each of the two routes – fast and slow, according to the design in Table 1. A small degree of variation was programmed (between 0-5 minutes around the mean) so that the messages were not seen as constant throughout the entire scenario. This simulated a simple variable message sign (VMS). Following the choice, a feedback was received regarding the ‘actual’ travel time on the chosen route but not of the alternative one. This travel time was randomly drawn from the distribution of the travel time range to encourage user confidence in the information. The control group \((N=25)\) received the same feedback but no other information regarding the possible travel times on the routes. No foregone payoff (i.e. feedback on the non-chosen route) was provided.

### 3.2 Longitudinal analysis

Longitudinal analysis, including analysis of variance, was conducted to identify significant effects. Separate analysis was also conducted for the short run (first 10 trials) and the long run (last 50 trials). Figure 1 portrays the maximization rate (average of Fast shares), exploration rate (average route switching share) and the payoff (average obtained travel times). Results of ANOVA for within (F-test) and between-group (t-test) significance are depicted in Table 2.

As expected, for both groups and under all three scenarios, maximization was higher than 50% and this increased gradually over time towards convergence. However, there were remarkable differences between the behaviors of the two groups.

Behavior without information was similar in all of the three scenarios. In line with the power law of practice (Blackburn, 1936), the maximization rate increased in a typical learning curve style as participants gradually learned that the FAST route is on average better off. Exploration rates gradually declined. Travel times were also slightly lower. These results are close to Selten et al., (2007) for the feedback treatment. It is evident that maximization is higher when the level of variability is low on both routes (Low Risk), whereas in the other two scenarios higher variability impedes the learning ability, resulting in more exploration and route switching. These results assert that without information, behavior is quite sensitive to the payoff variability effect. The results pertaining to Fast & Safe also appear in accordance with the findings of (Avineri & Prashker, 2003).
Behavior with information, in contrast, is quite different. First, FAST shares are significantly higher (in two out of three scenarios) in the short run (first 10 trials). In addition, the rate of exploration is lower. This suggests information contributed to expediting learning and reduced the need for further trials devoted to exploration (switching). Second, standard deviations were higher which suggests information contributed to more heterogeneity in the response. Third, information seems to encourage risk-seeking behavior in the short run (notice the higher share in Fast & Risky compared to Fast & Safe). This result fits well with the prediction of Prospect Theory (i.e. risk-seeking or gambling in the face of potential losses). However, in the long run, as participants gain more experience on the distributions of travel times, the affect of information decreased and the differences between the informed and non-informed groups loose significance. Fourth, despite higher maximization and lower exploration rates, when variability is high in one of the two routes (Fast & Safe, Fast & Risky), information does not contribute to a higher payoff in terms of lower travel times whereas when variability is low (Low Risk) the advantage of information is clearly visible. Thus even with information there is still a degree of sensitivity, albeit weaker, to payoff variability. For further discussion of these results see (Ben-Elia, 2007; Ben-Elia et al., 2008).

4. Modeling approach

The longitudinal analysis identified two important effects on route-choice behavior: the effect of information and the effect of learning from experience. It seems that these two effects also combine over time. As such, we would like to formulate a positive (i.e. realistic) behavioral choice model incorporating these effects to improve our understanding and to provide, at a later stage, a forecasting tool to drivers' behavior under such situations. This endeavor is a main aim of this paper.

4.1 Choice of modeling approach

Both the theories of Prospect Theory and Reinforced Learning appear to have important contributions to any possible model. However, from an analytical perspective both pose difficulties. Reinforced Learning-based models are better suited to predict behavior without a description of the payoff distribution i.e. without information. Prospect Theory, despite its better abstraction of behavior under uncertainty, is quite difficult to apply analytically. One problem is that it typically describes choice alternatives in terms of a small number of outcomes with particular probabilities, and not in terms of continuous probability density functions, which usually apply to travel time distributions (Ettema & Timmermans, 2006). A second problem is determining the reference point for the valuation of gains and losses. Different types of reference points could be assigned for our route-choice problem. Furthermore, in a repeated choice experiment this problem can become very complex as
reference points can shift or even become time-dependent. Third, Prospect Theory is based on one-shot decisions, i.e. without repetitions. However, this condition can be easily relaxed (Barron & Erev, 2003).

Given the above considerations and acknowledging that these issues deserve more attention, we decided to use random utility models (RUM) as a starting point. Recent advancements in discrete choice allow RUM to be modified in a way than can partially account for these two alternative theories. First, utility functions can be modified to represent concepts of risk-aversion and risk-seeking (Polak, 1987). Second, gains and losses in attributes can be accommodated in RUM by setting the reference point exogenously. Polak (2008) actually presents a possibility to estimate reference points endogenously. However, this depends to a great extent on the type of data and the design used. Senbil & Kitamura (2004) for example and Jou et al. (2008) specify a gains/losses model within the context of discrete choice for the modeling departure time. In addition, some of the dynamic aspects of reinforced learning can be dealt within RUM by specifying a dynamic structure (panel) or using lagged dependent and explanatory variables.

Following the decision to apply RUM to our data, the possible model specifications available were evaluated. Given that our data is longitudinal and consists of repeated choices, our observations are dependent and correlated. Thus, the classic assumption of identically independent distributed (i.i.d) error terms applied in standard discrete choice applications is violated. Specifying panel effects can accommodate for this deficiency. In the context of RUM, this can be accomplished using a mixed discrete choice model such as the Mixed Logit model described below.

4.2 The Mixed Logit model with panel data

Mixed Logit (also referred to as Logit Kernel) is an advanced and highly flexible discrete choice model that can approximate any random utility model. Although the derivation assures that the model is consistent with utility maximization, it does not preclude the model from being consistent with other forms of behavior (McFadden & Train, 2000). The very first application of Mixed Logit (MXL) was used for estimating automobile demand (Boyd & Mellman, 1980; Cardell & Dunbar, 1980). MXL accommodates random taste variation, substitution patterns, and correlation in unobserved factors unrestricted over time. Unlike Probit models, it is not restricted to normal distributions, but can be derived under a variety of different specifications, whereby each derivation provides a particular behavioral interpretation (Ben Akiva & Bolduc, 1996; Bhat, 1998; Brownstone et al., 2000; Hensher & Greene, 2001).
In formal terms of discrete choice modeling, the decision maker \((n)\) is assumed to have some level of utility \((U)\) regarding an alternative \((i)\) in his choice set which is defined generally as (eq. 1):

\[
U_{ni} = \beta X_{ni} + \varepsilon_{ni}
\]

whereby: \(\beta\) is a vector of parameters, \(X_{ni}\) is a vector of explanatory variables defining individual \(n\) and alternative \(i\) and \(\varepsilon\) is a random disturbance representing unknown factors.

MXL results from a particular assumption on the specification of \(\varepsilon\) in which there is an i.i.d Logit term and a flexible term (normally distributed but not necessarily) that captures the interdependencies between alternatives. MXL probabilities are the integrals of Logit over a density of parameters. The choice probabilities in MXL \((P_{ni})\) can be expressed in the form (eq. 2):

\[
P_{ni} = \int \frac{e^{\beta X_{ni}}}{\sum_{j=1}^{T} e^{\beta X_{nj}}} f(\beta) d\beta
\]

Thus an MXL probability is a weighted average of the Logit function evaluated at different values of \(\beta\), with the weights given by the density function \(f(\beta)\). That is, MXL is a mixture of the Logit function evaluated at different \(\beta\)'s with \(f(\beta)\) as the mixing distribution. Like all discrete choice model the log-likelihood function is used for estimation. Since the integrand of MXL has no closed form the probabilities are approximated using simulation. These simulated probabilities are then inserted in to the log-likelihood function to give a simulated log-likelihood (Train, 2000; Bhat, 2001; Bhat, 2003).

The MXL specification is easily generalized to allow for repeated choices i.e. panel (Bhat, 1999; Train, 1999; Revelt & Train, 2000). Thus, the utility from alternative \(i\) in response \(t\) by person \(n\) is defined as (eq. 3):

\[
U_{nit} = \beta_n X_{nit} + \varepsilon_{nit}
\]

The only difference between an MXL specification with a panel and one with a single choice per decision maker is that the integrand involves a product of logits, one for each time period (eq. 4):

\[
P_{nit} = \prod_{t=1}^{T} \frac{e^{\beta_n X_{nit}}}{\sum_{j=1}^{T} e^{\beta_n X_{nj}}} f(\beta) d\beta
\]
Accounting for the effect of repeated choices in the panel of each decision maker can be accomplished in two main forms. Both forms are virtually equivalent but differ in the interpretation. The first form (eq. 5) is error components ($\eta$), which are random factors with 0 mean (this can be relaxed to observe some other rule) and an unknown variance parameter ($\theta$). They represent different correlations or substitution patterns among alternatives:

$$U_{nit} = \beta_n X_{nit} + \varepsilon_{nit} + \eta_{nit}, \quad \eta_{nit} \sim iid(0, \theta)$$  \hspace{1cm} (5)

Note, in eq. 5 - $\beta_n$ represents a vector of fixed coefficients like standard logit.

The second form (eq. 6) applies random parameters, that is estimating for a certain variable's coefficient the parameters associated with its distribution (i.e its $f(\beta/\theta)$). In this case $\beta_n$ is a random vector with mean $\beta$ and a covariance matrix ($\Sigma_{\beta}$).

$$U_{nit} = \beta_n X_{nit} + \varepsilon_{nit}, \quad \beta_n \sim iid(\beta, \Sigma_{\beta})$$  \hspace{1cm} (6)

Correlations between random coefficients (i.e. $cov(\beta_n, \beta_{nj})$) can be specified to create more flexible structures. However, the dimensions of the integrand increase with each additional random parameter. With many parameters the model is computationally difficult to estimate. In this case error components are more general and convenient for model estimation. In both forms some assumption is set for this distribution of the random factors (e.g. normal, uniform, etc.). Greene (2008) provides a broader discussion of possible distributions and their implications.

The simplest specification treats the coefficients - $\beta_n$ that enter utility as varying over individuals but being constant over responses for each person. This assumption which assumes stability over time is defined sometimes as an 'agent effect'. In contrast to the stability over time assumption of the agent effect, a different specification defined as an 'autoregressive process' assumes time dependency by defining a serial correlation over time. However the latter is much more complicated to estimate and in the case of many repeated choices the size of the covariance matrix makes it virtually intractable. Notwithstanding, it is still possible to specify an agent effect and account for the series of decisions by estimating the outcome of previous choices with lagged dependent variables.

Lagged dependent variables (LDV) which represent state-dependence also referred to as stickiness or recency are straightforwardly estimated with MXL. The only problem is the necessity to assume initial conditions to compute the integrand. However, when the choice process is observed from the beginning - the issue of initial conditions does not arise. Experiment data is a clear example of observing the entire sequence of choice whereby, the initial condition can be assumed to be the first observation. It is worth noting that LDV's can be specified in MXL without adjusting the probability formula or simulation method (unlike
probit). With MXL, a LDV entering the utility function remains uncorrelated with the error terms for period $t$, since these terms are always independent over time. In addition, past and future exogenous variables can be added to the utility function to represent lagged responses, inertia and anticipatory behavior without changing the estimation procedure. For more detailed explanation see the detailed account provided by (Train, 2002) as well as others mentioned throughout this section.

4.3 Specification

The model specification applied for our case study, in the binary choice between the fast (FAST) and slow (SLOW) routes included two ‘linear in the parameters’ utility functions. The reference choice (i.e. choice = 0) was always the SLOW route and all alternative specific constants were specified in utility function of FAST. Since only two alternatives exist error components can account for only one general random effect. This limits the amount of heterogeneity that the model could capture. Consequently, panel effects were accounted for, using random parameters. These were specified only for certain variables with a normal distribution for the disturbance and applying an ‘agent effect’ i.e. stability over time. Furthermore, as discussed above, temporal differences (short and long run) were accounted for using LDV’s and dummy variables. Correlations between some of the random parameters had also been tested but were found to be, in retrospect, not significant. The formal specification (eq. 8) is followed by a description of the different factors included in the model. The definitions of the factors can be found in Table 3. Coefficients in parenthesis indicate standard deviations of the random parameter to the left. As noted, the main challenge in the model specification relates to the inclusion of learning and experience effects, sensitivity to variability as well as the behavioral impacts of real-time information.

$$
\begin{align*}
U_S &= \beta_{\text{MEAN}} \text{MEAN} + \beta_{\text{TIMES}} \text{TIMES} \\
U_F &= \beta_{\text{MEAN}} \text{MEANF} + \beta_{\text{TIMEF}} \text{TIMEF} + \beta_{\text{LOWRISK}} \sigma_{\text{LOWRISK}} + \beta_{\text{RISKY}} \sigma_{\text{RISKY}} \\
&+ \beta_{\text{LOW}} \text{LOW} + \beta_{\text{HI}} \text{HI} + \beta_{\text{STICK}} \sigma_{\text{STICK}} + \beta_{\text{CWA}} \sigma_{\text{CWA}}
\end{align*}
$$

Factors relating to the individual characteristics included: gender, age, holding a driving license, weekly frequency of car use and car ownership. These were gathered on screen as mentioned in the design’s description. However, in successive trial and error attempts it was found that neither these nor their interactions with route-specific variables were significant. This may be due to the small and homogenous sample (all undergraduate students) which resulted in a small variance. Consequently these variables were not included in the final results. The implications of personal-related variables are discussed in section 7.
Route attributes included: Feedback travel times (TIMEF, TIMES) and the mean travel time (MEANF, MEANS). Feedback travel times were the travel times displayed following each participant’s choice and were specified as a lagged exogenous variable, depending on whether the last choice was the fast or slow route. For example, if in a participant’s 1st trial FAST was chosen with resulting travel time of 29 minutes than this would appear in the data for his second trial under the feedback travel time of the FAST route. Mean travel times were specified as the average of the predicted travel times obtained in each choice trial and for each route from the simulated VMS according to the design described in Section 3..

Factors representing learning were specified separately for the short and long run in line with the different behavioral trends found in the longitudinal analysis (see Figure 1 and Table 2). Learning in the short run (STICK) represented stickiness or recency, i.e. repetition of previous behavior. This reflects the law of effect (Thorndike, 1898) which states that a positive outcome increases the probability of repeating the action that caused it. Formally it was specified as a LDV. (i.e. if in trial t-1 FAST was chosen then the value for trial t equals 1 and 0 otherwise).

In contrast, learning in the long run was defined as a function of all previous outcomes which reflects the effect of memorization. Formally (eq. 7) a cumulative weighted average (CWA) of the preceding choices of each participant (n) and in each scenario (s) was specified and computed with a harmonic average (based on the inverse rank of each observation)

\[
CWA_{nst} = \frac{1}{\sum_{i=2}^{t-1} \frac{1}{i-1}} \left( \sum_{i=2}^{t-1} Y_{nst} \times \frac{1}{i-1} \right), \quad t = 3,\ldots,100
\]

whereby: \( Y \) is 1 if choice was FAST and 0 otherwise, and \( i \) is the inverse rank of the preceding observations for trial \( t \) (in reverse order).

This computation is relevant from the 3rd trial since the 1st observation has no preceding choice (i.e. a missing value) and the choice preceding the 2nd trial is obviously the 1st. To clarify the computation we show an example for the 15th trial (t=15). We would include all 14 previous choices and weight them by their inverse rank: i.e. \( Y_{14} \) is ranked as \( i=2 \), \( Y_{13} \) is ranked \( i=3 \) etc. Note that if all choices are FAST than CWA equals 1, whereas if some choices are SLOW than CWA is less than 1. This computation is done for each trial in a recursive manner (using a programmed loop). As a result recent outcomes are weighted heavier than earlier ones. Retrospectively this specification out-performed (in terms of the final log-likelihood) other options such as specifying moving averages of different lengths or
using non weighted averages. Note that the weight could, in practice, have been endogenously estimated for each respondent. However, this would have made the estimation even more difficult than it already had been.

Since the longitudinal analysis also revealed very distinct behavioral tendencies in the short (first 10 trials) and long (last 50 trials) runs, two different levels of experience were also specified with dummy variables. Low experience reflects choices within the first 10 trials and high experience reflect choices in the last 50-100 trials. The reference value was the mid range between trials 11-49.

**Sensitivity to variability** of the travel times was specified using the combinations of travel time ranges which were available in each scenario. In each trial the possible ranges for each route were either Low (10 minutes) or High (30 minutes) depending on the specific scenario (see Table 1). The three possible combinations therefore are: Low-Low ($F=10$, $S=30$), Low-High ($F=10$, $S=30$) and High-Low ($F=30$, $S=10$). Each combination is unique per scenario. For the easiness of interpretation the coefficients are referred to by the scenarios names. The reference value was the Low-High combination which occurs only in the Fast & Safe scenario.

The main aim of this study is to understand the behavioral effects of real-time information. However, specifying information as an exogenous variable does not capture all the possible behavioral interactions and the cognitive implications involved in the availability of information. Given its importance it was decided to fully segment the sample by the two experimental conditions (or groups): with information and without information. A separate model was estimated for each group and the log-likelihoods of both segmented models were summed to obtain the total final log-likelihood of the estimation. However, to keep the two segmented models consistent one degree of freedom was sacrificed by normalizing one coefficient – the same in both segments. We chose to normalize the coefficient of the mean travel time to 1. In addition the scale of the Logit formula ($\mu$) was estimated for each segment. The result was two sets of estimated and scaled coefficients which permitted a consistent statistical testing of differences between the segments using the simple t-test. The segmented model is referred to as the **unrestricted model** - as it allow the parameters to vary across segments. This model was also compared to an alternative model – **restricted model** – restricting the parameters to be the same for both segments, thus assuming no differences between the two groups, i.e. information has no behavioral effect. We note in retrospect that the segmented model outperformed the restricted one which implied that choice behavior was indeed different between the two groups. However, to avoid overburdening of the paper we do not report the results of the restricted model. The interested reader is referred to Ben-Elia (2007).
4.4. Modified specifications

During the course of the research we tested two variants to the main model's specification. The first variant tested the sensitivity to framing gains and losses in travel time compared to absolute travel times specified in the main model. As described in section 2, Prospect Theory suggests that in the domain of losses, decision makers are expected to reveal risk seeking behavior. The results obtained from both the longitudinal analysis and support this hypothesis and assert that real-time information could well contribute to more risk-seeking. Consequently, we decided to test the sensitivity to framing gains and losses in travel time and whether loss aversion (i.e. the tendency to cognitively weight losses more than equivalent gains) prevails in our route-choice data.

A main drawback to applications of Prospect Theory is the assumption of the 'true' reference point. Various definitions can be applied and different assumptions will result in different outcomes. As no previous knowledge on the reference point exists, we applied the most basic assumption – setting the reference point to be the mean travel time (on each route). For each route and each trial the differences in travel times are computed as the mean of the range less the obtained feedback. Recall that the feedback is a randomly drawn result based on a defined travel time range in each scenario, while the mean is predicted as the mean of the range with a difference of up to 5 minutes between consecutive trials. When this difference is positive (i.e. the mean is larger) the result is marked as a gain and when it is negative (i.e., the mean is smaller) as a loss. As the mean is now incorporated into the gain and loss variables, it can not be used for the normalization as it was in the main model and since we do not want to sacrifice any travel time variable, we decided to normalize the disturbance parameter of stickiness ($\sigma_{STICK}=1$). In retrospect this decision did not seem to have any major impact on the overall results.

A second variant to the model's specification replaced the range combinations with the variance of travel time on each route. Similarly to the range values, this depended only on the experiment design. Each route had two values depending on the range attribute in the relevant scenario (High/Low). No other changes were made compared to the main specification.

4.5 Estimation and identification

Model estimation was conducted using the BIOGEME software (Bierlaire, 2003) version 1.4 (2005) and applying the CFSQP algorithm (Lawrence et al., 1997) for the simulated log-likelihood optimization. The model was run with 1,000 Halton draws. Halton sequences (Halton, 1960) or Halton draws are designed to cover the integration space in a more uniform way and unlike other methods, induce a negative correlation over observations.
which guarantee a lower variance and therefore can significantly reduce the number of draws required (Train, 2000; Bhat, 2003). The model was estimated twice once with 500 draws and again with 1,000 draws. The differences between the two sets were negligible. The results presented here are for the set of 1,000 draws.

Identification is another key issue with discrete choice models in general and MXL in particular. The issue of identification is determining the set of restrictions to impose in order to obtain a unique set of estimated parameters. Guidelines to identify any MXL model and specifically with panel data are provided by Walker, (2001) and Walker et al., (2004). We applied these guidelines to verify that our models were correctly identified. Consequently random parameters for the alternative specific constants were included only in the utility function of the fast route. A proof of this is given by Ben-Elia (2007).

5. Results

Table 3 presents estimation results of the main model and the two modifications described in section 4.4. The table includes the estimated coefficients’ values, definition of each parameter and significance, as well as goodness of fit measures (final log-likelihood and rho-square). Section 5.1 presents the results of the main model’s estimation. Section 5.2 and 5.3 present the results of the two modified variants. Section 5.4 discusses goodness of fit measures.

5.1 Main model

The results of the main model’s estimation appear in the left most column of Table 3. The coefficient for stickiness ($\beta_{STICK}$) is positive and significant for both groups. It reflects the effect of recency, that is, if the last choice was FAST it is more likely that the current choice would be FAST. In addition, since the coefficient in the non-informed group is almost twice the size of the informed group (a significant difference), the effect of recency is greater over non-informed drivers. The parameter’s s.d coefficient ($\sigma_{STICK}$) for both groups is significant and different. The parameter for the non-informed group is larger which is to be expected as more trials, through a trial and error process, were devoted to exploration.

The coefficient for the cumulative weighted average ($\beta_{CWA}$), is significant only for the informed group. It reflects the effect of long term learning and memorization. The positive sign suggests that, in the long run, a tendency to choose FAST increased the likelihood of repeating the same choice. These results assert that informed participants decisions are influenced by having a longer perception horizon, whereas non-informed participants rely more heavily on their recent choices, (in our case the last one).
The coefficient of low level of experience ($\beta_{\text{low}}$) is negative and significant for both groups (the difference is not significant). This result suggests that lacking experience and knowledge on the travel time distributions, will result in greater exploration, i.e. trial and error learning process. Conversely, high experience ($\beta_{\text{high}}$) is only significant for non-informed participants. This asserts there is no real difference in the response of informed participants between trials 11 and 100, implying that the learning rate of informed respondents was, on average, much faster compared to non-informed participants who required more time to explore and learn. The results also concur with those obtained from the longitudinal analysis which showed between-group differences were not significant in the long run (see Table 2).

The coefficients of the routes' travel times ($\beta_{\text{timesf}}, \beta_{\text{times}}$) are negative and significant as expected for both groups (the difference is not significant). The disturbance parameters ($\sigma_{\text{timesf}}, \sigma_{\text{times}}$) are also quite similar (the difference is not significant) this result asserts that information did not alter the way participants perceived travel time feedbacks.

In contrast, travel time ranges appear to have affected mostly the informed group. The coefficients ($\beta_{\text{low-risk}}, \beta_{\text{fast-risky}}$) for the informed group are significant, whereas for the non-informed group the effect is smaller and significant only for Low-Risk. These differences are significant. The results indicate that for both groups the coefficient of Low-Risk is positive. This is expected since under Low-Risk the maximization rate (i.e. FAST shares) was significantly larger than under Fast & Safe (the reference value). The results assert the strength of the payoff variability effect, whereby the increases variability in Fast & Safe resulted in lower maximization rates compared to Low-Risk.

A different picture appears in the results of the informed group. The coefficient of Fast & Risky is significant and positive. Thus, consistent with the predictions of Prospect Theory, informed participants show a tendency for risk seeking. Conversely, without information, only the coefficient of Low-Risk is significant asserting that no real differences appear in the choice behavior between Fast & Safe and Fast & Risky. Again this is a demonstration of the robustness of the payoff variability effect. Furthermore, estimated s.d.’s of the scenario’s coefficient ($\sigma_{\text{fast-risky}}$) are larger in the informed group which suggests a greater degree of heterogeneity in risk attitudes. Between-group differences are significant only for the response under Fast & Risky. This result concurs with our previous results that heterogeneity in risk attitudes is more evident with the provision of real-time information.

The group scales are significant and the difference between the group scales is also significant. Since without information choice behavior is more uniform, the scale of the non-informed group is smaller. With information, the response is more heterogeneous, resulting
in a larger variance and larger scale coefficient (the scale's t-test is in relation to 1 and not 0 as is custom).

5.2 Model Variant #1:

The results of the estimation of the first modified specification - with gain and losses in travel times - appear in the middle column of Table 3. The results of the parameter estimates of stickiness, experience levels and travel time ranges were similar to the previous estimation. We will relate only to the main differences in respect to the main model.

The coefficient of the cumulative weighted average ($\beta_{CWA}$) is significant for both groups and the difference between groups is also significant. The coefficient is positive for both informed and non-informed participants. The result is similar to the main model for informed participants i.e. learning in the long run has a positive effect. Conversely, unlike the main model the result is also positive for non-informed participants. However, without information, the effect size is significantly smaller compared to the informed group. Thus, recency remains more dominant without information.

Regarding travel time feedbacks, in the informed group, both gains ($\beta_{GAINF}$) and losses ($\beta_{LOSSF}$) of travel time on the FAST route have significant coefficients. The coefficient for gains is positive indicating that a relative gain on the FAST route encourages its choice, whereas the opposite is observed for losses. Similar results were obtained for the non-informed group. The difference between groups is not significant, suggesting that information did not have an impact on the perception of gains or losses in travel time. This result is contrary to our expectations. Furthermore, a within group F-test does not show significant differences between the coefficients of gains and losses in either group. This result suggests only the absolute difference is relevant but not the sign.

The results of the SLOW route for non-informed participants show similar results to those of the FAST route and no significance within group differences were found. In addition the between group difference is also not significant. For the informed group, only gains ($\beta_{GAINS}$) have a significant (positive) coefficient, however the within group test was found to be not significant. Thus the result does not confirm that informed participants are more sensitive to gains than losses. We note that even though the loss coefficient ($\beta_{LOSSS}$) is not significant – for informed participants it is non-negative. This could indicate a weak tendency for risk seeking in the choice of the slow route. Since the specification treated the coefficient of Fast & Safe as the reference value we could not verify its effect on the choice of the SLOW route.

Evidently, we could not find major significant differences between gains and losses in travel time feedbacks. This result is contrary to what we expected that gains and losses
would have a stronger effect on informed participants that are subjected to descriptive information. Neither did there appear differences between the groups in relation to travel time differences, nor there appear any substantial within-group differences. The only remaining significant effect seems the overall difference in travel times.

5.3 Model Variant #2

The results for the second modified specification - travel time variances substituting the range combinations - are presented in the right most column of Table 3. The results of the parameter estimates of stickiness, experience levels and travel times were similar to the previous estimation. We will relate only to the main differences in respect to the main model.

The coefficient of the cumulative weighted average ($\beta_{CWA}$) is only significant for the informed group (the difference between the groups is significant). Both groups have positive coefficients. However the size of the coefficient of the informed group is larger. This result suggests that long term learning is more dominant with information, whereas without information recency is stronger. In this sense the results concur with all previous models.

Some interesting observations can be made regarding the coefficients of the routes’ travel time variance ($\beta_{VARF}, \beta_{VARS}$). The coefficients of both the SLOW and FAST route are significant for informed participants. For non-informed participants only the FAST route is significant and the size is smaller compared to the coefficient of the informed group. However, the differences between groups (on both routes) are not significant. The coefficients of both groups of the FAST route are negative. This suggests that taking in account all 300 observations the dominant attitude in the choice of the FAST route is risk aversion. This result does not contradict the risk-seeking effect, of the informed group, found in the main model which differentiated between the three scenarios. Thus under Low-Risk and Safe & Fast - risk aversion is more prominent, whereas under Fast & Risky risk-seeking is evident.

In the choice of the SLOW route, for the informed group, the coefficient is positive and significant. This result asserts that with information the SLOW route appears more attractive when it is associated with a higher travel time variance (as in Fast & Safe). On one hand this is a demonstration of the robustness of the payoff variability effect. On the other hand it also demonstrates the tendency of risk seeking in the choice of the SLOW route. It is probable that this is the mirror image of the results in Fast & Risky for the choice of the FAST route.
5.4 Goodness of fit

In terms of goodness of fit, the results show that overall the main model specification outperforms the other two variants. The 'Gain/Loss' variant has more estimated parameters and a higher log likelihood than both the initial model as well as the second variant. The second variant has the same number of parameters as the main model. However, the latter has a smaller log likelihood and therefore has a better goodness of fit. Probably this has to do with the differentiation between the three scenarios which was not accounted for in the second variant.

6. Discussion

ATIS development and deployment is an important step in making transport more efficient and promoting policies toward sustainable mobility. However, despite of numerous studies, there is still lack of sound knowledge as to the behavioral impacts real-time information provided by ATIS will have on drivers' route-choice decisions, which is essential for the design of ATIS and for maximizing its benefits. Behavioral decision research suggests choices are not necessarily made from a rational perspective as assumed by EUT and adapted by most travel behavior models. Moreover, different behavioral generalizations such as the 'Hot Stove', Prospect Theory and Reinforced Learning, imply deviations from rationality in different directions, making it more difficult to implement in a modeling framework. Based on this knowledge, a route-choice experiment was designed and carried out. It presented in a laboratory setting, a simple yet nonetheless non-trivial, route-choice decision problem when real-time information (about the range of travel times) is provided. The results, some of which were presented in the paper by Ben-Elia et al (2008), demonstrate the importance of both information and learning from experience in route-choice decisions. In addition, there is evidence to suggest that both Prospect Theory and Reinforced Learning seem to be relevant to the underlying decision process.

Empirical findings from behavioral decision research were shown to violate some the key assumptions held by EUT and inherited by RUM. However, this study has demonstrated how state-of-the art behavioral research and their insights can enhance RUM-based travel behavior choice models. The presented results illustrate how choice modeling can be improved and provide elaborate behavioral interpretations by incorporating leading theories from behavioral decision research.

Moreover, recent developments in discrete choice modeling, such as the MXL model, provide greater modeling flexibility, allowing the inclusion of behavioral factors such as learning, experience and information. These factors are rather difficult to specify correctly in the context of more traditional RUM formulations, mainly due to assumptions concerning...
independence across observations or very strict correlation structures. Consequently, a MXL discrete choice model (accounting for repeated responses in a panel) was applied to analyze the data. This model accounts for existing correlations in repeated choice sequences conducted by the same group of participants. In addition, it is flexible enough to accommodate lagged dependent variables which are important for modeling learning effects.

Two main results can be asserted from the model estimation. First, that learning and experience are essential for understanding choice decision behavior. Second, providing information results in an interaction with other behavioral factors resulting in a distinct choice behavior process that is significantly different from the type of behavior associated without the same information. Furthermore, the results assert the relevance of both Prospect Theory and Payoff Variability in explaining choice behavior.

Learning associated factors – in the short-run (recency) and in the long-run (memory) - were found to have a strong and significant effect on choice behavior. Without information, participants based their decisions mostly on recent outcomes whereas, learning in the long run was more constrained due to the additional effort devoted to exploration. This result is in line with the law of effect which states that the probability to repeat an action depends on its outcomes. In comparison, with information, participants based their choices both on short and long-run learning. This result is important since it signifies that providing information forms some kind of cognitive anchor which contributes to the memorization of previous outcomes. This suggests that credibility of an information system is a crucial issue in ATIS implementation and operation. Moreover, information expedites the learning process, increasing the rate of maximization (i.e. share of FAST) and reducing the rate of exploration (i.e. switching). Conversely, lack of information will result in slower adaptation and a longer learning process. These results also support our initial analysis (Ben-Elia et al., 2008) which indicated a similar direction of effects.

A second important factor involved in the behavioral decision process is the sensitivity to variability and the attitude to risk. The results suggest that, without information an increase in the variability of the choice environment as configured in the two scenarios of Fast & Safe and Fast & Risky, limited the ability to learn and resulted in a lower rate of maximization compared to the outcome 'with information'. This result sustains the robustness of the payoff variability effect as suggested by (Erev & Barron, 2005). In addition, the prevailing risk attitude without information was risk aversion. These findings also corroborate the results of (Avineri & Prashker, 2003). The payoff variability effect is also apparent with information albeit weaker. Conversely, real time information (about the range) amplified the sensitivity to variability. However, the prevailing attitude was risk seeking. This was demonstrated by the positive coefficient of the Fast & Risky scenario as well as the preference for the SLOW
route in the second variant. These results provide some support to the generalizations of Prospect Theory regarding risk-seeking in the domain of losses. In addition it seems that information also increases the heterogeneity in the response.

The results of Avineri & Prashker (2006) which assert that information promotes more risk aversion contradict our findings. This contradiction can be explained by the different settings of the experiment. Whereas Avineri & Prashker (2006) provided static information which related only to the mean, we treated information as real-time, dynamic and relating to variability through the range. Further the only comparable scenario with their study is Fast & Risky, while our findings regarding risk seeking tendencies result from a wider range of realistic settings. As different projections of information may well result in a very different allocation of traffic on the road network, this difference in drivers’ response to information content has important implications for the design of ATIS.

Unlike the differences in learning and risk attitudes, there seem to be no apparent between-group differences in the perception of the feedback apart for the expected negative sign for the travel time coefficients (main model and second variant). In the second variant, framing of the choice problem using gains and losses did not provide concrete results. In line with the generalization of Prospect Theory informed participants responded to the most part positively to gains and negatively to losses in travel times. Unexpectedly, this result was also significant for non-informed participants who were not faced with a description of the alternatives or their distributions. However, lack of both within-group and between-group significance makes it difficult to make valid conclusions about the relevance of framing. Clearly travel time differences have some significant impact which will require more research efforts in the future.

7. Conclusions

This study showed that incorporating behavioral insights within an econometric application provide promising results in terms of developing positive travel behavior models that are based more on empirical behavior and have less restricting behavioral assumptions. This approach enables to confirm and expand our initial findings (Ben-Elia et al., 2008) identifying the effect of learning in two different time scales as well as the interaction of real-time information with learning and risk attitudes.

The results of this study can provide some important policy implications for the application of ITS on road networks. First, behavior is to a large extent dependent on the content of information. Information provided on average travel time will result in a different distribution of traffic compared to information which also provides an indication to the
variability of travel times. Which projection is more beneficial, from a welfare point of view, is still an open question. If traffic control would like to shift travelers to alternative routes, it seems variability cues are important as they induce more heterogeneity in the response. Second, the main impact of information is on situations where conditions are unfamiliar and travelers lack long term experience with traffic conditions. Under recurring congestion, information will only reinforce the existing knowledge of the familiar environment. Conversely, if a traveler is new to the road, has limited knowledge, or if there is an incident on the road such as an accident or an unscheduled event such as a big sport match, information could well have a key role by expediting the learning process of the drivers on the network. However, given that these types of non-recurring congestion are also associated with greater levels of variability, the risk-seeking tendency of travelers might lead to exploration and result in capricious route switching. From the perspective of traffic control, there may well be situations that it is then even better not to provide information in real-time. Another point to note is that given that informed travelers tend to memorize previous outcomes, the credibility of the information system is important. If the predictions of the system prove over time to be false, travelers might learn to ignore it and the capital investments in sophisticated systems may be lost. Therefore more research is needed to ascertain the behavior of informed travelers on road networks characterized by non-recurrent congestion.

Several directions for future research can be indicated: First, regarding the experiment design, there is added value to expand the analysis by evaluating choice behavior on a sequence of travel time ranges and not only two points of reference (high versus low). This would provide a clearer picture of the interaction between information and travel time variability. It would be possible also to vary the mean travel times for the routes or even explore more than two routes as other simulative studies have done (Bogers et al., 2005; Bogers et al., 2007). Second, although the sample size is equivalent to the common sizes applied in behavioral research, a further investigation of the choice process should be conducted with a larger and more heterogeneous sample. This might increase both statistical power and reveal possible individual and socio-economic characteristics’ effects that were insignificant in our setting. Third, important insights could well gained by incorporating revealed preference (RP) i.e. field data, in addition to the stated preference data acquired in the laboratory, providing them with more external validity. A recent field study conducted in the Netherlands regarding peak-hour avoidance strategies shows that provision of real-time information can influence departure time choices (Ettema et al., 2008; Ben-Elia & Ettema, 2009). Fourth, as noted by Hensher & Greene, (2001), despite its flexibility MXL, especially with multiple disturbance parameters, is a difficult tool to implement
and caution has to be given to the estimation and simulation procedures. Finally, the findings of Selten et al., (2007) regarding foregone payoffs provide interesting directions to widen the scope of study by investigating the impacts of real-time information on driver behavior while accounting for interaction between drivers competing on the same road network and incorporating the effects of congestion. This would eliminate one drawback of the current study which was that each participant was oblivious to outside interference.

And one final note. At the end, the Cheshire Cat did provide Alice with rather accurate information about her two possible routes ("in THAT direction," the Cat said, waving its right paw round, `lives a Hatter: and in THAT direction,' waving the other paw, `lives a March Hare". Visit either you like: they're both mad.). Alice, who already saw hatters before (previous experience), took the risk of choosing the route leading to the March Hare (exploration - "perhaps as this is May it won't be raving mad"). However, Alice discovered that real-time information does not always guarantee a positive outcome. The experience at the Mad Tea Party made her learn that it was the stupidest tea-party she ever was at in all her life.

Acknowledgments
The authors thank Dr. Tomer Toledo from the Techion's Transportation Research Institute for assisting in model estimation and providing valuable comments. The comprehensive comments and suggestions of two anonymous reviewers is also very much appreciated.

References


Ben Akiva, M., Bolduc, D. (1996), Multinomial probit with a Logit kernel and a general parametric specification of the covariance structure, working paper, Massachusetts Institute of Technology.


Bhat, C. (2000), "Incorporating observed and unobserved heterogeneity in urban work mode choice modeling." Transportation Science, 34, 228-238.


Bierlaire, M. "BIOGEME: a free package for the estimation of discrete choice models." proceedings of the 3rd Swiss Transportation Research Conference, Ascona, Switzerland.


Cascetta, E., Papola, A. "Implicit availability/ perception logit models for route-choice in transportation networks." proceedings of the *8th World Conference on Transport Research*, Antwerp


Table 1: Hypothetical travel time scenarios

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Travel Time Ranges (minutes)</th>
<th>Route F – 25 min.</th>
<th>Route S – 30 min.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fast &amp; Safe</td>
<td>±5</td>
<td>±15</td>
<td></td>
</tr>
<tr>
<td>Fast &amp; Risky</td>
<td>±15</td>
<td>±5</td>
<td></td>
</tr>
<tr>
<td>Low-Risk</td>
<td>±5</td>
<td>±5</td>
<td></td>
</tr>
</tbody>
</table>

Table 2: ANOVA results (between and within group differences)

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Avg. FAST shares*</th>
<th>Avg. switching rate</th>
<th>Avg. Travel Time</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>info</td>
<td>No info</td>
<td>info</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fast &amp; Safe</td>
<td>0.844 (.165)</td>
<td>0.838 (.128)</td>
<td>0.148</td>
</tr>
<tr>
<td>Fast &amp; Risky</td>
<td>0.881 (.217)</td>
<td>0.817 (.128)</td>
<td>0.074</td>
</tr>
<tr>
<td>Low-Risk</td>
<td>0.966 (.228)</td>
<td>0.88 (.187)</td>
<td>*</td>
</tr>
<tr>
<td>F</td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>short run</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fast &amp; Safe</td>
<td>0.729 (.219)</td>
<td>0.636 (.173)</td>
<td>0.328</td>
</tr>
<tr>
<td>Fast &amp; Risky</td>
<td>0.85 (.228)</td>
<td>0.66 (.187)</td>
<td>*</td>
</tr>
<tr>
<td>Low-Risk</td>
<td>0.912 (.136)</td>
<td>0.672 (.196)</td>
<td>*</td>
</tr>
<tr>
<td>F</td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>long run</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fast &amp; Safe</td>
<td>0.873 (.168)</td>
<td>0.914 (.137)</td>
<td>0.106</td>
</tr>
<tr>
<td>Fast &amp; Risky</td>
<td>0.885 (.226)</td>
<td>0.886 (.154)</td>
<td>0.067</td>
</tr>
<tr>
<td>Low-Risk</td>
<td>0.986 (.046)</td>
<td>0.954 (.091)</td>
<td>0.037</td>
</tr>
<tr>
<td>F</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* standard deviations in parenthesis, (-) not sig., (*) sig at 0.05, (**) sig at 0.1.
Table 3: Estimation Results

<table>
<thead>
<tr>
<th>Coef.</th>
<th>Definition</th>
<th>Main Model</th>
<th>Variant 1: Gains/Losses</th>
<th>Variant 2</th>
<th>t**</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>With Information</td>
<td>Without Information</td>
<td>With Information</td>
<td>Without Information</td>
</tr>
<tr>
<td></td>
<td>Value</td>
<td>t</td>
<td>Value</td>
<td>t</td>
<td>Value</td>
</tr>
<tr>
<td>βSTICK</td>
<td>Stickiness</td>
<td>8.15</td>
<td>7.14</td>
<td>16.46</td>
<td>6.54</td>
</tr>
<tr>
<td>βCWA</td>
<td>Weighted average of past</td>
<td>4.68</td>
<td>5.1</td>
<td>-1.6</td>
<td>-1.65</td>
</tr>
<tr>
<td></td>
<td>choices</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>βHIGH</td>
<td>High experience (last 50</td>
<td>0.24</td>
<td>1.22</td>
<td>1.64</td>
<td>4.49</td>
</tr>
<tr>
<td></td>
<td>trials)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>βLOW</td>
<td>Low experience (first 10</td>
<td>-1.24</td>
<td>-3.41</td>
<td>-2.3</td>
<td>-4.42</td>
</tr>
<tr>
<td></td>
<td>trials)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>βTIMEF</td>
<td>Travel time feedback -</td>
<td>-0.13</td>
<td>-4.48</td>
<td>-0.24</td>
<td>-4.86</td>
</tr>
<tr>
<td></td>
<td>FAST</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>βTIMES</td>
<td>Travel time feedback -</td>
<td>-0.09</td>
<td>-5.88</td>
<td>-0.21</td>
<td>-10.48</td>
</tr>
<tr>
<td></td>
<td>SLOW</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>βMEAN</td>
<td>Mean travel time (normalized</td>
<td>1</td>
<td>1</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>to 1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>βGAINF</td>
<td>Travel time gains on FAST</td>
<td></td>
<td></td>
<td></td>
<td>0.07</td>
</tr>
<tr>
<td>βLOSSF</td>
<td>Travel time losses on</td>
<td></td>
<td></td>
<td></td>
<td>-0.08</td>
</tr>
<tr>
<td></td>
<td>FAST</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>βGAINS</td>
<td>Travel time gains on</td>
<td></td>
<td></td>
<td></td>
<td>0.16</td>
</tr>
<tr>
<td></td>
<td>SLOW</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>βLOSSS</td>
<td>Travel time losses on</td>
<td></td>
<td></td>
<td></td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>SLOW</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>βFAST-</td>
<td>FAST &amp; Risky scenario</td>
<td>1.75</td>
<td>3.6</td>
<td>0.13</td>
<td>0.51</td>
</tr>
<tr>
<td>RISKY</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>βLOW-RISK</td>
<td>Low Risk scenario</td>
<td>4.73</td>
<td>4.3</td>
<td>0.76</td>
<td>2.33</td>
</tr>
<tr>
<td>βVARF</td>
<td>Variance on FAST</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>βVARS</td>
<td>Variance on SLOW</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>σSTICK</td>
<td>s.d Stickiness</td>
<td>2.14</td>
<td>4.51</td>
<td>3.83</td>
<td>4.88</td>
</tr>
</tbody>
</table>

* Table 3: Estimation Results
<table>
<thead>
<tr>
<th>Coef.</th>
<th>Definition</th>
<th>Main Model</th>
<th>Variant 1: Gains/Losses</th>
<th>Variant 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>With Information</td>
<td>Without Information</td>
<td>t**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Value</td>
<td>t</td>
<td>Value</td>
</tr>
<tr>
<td>σCWA</td>
<td>s.d Weighted average of past choices</td>
<td>2.4</td>
<td>3.97</td>
<td>1.67</td>
</tr>
<tr>
<td>σTIMEF</td>
<td>s.d Travel time FAST</td>
<td>0</td>
<td>0.34</td>
<td>0.06</td>
</tr>
<tr>
<td>σTIMES</td>
<td>s.d Travel time SLOW</td>
<td>0.06</td>
<td>3.88</td>
<td>0.08</td>
</tr>
<tr>
<td>σGAINF</td>
<td>s.d for gains on FAST</td>
<td>0.06</td>
<td>3.88</td>
<td>0.08</td>
</tr>
<tr>
<td>σGAINS</td>
<td>s.d for gains on SLOW</td>
<td>2.79</td>
<td>4.05</td>
<td>0.88</td>
</tr>
<tr>
<td>σLOSSF</td>
<td>s.d for losses on FAST</td>
<td>3.36</td>
<td>4.23</td>
<td>2.32</td>
</tr>
<tr>
<td>σLOSSS</td>
<td>s.d for losses on SLOW</td>
<td>0.02</td>
<td>1.36</td>
<td>0.12</td>
</tr>
<tr>
<td>σVARF</td>
<td>s.d Variance on FAST</td>
<td>3.36</td>
<td>4.23</td>
<td>2.32</td>
</tr>
<tr>
<td>σVARS</td>
<td>s.d Variance on SLOW</td>
<td>0.06</td>
<td>3.06</td>
<td>0.04</td>
</tr>
<tr>
<td>μ</td>
<td>Group scale***</td>
<td>0.6</td>
<td>-4.01</td>
<td>0.48</td>
</tr>
</tbody>
</table>

| No. of Halton draws: | 1,000 | 1,000 | 1,000 |
| No. of parameters:  | 30 | 30 | 30 |
| No. of observations:| 14,553 | 14,553 | 14,553 |
| No. of individuals: | 49 | 49 | 49 |
| Null log-likelihood:| -10,087.40 | -10,087.40 | -10,087.40 |
| Final log-likelihood:| -3,512.90 | -3,512.90 | -3,543.30 |
| Rho-square:         | 0.65 | 0.65 | 0.65 |
| Adjusted Rho-square:| 0.65 | 0.65 | 0.65 |

*t values in *italics* – not significant at 0.05 level. ** between-groups test. ** tested is in relation to $H_0: \mu=1$. 32
Figure 1: Longitudinal results