UNIVERSITÀ DEGLI STUDI DI NAPOLI "FEDERICO II"
FACOLTÀ DI INGEGNERIA

Dipartimento di Ingegneria dei Trasporti "L.Tocchetti"

DOTTORATO DI RICERCA IN
INGEGNERIA DEI SISTEMI IDRAULICI, DI TRASPORTO E TERRITORIALI

INDIRIZZO INFRASTRUTTURE VIARIE E SISTEMI DI TRASPORTO
XXI CICLO

Coordinatore: Prof. Ing. Bruno MONTELLA
Coordinatore di Indirizzo: Prof. Ing. Renato LAMBERTI

TESI DI DOTTORATO

ANALYTICAL TOOLS FOR ATIS
(STRUMENTI ANALITICI PER APPLICAZIONI ATIS)

Tutor:
Prof. Ing: Gennaro Nicola BIFULCO

Dottoranda:
ing. Roberta DI PACE

Anno Accademico 2007 - 2008

A chi mi ha trasmesso il Sogno, mio padre, a Pasquale la cui forza e determinazione sono per me un continuo stimolo, a mia madre e Bruno che con Amore mi sono sempre vicini…

Table of Contents

"… alla ricerca di un ago in un pagliaio, in una sfida in cui l'unico vero *avversario* sei te stesso!"

1 BACKGROUND

1.1 INTRODUCTION: ATIS RESEARCH AREAS

The thesis focuses on analytical tools for ATIS (Advanced Traveller Information Systems) implementations. ATIS implementations belong to the wider area of Intelligent Transportation Systems (ITS) applications. ITS can be described as "systems consisting of electronics, communication information processing used singly or integrated to improve efficiency or safety of surface transportation" (Tindemans et al., 2003). ITS can influence several parts of the transportation network at different levels, including infrastructure applications, vehicle applications or a combination of both.

The research areas in ITS can be grouped in different fields, main of these are:

- **ADAS** (Advanced Driver Assistance Systems); supporting the drivers in controlling their vehicles;
- **ATIS** (Advanced Traveller Information Systems); aiming at dispatching to travellers information about traffic conditions, accident delays, road works etcetera in real time so that their travel choices could be more efficient and aimed at the maximization of perceived utilities;
- **ATMS** (Advanced Traveller Management Systems); aimed at controlling and managing the transportation system by dynamically setting-up main network rules and parameters;

Examples of ADAS are Intelligent Speed Adaptation (ISA) and Adaptive Cruise Control (ACC): ISAs are in-car systems that assist or control vehicles in order to adjust its speed according to actual speed limits; if a car is equipped with ISA and it runs an a link where the adopted speed is 80 Km/h, automatically the car slows adjusting the speed by getting a signal from infrastructural devices or by using the digital map in the vehicle on which the speed limited has been loaded. ACCs are in-vehicle systems that automatically regulate vehicles speed by adopting the programmed desired speed if in free-flow conditions or by adapting the speed to the one of the leading vehicle if in a car following conditions.

It worths briefly discussing about the main differences between two very popular fields of research in the ITS areas, ATMS and ATIS. Both of these kind of applications, in fact, are often viewed as different tools for reaching a common aim, which is to control traffic systems. In Figure 1 the generic scheme of a control system is described.

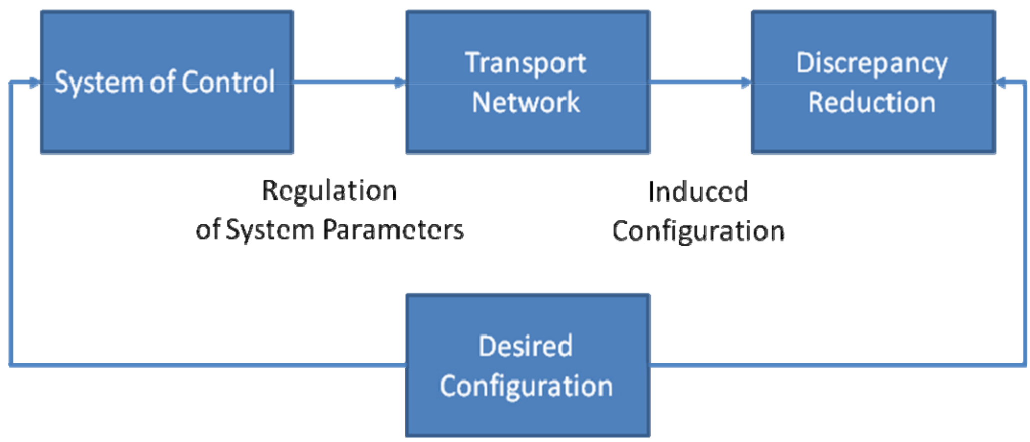


Figure 1: Logical Scheme of Control Systems

Control systems are mainly used to reduce the discrepancy between the desired traffic patterns on a network and the actual patterns. This straightfully applies to ATMS. In fact, on the base of the given objectives (network capacity maximization, travel times minimization etc.) ATMS change (in a dynamic and adaptive way) the rules of the transport network (e.g. traffic lights regulation and adaptive synchronization). By doing this, ATMS directly effects traffic flows propagation and in turn, network performances. The network performances induced by ATMS affect the day-to-day travellers' learning mechanism of network travel costs and so ATMS have side-effects also on travellers' choices. These side-effects are as much predictable and prone to travel-choice-level control application of ATMS as more the induced daily network performances profiles became similar over days. In other terms, ATMS have direct influence on traffic propagation and indirect (arguable) effect on travel choices.

Differently from ATMS, ATIS directly influence travellers' travel choices, of course, if travel choices (e.g. route choices) change, also flows propagation (and in turn network performances) result to be modified. This has induced some analysts to classify also ATIS as traffic control tools, even if this still is controversial (see Bifulco et al., 2007).

In the ATIS research area we can identify two kinds of problem: the first one is technological[1], the second one is related to the design of the information to be dispatched to drivers and to the effects that ATIS can have both on individuals drivers' behaviour and on traffic patterns . The research on ATIS dates back to the 1950s and has evolved over the past 40 years. From the technological point of view it's possible to identify two main periods. The first generation was finalized to make travellers aware in a case of not recurring congestion when special events or accidents happen and to prevent or mitigate the disruption of flow propagation abilities. In particular, starting from late 1960s and early 1970s, variable message signal (VMS) and highway advisory radio (HAR) have been world wide implemented as 1st generation systems. The second generation has been aimed to improve the personalization/customization of the information; in particular, the advances in technologies have allowed the systems to implement application such as routing, way finding, etcetera.

[1] The technological –side of ATIS is strictly related to researches in Intelligent and Communication Technologies (*ICT*) which actually are enabling tools for gathering and dispatching to drivers the information on the prevailing traffic patterns.

In practice the 2nd generation of ATIS is founded from a technological point of view on navigation application. Characteristic 2nd generation communication devices include mobile phones, GPS navigators, cable television, information kiosks, internet, etcetera. The information provided by technologies of 2nd generation embed several features: interactive user interfaces, highly mapping characteristics, management of graphical network datasets, automatic positioning by GPS, customization of path searching, multimodal information, tracing and dynamic route guidance.

From the not technological point of view, the main questions regard:

- if the dispatched information should be only based on instantaneous prevailing traffic conditions or on (short/ medium term) traffic predictions;
- if the dispatched information should be properly (and usefully) modified in order to induce desired traffic patterns;
- which is the effect of different levels of information accuracy/reliability on drivers' reaction ad on traffic patterns.

Several studies have been carried out about the ATIS effect on traffic networks, but a comprehensive about the information impact as well as about the users' reactions to the information still are the main unsolved (or at least, controversial) problems. Many researchers have studied the information effect on the travellers' route choices in case of monomodal networks and both in recurrent traffic conditions (Mahmassani and Jayakrishnan, 1991; Arnott et al., 1991; Ben- Akiva et al., 1991; Khattak et al., 1999; Polydoropoulou and Ben- Akiva, 1999) and not recurrent ones (Khattak et al., 1993; Abdel- Aty, 1998) and some of these studies have been specifically focused on the effect of information accuracy/reliability on drivers' behaviour.

It is worth noting that ATIS are becoming an emerging issue, also because of the technologies of the 2nd generation represent an appealing commercial facility per se. However, the expected diffusion of ATIS application on large scale and across large demand segments suggests a deeper analysis on ATIS related issues. One of the key points that animates these controversial analysis depends on the fact that the impact of (possibly voluntary) inaccurate/unreliable information on travellers' behaviours still requires a deeper analysis. This work will focuses on these kinds of analysis, particularly on these related to the effect of information accuracy.

1.2 A GENERAL CONCEPTUAL FRAMEWORK FOR ATIS ANALYSIS

Several kinds of information can be dispatched to travelers: the length of queue, the estimated travel times by ATIS, the estimated shortest route. When the information provides travelers with the networks performances is defined as "descriptive" while in the case in which the information suggests to the travelers the behavior (or the choice) to be adopted, the information is defined as "prescriptive". Information can be classified also on the base of its accuracy level, (information can be considerate as much accurate as much it is consistent with the network performances) and on the base of its levels of detail (Dell'Orco et al., 2006). In this work we will only consider the inaccuracy of the information as previously defined with respect to actual network performances.

The information kinds considered in carrying out this work are both the descriptive and prescriptive ones. In case of descriptive information we will only consider that estimated travel times are dispatched to travelers.

The suggested information can have significant effects on travelers' choices and particularly it induces a different flow propagation and also a different configuration of network parameters (speed, travel times, queues, etcetera).

An interesting contribute for ATIS modeling is to analyze travelers' compliance[2] both at a "disaggregate level" (the choice behavior of a given decision-maker) and at an "aggregate level" (the cumulate effect on the network of all choices made by

[2] Compliance can be defined as the attitude of the travelers in behaving according to ATIS information and/or suggestions.

all decision makers, taking into account the presence in traffic network of congestion phenomena).

At a disaggregate level, under repeated choices (consistently with a *day to day dynamics* [3]process), the traveler can update its expectations in two different ways, depending on if the context of choice is *with information* or *without information*.

In case of absence of information (Figure 2) the day-to-day dynamic updating of the perceived utilities is based only on the network performances. Every day the traveller makes his/her choice on the base of the experienced travel times at previous day and depending (also) on the new experienced travel times he/she updates his/her expectations.

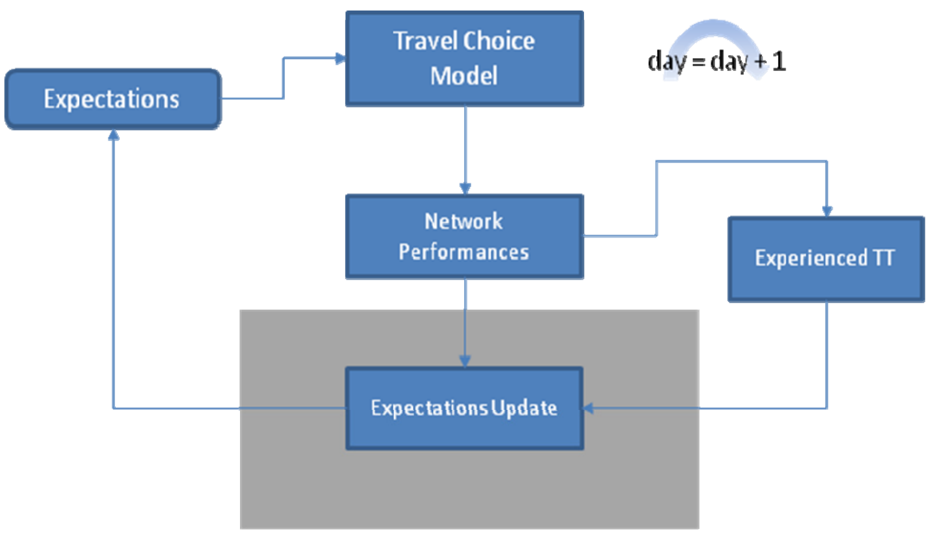


Figure 2: Day- to- day dynamics without information

[3]Transportation systems can be represented by considering two different approaches for the dynamically descriptions of the variables: the first one is the *intra-period* or *within day* approach the second one is the *inter-period or day to day* approach. In the first case flow propagation and the flow simulation are the main goals (simulation of queues is one of the most important aspects) in the second case the system evolves over time (successive reference periods can be considered) and the mechanism of decision makers updating in choosing is the main focus of the model.

In a context with information the route choice is influenced not only by the network performances, but also by the provided information (see Figure 3) and particularly by its credibility (accuracy).

Differently than the context without information, in a context with information, every day the traveller updates his/her expectations not only on the base of the experienced travel times, but also by considering the information accuracy level; at the same time the expectations updating influences the travel choices made by travellers and the network performances, on the base of which information has been designed the network. Obviously it is expected that an accurate ATIS is able to reduce the more is possible the discrepancy between the designed (and dispatched) information and the performances that actually are exhibited by the network.

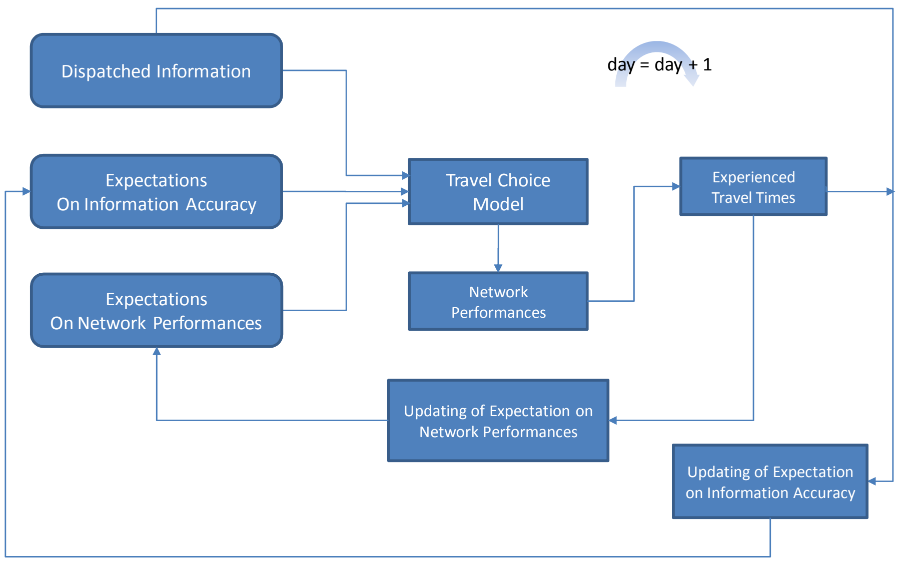


Figure 3: Day- to- day dynamics with exogenous information

A more detailed view of a simulation model in absence of information is described such as in Figure 4 while in Figure 5 the presence of an ATIS is also considered.

In a context without information the simulation model is composed by:

- **The supply model, in turn composed by**:
    - The network flow propagation model;
    - The link performance model/ congestion model;
    - The path performance model/ cost model;
- **The demand model, in turn composed by**:
    - The travel choice model;
    - The utility model;
- **The assignment model** (interaction between Supply and Demand).

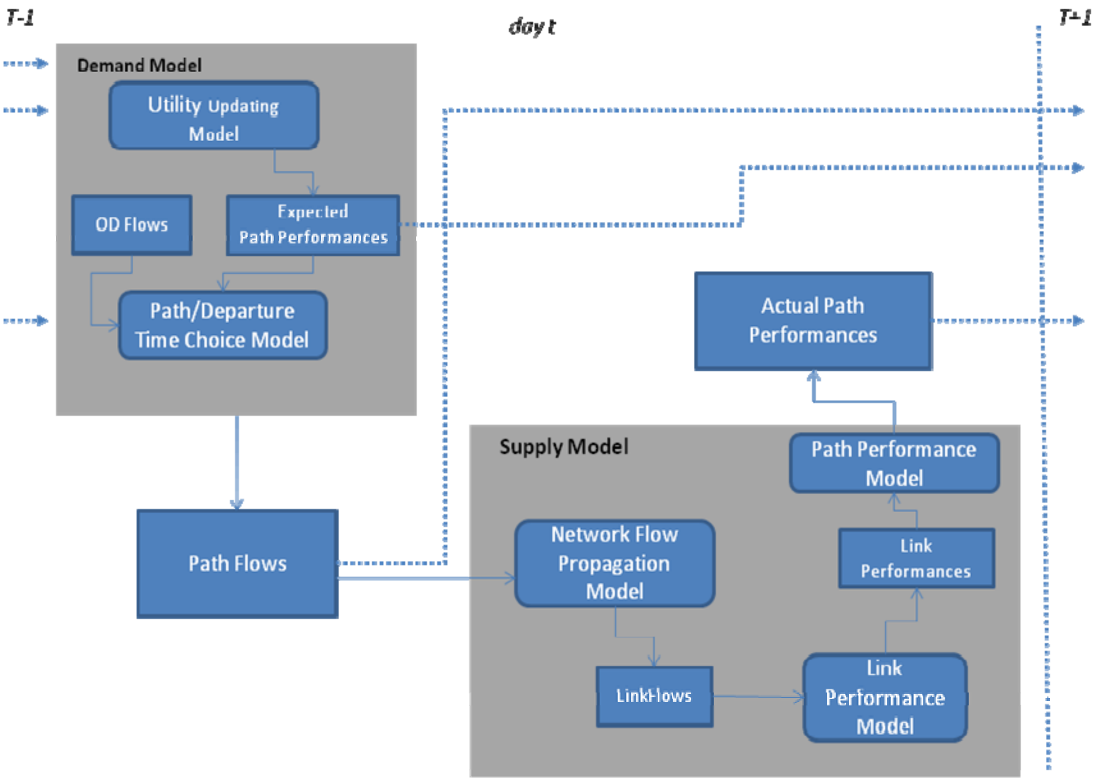


Figure 4: Dynamic Process Assignment model (without Information)

If an ATIS is on place in addition to the previous models we have also to consider a model for the simulation of travellers' reaction to the information and an other for the simulation of the variability of user's tendency to follow/ reject the suggestions received by the ATIS. Moreover, to complete the theoretical framework for the information design, it's necessary to introduce the *Information Design Model.* In Figure 5 the proposed model describe the interactions among demand model, supply model and the model for providing the information. In this case the suggested information has been supposed to be available both pre-trip (information available via internet or television) and en-route (information available via variable message signs, mobile phone, etc).

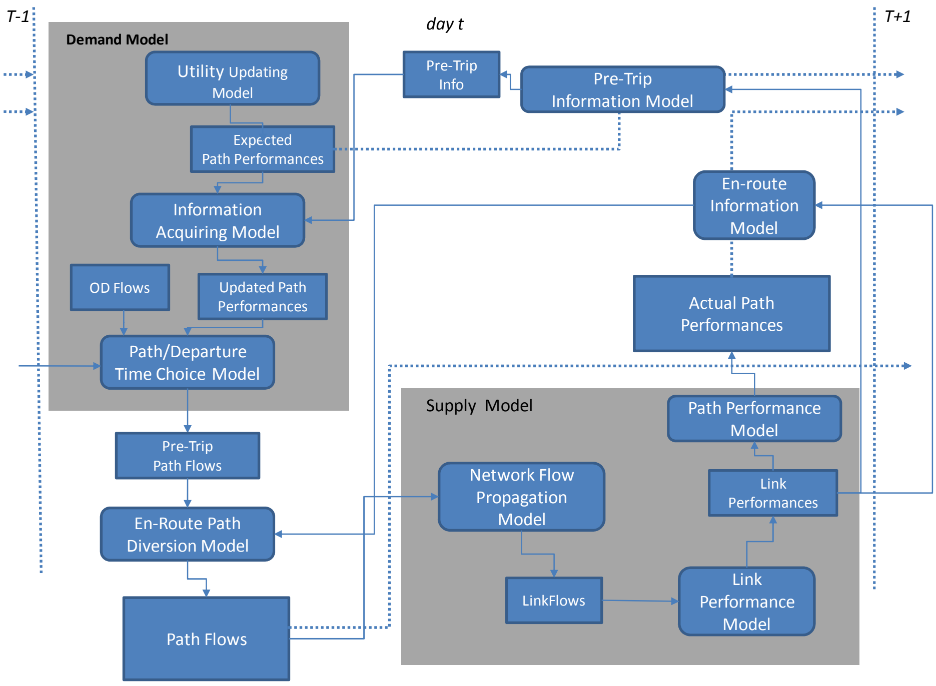


Figure 5: Dynamic Process Assignment model (with Pre-Trip and En-Route Information)

1.3 RESEARCH GOALS AND METHODOLOGY

With reference to the general framework previously presented, the main research area that has been identified for this work has been to better understand how travelers update their choice utilities and, in particular, which is the role in such an updating mechanism of the travelers' compliance with ATIS information.

This work has:

1. investigate the effect of kind of the information (prescriptive/descriptive) on travellers' compliance, given an information inaccuracy level;
2. analyze the effect of different inaccuracy levels on drivers' tendency to chose the best route (the shortest route), to chose the most reliable route (the route with the smallest standard deviation of the day-to-day profile of actual travel-time) and to be compliant;
3. build a model specifically devoted to travellers' compliance modelling.

All the research has been carried out in five step:

1. literature review;
2. data capture by Stated Preferences investigation methodology;
3. preliminary statistical aggregate analysis of gathered data;
4. specification, calibration and validation of a disaggregate model;
5. conclusions and reflections.

The previously described research methodology is here suggested as a reference pattern in carrying out analysis on travellers' behaviour in presence of ATIS systems.

The literature review has mainly concerned the analysis on ATIS effect on drivers' behavior and the analysis of ATIS effect on networks. A critical review has been also performed on the role assigned in the literature to the travelers' compliance with ATIS.

In the data gathering phase, the methodology for data acquisition has been fixed, in particular the implementation advantages of different data gathering approaches have to be analyzed. The technology of the survey has been identified by choosing to carry out it by means of an internet tool. Obviously, need of carefully design the experiment has also been fully addressed and the experiment details fully developed. The obtained data have been submitted (step 3) to some preliminary non parametric statistical tests, aimed at verifying the suitability of more detailed analyses and modeling efforts. The specification, calibration and validation of a disaggregate compliance model (where the information accuracy plays a crucial role) has been carried out; the discrete choice theory has been a pre-defined choice in our approach, but the type of model specification has been carefully identified. The modeling structure has been specified with reference to the two main alternative possible approaches: holding and switching modeling architecture.

Explaining variables have been carefully selected, also by using the results of the preliminary disaggregate tests. The calibrated models have been verified against their ability to reproduce the data sample both in case of considering or not considering the inherently different dynamic characteristics of the switching and the holding approach. Finally, the obtained results have been discussed and some perspectives for future analysis and enhancements introduced.

1.4 OUTLINE OF THIS THESIS

The research here described is the result of a three years work. In Figure 6 a brief overview of the thesis has been described; it is aimed to present in detail the research phases as described in the previous paragraph.

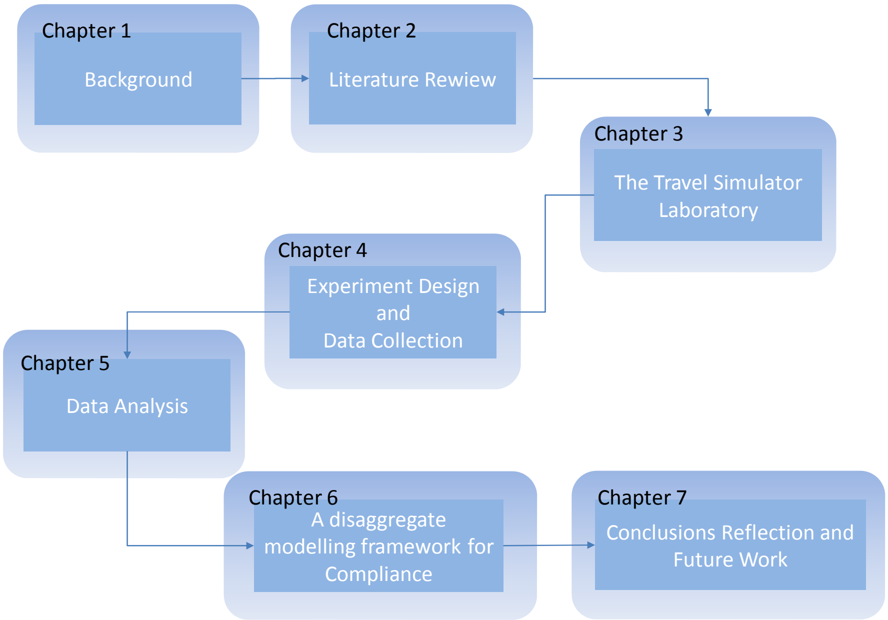


Figure 6: Thesis Overview

Chapters 1 and 2 are related to the first year of work, which has been mainly devoted to understand the key issues in ATIS modelling and to fix the reference scientific context of the work. In particular, the critical analysis of chapter two has been also oriented to identify the most promising research approaches and directions for addressing the thesis objectives. Chapter 3 and 4 are related to the work mainly carried out during the second year of the thesis. Part of this work has been done at the Delft University of Technology within an Erasmus exchange. In particular, in Chapter 3 has been described the job of understanding the capabilities of the Travel Simulator Laboratory (The TSL) kindly made available by the Technical University of Delft, identified during the work described in Chapter 2 as one of the most promising operative tools for the experiments required by our research. The functionalities of the TSL have been adapted to the requirements of our research and chapter 4 describes the set-up of the experiment.

The whole experiment design has been also described and theoretically justified in Chapter 4, as well as the run of the survey and the obtained experimental data base.

The third year of research has been devoted to finalize the research and to exploit the gathered data; the work that has been done is described in Chapter 5 to 7.

The survey data has been submitted to two kind of analysis: an *aggregate analysis* and a *disaggregate analysis*. The aggregate analysis is described in Chapter 5, and the disaggregate one is described in Chapter 6. All conclusions, reflection and considerations on future work are described in Chapter 7.

2 LITERATURE REVIEW

Not negligible efforts have been made in the scientific literature for understanding and modelling the travellers' response to information. In the following a brief state of the art is discussed. Impacts of ATIS can be analyzed with reference to their network effects by considering the effects on individual drivers' behaviour as depicted in Figure 7.

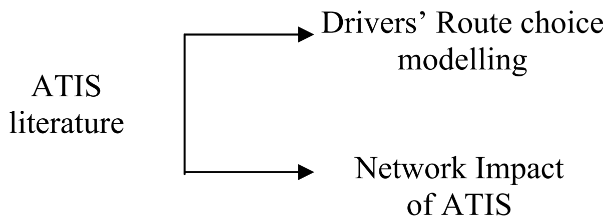


Figure 7: Overview on the ATIS literature

2.1 DRIVERS' ROUTE CHOICE MODELLING

The aim of this area of research is the drivers' behaviour modelling in terms of route choices in ATIS contexts. ATIS dispatches information that can modify the way the travellers behave with respect to route choices; for this reason is worth analysing how route choices are generally modelled in absence of information, then the presence of information is considerate as well as how this modifies the modelling approaches. In fact modelling aspects in absence of information still are valid in presence of information, even if they have to be modified and/or integrated in order to take into account additional phenomena.

2.1.1 Drivers' behaviour without Information

The drivers' route choice is directly affected by the traffic network flow performances and by travel costs. Several researches have been addressed to understand and modelling the mechanism of route choices. According to the literature review it's possible to identify different approaches.

In particular two main approaches can be distinguished: a statistical approach and a disaggregate choice analysis method. The statistical approach (Huchingson, McNees et al. 1977; Ratcliffe, 1972; Heathington, Worral et al., 1971; Wachs, 1967; Pederson, 1998) investigates the relationship between drivers' behaviour and some factors; in particular for selecting routes, drivers' attitude to divert from straight routes in order to avoid delay is investigated by adopting statistical methods. The disaggregate (i.e. individual- level) choices analysis methods is based on the application of discrete choice (mainly random utility) models. The route choice of (potentially) each traveller can be explained by his/her perceived utility in making each available choice and the probability that each choice is made can be computed. Several modelling frameworks have been developed within the approach of the discrete choice models.

Route choices inherently are (at least in great part) repeated (over days) choices phenomena. Different approaches have been suggested for taking into account the learning updating mechanism which is typical of repeated choices and while should take into account mechanisms such as experience and learning.

Two main approaches can be identified: the first one is the weighted average approach (Horowitz, 1984) in which the perception of travel times is based on the weighted average of travel times in previous time period); the second one is the adaptive expectation approach (Cascetta and Cantarella, 1991; Van der Mede and Berkum, 1993), in which it's assumed that the travellers update their perception of the utilities of choice alternatives on the basis of both the actual travel times and the perceived utilities of the previous day. Another model for the analysis of the learning process, has been proposed by Erev et al. (1999); they have proposed a reinforcement learning model (*REL*) in which every day the propensity to choose the given route can be updated on the base of the propensity value at previous day and on the base of the number of times in which the given route has been chosen and the average of the payoff[4]. Of course, another trivial approach is to assume that learning updating mechanism are negligible or in any cases to do not consider choice dynamics.

In summary three models groups could be identified: *cross sectional models* in which the utility values doesn't exhibit any dependency by a given time *t* (Bogers et al. 2006), *weighted model* in which attributes of the current utility are modelled as function of weighted values of previous attributes (Horowitz, 1984) and finally explicit dynamic models characterized by *the adaptive expectation approach* in which the current utility is updated by considering the utility value at previous day and the experienced attribute values (Cascetta and Cantarella, 1991; Cascetta and Cantarella, 1993) or the current propensity to chosen a given route is update on the base of the attributes at previous days (Erev et al., 1999).

Some studies have also formalized the choice paradigm by going beyond the classical utility theory. This has lead to some alternatives frameworks, all of them aimed to address more explicitly the fact that decision makers exhibit different attitudes under unreliable choice context, so that the utility maximization paradigm could result to be inadequate. In other terms, in case of travel times uncertainty (unreliability) it could be judged to be unrealistic that users make their choices in a perfect rational way.

In the last few decades the study of strategies and decision-making has been also the object of the Game Theory (Nash, 1950), and of the Prospect Theory (Khaneman and Tversky ,1979) both of the oriented to analyze the people attitude under risk and uncertainty context of choice.

The Prospect Theory allows to describe how people make choices in situations where they have to decide between alternatives that involve risk (e.g. in financial

[4] Is the difference between the travel time of the estimated reference route (RP) and the observed travel time.

decisions). At the base of the theory there is an evaluation of the potential *losses* and *gains* as perceived by individual. In the original formulation the term *prospect*, referred to a lottery, is related to the relationship between the probability of losses and gains and the phenomena of *risk aversion* and *risk seeking*. The theory consists of two stages, *editing* and *evaluation*. In the first, people decide which outcomes they see as basically identical and they set a *reference point;* the domain of the gains and the losses is defined by establishing that lower outcomes can be considerate as *losses* and larger outcomes can be considerate as *gains*. In the second phase people compute a value (utility), based on the potential outcomes and their respective probabilities, and then choose the alternative having a higher utility.

By referring to the expected utility theory the utility is obtained by summation of the outcome multiplied by its respective probability: $EU(\underline{x},\underline{p}) = \sum_i p_i u(x_i)$. The formula that Kahneman and Tversky assume for the evaluation phase the utility is somewhat different and is computed as $U = w(p_1)v(x_1) + w(p_2)v(x_2)$ where $x_1, x_2 ...$ are the potential outcomes and $p_1, p_2 ...$ their respective probabilities, *v* is a so-called *value function* that assigns a value to an outcome.

An example of value function is depicted in Figure 8, the function passes through the *reference point,* is s-shaped and, as its asymmetry implies in the depicted case, there is a bigger impact of losses than of gains (loss aversion). In particular the value function is assumed to be concave in gains and convex in losses, pattern which is consistent with experimental evidences obtained by analyzing the risk sensitive preferences.

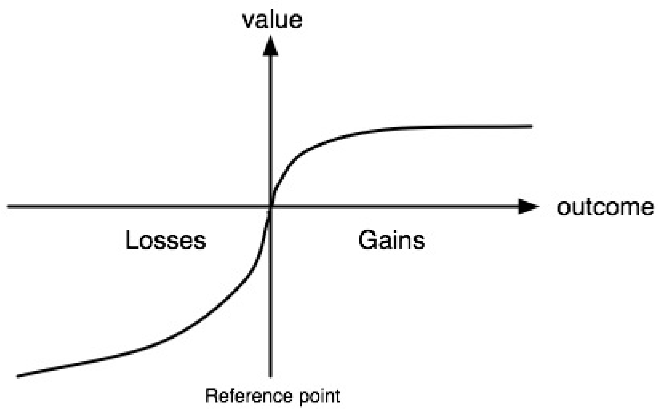


Figure 8: Example of value function based on Kahneman and Tversky

The function *w* is called *weighting probability* function; as is depicted in Figure 9 it expresses that people tend to overreact to small probability events and to under-react to medium and large probabilities.

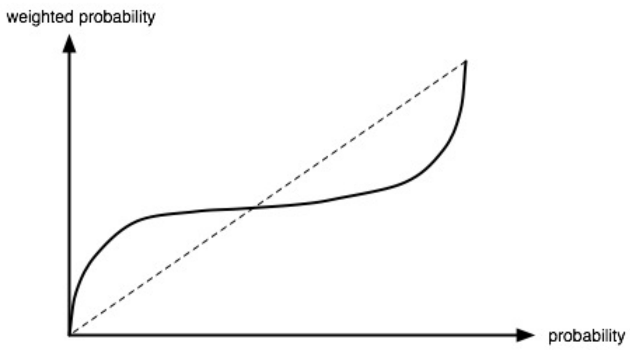


Figure 9: A typical weighting function in Prospect Theory

Tversky and Kahneman later developed a new version of prospect theory that uses cumulative rather than separable decision weights; the *Cumulative Prospect Theory* (*CPT*) is an improvement of the Prospect Theory. This version applies the cumulative functional separately to gains and to losses. In case of CPT the utility is computed as $U(\underline{x},\underline{p}) = U^{+}(\underline{x},\underline{p}) + U^{-}(\underline{x},\underline{p})$, considering separately gains (U^{+}) and losses (U^{-}). The main contribution of the Prospect Theory and the Cumulative Prospect theory is the introduction of the concepts of risk averse and risk seeking. In a context of travel choices, risk aversion is referred to the cases where, for instance, many routes have the same average expectations and the users more often choose the route with more reliable travel time (say the one with a lower dispersion with respect to its average travel time); at the contrary, risk seeking is referred to the cases when the users more often chooses the route with smaller reliability (higher variability) thus aiming at great gains but also to lose more with respect to the conservative behaviour. The unreliability effect on users' behaviour has been also analyzed by several authors in particular by Avineri and Praskner (2004). They have shown that in case of high payoff variability the tendency of the users' choices is to move toward random choice; unreliability of travel times also increases the travellers' inability to perceive the actual differences among the travel times.

An important study about the users' attitude to risk was carried out by Kastikopoulos et al. (2002). For every route *(j)* some variables are defined in particular: the range *(r_j)* defined as the absolute value difference between the maximum and the minimum values (across days) of travel times and the expected travel times *(e_j),* in turn defined as the average between the minimum value and the maximum value of travel times. In a network one of the routes is defined as the *reference route* (the main route – say the more straight- that connects the considered origin and destination), and the others are defined as the alternative routes. According to these preliminary considerations, the estimated travel times are computed as: $ETT = e + \lambda r$ where λ is a random variable distributed like a Normal function of distribution with average μ and variance σ^2. The diversion probability towards alternative route is computed as *P(div) = P[(e+r) <c]=*

P[λ <(c-e)/r]; it is also interpreted as the probability to exhibit a *risk seeking* attitude. The model proposes that drivers choose the route with the smallest value of *ETT*. Therefore when *c>e, P(div)* is decreasing in *r* (i.e., in the domains of *gains*) and *P(div)* is increasing in *r* when *c<e* (i.e. on the domains of *losses*).

In previous case the travel time of reference route was considered not distributed (for instance the travel times is equal to 33 minutes, moreover for the alternative route on the base of the minimum and maximum values of travel times parameters *e* and *r* can be calculated), but if for every route can be calculated the values of *e* and *r*, to be more precise to both routes are associated the minimum and the maximum values of travel times, consequently the value of probability diversion can be computed as following:

$P(div) = P[\lambda < (e_R - e_A)/(r_R - r_A),\ r_R > r_A]$;

$P(div) = P[\lambda > (e_R - e_A)/(r_R - r_A),\ r_R < r_A]$;

the computed probabilities refer to the defined conditions of risk as described in Figure 10.

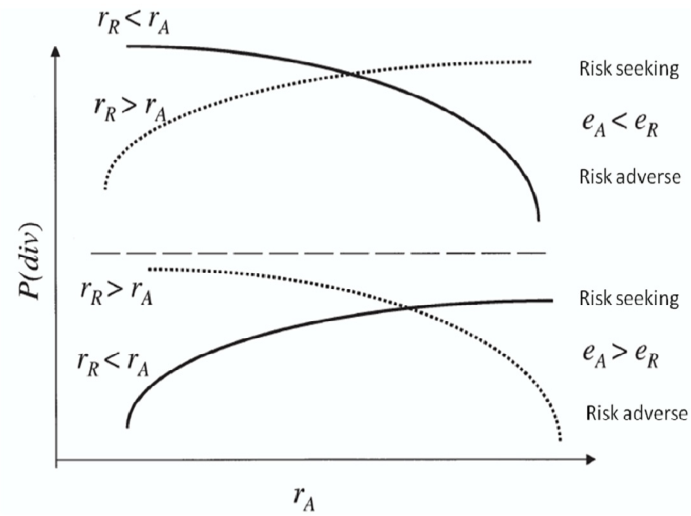


Figure 10:Computation of the Diversion Probability

In several cases the travel time unreliability on route choice is incorporate in the modelling (Small et al., 1982; Bogers et al.,2004).

Arentze and Timmermans (2005) have proposed a Bayesian belief updating model considering a dynamic mental maps for the choice under conditions of uncertainty and learning. They try to enhance and complement some aspects of the prospect theory. The prospect theory doesn't' capture that the travel behaviour rationality is limited due to the uncertainty that influences previous users' choices, moreover users' choices could change if the information reduce the uncertainty.

In their study the so called *expected information gain* is considerate as an attribute of the utility function and it may compensate for the risk of choosing a less attractive alternative. Viti et al. (2005) studied a day- to- day model of learning under uncertainty by applying the adaptive learning model to quantify how users learn and which is the information role with respect to the learning velocity and the past experiences updating. The enhancement of the drivers' behavior modeling is also considered; mainly it's proposed to take into account that (often) decision makers doesn't maximize the perceived utility, rather they make their choices like habitual choices (see also Bogers et al, 2006). In fact, according to other research areas (particularly the psychology) different studies deal with the *bounded human rationality* (Simon, 1957). In a case of bounded rationality it's considered that the users maintain their choices from previous day as long as the outcomes do not exceeded some threshold.

In summary drivers' behaviour (in contexts without ATIS information) has been modelled by different approaches, in all cases the aim has been to reproduce as better as possible the drivers' choices and in great part of the cases the dynamic trajectory (across days) of the users' choices have been explicitly taken into account, consistently with a context of repeated choices. Among the most interesting innovative attributes and concepts introduced in recent literature developments the inertia to change (habit) and the risk perception (in unreliable context) should be quoted as the most significant.

2.1.2 Drivers' behaviour with information: the accuracy role

Several studies have been carried out for ATIS context and in many cases the impact of information reliability on drivers' route choice has been considerate. Some considerations on the relation between the information and the users' chooses have been made by Hogarth (1987). He has considered four main aspects of limited information on the travellers' judgement ability: (1) the human have a selective perception of information, (2) generally, the nature of human processing is sequential (3) humans have a limited capacity to evaluate the information and (4) the humans have a limited memory. For these reasons he suggests that the presentation of the information is more important than the amount of presented information (see also Adler, 1993).

Ben- Akiva (1991) has proposed a convex combination between the historical perceptions of travel times and the ATIS provided information; the process of updating estimation of travel time has been described by means of the convex combination.

Emmerink et al. (1994) studied the relationship between the users' reaction to the information and the network effects. In particular they consider several choice contexts with and without information and how the utility updating mechanism is influenced by the travellers' experience, the considered types of information were the pre-trip information, the en- route information and the ex- post information.

Jha et al. (1998) developed a Bayesian model aimed to capture the perception by travellers at both the travel time and the information; they analyse also the role of the travellers' experience. In their research a theoretical framework has been developed to study the day to day dynamics and to incorporate in the route-choice model the effect of the perceived total travel-time uncertainty. The framework is composed by two main models: the first model is for the dynamic traffic network analysis; the second model concerns the travellers' day- to- day choice dynamics and it is in turn composed by two sub- components, one sub-model for the updating of the travellers' perceptions of network performances (influenced by traffic information and travellers' historical perceptions) and another sub-model for the simulation of the travel choices. The fulcrum of this study is the process of day-to-day updating of perceived travel times by using in the same time the dispatched information and the experienced travel times.

Very interesting is the research made by E. Van Berkum and P. Van der Mede (1999). The information impact has been analyzed by comparing two context with and without information. The effect of two different kind of information (descriptive and prescriptive) have been also observed. In both cases of prescriptive and descriptive information the respondents have been provided with a variable message sign (VMS). The research has been carried out by combining two different approach for the data acquisition: the Stated Preferences and the Revealed Preferences techniques. In the first case the respondents have been submitted to a laboratory experiment, while in the second case data have been gathered with reference to an existing VMS. In a context without information, decision maker can be recognized by the tendency to chose according to their own utility maximization or by habitual choice tendency. In case of ATIS, two different information effects have been identified, according to the descriptive or prescriptive information context. If case of prescriptive information, travellers can be compliant or non compliant (also depending on information accuracy), but information doesn't play a role in the expected utility of travel choices.

Rather, the dispatched information explicitly plays a role in the expected utility in case of descriptive ATIS. It is worth detailing the role that the authors assign to the prescriptive information. For a given day *(t)* and for a given individual *(i)* if the context of choice in descriptive, the travel choices depend on the probability that the travellers are habitual decision makers (H_{it}, Figure 11);in fact, the probabilities of choosing a generic route $r \in \{1....n\}$ is computed in different ways for habitual travellers *(*P^{h}_{i1t} *...*P^{h}_{int} *)* and not habitual ones *(*P^{um}_{i1t} *…….*P^{um}_{int}*)*. In such a computation information as an input into the expected utility maximization process, in particular for traveller that does not choose in an habitual way (*1-* H_{it}, in Figure 11).

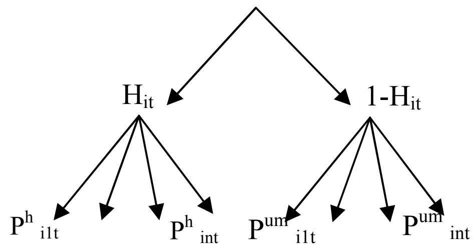


Figure 11: Route choice mechanism without descriptive information

In cases where prescriptive information is dispatched the route choice mechanism became more complex (see Figure 12). In this case the previous mechanism (in case of descriptive information) is modified with the preliminary evaluation of travellers' probability to be compliant (C_{it}) or not compliant (*1-* C_{it}).

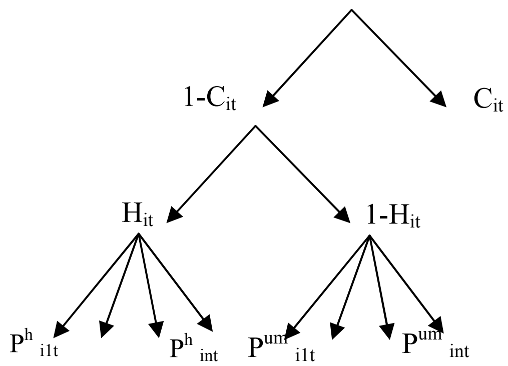


Figure 12: Route choice mechanism in case of prescriptive information

Choice probabilities in Figure 12 are not affected by the information, provided that its role has been already considered with respect to the compliance.
Bogers et al. (2006) have develop a mixed logit model for panel data which also takes into account human factors and the experience effects in the learning process. In the model the perceived travel time is updated at each time also

depending on previous perceived travel times and on previous experienced ones. Hato et al. (1999) have developed a conceptual framework (referred to route choices) for analyzing the effect of different kinds of information devices (e.g.: map signs, travel time signs, variable message signs, radio traffic reports) on travellers' behaviour. The processes of both information acquisition and use have been considered. Latent psychological factors are extracted by considering external and endogenous variables of users; this factors (information process capability and cognitive involvement) are used for information acquisition and reference model. The model has been validated against data related to the Tokyo Metropolitan Expressway, where four different information kind are adopted (the graphical maps displaying, the text messages dispatching about the length of the queues, the notification of accidents, and the dispatching of travel times for certain destinations).

Gou et al. (2004) have proposed a framework for investigating the users' behaviour under information. This framework is composed by two sub components: a within-day model aimed at analyzing the network stability and a day- to-day model aimed at analyzing the users' response to the information.

The proposed day-to-day network simulation framework consists of the following sub-models: *i)* the stochastic route choice model; *ii)* the departure time and users behavior adjustment model; *iii)* the real time information supply model*; iv)* traffic flow dynamics and within day network assignment model.

Avineri et al. (2003) have shown that as higher is the variance in actual travel times (higher uncertainty), as lower is the drivers' sensitivity to travel time differences. They predict the Payoff Variability Effect within a learning model by analyzing the effect of providing travel-time information on a learning process (under uncertainty). They also analyse the effect of two different information strategies: a "static" one, where pre-defined travel-times are supplied, and a "dynamic" one, where real-time travel-times are supplied to the travellers. The main finding is that the effect of dynamic information seems to be less positive with respect travellers ability to make best choices (in terms of increasing of utility). Ettema and Timmermans (2006) have studied a model to reduce the negative effect of travel time uncertainty. This model is based on the expected utility and includes the variation of travel time, the quality of travel time information and the travellers' perception of travel time.

Several researchers have studied the information effect on improving the quality and rapidity of learning. Fujii and Kitamura (2000) have investigate the information effect on the prediction of travel times. Other studies have been carried out by Polak and Oladeinde (2000): they examine the effect on the learning mechanism of both the actual travel time unreliability and of the dispatched information.

In summary, all approaches to route choice modelling in presence of information extend route choice models developed in absence of information. The main critical issues are related to understand how the dispatched information influences the choice utilities and/or the choice updating mechanism, as well as properly modelling the role and the effect of information reliability.

2.2 NETWORK IMPACT OF ATIS

ATIS directly have effects on travellers' choices, this means that they can influence traffic patterns and, in turn, network performances. This consideration has induced several analysts to assess that information to be dispatched can be designed in order to induce desired network effects. In other terms, some authors have assessed that ATIS can be used as traffic control tools able, for example to move user- optimum traffic patterns toward system optimum ones, similarly with what allowed by ATIS applications (Yang and Meng, 2001; Lo and Szeto, 2002a).

Consistently with the previous approach, several analysts have quantified the potential advantage of ATIS implementations as the travel- time saving induced by moving from user-optimum to system-optimum (Mahmassani and Peta, 1993). Other analysts (Mauro, 1998 and De Florio, 2003) have studied the presence of an ATIS aimed to *control/regulate* traffic patterns in order to preserve a reached (stochastic) equilibrium state with respect to the effects of (small) stochastic perturbations. In this case, the ATIS works in conjunction with an extensive monitoring system and any deviation from expected equilibrium traffic pattern is detected and resolved by fine-tuning the information dispatched by the system.

Clearly, the magnitude of the network-level ATIS impact depends first off all on the number of travellers that receive the information. If only a few travellers are provided with information, the cumulative impact of their changed travel choices on network conditions is negligible.

As more travellers receive ATIS messages, the aggregate effect of their reactions becomes important; this aggregate effect depends both on what particular guidance messages are dispatched (where, when e what!) as well as on how drivers react to the messages. Ben Akiva et al (1991) have identified some of the possible adverse network effects that can result by the guidance dissemination. One of the main adverse impacts is the overreaction phenomena, that arises if travellers react roughly to the information by inducing some oscillations in path flows. Another phenomenon arises in cases when *too much* travellers react in the same way to the same information and the congestion is moved from one route to the suggested route, by inducing a concentration phenomena. Moreover an adverse phenomena can be induced in travellers' reactions if the amount of dispatched information is too big; in fact, this leads to the problem of individual saturation, because travellers became no more able to process messages as properly as required for decision making.

2.3 MODELLING THE COMPLIANCE

In a lot of cases the relationship between the quality/accuracy of the information and traveller's compliance is neglected in the scientific literature (Thakuriah et al, 1996) and in some cases the information strategies are evaluated on the base of statistical and descriptive considerations, without incorporating an explicit model of travellers' reactions. Moreover, it's often neglected that the guidance messages have to be consistent with the traffic condition after travellers' reactions.

The travellers' compliance is often considered such as an exogenous variable and frequently its role is confused with the one of the market penetration. Even where compliance and market penetration are separately considered, the elasticity of the compliance is related to attributes such as the (potential or actual) saved travel time and the monetary cost of both equipments and access to information services (Yang, 1998; Yang e Meng, 2001; Lo e Szeto, 2002b), even if these attributes seem to be more realistically related to market penetration rather to compliance.

Yin and Yang (2003) consider that the objective of ATIS is to reduce the travellers' uncertainty about travel times. In their work the inaccuracy/ uncertainty has effect on the dispersion of travellers' behaviours (in a more or less deterministic way). The level of compliance isn't considered to be dependent on the information accuracy but on the total saving of travel time.

In several cases researches on travellers' behaviour in ATIS context have been carried out by acquiring the data using a travel simulator. In a lot of cases the effect of the prescriptive information has been investigate, because the prescriptive information is a specific recommendation to do a particular thing and is often assumed to potentially enable a traffic control system with a more direct influence on travellers' decisions. In studies carried out in order to determine how travellers respond to route guidance and other ATIS information data are collected in two main ways. One is to analyze decisions under ATIS information using PC-based simulation of the travel environment another is to directly observe behaviour in real contexts.

An example of the second kind of approach is represented by series of studies carried out by using multiphase panel surveys of commuters in the Los Angeles metropolitan area (Vaughn et al. 1992, 1993b, 1994a, 1995a; Abdel- Aty, 1995). The major findings of these studies were that the effect of the information effect is also related to the commuters' characteristics for instance female are more likely to receive information pre- trip but not en- route and commuters with college education are more likely to receive information, either pre- trip and en- route. Moreover the compliance increases with the accuracy of the information, and that by combining prescriptive and descriptive information very effective results are reached, especially in term of improving the user's perception of information accuracy that means in turn that the travellers became more predisposed to comply with route guidance information.

A model for the simulation of travellers' reaction to the information has been calibrated also by Bierlaire et al. (2006). They captured the travellers' reactions under different information kinds. The data were obtained by combining a Stated Preference survey and a Revealed Preference one. The models that have been estimated are advanced random utility models. The en-route-choice model is a mixed logit for panel data while the pre-trip-route-choice model is a nested logit.

Perhaps the most comprehensive studies specifically related to modelling the compliance have been carried out by Mahmassani et al (1999) and by Srinivasan et al. (2002). They used a travel choice simulator interfaced to the Dynasmart mesoscopic traffic model, the aim of their research was to try to observe the relation between the compliance and some parameters by investigating the effect of these parameters on the utility to *be compliant*.

It's worth noting that in this case the analysis directly relates to the compliance phenomenon, rather than to the information effects on route choices. The experiment consider the effects of two different types of information (descriptive and prescriptive), as well as the effect of the quality of the information and of the feedback provided to the users on the information accuracy. At first, they studied the effect of some parameters characterising the information, and then they calibrated a model of compliance utility. Two important aspects were focused: the study of the compliance not only for the prescriptive information but also for the descriptive one, and the difference between the accuracy of the information and its reliability.

The authors considered as parameters: the information kind (descriptive and prescriptive), the quality of the information (six levels of accuracy of information) and the feedback (feedback on own experience, feedback on the path recommended by the system and feedback on the actual best path).

The definition of *compliance* adopted by Mahmassani is different than the one (for instance) in the Van Berkum and Van der Mede's research; the *complianc*e has not only been defined as the decision to follow (or to reject) a routing advice (prescriptive information), but is extended to the descriptive information as the decision to use (or not) the supplied information in order to compare travel alternatives and make the choice that system implicitly suggests to be more advantageous. In Mahamassani's research a compliance model has been calibrated as a dynamic kernel logit (discrete choice binary model). For the model calibration several attributes are identified, among others the *accuracy* and the *reliability*. The *accuracy* is defined as the discrepancy between the information dispatched by the ATIS and the experienced travel time.

The *reliability*[5] of the information is defined as the probability that the *relative error of the accuracy* (the difference between the dispatched travel time and the actual travel time, divided by the actual travel time) exceeds a given threshold (five values of thresholds have been considerate) in the previous experiences,

[5] The reliability indicates how frequently the information is accurate.

divided the number of previous experiences. It's worth noting that the accuracy measures how much the information has been accurate, while the reliability measures how frequently the information has been accurate enough.

2.4 ATIS TERMINOLOGY

One of the main terminology issues has been already introduced in the previous paragraph. It's usual in ATIS to distinguish between prescriptive information (e.g. a route recommendation) and descriptive information (e.g. data about traffic conditions). In case of the *prescriptive information* it's usual to say that *route-guidance is provided*, while the descriptive case is referred to as *information providing.* Information can be studied and modelled in two different traffic contexts: *recurrent* traffic conditions and *non-recurrent* ones. In the previous paragraphs only the scientific literature referring to use of information in recurrent traffic conditions has been analysed. Similarly, the following of this work will refer to recurrent traffic conditions, if not explicitly differently stated.

Recurrent conditions refers to the fact that the supply sub-system (network) and the demand subsystem (O/D data) do not change over days; these conditions could lead to steady-state traffic patterns through dynamic processes (day-to-day evolution of traffic patterns). It is worth noting that *recurrent* traffic conditions are not necessarily *equilibrated* traffic conditions. Equilibrium can be probably reached but recurrence only is a necessary (but not sufficient) condition for equilibrium.

Some terminological consideration also worth regarding the nature of the information .

Two main different nature of the information can be distinguished. The fixed information (or static) which is related to things that rarely change (e.g. location of point of interest) or basic way findings directions that are not tied on the actual traffic condition (or the influence of traffic conditions on route is considered to be negligible). In contrast, information can be dynamic, here considered as an information that changes as traffic conditions change.

One can assume that the supplied information does not drastically change the distribution of (actual) travel-times on the network and across the days, basically because the network is not heave congested and/or because the percentage of travellers that have access to information (and react to it) is relatively small and/or because the information system is aimed to stabilize a steady-state traffic pattern against stochastic fluctuation ("control" approach to information systems). In this case, the system can be aimed to dispatch *instantaneous* information based on the prevailing (instantaneous) travel times on the network. In contrast, a predictive ATIS can be considered, consisting in dispatching the predictive information based on the travel times that actually the travellers will have experienced once the destination will have been reached. Obviously, the instantaneous travel-times and the actual travel times (as well as the instantaneous information and the predictive one) coincide only in case of within day-static traffic patterns. In case of predictive information one can assume that the travellers' reaction (in terms of redirection of route choices) to the dispatched information can drastically change traffic patterns (and so travel times); in this case, the information has to be not only predictive, but also *consistent* with the travellers' reactions it will produce. The last is also known as the *anticipatory route guidance* problem (Bottom, C and Bierlaire, M., 2001). In case of instantaneous information "*accuracy*" can be intended as the ability of the information system to induce in the travellers a travel-times perception the more "correct" as possible with respect to the "prevailing" travel-times on the network. In case of predictive information the accuracy has to be measured with respect to actual travel times. In case of consistent information the accuracy also refers to the correct solution of the anticipatory route guidance problem. It is worth noting that dispatching instantaneous, predictive or consistent information is not a free choice, that this depends on the traffic context in which the ATIS solution is implemented and, finally, that an accurate instantaneous information has to be considered as actually inaccurate if the traffic context requires a predictive and/or a consistent information.

2.5 SUMMARY

Analysis of previous paragraphs have evidenced that for ATIS contexts several questions still worth enhancements and refinements. In particular, relatively few studies have been carried out with direct reference to the modelling of compliance phenomena, most of the studies having be addressed to how ATIS influence travel choices in terms of changing route choices.

In this paper the approach by Srinivasan et al (2002) will be followed (and possibility enhanced) with reference to the behaviour of Italian travellers. The effects of the quality of the dispatched information (in terms of information accuracy) on travellers' compliance will be fully investigated and an explicit compliance modelling framework developed.

It is also expected that an explicit compliance model, developed at disaggregated/individual level, will contribute to allow new approaches and perspectives toward a consistent development of modelling tools able to correctly simulate the effect of ATIS at a network level.

3 THE TRAVEL SIMULATOR LABORATORY (TSL)

3.1 STATED PREFERENCE AND THE REVEALED PREFERENCES APPROACHES

The *Stated Preference* and the *Revealed Preference* approaches are the most diffused methods in data collection. SP and RP approaches are mainly distinguished on the base of the adopted methodology: the RP strategy is referred to the case in which respondents have really made their choice; differently, SP strategy is referred to the case in which respondents are submitted to an hypothetical context of choice. The main problem in economic contexts (at the beginning these approaches were introduced for market investigations) and also in transport problem solving is the ability in forecasting the demand of market/transport, but as described in the following, several other issues are related to the RP and SP approaches.

Once the SP approach is chosen, also other issues characterizes how the survey is carried out: if the context of choice is considerate well know by respondents (familiar), standard methods can be adopted in surveying data, otherwise more complex and multimedia approaches have to be adopted.

Economists and researchers, have identified different reasons on the base of which, for dealing with the real “market”, respondents are involved in hypothetical context of choice by adopting the SP methodology:

1 It is needed to introduce in the market a new product (therefore the product is not real in market) or new variables/attributes are introduced; in case of transport, a new transportation mode or a new service (ATIS, for instance) is planned to be introduced;

2 Explanatory variables have a little variability in the marketplace; this happens also in transportation, for instance fares or travel times could be too much flat in the actual system, so that their influence on travel choices can’t be isolated

in the actual contexts; in ATIS, for instance accuracies of actual systems could be inherently low, so that their influence can't be stated;

3 Explanatory variables are highly collinear in marketplace; this also can happen in transportation, the distribution in the real context of the attributes can actually show an apparent collinearity, even if this is not a inherent propriety, so that calibration should also involve these parameters;

4 Observation of data take a long time and data are expensive to be collected; this is typical in transportation contexts, especially when disaggregate travel choices should be observed.

In conclusion, RP data availability could be limited or problematic, often because of technological frontiers. With reference to Figure 13, even if RP data are useful in understanding an existing market, technology or service, the SP approach result to be complementary to RP ones in case of new technological frontiers investigations.

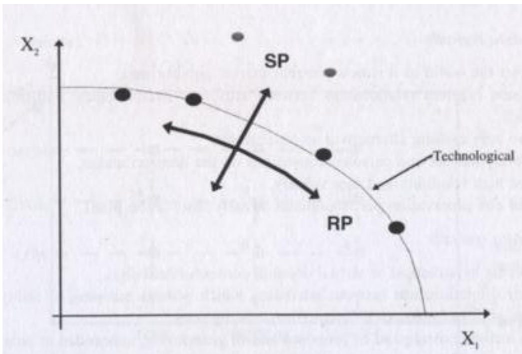


Figure 13: The technological frontiers and the roles of RP and SP data

Based on the above preliminary considerations in the following table (Table 1) the differences between the SP and the RP approach are synthetically shown.

Table 1

Comparison between RP and SP approaches

RP approach		SP approach	
☹	Is able to analyze current market contexts	☺	Is flexible in hypothetical contexts describing
☹	Some technological constraints are fixed	☺	Allow to test unexplored relationship between some attributes
☹	It's only possible to observe existing alternatives	☺	Allow for testing innovative alternatives and services and to test innovative attributes
☺	Embody market and respondents constraints	☹	Cannot easily embody market and personal constraints
☺	The acquired survey are reliable, realistic and valid	☹	The availability is related to the experiment clarity and to respondents' capability in understanding the choice context; moreover, stated choices are not actual ones
☹	It could be cumbersome to track repeated choice contexts	☺	It's relatively easy to capture multiple observations for each respondent in repeated choice context

Moreover, the SP approach is cheaper, the variables are more controllable, several attributes can be introduced (qualitative and quantitative).

To be more precise, as following described, tree main different aspects can be identified by comparing the RP and the SP strategies:

- The specification model; by adopting the SP approach, the respondent answers on the base of the designed factors without considering other aspects; on one hand this ensure that the tested attributes are not biased by other undesired effects, but, on the other hand, this means that the observed choices could be different in a real context where other factors could influence the choices;

- Statistic estimation; the data distribution in the RP approach cannot be controlled and in particular it isn't possible to verify the effect of the related attributes;
- Range of applicability; the RP strategies cannot be adopted to estimate models in case of innovative scenarios of choice.

It is worth noting that strong and weak point of the two approaches could be integrated within a joint use of RP and SP techniques. The theoretical integration between RP and SP surveys (*data enrichment*) has been defined by Morikawa (1989). Other researchers (Ben-Akiva and Morikawa, 1990; Ben-Akiva, Morikawa and Shiroishi, 1991; Bradley and Daly, 1994; Hensher and Bradley, 1993; Adamowicz, Louviere and Williams , 1994; Adamowicz et al., 1997; Hensher, 1998) have discussed the *data enrichment paradigm.* Mainly, they fixed as the goal of the analysis to produce a model that can forecast to real market the future scenarios. The *data enrichment paradigm*, as in Figure 14, suggests an integration of RP survey (mainly based on the analysis of the current equilibrium) with the SP collected data, often addressed to the attributes trade-off understanding.

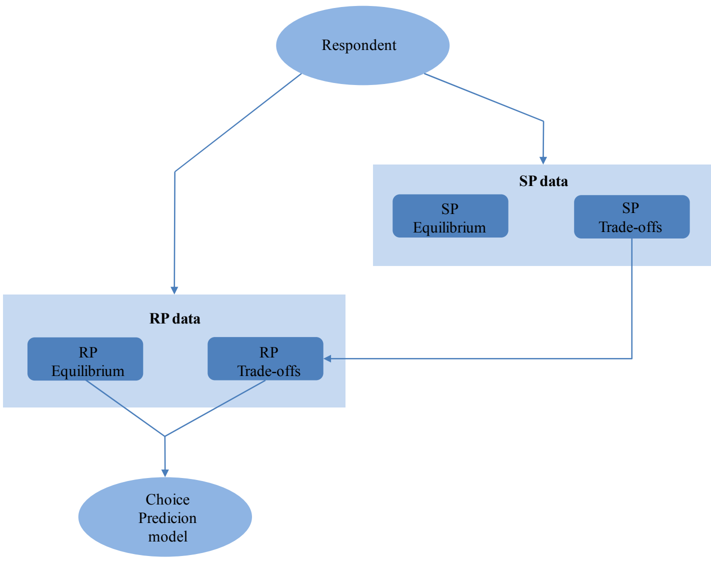


Figure 14: Data enrichment paradigm-1

Therefore Swait, Louviere and Williams (1994) (as described in Figure 15) propose to use each data source to capture those aspects of the choice process for which it's superior. RP are efficiently used for market equilibrium understanding and SP data can be used for the trade-off analysis. Moreover SP data provide more information and enriches the market equilibrium in case of large range of situations.

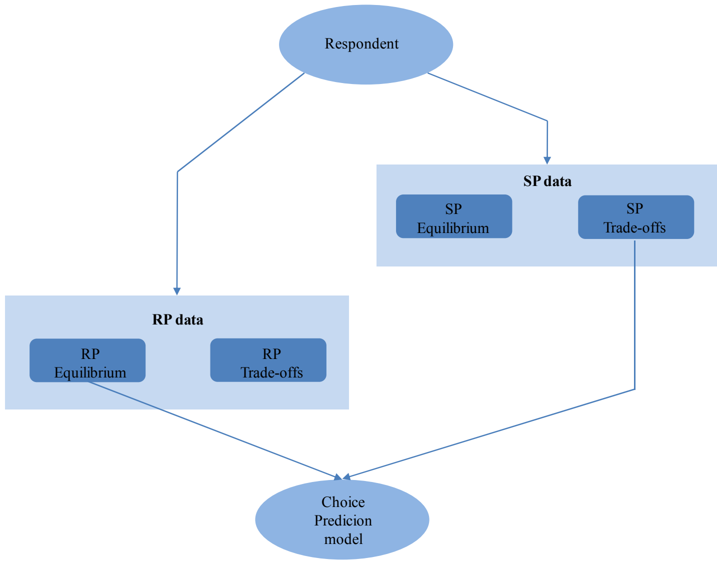


Figure 15: Data enrichment paradigm-2

The adopted methodology in *data enrichment paradigm*, is currently constrained by the enhancement of the state of the art, in these terms the most applied methodology is referred to the *paradigm 1*.

3.2 STATED PREFERENCE WITHIN THE TSL

As described in the previous paragraph, in order to observe travelers' behaviors two approaches can be adopted: the Stated Preferences method (SP) and the Revealed Preferences (RP) method. Both the approaches have some advantages and disadvantages. The SP method is often preferred by the analysts for its ability to allow for fully controlled experimental environment and data acquisition processes. However, the SP method does not refer to actual choices and choosing contexts and sometimes it's hard to determine the validity of acquired data.

Moreover in case of transport studies several researchers have adopted the Stated Preference approach for data surveying in case of ATIS contexts (Polak and Jones, 1993; Khattak et al., 1993; Abdel- Aty et al., 1996; Wardman et al., 1996; Mahmassani et Jou, 1998; Bogers et al., 2006).

An important aspect at is to define the instruments for data collection. Traditionally the most diffused task for collecting data is presented to respondents the test in paper and pencil surveys. Moreover more advanced technological tasks are also developed. In case of transport investigations, for data surveys generally two kind of simulators can be adopted: travellers' simulator and driving simulators. In case of driving simulators travellers are submitted to an immersive/virtual reality while in case of travel simulators they are submitted to a network simulated by scenarios on a PC. The driving simulator are naturally configured as a more realistic scenarios but in the same time they are too expensive and moreover the set up step and the calibration require long time.

Moreover driving simulators well fit for researchers in the driving-behaviour field (for instance ADAS, in case of ITS), while travel simulators are more adequate in the travel behaviour filed.

In this research the SP approach has been adopted and in particular the Travel Simulator Laboratory, developed at Delft University of Technology, has been used in collecting data. The goal of the TSL (Hoogendoorn, 2004) project is to develop a research tool for empirical studying the people's behavior with respect to dynamic travel information. The TSL is an internet toolbox aimed to observe users' choices in order to reproduce the travellers' behaviours in a Stated Preference data context. Respondents can answer a number of questions by means of an internet site. Results are stored in a database. These results are accessible for the research worker by using a standard, external SQL tool.

To apply an SP approach, it is fist of all needed to define:

- The alternatives of choice;
- The variables or attributes to be tested in order to explain respondents' choices;
- The scenarios by that compose the experiment, differing with respect to the choice context which the respondents are submitted, as well as distribution of the respondents across the scenarios ;
- The values of attributes assumed in the experiment and how these are distributed across scenarios and simulated choice contexts.

Several methods can be adopted in design the experiment and formalizing the questionnaire to submit to the respondent. According to the literature SP, can be implemented by tree possible approaches:

- *Rating* (or preference estimation): for each alternative the respondent is asked to state a preference estimation (like, dislike; consider, not consider, etc);
- *Ranking* (or ordered preference): the alternatives are ordered by the respondent on the base of his/her preference;
- *Choice*: among all alternatives we ask to the respondent to choose only one alternative.

In the following, the choice method will be implicitly referred, provided that has been the one adopted in this thesis.

In the specifically case of transportation research application, the alternatives can be distinguished in within mode or between mode, such a notation refer to the case in which mode choices are deal with in the experiment; however, the concept can be easily extended to different kinds of choices contexts. In case of within mode the alternatives are generated by varying the values of the attributes in the same transportation mode; in the second case the alternatives are obtained by considering different modes of transport by varying the attributes and their values for each mode.

It's also appropriate, for a correct application of the SP method, that:

- The respondents are selected and asked to join to the sample by the analyst (uncorrelated access by voluntary answering to generic-impersonal calls should be avoided);
- The alternatives are unknown by the respondent without inducing some influences on the respondents' flavor;
- Some extreme alternatives introduced the *best alternatives* and the *worst alternatives* are

In any case, one of the main critical issue in designing a SP experiment is to properly identify the distribution across scenarios and respondents of the values of the attributes that define different choice contexts.
Two main characteristics are required for the values of the attributes:

- they should be as more realistic as possible;
- they should have a perceivable and not negligible variance.

It is worth noting that several attributes participants to the definition of a sample context of choice. In theory all possible values of each attribute should be tested against all possible values of each other. The number of choice contexts defined in this way could very quickly increase as the number of attributes (and of possible attributes values) increases. This should lead to a combinatorial problem in which the number of combination is much higher than the number of respondents that can surveyed with reasonable effort.

3.3 THE TRAVEL SIMULATOR LABORATORY PROTOTYPES

The Travel Simulator Laboratory has been developed at the Delft University of Technology. It allows for carrying out surveys aimed to data gathering for transportation and traffic related analysis. The TSL has been employed in several researches[6].

3.3.1 The Prototypes

Four main prototypes of the TSL have been released. The Prototype 0, 1 and 2 have been implemented using HTML and PHP, while the Prototype 3 has been implemented by using Java.

Prototype 0, 1 and 2 have been developed to test the effects of travel information on route travelers' behavior (in prototype 2 the role of the information accuracy is introduced) while prototype 3 was mainly intended to test the effects of information in a multimodal context of choice (private and public transport).

In Prototype 0 and 1 decision makers were submitted to route choice scenarios with and without information. The network deal with by the experimental environment was composed by three routes in prototype 0 and by two routes in prototype 1, the travel time on every route was obtained by appropriate distributions.

In prototype 0 the three routes had different characteristics: travel times of route 1 and 2 had uniform distribution. Travel time of route 1 had a lower average and higher standard deviation than route 2. Travel time of route 3 was normally distributed and the value of the average was between the ones of route 1 and route 2. In Prototype 1 travelers were asked to choose between two routes and both *day to day dynamics* and the *within day dynamics* issues were tested (the latest with reference to the choice of the departure time within the day).

---

[6] The main programs are: AMICI (Advanced Multi-agent Information and Control for Integrated multi-class traffic networks), finalized to find solutions to efficiently manage traffic congestion problem; MD-PIT (A Multi-Disciplinary study of Pricing Policies in Transport), studying the effect of pricing policies; PITA (Personal Intelligent Travel Assistant), aimed at understanding the information effect on travelers' behavior.

Travelers were asked to make their choices repeatedly for several days, different travel time values (describing different degree of congestion) were experienced in accordance with the departure time chosen by respondents.

The respondents were submitted to scenarios where two different travel reasons were proposed: the arrival, at destination was identified with an important meeting place (and the trip reason was a job meeting) while the alternative reason was an arrival with fewer reason of marginal obligations time.

In prototype 1 (as well as in prototype 2, where three routes were once again considered), the running of the experiment was similar to prototype 0. During the experiment the daily choices were simulated. Every day the traveler was asked to choose the departure time (see Figure 16, step 1) and consequently was provided with en-route suggestions (step 2) and asked to make his/her travel choice.

Different types of suggestions were considered for step 2. In some cases it was simulated the presence of a dynamic route information panel displaying estimated queue lengths on routes; as an alternative, estimates of route travel times were displayed instead of queue lengths.

In any case (step 3) the respondent was notified about the consequences of his/her choice, in terms of congestion actually occurred and actual travel times.

One of the differences between Prototype 1 and Prototype 2 was that in the second one the respondent had a predefined time within which choosing, but generally Prototype 2 can be considered as a slightly modified version of Prototype 1.

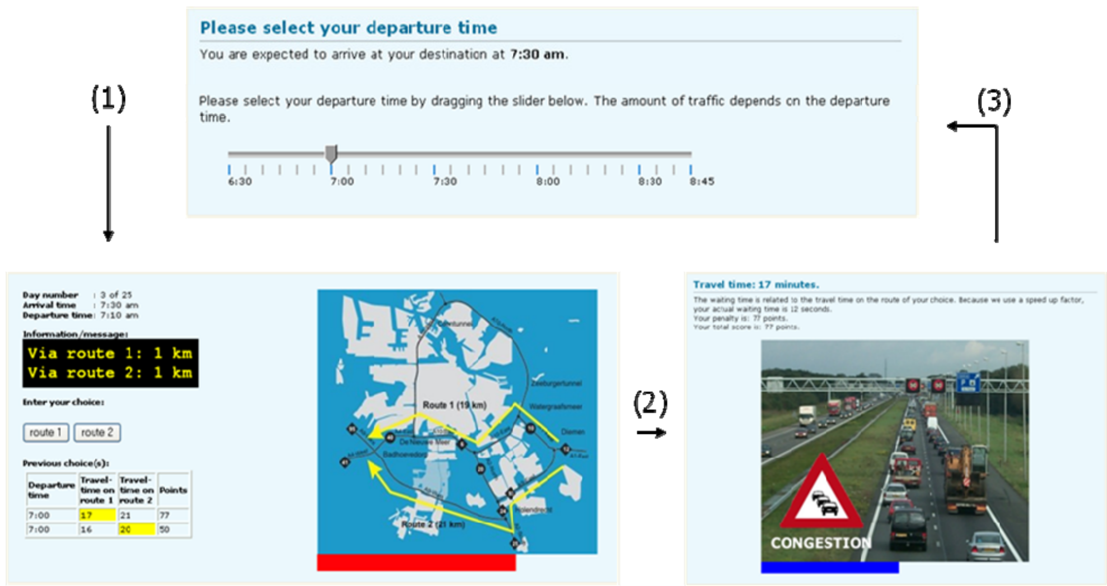


Figure 16: Prototype 2

The aim of Prototype 3 (see Figure 17) was to study the effect of the information on travelers' choices among different modes of transport. Travelers were provided with the en-route information at an intermodal facility where they were allowed to switch to a different transportation mode. The information dispatched at the intermodal node was referred to the estimated travel times to destination (for each choice alternative). After the choice the respondent was notified about the consequences of his/her choice, expressed in terms of travel times. These notifications regarded not only the chosen alternative but also the other one; this was mimic of the travelers' (multimodal) network performances awareness mechanism and allowed for the respondents' awareness of information accuracy. During the simulation, like in other prototypes, travelers were asked to fix their departure time. Respondents were not only informed about estimated travel times (estimated arrival at destination) but also about travel prices.

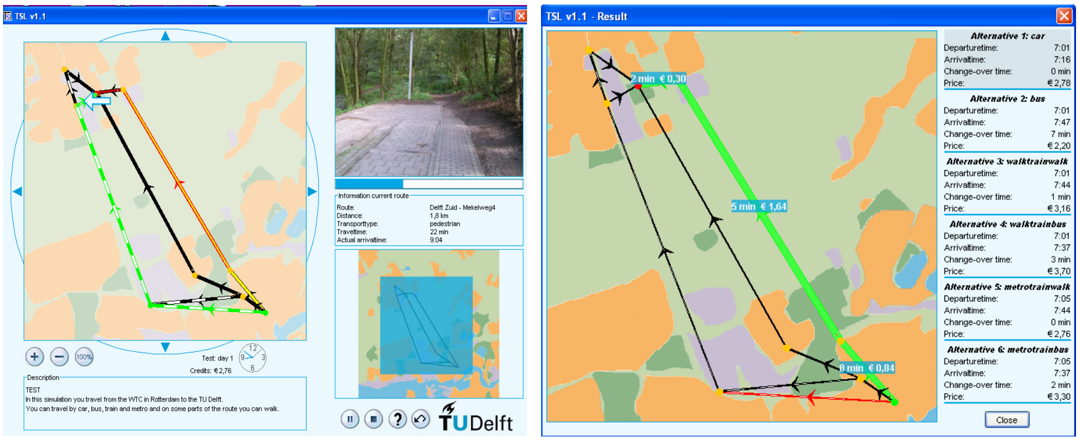


Figure 17: Prototype 3

3.3.2 The TSL web application: technical considerations

In the TSL web application respondents are able to answer questions through use of a web site and, by means of that, to participate to Stated Preferences surveys; their answers are stored in a database that constitutes the sample database of the analysis phase. The questions are designed basing on an information scenario and on experiment settings.

In the figure below we can see the complete system in very simplified scheme (figure 18): the whole application is defined by the interaction between the External SQL Tool and the TSL platform composed by the TSL web application, the I/O files in XML and the Database in which results are stored.

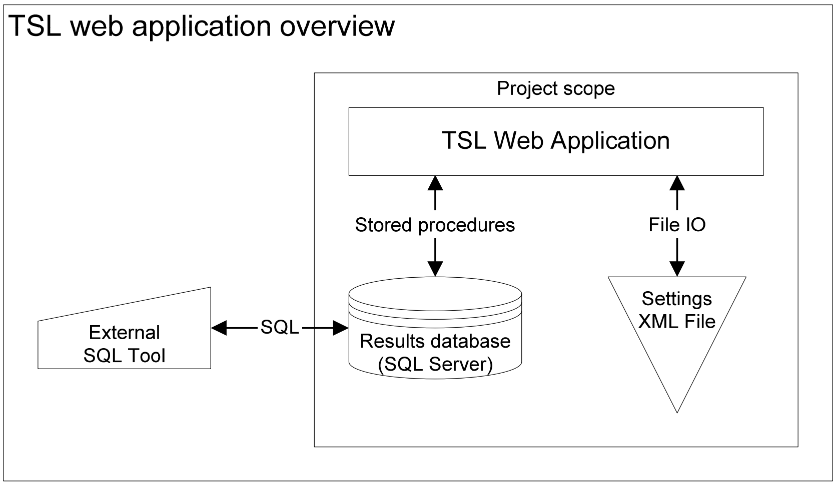


figure 18: The TSL web application overview

Several actors[7] will be involved in the experiment design, in particular we can distinguish the following actors:

- *The Research worker:* a person interested to data acquisition for research aim (or to carry out the projects related to the research), who wants to configure and run an experiment;
- *The Respondents*: the invited persons that will connect on the webpage and will fill the answers to questions in an experiment;
- *The Web master*: a technical person that makes the experimental environment available by installing the software components and configuring the experiment settings.

The experiment design can be schematized by considering three different step:

- The configuration of the experiment;
- The configuration of the scenarios;
- The configuration of the screens (see Appendix A).

Each scenario is configured congruently with the configuration of the experiment steps: the information to be dispatched (in the simulated environment) is

[7] Actors: referred to the technical people that will be involved in the configuration experiment and to the respondents that will be invited to make their answers.

identified, the actual travel times and the information (as they evolve over time) are designed.
A file code must be upgraded similarly to what has been done in our specific experiment (see Appendix A).

3.3.3 Running the experiment

Each experiment with the TSL is characterized by successive steps, the sequence of the steps is described in the following:

1) Start;
2) Open questions;
3) Example question;
4) Start simulation;
5) Questions 1 to n;
6) Roundup screen with questions;
7) Results and roundup screen.

During step 1 the experiment is presented. In general, all the steps that respondents have to follow are introduced, as well as some detail about the aims of the experiment; in such a way the respondents are informed on which is their role in the research and on what they should do to contribute to the success of the research. Of course, behaviors (that are the subjects of the analysis) should be not suggested or induced in any way to respondents. After step 1, in step *2* the respondent can fill the information referred to his/her identification (for instance the age, the gender and the educational level) by answering to some preliminary questions. In the same step other information are acquired typically on travelling behavior, like for instance the feeling in arriving bit late, more late, bit early, more early[8]. At the end of this preliminary investigation the captured information

[8] Traveler can define its perception by selecting these attributes: very pleasant; pleasant; neutral; unpleasant; very unpleasant

allows for characterizing the sample composition , for instance in terms of gender, age, educational levels and dislike in arriving late or early. In *step 3* the respondent is allowed for a trial of the asking/answering procedure of the experiment. At *step 4* respondent is notified that the simulation is going to start and *step 5* (which is repeated n times) actually asks for travel choices. A typical decision making context, presented to respondent within step 5 in one of the ATIS related prototypes of the TSL is described in the following; the choice context is referred to a recurrent trip, that is one of the n times step 5 is repeated.
First of all, the traveler, by using a bar, can select the departure time; then it is simulated the arrival of the traveler to a VMS (variable message signal) where he/she is provided with a prescriptive or descriptive information; on the base of the information (and of his/her previous experiences) the traveler is asked to make a route choice. The result of the choice (for the current day) is notified to the respondent in step 6 (which is repeated n times together with step 5); actual travel times and arrival time are notified. Step *7* is displayed at the end of the repeated choices; it is intended to ask some open questions related to the general impact that he/she has received from the simulated information system. On the same page also a report for the respondent is displayed; this refers to the performance of the simulation: the most chosen route, the average travel time, the number of times too late and number of times he/she arrived at destination too soon. Some other closing questions can be displayed, like as if the respondent would like to receive some information on the preliminary research results, or if he/she would like to be involved in future researches.

3.4 SUMMARY

This chapter has presented a general overview of the TSL, which is the experimental tool that will be used in the following in order to capture stated preferences for our analysis. Its functionalities have been described from the basic functionalities implemented since the beginning of the TSL project until the improvements made in successive research applications.

Some research projects have already tested the applicability of the TSL and in this chapter the implemented prototypes have been presented and described. Also a few of the technical aspects of the platform have been here described. Finally have been also described a typical configuration of one experiment, carried out in developed prototypes oriented to study ATIS application.

The same experiment configuration will be adopted in the following to carry out the research described in this PhD Thesis.

This adoption process, as well as the actual experimental design (in terms of definition of the variables defining the experimental scenarios) will be presented in the next chapter.

4 EXPERIMENT DESIGN AND DATA COLLECTION

4.1 EXPERIMENT DESIGN

During the experiment travellers are asked to make their route choices as these could happen in a real context, because of that a route choice context has to be reproduced in the simulation and the distribution (across the simulation days) of the actual travel times represents the *first control variable* of the experiment. Provided that the experiment aim is to study the effect of the information inaccuracy on travellers' behaviour the *second control variable* of the experiment are the *ATIS estimated travel times*. On the base of the inaccuracy of travel times estimation, several scenarios can be obtained, each of these associated to a given inaccuracy level.

Information can be also dispatched to the traveller in different way according to the information kind that we are considering. Generally travellers can be provided with *descriptive information* (if the information is the estimated value of travel times for every routes) or with *prescriptive information* (if the information is the suggested route) and in this case the prescriptive information is computed on the base of the estimated travel times by ATIS. In summary scenarios can be differentiated by two main parameters: the first one is the information kind, the other the information accuracy.

The combination of information kind and information accuracy has lead to 8 scenarios on which the stated preference survey has been carried out.

For all the scenarios actual travel times (the ones that respondents will actually experiment during the simulation) change over days according to the same random distributions, even the sequence of actual travel time draws across the 40 days is the same for all scenarios. Four of the scenarios are related to descriptive ATIS and four to prescriptive. Descriptive and prescriptive ATIS were coupled to four information accuracy level. Information provided by ATIS, in fact, can be inaccurate because, at a given day, supplied route travel times (in case of descriptive information) could be different from actual travel times or because the suggested route (in case of prescriptive information) actually is not the one with minimum travel time. In practice for each level of accuracy, it has been designed the distribution across days of the ATIS error in estimating travel times. Such an error is used to directly compute ATIS estimates supplied to travellers in case of descriptive information or as the base for ATIS estimation of the best route in the prescriptive case. Provided that, for each level of accuracy, the sequence of draws of ATIS errors across days has been unique, the descriptive and prescriptive scenarios associated to the same accuracy level are identical in terms of modelling accuracy and only differ on how such an accuracy actually manifests it-self and can be inferred/perceived by travellers. It is worth noting that at the end of each daily travelling simulation, respondents are notified about travel times actually occurred (in the simulation context) on the network for all 3 alternatives. It could be deeply and controversially discussed about how realistic is the ex-post information mechanism on all routes. We observe that if an ATIS is on-place an ex- post information facility is a so straightforward service that unlikely it can be not requested by the travellers (so, supplied by the system). Moreover, ex-post information can be considered as a (very rough and naive indeed) approximation of the natural phenomenon of experienced travel times sharing, which has been widely discussed (and employed for modelling purposes) by transportation system analysts. In any cases, in our experiment, the ex-post information mechanism is the only tool available for enabling respondents' perception of ATIS accuracy (as well as actual as actual travel times reliability). The experiment has been implemented within the TSL platform. This has hallowed the selected respondents to participate to the stated preference survey via WEB.

The respondent have been properly pre-selected and invited by e-mail to connect to the TSL website in order and to answer to the questionnaire. The questionnaire was composed by three parts (according to a typical structure of a TSL experiment, as described in the previous chapter): in the first part the experiment is described to the respondent who is also asked to answer to some opening questions (for respondent's characteristics acquiring); the second and main part of the experiment ask for decision-making on route choices (assisted by ATIS) for 40 consecutive simulated days; the third and last part of the experiment asks for some closing (main courtesy) questions.
In the following the design of the controls variables of the experiment is described, as well as the sample composition.

4.2 EXPERIMENTAL VARIABLES

4.2.1 The Actual Travel Times

The network used for the stated preference experiment has been considered to be composed by three routes. Actual travel times of the 3 simulated routes are considered to be distributed across days independently for each route, and for each route 40 draws are generated in order to represent a sequence of actual travel times instances over 40 successive days. The distribution parameters of the draws are reported in Table 2 and the resulting actual travel-time instances are depicted in Figure 19.

Table 2

Sample means and standard deviations of route actual travel times

Route (j)	Mean (minutes)	Standard Deviation (minutes)	Coefficient of Variation (CV_j)
1	44	15.1	0.3411
2	53	0.8	0.0015
3	47	11.9	0.2360

Actual travel times instances show that route 1 (mainly) and route 3 (partially) are the best performing ones while route 2 is the worst one. However, travel times on routes 1 and 3 are not reliable, in the sense that sometimes (for instance 10 times over 40 for route 1) they are much greater than the average and much greater than the one of route 2.

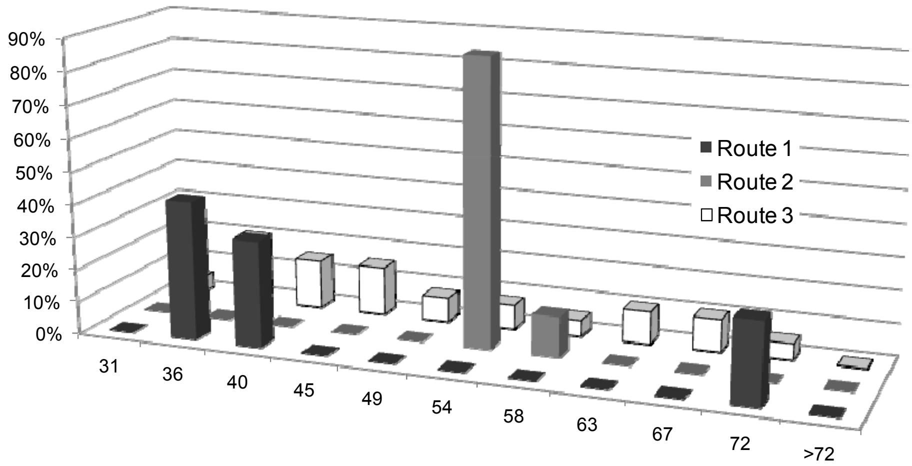


Figure 19: Sample distribution of actual travel times

At the opposite, never happens that the actual travel time of route 2 dramatically increases. In other terms, route 2 is such that the risk of extremely high travel times is minimum, even if seldom it is the best route. It is worth noting that the distribution of actual travel times is of course known to the analyst but unknown to respondents, they can only infer such a distribution during the experiment, as long as it runs.

4.2.2 Travel Times Estimation: computation of ATIS Error

The ATIS error is super-imposed on the variability of actual travel times in order to produce ATIS- estimation of travel times; this is made with reference to four different simulated travel contexts with different standard deviations of the ATIS error. In other terms, the actual distribution of the ATIS error depends on the level

of the accuracy of the ATIS and four cases of accuracy are considered. Within each case, different distribution of the error are considered for each route.
The estimation error with respect to actual travel times. this has been designed to be increasing from accuracy level 1 to accuracy level 4. For each accuracy level, the ATIS error is considered distributed across days, independently across routes, according to a normal distribution (for the first three accuracy levels) or according to a uniform random distribution (for the last accuracy level). In case of uniform random distribution (accuracy level 4), the error is such that the resulting draws of ATIS travel time estimates are between 70% of the minimum actual travel time and 130% of the maximum one, where minimum and maximum are computed over all routes and all days. Means and standard deviations of ATIS errors for the normally distributed accuracy levels are reported in the following Table 3.

Table 3

ATIS error means and standard deviations for normally distributed accuracy levels

Inaccuracy Level	Mean	Std. Dev. Formula	Standard Deviation Value		
			Route1	Route 2	Route 3
1	0	0.25 CVj $\forall$ j$\in${1,2,3}	0.0853	0.0004	0.0590
2	0	0.70 CVj $\forall$ j$\in${1,2,3}	0.2388	0.0011	0.1652
3	0	0.25 $\forall$ j$\in${1,2,3}	0.2500	0.2500	0.2500

For accuracy levels 1 and 2, the standard deviation of the ATIS error for a generic route (*j*) is proportional to the coefficient of variation of the actual travel time of such a route (*CVj*). For accuracy level 3, the standard deviation of the ATIS error is independent on the actual travel time distribution parameters. In practice, accuracy of level 1 is high , while is very low in level 4; in levels 2 and 3 it is intermediate but in level 2 the (in)accuracy of the ATIS error is appreciable for route 1 and route 3 and negligible for route 2; at level 3 the error for routes 1 and 3 is a little bit higher than at level 2, while for route 2 is much higher than at level 2. In other terms, if ATIS estimates on travel times are directly supplied (descriptive information) at accuracy level 2, the travellers receive on route 2 information almost as precise than at level 1 while at level 3 they receive on route 2 considerably imprecise information. The values of error dispersion distribution over days under scenarios 1, 2 and 3 can be observed by referring to Figure 20, Figure 21 and to Figure 22.

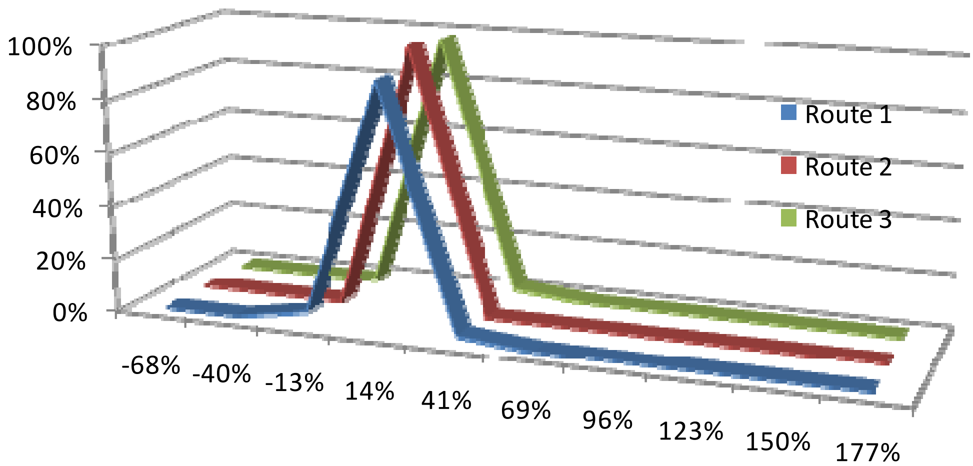


Figure 20: Error dispersion in scenario 1

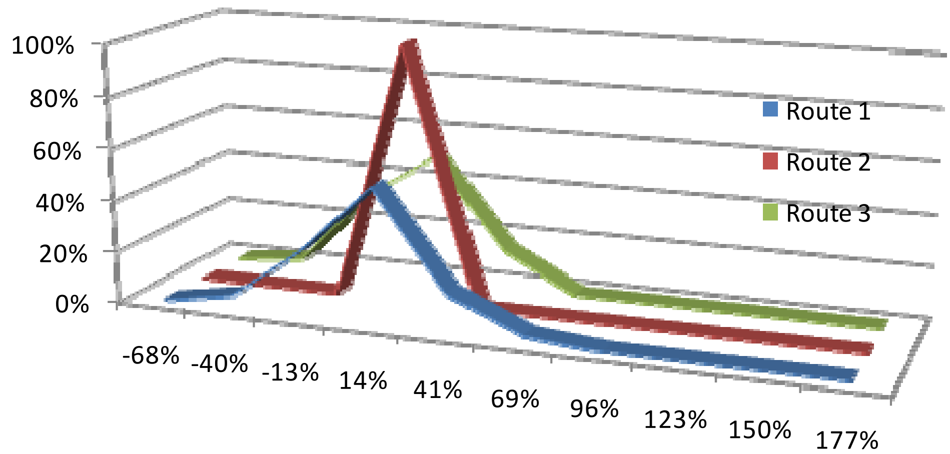


Figure 21: Error dispersion in scenario 2

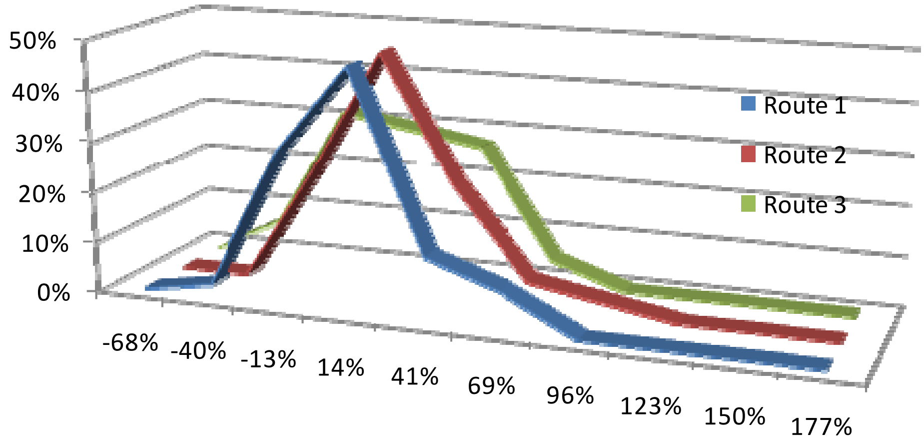


Figure 22: Error dispersion in scenario 3

It's also possible to observe for the descriptive scenarios the cumulative distribution of the reliability over days. By referring to figure 23 it's possible to observe that as the accuracy level decreases (scenarios from 1 to 4) the cumulative value of the information reliability[9] also decreases.

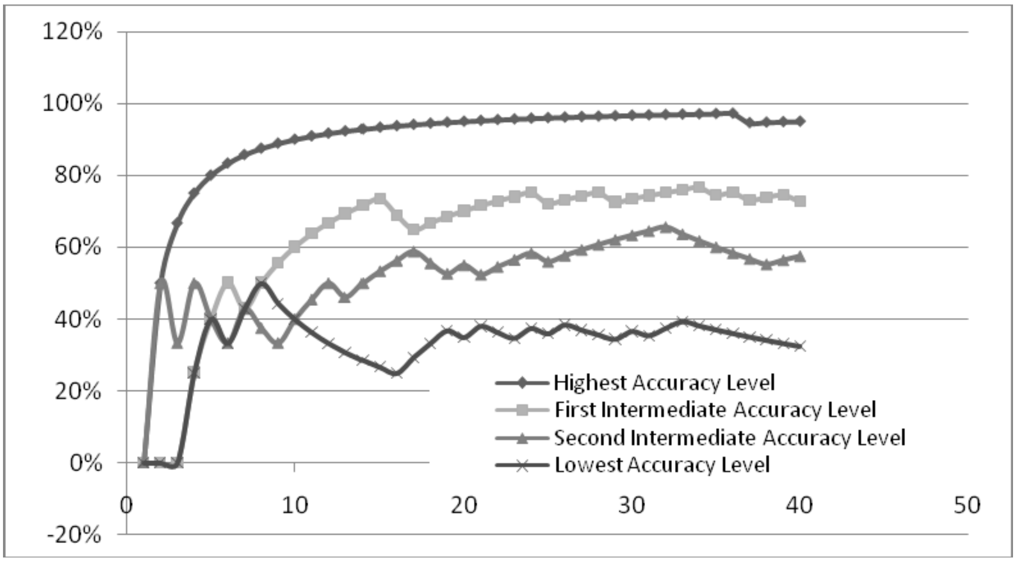


figure 23: Cumulative distribution of information reliability in descriptive scenarios

[9] Is the cumulative number of times over days in which information system is reliable.

In case of prescriptive information, ATIS estimates of travel times are employed to compute the ATIS-estimated best route, which is suggested to travellers; thus the ATIS reliability perceived by the travellers depends on if the suggested route actually has been the best one. In our experiment, reliability performances of scenarios 1, 2, 3 and 4 respectively are 35/40, 28/40, 21/40 and 12/40 (see Figure 24 and Figure 25).

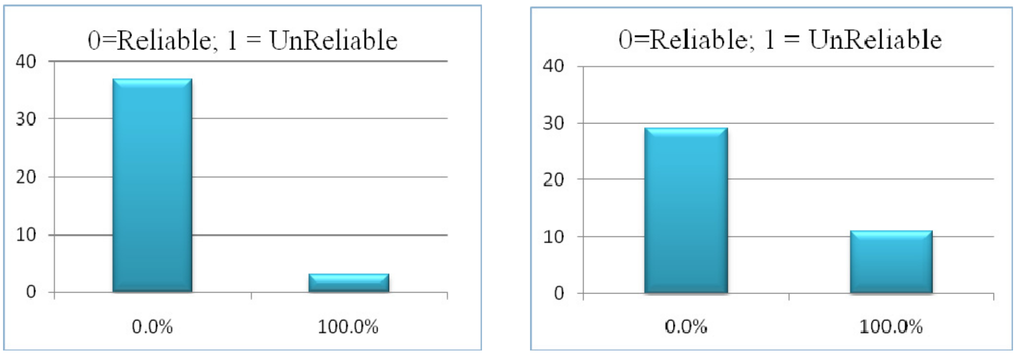


Figure 24 : Reliability of information in scenarios 5 and 6

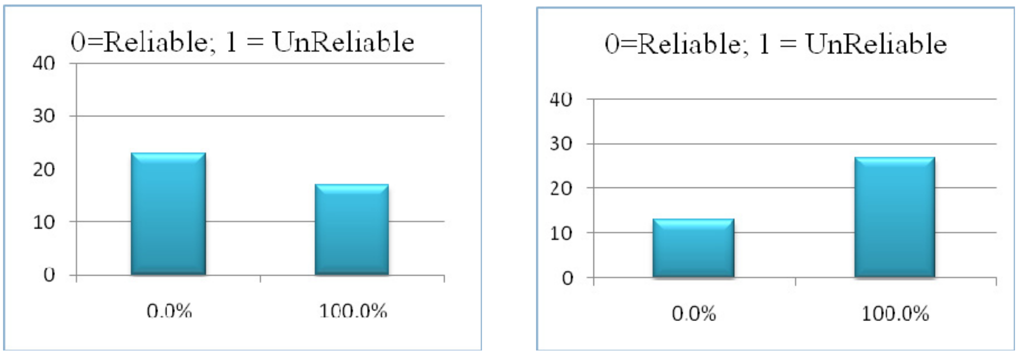


Figure 25 : Reliability of information in scenarios 7 and 8

4.3 THE SAMPLE

The experiment has been made by considering about 160 respondents. The sample was composed by the 40% of students, the 30 % of professors, researchers and free lancers and finally by the 30% of administrative and technical employers.

The respondents have been uniformly distributed over scenarios. In total the scenarios are eight of which four referred to descriptive information and the other four referred to prescriptive information. The 160 respondent are divided among the scenarios by considering 80 respondents for descriptive scenarios and 80 respondents for prescriptive scenarios; for each information kind 80 respondents are divided for each inaccuracy level, so that 20 respondents have been assigned to each scenario. In practice, the dimension of the sample has been a few greater than 160 respondents in order to have chance to purge the final database from manifestly inconsistent respondents.

4.4 THE COLLECTED DATA

At the end of the simulation the final data base has been obtained. In each record are stored some information relative to the respondents, and for every respondent will be available the chooses he/she made over 40 days. In particular the data base is composed by 6400 records; the registered fields have been:

- the scenario id;
- the respondent's name;
- the respondent's age;
- the respondent's educational level;
- the respondent's gender;
- the respondent's job;
- the choice day;
- the ATIS estimated travel time (relative to each route) for descriptive scenarios;
- the suggested route in case of prescriptive scenarios;
- the route chosen by respondent;
- the punctuality (lateness or earliness in arriving at destination);

- the actual travel time;
- the reaction time (how much time is employed by respondent to make his/her choice).

Other information are stored in the data base referred to the answers of the respondents at the end of the simulation, like as the reason the respondents as chosen more frequently the actually most chosen route, the perceived information quality and the perceived utility of the information.

4.5 SUMMARY

By using the Travel Simulator Laboratory an experiment has been prepared according to the adopted Stated Preference approach. In this experiment the effect of two different information kinds (prescriptive and descriptive) has been tested; moreover different inaccuracy levels have been considered. Four inaccuracy levels have been tested, from the most accurate level to the most inaccurate.

Therefore the experiment has been designed by considering four scenarios for descriptive information and four scenarios for prescriptive information.

Several respondents have been involved in the experiment, 20 for each of the 8 scenarios. The design of the experiment took about 3 months (comprehensive of technical phase of implementing the experiment within the TSL platform). The running of the experiment took around 1 month. At the end all data were stored in the database.

5 DATA ANALYSIS

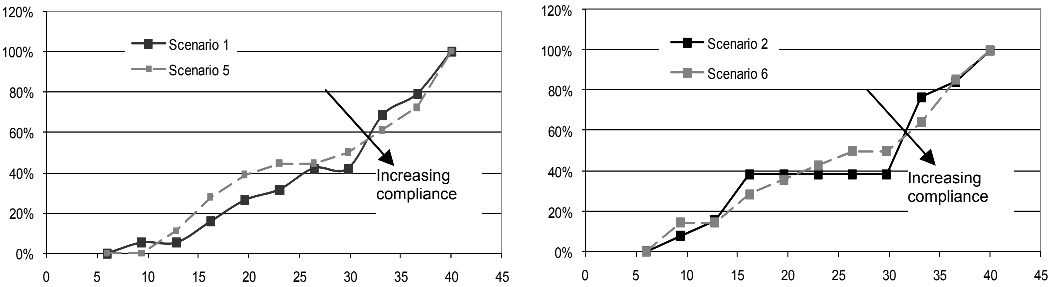
5.1 AGGREGATE ANALYSIS ON COLLECTED DATA

At the end of the survey the final database has been inspected in order to discover and remove manifestly inconsistent respondents. Finally, the target of 160 valid respondents (20 per scenario) has been, as desired, reached, summing up to 6400 valid records (each respondent has been asked to simulate a sequence of 40 successive simulated days). The resulting database has been subject to some preliminary analyses, aimed at understanding aggregate respondents' behaviours.

Here and in the following a traveller/respondent is said to be *compliant*, at a given day, if his/her behaviour with respect to route choice is *concordant* with ATIS information; concordance, in turn, is differently defined in case of descriptive or prescriptive ATIS. In case of descriptive ATIS the traveller is here defined to be concordant if the route he/she actually chooses has an ATIS-estimated travel time not higher than other routes; in case of prescriptive ATIS, the traveller is said to be concordant if he/she actually chooses the route suggested by the system.

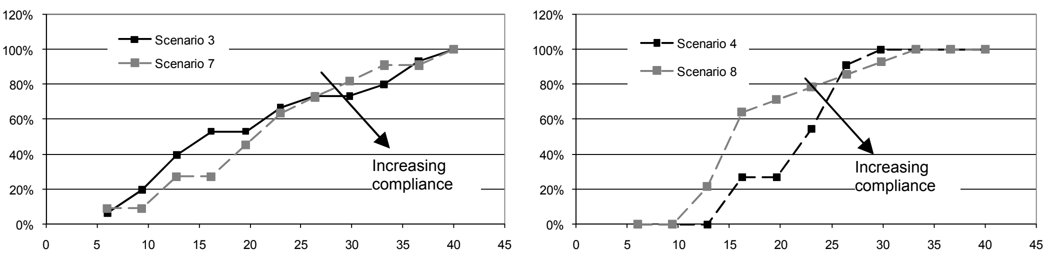
One may argue that compliance and concordance should be considered as two different things; in fact, a traveller could be observed to be concordant not only because he/she *trusts* in the system (is compliant) but also because he/she have had chosen that route in any case, because of his/her own considerations, independently on ATIS indications. Here we are not interested in a discussion on differences and similarities between concordance and compliance and we will consider these two terms as reciprocally substitutable; this is justified by the fact that the effect of both concordance and compliance in terms of travel choices are exactly the same, even if induced by different causes.

A first analysis has been carried out in order to assess if different kinds of ATIS (descriptive or prescriptive) can have effect on travellers' compliance. This has been made by comparing the cumulative compliance distributions of coupled descriptive and prescriptive scenarios. It has been already noted that descriptive scenarios 1, 2, 3 and 4 respectively correspond to prescriptive scenarios 5, 6 7 and 8, provided that they refer to the same levels of accuracy and to a consistent way of designing the information. Comparisons are shown in following Figure 26 and Figure 27, where cumulative compliance distributions are depicted.



(a) accuracy level 1 *(b) accuracy level 2*

Figure 26. Effect of ATIS type (descr. vs. prescr.) on compliance, accuracy levels 1 and 2.

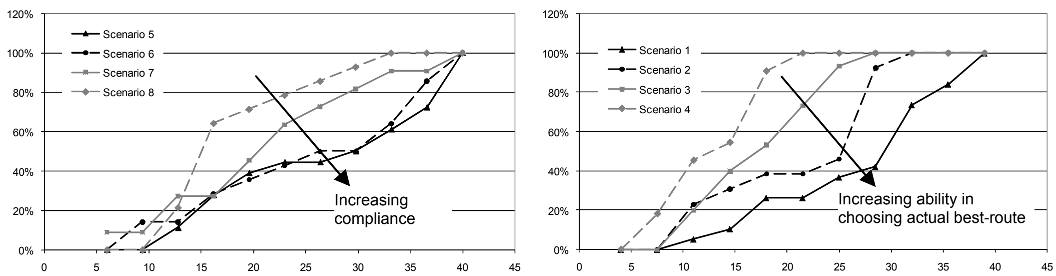


(a) accuracy level 3 *(b) accuracy level 4*

Figure 27. Effect of ATIS type (descr. vs. prescr.) on compliance, accuracy levels 3 and 4.

The cumulative compliance distribution at a given point (x, $F(x)$) represents the fact that no more that $F(x)$ of the respondents of the considered scenario are compliant x times (over 40) and $1\text{-}F(x)$ of the respondents are compliant more than x times. To fix the ideas on how cumulative distributions should be interpreted, consider two distributions A and B; assume, for instance, that $F_A(x=15)=30\%$ and $F_B(x=15)=65\%$, this means that more than *70%* of the travellers are compliant at least 15 times for distribution A and only *35%* of the travellers are compliant at least 15 times (over 40) for distribution B. In other terms, distributions that later (closer to *x= 40*) reach *100%* are representative of more compliant overall behaviours. Figure 26 and Figure 27 show that an effect of ATIS type cannot be excluded but is not particularly evident.

The second aggregated analysis has been carried out with reference to a preliminary (coarse) investigation of the impact of the accuracy on the compliance. Once again, the analysis has been performed by using the cumulative compliance distributions (as defined before); the analysis is here reported with reference to the prescriptive case (the descriptive one behaves similarly). It has been already noted that the scenarios only differ because of the ATIS accuracy level, so that the *scenario ID* can be used as a proxy of the accuracy level. Figure 28.a reports the result of the analysis; it can be noted a not negligible effect of the accuracy and, as expected, less accurate scenarios exhibit curves which are above the ones related to more accurate scenarios, this means that the compliance increases as the accuracy increases.



(a) effect on compliance (prescr. cases) *(b) effect on best-route (descr. cases)*

Figure 28. Comparison of effects of accuracy levels.

Figure 28.b shows the result of the third aggregate analysis, aimed to gather indication on if ATIS accuracy has an effect on the ability of travellers in choosing the actual minimum travel time route (which is unknown to travellers at the moment of their choices). The result here reported refers to the descriptive case (the prescriptive case gives the same indications); the effect can be said to be evident and it is as expected: the ability in choosing the actual minimum travel time route increases according to accuracy.

Some more formal aggregate analyses have been carried out also by applying non parametric statistical test[10]. In particular, a *Kruskall-Wallis* testing approach has been applied to the following null-hypotheses on the only casual accuracy level (approximated by the scenario ID) influence on: (i) the travellers' compliance; (ii) on the ability to choose the route of minimum actual travel time; (iii) on the propensity to choose the most reliable route (route 2). In all cases, accuracy can have a only casual effect if the null-hypotheses can be rejected; this means that we hope to obtain by the tests as high as possible chi-square values and as low as possible asymptotic significance values. The results of the tests are shown in Table 4, they suggest that the accuracy plays a role (even if not so evident with respect to the propensity to choose the most reliable route in descriptive scenarios) and that further disaggregated analyses worth.

Table 4

Kruskall-Wallis tests

Effect	Descriptive ATIS		Prescriptive ATIS	
	Chi-square	Asymp. Signif.	Chi-square	Asymp. Signif.
Compliance	7.312	6.0%	5.666	13.0%
Choose Best Route	620.356	0.0%	573.748	0.0%
Choose Reliable Route	3.400	33.4%	8.035	4.5%

5.2 SUMMARY

The research aims at investigating and modelling the effect of the inaccuracy information on travellers' behaviour. For these reason the obtained data have been submitted to some preliminary analysis.

These have been carried out in aggregate terms and have been based both on qualitative judgment of collected data and on more formal non –parametric statistical test. In particular, non-parametric tests have been aimed at checking if data exhibit a non casual effect of information accuracy on travellers' route choice.

[10] See in Appendix A.2 for more information on the applied test.

The preliminary analysis have shown that it worth a deeper disaggregate investigation of the phenomena and, in particular, the calibration of an explicit compliance model is likely to be an useful and successful activity.
The disaggregate calibration will be the subject of the next chapter.

6 A DISAGGREGATE MODELLING FRAMEWORK FOR THE COMPLIANCE

Aggregate analyses have shown that it worth a deeper disaggregate understanding on how (not only if, as in the aggregate analysis) ATIS accuracy influences travellers' behaviour, in particular with respect to compliance, which is the main topic of this work.

A model is specified and calibrated in which travellers' compliance is the target variable and ATIS accuracy one of the independent (explicative) variables. The model is framed within the theory of discrete choices, where the choice alternatives are defined as to be or not to be compliant.

Provided the nature of the experiment and of the gathered data, the calibration of the discrete-choice model is carried out within a *panel data* approach; in particular, the choice utilities of the alternatives are assumed to be distributed according to identical and independent Gumble variables (as in traditional multimomial-logit models), but a further error component term (normally distributed) is added in order to capture the correlation across days of the respondents' answers (see McFadden and Train, 2000).

6.1 MODELING APPROACHES: SWITCHING AND HOLDING

Two alternative modelling approaches have been tested; the first one refers to a *switching* formulation (see Figure 30), while the other refers to an *holding* one (Figure 28).

The holding approach is the most straightforward; the compliance status at day *t* is directly computed as a function of the model attributes. Each day the traveller chooses if to be or not to be compliant; this depends on the attributes observed in previous days and/or on characteristics of the dispatched ATIS information at the current day

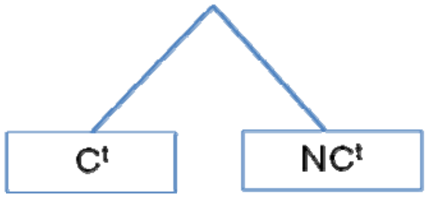


Figure 29. Holding approach to compliance modelling

Figure 30. Switching approach to compliance modelling

The formulation of the holding approach is:

$$\lambda^t = Prob[C^t = 1] = 1 - Prob[C^t = 0] = P(C^t) = 1 - P(NC^t) \qquad (1)$$

where:

- λ^t; is the probability to be compliant at day *t*.
- C^t; is the compliance status at day *t* (*1* if compliant, *0* otherwise) and, obviously $NC^t=1-C^t$;
- $P(C^t)$ and $P(NC^t)$ are respectively the probability to be compliant and to be not-compliant at day *t* and obviously $P(NC^t)= 1-P(C^t)$.

A deeper introduction worth for the switching approach; the probability a traveller is compliant [not compliant] at a given day *t* can be computed through the probability he/she is compliant [not compliant] at previous day *t-1* and he/she remains in his/her compliant [not compliant] status or switches to the opposite one.

Following equations from 2 to 6 formalize the switching approach:

$$P(FCTDt) = 1\text{-}P(SCt) \quad (2)$$

$$P(FDTC^t) = 1\text{-}P(SD^t) \quad (3)$$

$$P(SW^t) = P(FCTD^t) + P(FDTC^t) \quad (4)$$

$$1\text{-} P(SW^t) = P(SC^t) + P(SD^t) \quad (5)$$

$$\lambda^t = Prob[C^t = 1] = C^{t-1}\,(1\text{-}P(SW^t)) + (1\text{-}\,C^{t-1})\,P(SW^t) =$$
$$= C^{t-1}\,[P(SC^t) + P(SD^t)] + (1\text{-}\,C^{t-1})\,[P(FCTD^t) + P(FDTC^t)] \quad (6)$$

where

- $FCTD^t$ $[FDTC^t]$; is the choice to switch to not-compliant [compliant] at day t being compliant [not-compliant] at day $t\text{-}1$;
- SC^t $[SD^t]$; is the choice to stay compliant [not-compliant] at day t being compliant [not-compliant] at day $t\text{-}1$;
- $P(FCTD^t)$, $P(FDTC^t)$, $P(SC^t)$, $P(SD^t)$; are the probabilities to switch to not-compliant, switch to compliant, stay compliant and stay not-compliant at day t;
- $P(SW^t)$; is the probability to change the compliance status (from compliant to not compliant or from not-compliant to compliant) from day $t\text{-}1$ to day t.

In practical terms, in order to specify and calibrate the compliance model within the switching approach, two separate binary-choice sub-models have to be specified and calibrated, one relates to the choice of staying compliant or switching to not-compliant (if compliant at previous day), the other to the choice to stay not-compliant or to switch to compliant (if not-compliant at the previous day).

6.2 DISCRETE CHOICE MODELS

Both the holding and the switching approaches of the previous paragraph can be framed within the discrete choice theory where the models represent the choice of individual travellers. The discrete choice methods is the most diffused methodology in analyzing and predicting travellers' behaviours.

In particular, best results are obtained if the discrete choice theory is implemented through random utility models. Random utility methods are based on the theory that decision makers have a perfect discrimination capability and the choice utilities can be represented by random variables.

The utility that the individual *n* associates to the alternative *i* of the choice set is generally modelled as $U_{in} = V_{in} + \varepsilon_{in}$, where V_{in} is the deterministic part and ε_{in} is the random term. According to the utility theory, traveller will chose among the alternatives by preferring the one of highest utility.

The probability to choose the alternative *i* in the choice set I_n is computed as: $P(i/I_n) = P[U_{in} \geq U_{jn}, \forall j \in I_n, j \neq i] = P[U_{in} = \max U_{in}]$. The deterministic term of the utility (V_{in}) is expressed as a function of the *attributes* (which are the independent variables used to describe the utilities) and by the decision maker' s characteristics. In almost all of the practical implementations of the discrete choice theory the systematic part of utility is computed as a linear combination of attributes: $V_{in} = \sum_k \beta_k x_{ink}$, where $\boldsymbol{X_{in}} = [x_{in1}, x_{in2} .... x_{ink} ...]^T$ is the vector of the attributes of alternative *i* for decision maker *n* and $\boldsymbol{\beta_{in}} = [\beta_{in1}, \beta_{in2} .... \beta_{ink} ...]^T$ is the vector of the so called reciprocal substitution parameters.

Different hypothesis about the random residual distribution (Cascetta, 2001) lead to different well know random utility models, like as the Multinomial Logit (MNL), the Nested Logit (NL) and the Multinomial Probit (MNP) (see Appendix A).

Random utility models can be interpreted by considering that the utility (***U***) associated by decision makers to choice alternatives cannot be fully observed by the analyst. The analyst is able only to observe a part of the utility (***V***) as well as to explain how the observed part depends on an appropriate set of attributes (***X***); parameter ($\boldsymbol{\beta}$) account for preferences of decision makers in weighting the attributes within the perceived utility. The unobserved part of the utility is taken into account by the random residual (ε). The random residual (or dispersion term or, also, error component) accounts not only for unobservable discrepancy of the perceived utility with respect to its explained par but also, in case of appropriate use of the model, for the dispersion of decision makers attributes values and preferences with respect to the explained average part of observable utility.

Each of the previous modelling approaches (MNL, NL, MNP) is characterized by strength and weak points. In the Logit family the random residual is considerate to be distributed according to Gumble based specifications while in case of Probit models the random residual is distributed according to a Multivariate Normal specification.

The main advantage of logit-type approaches is their analytical treatability, while Probit type specifications have to be solved by means of numerical simulation methods.

In the multinomial logit (MNL) tree main restrictions are generally identified: the first one is related to the property of independence from irrelevant alternative (IIA); the second one is related to the model incapability in describing the random taste variation while only the systematic taste variation can be handled with respect to the observed variable; the third one is related, in repeated choices contexts, to the fact that the model can be only adopted if the choices are independent over time, otherwise IIA induce some trouble in properly modelling across-time phenomena.

The nested-logit approach can be framed in the more general family of generalized extreme value (GEV) models. In this case a covariance structure can be imposed among alternatives, allowing to take into account (at least to some extent) quite complex phenomena in real world correlations among alternatives; the IIA property and its consequences results to be mitigated.

The most advanced model is represented by the MNP. In fact the nested logit introduces the correlation among the alternatives but the structure of the variance covariance matrix, is flexible only to some extent. The variances are identical and the covariance are obtained by framing the alternative within a (potentially cumbersome) tree-of-choices architecture. Rather, the Probit model allows for an as general as needed structure of the variance covariance matrix. However, in case of Probit, as already stated, numerical simulation techniques have to be adopted in order to compute the choice probabilities of the alternatives (often the Monte-Carlo method).

A further modelling approach can be found in literature; it's the so-called dynamic kernel logit (or mixed logit) one. The peculiarity of this model is a mixed structure for the random residual. By following the so called *error component* interpretation of the mixed logit approach, the random residual can be considerate as to be composed by two parts: a Gumble i.i.d. one (such as in the MNL) and a normally distributed term. A typical implementation of the mixed logit approach is in case of repeated choice contexts, where the unobserved part of utility is needed to be no more independently distributed across time (as in case of MNL).

It is evident that the mixed-logit approach well fits our experiment where a classical panel data context holds. For this reason the mixed-logit approach is adopted for our modelling framework, as more deeply discussed in the following paragraph.

6.2.1 The covariance structure of the proposed model

Both the switching and the holding models previously introduced are here framed within the (random) discrete choice theory. As widely known (and as already stated) this theory assumes that the probability of choosing between different (discrete) alternatives can be represented as the probability that the *utility* associated by the decision-maker to the chosen alternative is greater than the ones associated to the others. The probabilistic approach accounts for the fact that the analyst is not able to observe (and to use in simulation) the utility associated to the alternatives. Rather, the analysts is able to include in the model in a deterministic way only the *observable* count-part of the utility which, in turn, is *explained* by some *choice attributes* (playing in the choice model the role of the independent variables). These attributes are combined in order to build-up the observed part of the utility with appropriate parameters that can be referred to as *reciprocal substitution parameters* (or *weights*) of the attributes of the models. The unobserved count-part of the utility is included in the model in the form of a random dispersion (with zero mean) that can be referred to as *the error-component* of the model. The most accepted way to formalize the previous, at least in the case that the reciprocal substitution parameters combine the attributes in a linear way and are identical for all decision-makers, is:

$$U_j^n = \beta^T{}_j X_j^n + \varepsilon_j^n \qquad (7)$$

Where U_j^n, X_j^n and ε_j^n are the unobserved (perceived) utility, the vector of attribute values and the error-component (that is a random term) associated to the generic decision-maker n and to the generic alternative j and β_j is the vector of reciprocal substitution parameters related to the alternative j; obviously, $\beta_j\ X_j^n$ is the observed count-part of the utility, while ε_j^n is the unobserved one.

The calibration process is generally intended as the evaluation of the values of the model parameters (βs) and of the parameters of the error-component distribution; it is typically performed (maximum likelihood approach[11]) by maximizing the probabilities estimated by the model for the observed choices (choices are observable, not utilities) of a proper sample of decision-makers; this process will be dealt with in the next sub-section against the sample collected by way of the experiment described in the previous chapters.

---

[11] The maximum likelihood approach is more clearly detailed in Appendix F

In our case, the specification of the model requires some more details, related to the fact that each decision-maker is observed to make a sequence of choices, simulating his/her compliance (uncompliance) across a sequence of days. In such a case the formalization of the utility for each choice can be:

$$U^n_{j,t} = \beta^T_j X^n_{j,t} + \varepsilon^n_{j,t} \qquad (8)$$

where *t* represent the generic day of the sequence.

The specification we adopt for the error-component count-part of the utility is:

$$\varepsilon^n_{j,t} = \phi^n_j + \eta^n_{j,t} \qquad (9)$$

where $\eta^n_{j,t}$ is assumed to be identically and independently distributed (across decision-makers, days and choice alternatives) according to a Gumble distribution with zero mean and ϕ^n_j is assumed to be identically and independently distributed across respondents and alternatives as a normal variable (with zero mean). The previous structure of the unobserved utility introduce a (simple) structure of covariance between different choices of the same respondent in different days (it depends on the variance of the ϕ^n_j distribution.

The previous specification deals with the so called panel-data structure which is implicit in our experiment, for this reason the term ϕ^n_j of the unobserved part of the utility will be referred to in the following as the *panel-data* count-part of the error-component. The main reason for which we have explicitly considered the panel-data correlation is that we will specify in the following some attributes which have a sort of dynamic structure (in the sense that their values depend, for instance, on the ATIS accuracy of several consecutive days); in such a condition it make sense to assume (and test in calibration) that also the unobserved part of the utility has an inherent dynamic structure.

Our hypotheses on the structure of the error-component leads to one of the possible forms of the so called (an already introduced in paragraph 6.2) mixed-logit models. A further hypothesis could be introduced within a mixed-logit framework by considering that the *βs* actually are not deterministic parameters but are random variables distributed across decision-makers; this allows for introducing a sort of *taste-dispersion* across decision-makers. In our specification we have also done this assumption but it has been rejected during the calibration phase, provided that the distribution parameters of *βs* were not statistically acceptable in terms of significance.

6.3 SPECIFICATION, CALIBRATION AND EVALUATION OF THE PROPOSED MODEL

6.3.1 Specification: the attributes

Different sets of attributes have been tested in order to specify the binary switching sub-models and the holding model. In the following the description of the selected ones (on the base of the best performances in calibration) is enlisted; some of the attributes make sense only in case of descriptive ATIS, while others make sense also in case of prescriptive ATIS.

Attributes that make sense only in the descriptive case are:

- *DescriptiveInaccuracy*[12], is the sum of the square relative differences between the actual travel times and the ATIS-estimated travel times; it measures how much the supplied ATIS information is inaccurate;
- *ReliabInacc*, is the mean on previous two days of the Descriptive Inaccuracy, considered only if the informative system has been reliable in previous two days; it measures the fact that even reliable information can be accurate at different levels;
- *ProspectedGain,* is the relative difference between the shortest ATIS-estimated travel time and the second-best one; it measures the gain in terms of travel time prospected by the ATIS in following its information;

[12] How is computed the Descriptive Inaccuracy is explained in Appendix D.1

- *HighUncomplRisk,* is the ProspectedGain over the frequency in last 5 days the suggested route has been the best; it measures the fact that the systems induce (at a greater or lower extent, depending on the ProspectedGain) a choice which has been viewed by the travellers to likely have been the best one (so that being not-compliant is risky);
- *TooOptimisticInfo,* is the relative difference between the ATIS-estimated travel time of the shortest route and the average of the actual travel times of all suggested routes in previous 5 days, computed only if the average of the actual travel times is greater than the actual travel time of the route suggested today and if the suggested route is chosen at the previous day, otherwise the value is 0; it measures the fact that the estimated travel time of the suggested route is too much optimistic with respect to the experience of the traveller;

Attributes that make sense also in the prescriptive case are:

- *Reliability*[13], the computed value is 1 if the suggested route actually is the best route with respect to the actual travel times ; 0 otherwise; it measures if the information has been accurate enough to suggest the actually best route;
- *PrescriptiveInaccuracy*[14], is the relative difference between the actual travel times of the suggested route and the one of the route (shortest) that have had chosen by the respondent if he/she have had knowledge of the actual travel times; it measures how relevant have had (or has) been the error (if any) in choosing the route suggested by the ATIS;
- *PrescrInacc_In5*, is the average PrescriptiveInaccuracy computed over previous 5 days;
- *Discrepancy*, for each it is computed as the relative difference between the actual travel time of yesterday and the average of the actual travel times of the

[13] How is computed the Reliability is explained in Appendix D.3
[14] How is computed the Prescriptive Inaccuracy is explained in Appendix D.2

previous 3 days; it is the measure of a sort of on-average-perceived ATIS inaccuracy;

- *SuggRouteDiscrep*, is the *Discrepancy* of the suggested route over the sum of the discrepancies of all route;
- *SuggRouteIncr*, is the square of the relative difference between the yesterday actual travel time of the suggested route, and the average of the actual travel times of all suggested routes in previous 5 days;
- *NearInacc*, is the yesterday *PrescriptiveInaccuracy,* computed only if the system has been unreliable in the day before yesterday, otherwise the associated value is zero; it measures that two consecutive negative performances effect the traveller and, more precisely, an unreliability occurrence alerts the traveller who is induced the day after to carefully look at inaccuracies;
- *Is2AndReliab*, its value is 1 if the suggested route is the most reliable route (route 2 on the base of our actual travel time) and the system has been reliable in all the previous 4 days; it accounts for the joint occurrence that the more reliable route is suggested and the ATIS accuracy level is high (so that the more reliable route, if suggested, also is likely to be the best one);
- *RecovReliab*, its value is 1 if the informative system has been unreliable two days before yesterday but then has recovered its reliability for the two successive days, 0 otherwise;
- *AtLeastOneUnrel*, its value is 1 if the informative system has been unreliable one or more times in the previous 3 days, zero otherwise;
- *Consec*, is the number of times in previous 5 days in which traveller chooses the same route if this route is the one suggested today by the system, zero otherwise; it measures the fact that the suggestion fits traveller's consolidated preference;

- *FreqChosen,* is the frequency the suggested route at current day has been chosen in previous 5 days; note that this attribute differs from the previous *Consec* attribute both because it is computed also if the route suggested today has not been chosen in the previous days and because it accounts also for not necessarily consecutive identical route choices;
- *FreqCompl,* is the frequency with which the traveller has been compliant in the previous 5 days; it measures a sort of habit/inertia;
- *NotPreferredSugg*, if the suggested route is not the one chosen yesterday, it is the frequency over previous 5 days the suggested route has not been the actually best one; it measures the average poor performances experienced for the suggested routes if different from the route the traveller has chosen yesterday (likely, it is not the suggestion the traveller would have preferred).

6.3.2 Calibration: procedure and results

The calibration has been performed for both modelling approaches (switching and holding) and separately for the two ATIS-information type contexts (descriptive and prescriptive); it has been carried out by means of the BIOGEME software (Bierlaire, 2007); all calibrations have been performed with 1000 draws (per algorithm iteration) of the normally distributed panel-data-related error component.

A particular question arises in the calibration procedure for what concerns the switching approach. By looking at equation 6, it is evident that the probability to be compliant or not-compliant at day t depends on the compliance (not compliance) status at day t-1; the status at day t-1 can be referred to as the *acquired compliance* at (the beginning of) day t (AC^t, which value is 1 if the traveller starts his/her choice from a being compliant status and 0 otherwise); in such a way equation 6 can be rewritten as:

$$\lambda^t = AC^t\,[P(SC^t) + P(SD^t)] + (1 - AC^t)\,[P(FCTD^t) + P(FDTC^t)] \qquad (10)$$

In theory, the computation of the acquired compliance should be consistent with its definition, according to the following equation 11:

$$AC^{t} = C^{t-1} \qquad (11)$$

In practice, the following equation 12 is advantageous for the suitability of standard discrete-choice-theory calibration processes.

$$AC^{t} = OC^{t-1} \qquad (12)$$

Where OC^{t-1} is the observed compliance (the one revealed in the survey) for day *t-1*. By using equation 12 the calibration consists in maximizing probability λ^t predicted by the model for all respondent for which the compliance status has been revealed at day *t* ($OC^t = 1$) and maximizing the probability ($1-\lambda^t$) for all respondent for which the not-compliance status has been revealed at day *t* ($OC^t = 0$). Use of equation 12 implies that the calibration process is performed through a so called *Input/Output* (I/O) *approach*, this is inherently inconsistent with the dynamic-process nature of the specified switching framework. It is evident that the inconsistent (I/O) calibration is here adopted for sake of simplicity and in order to allow for using standard calibration tools for discrete choice models (as BIOGEME). It is worth noting that the consequences of the I/O calibration should be carefully evaluated after the calibration process. The previous considerations on calibration inconsistency applies only to the switching model, being the holding one unaffected by the acquired compliance.

It is also worth noting that our calibration (as well as the whole modelling framework), both for the switching and holding approach, only considers a context in which traveller's preferences (and behaviours) are *consolidated*; because of that, we have employed in calibration only a sub-set of the survey data, which is the part related to responses given starting form day eleven ($t>10$). In fact, we assume that the first ten days are employed by travellers/respondents in order to build-up their behaviour. In other terms, we exclude from our analyses any behaviour-learning/forming phase and assume that the model formula and its parameters (βs) are constant over days (even if choices can change, changing the related attributes).

Table 5

Calibration results, descriptive ATIS (t-test values in brackets and italic)

	Switching Model				Holding Model	
	From Compliant		From Discordant			
ρ^2	0.358		0.309		0.333	
P-data parameter (Var[ϕ^n])	0.771 *(4.84)*		0.350 *(1.47)*		0.882 *(6.34)*	
	FCTD	SC	FDTC	SD	C	NC
ASA	1.12 *(3.15)*			1.56 *(5.24)*		0.85 *(3.07)*
ReliabInacc	5.20 *(2.51)*			5.42 *(2.20)*		5.76 *(3.59)*
TooOptimisticInfo	2.62 *(2.80)*			2.81 *(3.24)*		2.84 *(4.37)*
NearInacc	2.26 *(3.68)*			0.99 *(1.67)*		2.04 *(4.57)*
SuggRouteDiscrep	0.34 *(3.69)*					0.19 *(2.89)*
SuggRouteIncr	5.79 *(3.97)*					3.65 *(3.18)*
NotPreferredSugg				2.01 *(4.25)*		0.93 *(2.89)*
HighUncomplRisk		3.97 *(7.15)*	3.25 *(1.95)*		3.72 *(9.13)*	
FreqChosen		3.06 *(9.42)*	1.95 *(4.56)*		2.11 *(7.08)*	
FreqCompl		2.52 *(5.65)*	2.28 *(4.74)*		1.96 *(5.47)*	
Is2AndReliab		1.21 *(3.00)*	1.08 *(2.15)*		1.37 *(4.38)*	
RecovReliab		1.70 *(1.90)*	1.13 *(1.15)*		1.56 *(2.43)*	
Consec			3.33 *(4.28)*		3.25 *(5.47)*	

Table 5 shows the calibration results for the case of descriptive ATIS; the t-test values for the calibrated parameters (βs) are shown in italic and between brackets. Results are shown for both the holding and the switching approach, in this second case both the sub-models have been calibrated so that the results refer to the sub-models originated by an acquired compliance status of being compliant and an acquired compliance status of being not-compliant.

The ro-square statistics are not exciting but fully acceptable; the panel-data parameter represent the estimate of the variance of the error-component count-part (ϕ^n) used to simulate the correlations generated by the panel data structure of the experimental context; it is significant in all cases (almost significant for the From-Discordant sub-model), thus suggesting that the hypothesized structure of the unobserved utility makes sense. As already stated, in all specifications (also for the prescriptive ATIS case discussed in the following) a further attempt has been made to consider all the parameters (βs) as to be randomly distributed across respondents, but parameters of the distributions have been found in all cases to be not statistically significant.

Results for the descriptive case show that all βs of the models are significant; the Alternative Specific Constant (ASC) is consistent in magnitude with other parameters and does not explain to much part of the utility.

As expected, the choice to be not-compliant in the holding approach depends more or less on the same attributes that influence the choice to switch to not-compliant for the being-compliant acquired status and the choice to stay discordant for the being not-compliant acquired status; analogous considerations can be applied to the choice of being not-compliant in the holding model. The main attributes that induce a not-compliant choice (directly in the holding model or indirectly in the switching approach) relate, as expected, to the experienced (in)accuracy in past days (ReliabInacc and NearInacc), to the suspect the respondents has that the current ATIS information is erroneous (TooOptimisticInfo, SuggRouteDiscrep, SuggRouteIncrement) and to how much could be risky to follow (or not to follow) the supplied information (NotPreferredSugg). Among all the significant attributes that induce not-compliant choices, note that the experienced inaccuracy at the day before the choice day (NearInacc) and the average experienced inaccuracy related to two days before the choice one (ReliabInacc) both influence the choice but with different impacts; in particular, the respondents seem to be less sensitive to the nearest inaccuracy with respect to the relatively longer-term inaccuracy; in any case, as expected, the two inaccuracy attributes play in an additive way. In case of choice related to not-compliance, attributes RecovReliab and Is2AndReliab refer to the inaccuracy in past days, attributes FreqChosen and Consec refer to the suspect toward the current ATIS information (or, in this case, to the confidence toward the information) and attribute HighUncomplRisk refers to the risk for not following the supplied information; moreover the attribute FreqCompliant can be interpreted a sort of inertia (or habit) in being compliant.

Table 6

Calibration results, prescriptive ATIS (t-test values in brackets and italic)

	Switching Model				Holding Model	
	From Compliant		From Discordant			
ρ^2	0.435		0.340		0.391	
P-data parameter (Var[ϕ^n])	1.060 *(4.90)*		0.554 *(3.25)*		1.550 *(7.75)*	
	FCTD	SC	FDTC	SD	C	NC
ASA	1.87 *(4.66)*			1.47 *(5.77)*		1.07 *(4.17)*
NotPreferredSugg	1.08 *(2.76)*			1.94 *(4.31)*		1.48 *(4.78)*
AtLeastOneUnrel	1.10 *(1.19)*					0.87 *(1.50)*
PrescrInacc_In5	2.12 *(1.99)*					
SuggRouteIncr				1.67 *(1.96)*		0.80 *(1.16)*
FreqChosen		3.22 *(7.33)*	1.88 *(4.17)*		2.51 *(8.09)*	
FreqCompl		2.84 *(5.50)*	1.47 *(3.21)*		1.47 *(4.88)*	
Consec		1.52 *(2.00)*	3.12 *(5.07)*		2.64 *(5.42)*	

Table 6 shows the calibration results for the Prescriptive ATIS case. Also in this case the ro-square values are acceptable, as well as the significance (t-test) of the calibrated parameters. Parameters that are relevant for the not-compliant choice of the holding approach play a consistent role in the switching sub-models. Parameters PrescrInacc_In5 and AtLeastOneUnrel relate to inaccuracy; parameters NotPreferredSugg, SuggRouteIncr, FreqChosen and Consec relate to the suspect toward the current suggestion (or to the confidence toward it); also in this case, the habit/inertia toward the compliance status is taken into account by the parameter FreqCompl.

6.3.3 Evaluation of the calibrated models and further refinements

The ability of the calibrated models to reproduce the choices observed in the sample have been considered as a first issue in order to assess the performances of the calibration process. This has been done by using the *Percent-Right* indicator;

Table 7 and Table 8 show the result for to the switching approach, both in the descriptive and prescriptive case. For each of the information types the indicator has been computed with reference to the two sub-models (Switching From Concord and Switching From Discord) of the switching approach and also with reference to the Resulting Compliance that can be obtained via equation 10.

Table 7

Sample reproduction, Descriptive information, Switching Approach

		Observed Choices					
		Sw. From Concord		Sw. From Discord		Resulting Compliance	
		FCTD	SC	FDTC	SD	C	NC
Predicted Choices	FCTD/FDTC/C	57%	10%	69%	16%	83%	27%
	SC/SD/NC	43%	90%	31%	84%	17%	73%
% Right		57%	90%	69%	84%	83%	73%
Overall % Right		80%		78%		79%	

Table 8

Sample reproduction, Prescriptive information, Switching Approach

		Observed Choices					
		Sw. From Concord		Sw. From Discord		Resulting Compliance	
		FCTD	SC	FDTC	SD	C	NC
Predicted Choices	FCTD/FDTC/C	61%	10%	59%	12%	81%	22%
	SC/SD/NC	39%	90%	41%	88%	19%	78%
% Right		61%	90%	59%	88%	81%	78%
Overall % Right		81%		77%		79%	

The percent-right computation is based on use of equation 12 for the acquired-compliance in equation 10; this means that this task exactly fits the calibration and that an inconsistent I/O procedure has been also adopted here for the assessment of the calibration performances. Table 7 shows that the observed choice of switching from compliant to not-compliant (FCTD) is reproduced by the sub-model in 57% of the cases while is incorrectly estimated as a staying-compliant (SC) choice in 43% of the cases; the observed choice of staying-compliant is correctly reproduced in 90% of the cases and incorrectly estimated in 10% of the cases; it results that the percent-right is 57% for FCTD and 90% for SC; by considering the relative assortment of the observed choices, the overall sample reproduction of the switching approach in the descriptive case can be computed as to be 80%.

Other values in Table 7 and Table 8 can be similarly interpreted; they show that, both in case of descriptive and prescriptive ATIS, the switching approach reproduce with some difficulty the switching from one compliance status to the other (in both directions), while the staying phenomenon is very well reproduced. The authors' feeling is that this is due by the fact that almost all the explanatory attributes for the staying phenomenon have been revealed during the specification and calibration process, while one (or more) explaining attribute still is missing for the switching phenomena; however, the authors have not been able to discover the missing attribute(s), despite the great effort and concentration devoted to this task; future work will be surely done to this aim and suggestions and contributes will be welcome. This, if successful, will surely contribute to further improve the ro-square statistic of the calibration process. However, the results reached so far are not negligible and can be judged to be satisfying, in particular considering that the percent right of the compliance resulting by the joined application of the two switching sub-models is adequate (columns 5 and 6 of Table 7 and Table 8) in terms of reproducibility of the compliant (C) and not-compliant (NC) status.

Also the ability of the straightforward holding approach in reproducing the sample has been computed via the percent-right, for both the descriptive and the prescriptive case; the results are respectively shown in the first and second columns (Holding Compliance) of Table 9 and Table 10. The reproduction of the sample does not indicate substantial differences between the switching and the holding model; in case of prescriptive ATIS a little bit more accentuated tendency of the holding approach in overestimating compliance and underestimating not-compliance (with reference to the switching approach) can be noted. In other terms, it seems that the switching and the holding approaches can be considered to be equivalent; however, the switching approach has an inherent more explaining nature, as confirmed by the fact that it has allowed the authors to identify that the main problem is in reproducing the change of compliance status rather than in reproducing the staying phenomenon.

Table 9

Sample reproduction, Descriptive information, consistently evaluated resulting compliance, direct compliance model (holding model) and consistently calibrated resulting compliance

		Observed Choices					
		Holding Compliance		Consistent Evaluation		Consistent Calibration	
		C	NC	C	NC	C	NC
Predicted Choices	C	84%	29%	84%	30%	84%	25%
	NC	16%	71%	16%	70%	16%	75%
% Right		84%	71%	84%	70%	84%	75%
Overall % Right		78%		78%		81%	

Table 10

Sample reproduction, Prescriptive information, consistently evaluated resulting compliance, direct compliance model (holding model) and consistently calibrated resulting compliance

		Observed Choices					
		Holding Compliance		Consistent Evaluation		Consistent Calibration	
		C	NC	C	NC	C	NC
Predicted Choices	C	84%	29%	81%	23%	82%	21%
	NC	16%	71%	19%	77%	18%	79%/
% Right		84%	71%	81%	77%	82%	79%
Overall % Right		78%		79%		80%	

Another validation test has been carried out by computing the ability in reproducing the sample within a consistent simulation approach. In this case the computation of the percent-right indicator has been performed by using equation 11 for the evaluation of the acquired compliance; this means that the acquired compliance at a given day is the model-estimated compliance at the previous day and not the observed one, consistently with the dynamic nature of the switching approach. Moreover, the consistent simulation allows to evaluate the performances of the proposed models versus a sample that is, in some sense and at a limited extent, different from the one employed in the calibration process. In fact, the acquired compliance of the consistent simulation at a given day (C^{t-1}) can be different from the acquired compliance in calibration for the same day (OC^{t-1}). Results of the consistent evaluation are reported in the third and fourth columns of Table 9 and Table 10, respectively for descriptive and prescriptive ATIS; in both cases the model performances are quite similar (if not practically identical) to the inconsistently evaluated sample reproduction resulting from the joint application of the switching sub-models (see columns five and six of, respectively, Table 7 and Table 8). This comforts the authors in the opinion that the specified and calibrated models are robust and that the fundamentals of the travellers' behaviour have been understood; moreover, the impact of the inconsistency of the I/O calibration seems to be negligible.

In order to further confirm that the I/O calibration has been a not biasing approach, a further test has been made; the values of the parameters have been adjusted by re-calibrating the already calibrated models in a consistent way. This means that equation 11 has been applied to the evaluation of the acquired compliance of equation 10 during re-calibration. The re-calibration process has been performed by using the utilities of a commercial spreadsheet and with reference to a likelihood maximisation technique (but keeping the value of the variance of the panel-data-related error-component as to be fixed). The results of the re-calibration process in terms of parameters values (βs) are so negligibly different from the case of inconsistent calibration that are here not reported; the performances of the consistent recalibration in terms of reproduction of the sample are shown in columns five and six of Table 9 (descriptive case) and Table 10 (prescriptive case); as expected the percent-right increases, but at a negligible extent.

6.4 SUMMARY

In the chapter the final phase of the research is presented. In particular, it is proposed a modelling architecture for the simulation of the effects of the information accuracy on travellers compliance. The proposed architecture is carefully calibrated and research results are described and discussed. A modelling architecture has been defined by considering two different approaches: the first one (holding model) refers to the current state (compliant or not complaint), the second one (switching state) is obtained by combining the observed compliance status at the previous day and the choice made at the current day. The models have been developed according to the discrete choice theory. In particular, a mixed logit for panel data model has been specified and calibrated. The model has been chosen in accordance to the type of data gathered and to the experiment configuration. The database is composed by the answers made repeatedly by the same respondents, so that the answers can be considered as to be correlate over time for every respondent. For this reason the adopted modelling approach is the error- component interpretation of the mixed logit. The proposed models are also calibrated for the two different kinds of information, according to the results obtained in the preliminary step of the non parametric statistical test analysis (different information kind have a different effect on travellers' compliance). In the model specification the inaccuracy attributes are involved and preliminarily defined and formalized. In particular the attributes of inaccuracies can be distinguished in four groups: one referred to the ability of the informative system to dispatch the correct information (to the respect of the network performances the suggestions can be considerate congruently or not with the actual travel times and the Reliability has been defined in this way); the second group is referred to the ability of the informative system (in case of descriptive information) in making a correct estimation of the actual travel times therefore a mathematical expression of the eulerian distance has been adopted and a descriptive inaccuracy has been computed; the third group is the prescriptive inaccuracy evaluation, referred only to the prescriptive information, in which the different probability to chose every route to the respect of the actual travel times and to the estimated travel times are compared; the last group is identified with some frequency attributes referred to the reinforcement of the respondents' flavour, on the base of the good/bad experiences made by travellers' in previous choices.

The statistics referred to the calibration results are analyzed and discussed and a validation procedure has been made and described. The validation has been carried out with reference to both an approximate (static) interpretation of the proposed switching approach and a rigorous dynamic interpretation. Moreover, both the cases of carrying out the validation by considering and not considering the panel- data related error component have been deal with.

Results can be considerate fully satisfying for all performed validations.

7 CONCLUSIONS, REFLECTION AND FUTURE WORK

The chosen topic of research has been identified staring from the existing analysis of the assignment model frameworks in case of network in which travellers are provided with information.

The following Figure 31 depicts a comprehensive modelling framework for assignment models in which all components (and resulting interaction) are fully exploded in order to deal with ATIS application. The figure can be considered as an holistic enhancement of Figure 5 in chapter 1.

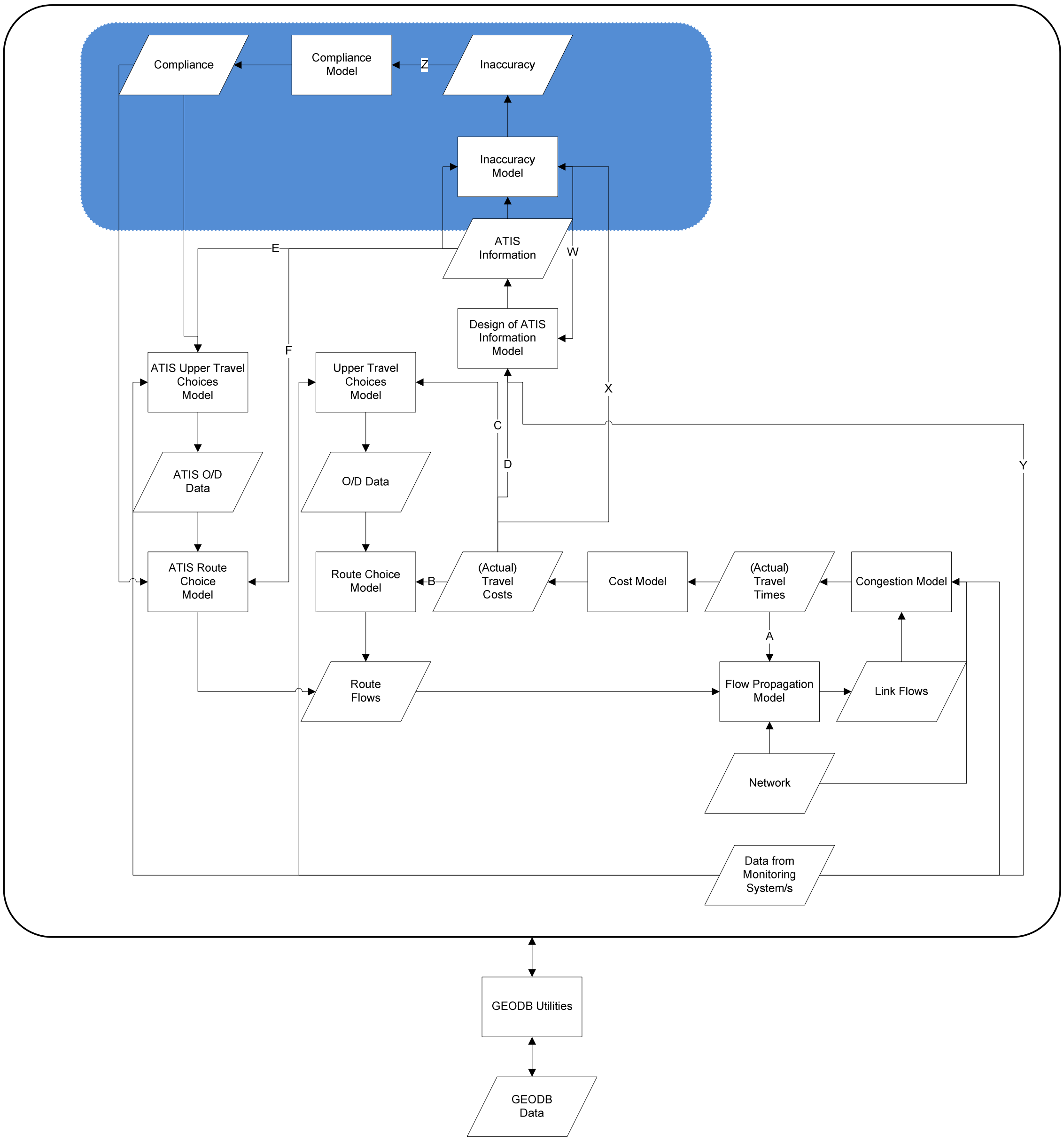


Figure 31: Framework of dynamic assignment model in ATIS context

Figure 31 can be explained by considering the following relations expressed by the arrows, to be more precise:

- **Arrow A**: its presence implies that flow propagation model is within-day-dynamics (with or without en-route guidance);

- **Arrow B**: its presence implies a supply problem (assignment), that can be both of equilibrium or day-to-day dynamics;
- **Arrow C**: its presence implies the elastic demand in assignment model;
- **Arrow D**: its presence implies that ATIS provides dynamic information;
- **Arrow E**: its presence implies that ATIS affects the depart time, the destination;
- **Arrow F**: its presence implies that ATIS affects path choices;
- **Arrow X**: it's present for planning models of ATIS information (predictive and/or reliable);
- **Arrow W**: it's present for planning models of ATIS information (predictive and/or reliable);
- **Arrow Z**: it's present for models in which the compliance is considered elastic (it's a function of the reliability of the information);
- **Arrow Y**: it indicates the presence of an ATIS system transmitting instantaneous information (by sensor devices).

The dashed box in Figure 31 focuses to a proper modelling of the accuracy role in ATIS applications. The review of the literature has shown that the dashed box in Figure 31 is one of the less mature and enhanced parts of the whole modelling framework, even if it is crucial for a proper simulation of ATIS impacts on transportation systems. All the previous has induced to focus this PhD thesis on the development of a detailed model aimed at properly defining (and computing) (in) accuracy attributes and evaluating their effects in terms of travellers' compliance with information.

The integration of the proposed modelling framework for accuracy and compliance within the whole architecture of Figure 31 is one of the more relevant tasks for further researches. However, some further enhancements (or at least some further consolidations of the obtained results) related to the compliance modelling issue still worth.

By the work already done we can make a conclusion that it exists an effect of the ATIS accuracy on travellers' choices and, particularly, on their compliance. Modelling such an effect or, more generally, modelling the travellers' compliance in an elastic way is not a trivial task.

Here we have presented some attempts that have lead to satisfactory results and that encourage toward further analyses and refinements.

The proposed modelling framework has been specified and calibrated against a SP survey and all the findings show that it is consistent and robust. Validation of the results has been performed under different modelling hypotheses; in particular, issues related to the consistency with the inherent dynamic nature of one of the approaches here presented have been appropriately addressed.

Some future research directions can be identified among others. For sure, the proposed modelling framework should be further confirmed (and possibly enhanced) against more experimental data (possibility, RP ones). Moreover, the presented approaches refer to a consolidated travellers behaviour and some attempt should be made for understanding and modelling also the phenomena related to the learning-phase of the travellers' behaviour. Finally, some further analyses could worth in order to explicitly understand travellers' attitude or aversion to risk, provided that being compliant with ATIS represents a risk in an uncertainty context.

But, as already stated, the pretentious future task can be identified as the integration of the calibrated model in a framework of assignment model as described in previous Figure 31.

ACKNOWLEDGEMENTS

Quando penso alla vita, penso ad un lungo viaggio in moto, fatto di soste in luoghi d'occasione, di incontri casuali, di profumi e sapori nuovi che il vento ti porta quando sei in sella e sfrecci senza meta o senza sapere neanche con precisione il nome dello scorcio che ti ritrovi all'improvviso ad ammirare ammaliato. Ogni tappa ha un senso diverso che ti si radica dentro e che aggiunge alla tua persona un nuovo tassello. Nel tempo ho maturato che buona parte "di noi" è fatta delle persone che abbiamo avuto la fortuna di incrociare lungo il viaggio, di ascoltare e che hanno avuto il potere di regalarci i loro occhi per guardare al mondo. Un'altra delle tappe del mio viaggio volge al termine e oltre a lasciarmi un prodotto scientifico, una consapevolezza intellettuale diversa, una figura professionale molto più matura e critica, mi lascia dentro molti rapporti umani importanti, ciascuno con un ruolo ed una identità diversa dentro di me.

Un cammino rassicurato dalla presenza della mia famiglia: dei miei genitori, dei miei due angeli custodi Pasquale e Bruno, e di Marghe tutti dotati di sovrumana e straordinaria pazienza, che mi hanno sostenuta, accettando e condividendo l'insensatezza di questa mia scelta; ringrazio mia madre a cui devo tutto, e molto più di tutto, ringrazio mio padre con cui avrò sempre piacere di affogare in profondi litigi ed incomprensioni, e soprattutto mi lascerò raccontare di politica; ringrazio mia nonna per "ogni cosa che mi ha lasciato e per ogni cosa in cui é rimasta".

Nel mio Cuore porto dei riferimenti importanti: gli amici di sempre, in particolare Rosalba che con la sua dolcezza ed il suo carisma mi ha fatto da sostegno nei momenti peggiori dimostrandosi una vera cara amica; una complicità nata oltre confine, Barbara che con me ha condiviso gli interminabili 7 mesi di Van Hasseltlaan 314 ed una folle notte in bicicletta; *tutti* i preziosissimi amici dell'università che con me hanno trascorso queste interminabili giornate: davvero

grazie per esserci *sempre*, in particolare sono grata "all'ing. Fulvio" nei riguardi del quale provo profonda stima e ammirazione, per la discrezione, la riservatezza, la concentrazione e quella "spanna in più" che ha fatto di lui la mia seconda guida scientifica in questo percorso di formazione!!
Porto nel Cuore tutti gli "amici" che hanno avuto la sfortuna di incontrarmi e che hanno invece la fortuna di vedermi molto raramente e non mi basterebbero pagine per elencarli: quelli che mi hanno dato più di una spalla nei momenti peggiori, quelli che con me condividono "l'ideologia", quelli con cui condivido "l'inanimato ottone"ed il jazz e che il jazz lo intonano e lo plasmano, quelli con cui condivido *l'Amore* nelle sue molteplici sfaccettature…
Un tassello bello grande di me porta il Sogno, il Sorriso, l'Ottimismo e il senso critico e costruttivo nei riguardi della società che mi ha lasciato il prof. Luigi Di Lascio…
Un ringraziamento più professionale va al gruppo di ricerca della Technische Universiteit di Delft: ai proff. Henk Van Zuylen e Serge Hoogendoorn, ringrazio Francesco Viti per lo straordinario affetto oltre che per la professionalità, ringrazio infine Enide Bogers, che nonostante le incomprensioni ricorderò sempre per 10 minuti di magia sotto la neve!
Ed ora che credo di avere ringraziato più o meno tutti, ringrazio dal profondo dell'animo chi con me ha scritto questa tesi a "quattro mani" (odiandomi per la mia incapacità nello scrivere in inglese!!): Cino. Come al "solito" ringrazio il Professore per avermi fatto da *Maestro*, per avere avuto la pazienza di sopportare e modellare la mia cocciutaggine, per avere cercato sempre il tempo per me, per avermi dedicato tutto se stesso e ogni briciolo che ne rimanesse. Ammiro di lui la dedizione vera che ha per l'insegnamento e la ricerca, la correttezza, la bravura, la precisione… lo ringrazio per avermi insegnato a comprendere e risolvere i piccoli ed i grandi "ostacoli" senza mai perdermi d'animo, con concentrazione e pazienza, per avermi fatto capire che per fare questo "mestiere" non si può che

essere animati da una passione profonda e che anche senza straordinarie doti intellettive è possibile addentrarsi con temperamento nel mondo della ricerca; lo ringrazio per la fiducia (cosa della quale forse si sarà pentito!), per avermi guidata in questa sfida, insegnandomi l'umiltà prima di tutto e per avermi aiutata a tirare fuori una sicurezza ed una sfrontatezza che non avrei mai pensato di acquisire altrimenti, forse perché in fondo più passa il tempo e più riconosco che *è il peggior nemico che io potessi incontrare su un campo di battaglia*. Sono poi grata all'Amico e Compagno Cino, che con straordinaria pazienza si è sopportato davvero 1 metro e 58 di *pessimo* scherzo della natura!

APPENDIX A STATISTICAL HYPOTHESIS TESTING

A.1. INTRODUCTION

The statistical test implementation is a procedure adopted in order to verify some hypothesis. Each test is organized with the input data (the sample components) with the output (the results of the statistical test itself) and the decision considerations. One of the more powerful applications of statistical analysis is the *hypothesis testing* where the decision considerations are based on the possibility to accept or reject some hypothesis that, given the input data, leads to a specification outputs. In hypothesis testing two hypothesis are compared: the so called null hypothesis (H_0) and the alternative hypothesis (H_1). In this thesis the hypothesis testing has been employed in order to test if two or more groups of sample data (differenting from a contexts variable which influence is needed to be tested) can be considered to be as different draws of the same random distribution (null hypothesis). In this case, one of the main test statistics is the probability with which it would be possible to obtain the same data in the case in which the null hypothesis isn't right. The higher the probability the more unrealistic can be considerate the null hypothesis.

Table 11

Errors of first kind and second kind

	H0 Right	H0 Isn't right
Accept	$1-\alpha$ (correct decision)	Error of second kind β
Reject	Error of first kind α	$1-\beta$ (correct decision)

	H0 Right	H0 Isn't right
Accept	$1-\alpha$	P(accept/Isn't right) β
Reject	P(reject/right) α	$1-\beta$

The hypothesis can be verified by referring to the following procedure:

1. the null hypothesis is fixed;

2. the alternative hypothesis is fixed;
3. a level of significance is identified;
4. a selection is made of the text/criteria according to which make a decision;
5. a decision is made related to accepting or rejecting the null hypothesis.

The test result has be synthesized by making the following considerations: if the *calculate value>critical value* (x_α), the test can be considerate *significant* at 1-α[15] level of significance and the null hypothesis can be rejected (see Figure 32). If for instance the test aim is to understand if there is or not a relationship between two parameters, the hypothesis can be identified as following:

H_0 : There isn't a relationship (Not R);
H_1 : There is a relationship (R).

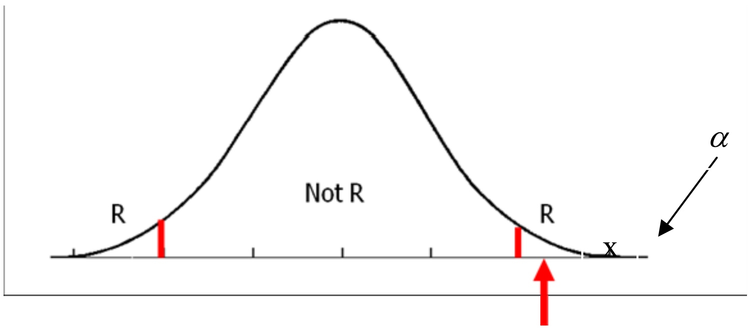


Figure 32: Rejecting regions related to the statistical test on the function of distribution

The statistical tests generally are divided in two main categories: the *non parametric test* and *the parametric test*. This difference is principally based on some hypothesis referred to the sample characteristics. The parametric test can be adopted when the variable that we are considering is a quantitative variable, and in all cases in which some hypothesis can be made on the sample (for instance the sample is supposed to be normally distributed). Otherwise the non parametric test can be adopted in all cases of variable and any hypothesis on the sample distribution. Parametric tests are generally more efficient, while non-parametric

[15] 1-α is defined as the test level of significance while 1-β is defined the test power.

ones are generally more flexible and more wildly applicable. Tests can be differentiated not only on the base of the variable that we are considering (generally qualitative, quantitative or based on the ranking) but also on the base of the sample correlation. Generally sample can be classified as dependent or independent: in the first case (if we are considering only two sample to compare) the sample are composed by the same subjects, otherwise (in case of independent sample) the sample that we are considering aren't the same. In Figure 33 an overview of the most diffused tests is presented.

By combining the characteristics of the variable and the relationship of dependency/independency among the sample some main statistical tests can be identified as described in Figure 33.

	2Groups different subjects	>2Groups different subjects	Same Subjects
Quantitative Variables	T-test (Student)	ANOVA	Paired t-test
Qualitative Variables	Chi Square	Chi Square	McNemar
Rank Variables	Mann-Whitney	Kruskall-Wallis	Wilcoxon

Not Parametric Test

Figure 33. Test Overview

Some considerations can be made that suggest that the chosen test applied to the analysis of the chapter 1 of this thesis should be the Kruskall Wallis test. First of all no hypothesis can be made on the sample (not constrained to be normally distributed) and for this reason it is preferred the adoption of a not parametric test. Among the not parametric test, the Kruskall Wallis one is suitable for cases where there are more than two groups composed by different subjects (the groups to compare are independent among them, because the subjects aren't the same[16]). This kind of test is generally adopted to analyze the experiments in which at least three groups are submitted to different treatment (in our case the *treatment* is the level of inaccuracy); it is an immediate generalization of the Mann-Whitney in which the number of sample is equal two.

A.2. THE KRUSKALL WALLIS TEST

In the following a Kruskall Wallis test is described with reference to a simple example. The effect of three different treatments has to be investigated. Treatments are identified as A, B and C. The effect of the treatment on each element of the group subjected to the A or B or C treatment is measured; measures are reported in the following table for the seven elements subjected to treatment A, the five elements subjected to treatment B and the six elements subjected to treatment C, the effect is quantified by points.

A	B	C
3	4	5
5	5	4
1	6	3
4	3	7
4	2	5
7		6
6		

The measures are aggregate in the same vector, ordered indifferently by the treatment and associated to their position order, as described in following:

Measures: 3, 4, 5, 5, 5, 4, 1, 6, 3, 3, 7, 4, 2, 5, 7, 6, 6, 4
Ordered Measures: 1, 2, 3, 3, 3, 4, 4, 4, 4, 5, 5, 5, 5, 6, 6, 6, 7, 7
Position Order: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18

[16] In our case the test is aimed at comparing the choices made by travelers at different inaccuracies levels; to be more precise, the answers refers to different scenarios.

Then to each measure it is associated the *rank,* which the average position of the measure:

Ordered Measures:	1	2	3	4	5	6	7
Positions to average:	{1}	{2}	{3, 4, 5}	{6, 7, 8, 9}	{10, 11, 12, 13}	{14, 15, 16}	{17, 18}
Rank:	1	2	3	7.5	11.5	15	17.5

A	R1	B	R2	C	R3
3	4	4	7.5	5	11.5
5	11.5	5	11.5	4	7.5
1	1	6	15	3	4
4	7.5	3	4	7	17.4
4	7.5	2	2	5	11.5
7	17.5			6	15
6	15				
	64		40		67

Figure 34. Test Overview

As described in previous Figure 34, after the preliminary aggregate evaluation of the ranks a correspondence between treatment and ranks is formulated. The total values of the ranks are: $\mathbf{T}^{A}$=64; $\mathbf{T}^{B}$=40 and $\mathbf{T}^{C}$=67. The statistic of the test can be calculated such as:

$$K = (N-1)\frac{\sum_{i=1}^{g} n_i(\bar{r}_{i\cdot} - \bar{r})^2}{\sum_{i=1}^{g}\sum_{j=1}^{n_i}(r_{ij} - \bar{r})^2}$$ [17]

where

- N is the total number of subjects;
- n_g is the number of the subjects referred to each treatment;

[17] The result of this expression is the Chi square statistic as also adopted end shown in this research in Chapter 1.

- r_{ij} is the rank referred to j^{th} observation in group i
- $\bar{r}_{i\cdot} = \frac{\sum_{j=1}^{n_i} r_{ij}}{n_i}$,

The test result is the asymptotic significance (as also shown in chapter 1 in the application of the test for the aggregate data analysis) that estimates the probability of obtaining a *chi-square* statistic greater than or equal to the one displayed.

APPENDIX B DISCRETE CHOICE MODELS

B.1. MULTINOMIAL LOGIT

The Multinomial logit Model can be considered the simplest among the random utility models. Particularly this model has been derived by assuming that the error terms of the utility function (the elements of vector $\boldsymbol{\varepsilon}$) are independently and identically distributed as a Gumble random variables with mean zero and parameter ϑ (see Ben-Akiva and Lerman, 1985). The relationship between the parameter ϑ and the probability to chose the alternative A is shown in Figure 35 with reference to a binary choice where the systematic value of utility is the same for both alternatives.

The analytical expression of the probability density function and of cumulative distribution function are respectively:

$$f_{\varepsilon_j}(e) = \frac{1}{\vartheta}\exp\left(\frac{-e}{\vartheta} - \Phi\right)\exp\left[-\exp\left(\frac{-e}{\vartheta} - \Phi\right)\right]$$

and

$$F_{\varepsilon_j}(e) = \Pr[\varepsilon_j \le e] = \exp\left[-\exp\left(\frac{-e}{\vartheta} - \Phi\right)\right]$$

where Φ is the *Euler* constant parameter (0.577).

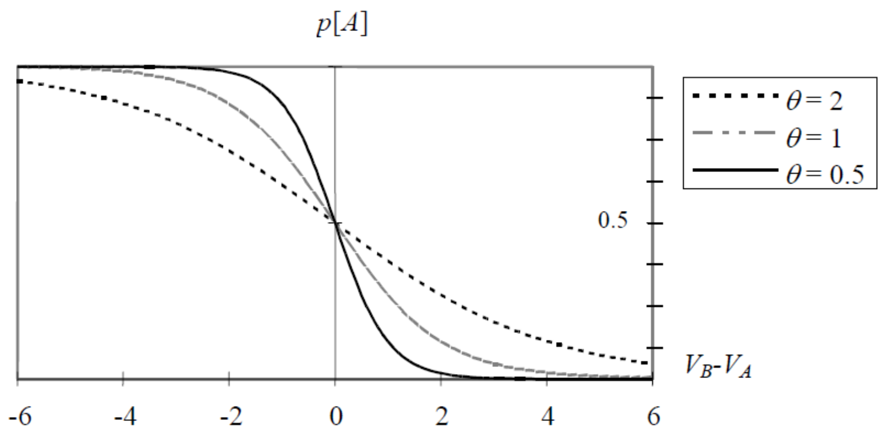


Figure 35. Relationship between the probability to chose the alternative A and ϑ

To be more precise the elements ε_j can be considerate *identically distributed* because each element is characterized by the same value of variance:

$$Var[\varepsilon_j] = \pi^2 \frac{\vartheta^2}{6} \quad \forall\ j \in I$$

Furthermore random residuals are considerate independently distributed because the value of covariance between two elements is null:

$$Cov[\varepsilon_j, \varepsilon_h] = \sigma_{j,h} = 0 \quad \forall\ (j,h),\ j \neq h$$

The main limitation of the multinomial logit model is related to the propriety of *Independence from Irrelevant Alternative (IIA).* According to this propriety, the ratio of the probabilities of any two alternatives doesn't change if the systematic utility of one or all other alternatives changes; the property can be easily verified as following:

$$\frac{P_n(i)}{P_n(l)} = \frac{e^{V_{in}/\sum_j V_{jn}}}{e^{V_{\mathrm{ln}}/\sum_j V_{jn}}} = \frac{e^{V_{in}}}{e^{V_{\mathrm{ln}}}} = e^{V_{in}-V_{1n}}$$

In particular this propriety can be interpreted by specifying that the relative share of any two alternatives is entirely unaffected by the systematic utilities of any other alternatives.

Moreover, another consequence of this propriety is that the ratio of probability choices of any two alternatives is not affected by the adding of one more alternative to the choice set.

The limitation induced by the IIA is generally discussed with reference to the example described in Figure 36. In this figure two alternative routes (Path 1 and Path 2) allow for travelling from the origin to the destination. Assume, for simplicity, so that the choice probability of each of the two alternatives is 50% so that the ratio is 1:1. Now assume that a short detour is introduced so that path 2 is now composed by two irrelevantly different routes (Path 2.a and Path 2.b). In this case, according to the multinomial logit model the ratio among the two main alternatives should be once again 1/1 and if the length of path 2.a is almost the same of path 2.b the choice probabilities of the MNL are 1/3 for all routes (*Path1, Path 2.a* and *Path 2.b*). This result is clearly paradoxical, provided that the intuitively expected probability values should be 1/2, 1/4 and 1/4 respectively for path 1, path2.a and path 2.b. A more realistic choice model can be obtained by introducing a covariance between two alternatives in the nested, provided that the main motivation of the IIA propriety is the independent distribution hypothesis of the component of the error vector.

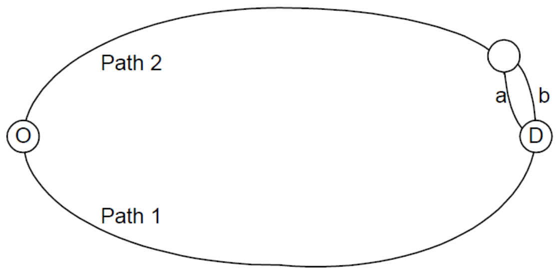


Figure 36.Path Choice Problem

B.2. LIMITATIONS OF THE MULTINOMIAL LOGIT

Multinomial logit can be considerate limited according the following three topics: the taste variation, the substitutions patterns and the repeated choices.

1. The multinomial logit can simulate the systematic taste variation to be more precise the taste variation that can be described by explicitly considering the decision makers characteristics, but not the taste variation that cannot be linked to the observed characteristics; the taste that vary with unobserved variables cannot be handled;
2. The multinomial logit model implies proportional substitution across the alternatives (IIA see in previous paragraph);
3. In case of repeated choice, the multinomial logit can be adopted to capture the dynamics of repeated choices only if such a dynamics is fully observable

(can be explained by the systematic utility); otherwise, if the unobserved factors are correlated over time, logit cannot be adopted.

B.3. PANEL DATA

In several cases of data acquisition travelers are submitted to a context of repeated choices. The decision-making behavior is calibrated against these data where repeated choices are made by the same respondent. In this case, data that represent the repeated choices are called *panel data*.

As already discussed, multinomial logit can be used only if the unobserved factors can be considerate independent across choices, so that each choice made by each decision maker can be considerate such as a separate choice. However, in case of repeated choices it is common to explain the observable part of utility by also using attributes related to the dynamics of the observed process (for instance, to be simple, the average of a given attribute over the last days). It should be a very restrictive hypothesis to assume that every dynamics is observable by the analyst and that the model could neglect correlations over time of the unobserved count-part of the utility. For this reason, in case of panel data it could be useful to use models such as the probit or the mixed- logit.

B.4. THE NESTED LOGIT

Even if the nested- logit approach can be derived from very general family of the GEV-based (Generalized Extreme Value) discrete choice model, the Nested Logit has been proposed such as an extension of the Multinomial Logit in order to capture the correlation among the alternatives (see Ben-Akiva and Lerman, 1985). A typical choice structure for a nested-logit model is depicted in Figure 37.

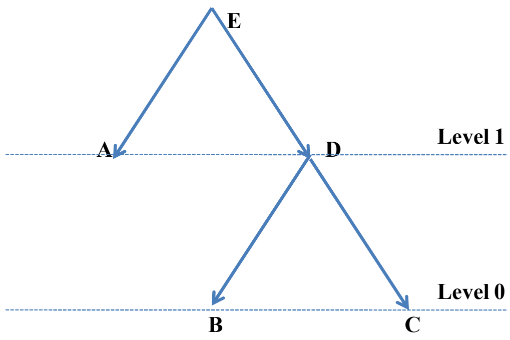


Figure 37.Example of Nested Logit

Consistently with the figure, the probability to choose any of the alternatives (A, B, C) can be computed by using conditional probabilities:

$P(A) = P(A)$

$P(B) \;=\; P(D)\; P(B/D); P(C) \;=\; P(D)\; P(C/D).$

Moreover by referring to the level 0 the utilities of alternatives can be calculated as:

$U_{B/D} = V_{B/D} + \tau_{B/D}; U_{C/D} = V_{C/D} + \tau_{C/D}$

where $\tau_{B/D}$ and $\tau_{C/D}$ are *i.i.d.* Gumble random variables with parameters ϑ_D: $G(0;\frac{\pi^2 {\theta_D}^2}{6})$. Note that covariance are null by definition $Cov(\tau_{B/D};\tau_{C/D}) = 0$ and the conditional probabilities can be computed according to a standard multinomial logit model:

$$p(B/D) = \frac{\exp\left(V_{B/D}/\vartheta_D\right)}{\exp\left(V_{B/D}/\vartheta_D\right) + \exp\left(V_{C/D}/\vartheta_D\right)}; p(C/D) = \frac{\exp\left(V_{C/D}/\vartheta_D\right)}{\exp\left(V_{B/D}/\vartheta_D\right) + \exp\left(V_{C/D}/\vartheta_D\right)}$$

Moreover by referring to the level 1 the utilities of alternatives *A* and *D* can be calculated such as:

$U_{A/E} = V_{A/E} + \varepsilon_{A/E}$; $U_{D/E} = V_{D/E} + \varepsilon_{D/E}$ where $\varepsilon_{A/E}$ and $\varepsilon_{D/E}$ are *i.i.d.* as Gumble random variables of parameters θ_E : $G(0;\frac{\pi^2 \theta_E{}^2}{6})$.

Further specifications are needed for the utility evaluation of alternative *D*; it has to be considered the contribution to the utility that comes from the alternatives in the *nest D*. The *composed utility* is expressed such as :

$V_{D/E} = \vartheta_D Y_D$.

The term Y_D is called *LOGSUM* or *Inclusive Variable* and is calculated as:

$$Vmx = E\left[\max_{j\in D}\{U_j\}\right] = \vartheta_D \cdot \ln \sum_{j\in D} \exp\left(V_{j/D} \big/ \vartheta_D \right) = \vartheta_D \cdot Y_D$$

It results that the probabilities at level 1 result to be:

$$p(A) = \frac{\exp\left(V_{A/E} \big/ \vartheta_E \right)}{\exp\left(V_{A/E} \big/ \vartheta_E \right) + \exp\left(\vartheta_D\, Y_D \big/ \vartheta_E \right)}$$

and

$$p(D) = \frac{\exp\left(\vartheta_D\, Y_D \big/ \vartheta_E \right)}{\exp\left(V_{E/A} \big/ \vartheta_E \right) + \exp\left(\vartheta_D\, Y_D \big/ \vartheta_E \right)}$$

Finally:

$$P(B) = P(D)\, P(B/D) = \frac{\exp\left(\vartheta_D\, Y_D \big/ \vartheta_E \right)}{\exp\left(V_{E/A} \big/ \vartheta_E \right) + \exp\left(\vartheta_D\, Y_D \big/ \vartheta_E \right)} \cdot \frac{\exp\left(V_{B/D} \big/ \vartheta_D \right)}{\exp\left(V_{B/D} \big/ \vartheta_D \right) + \exp\left(V_{C/D} \big/ \vartheta_D \right)}$$

$$P(C) = P(D)\, P(C/D) = \frac{\exp\left(\vartheta_D\, Y_D \big/ \vartheta_E \right)}{\exp\left(V_{E/A} \big/ \vartheta_E \right) + \exp\left(\vartheta_D\, Y_D \big/ \vartheta_E \right)} \cdot \frac{\exp\left(V_{C/D} \big/ \vartheta_D \right)}{\exp\left(V_{B/D} \big/ \vartheta_D \right) + \exp\left(V_{C/D} \big/ \vartheta_D \right)}$$

It can be easily shown that the dispersion matrix of the nested logit model here presented is: $\frac{\pi^2}{6}\begin{vmatrix} \theta_E^2 & \theta_E^2-\theta_D^2 \\ \theta_E^2-\theta_D^2 & \theta_E^2 \end{vmatrix}$.

It results that the nested logit model allows to introduce a structure for the variance-covariance matrix; such a structure is more general then the one allowed by a multinomial logit model, provided that it is able to deal with not null covariances.

The formalization here discussed for a simple case and only with reference to a bi-level choice can be easily generalized.

B.5. PROBIT

The Probit model (Domenich and McFadden, 1975; Daganzo and Sheffi, 1982) is the more flexible, provided that it allows for any structure of the variance-covariance matrix. Among other things, the probit model allows for overcoming limitations such as the *IIA* propriety and the correlation over time of the unobserved part of the utilities (typical of panel data contexts). This model has been derived under the hypothesis that the error terms of the utility are normally distributed :

U = **V** + **ε,** where **ε**~MVN(**0**, **Σ**), and general variances and covariances. In this case the function density probability is expressed as following:

$$\phi_\varepsilon(\boldsymbol{\varepsilon}) = \frac{\exp\left(-\frac{1}{2}\boldsymbol{\varepsilon}^T\,\boldsymbol{\Sigma}^{-1}\,\boldsymbol{\varepsilon}\right)}{\sqrt{(2\pi)^n\det(\boldsymbol{\Sigma})}}$$

The main problem of this model is in solving the integral which allows for the computation of the choice probabilities:

$$\forall\ j \quad p(j) = \int_{\varepsilon_{j=-\infty}}^{\infty} \dots \int_{\varepsilon_1<\varepsilon_j+V_j-V_1} \dots \int_{\varepsilon_1<\varepsilon_j+V_j-V_n} \phi_\varepsilon(\boldsymbol{\varepsilon})\, d\varepsilon_1 \dots d\varepsilon_j \dots d\varepsilon_n$$

This expression can't be solved (doesn't have a closed form) when the number of alternatives is major then two. In these cases some numerical simulation methods have to be adopted, typically the Monte-Carlo methods or the Clark approach, in order to compute choice probabilities .

B.6. MULTINOMIAL PROBIT FOR PANEL DATA

In this case each decision maker faces a choice among J alternatives in each of the T time periods.

The utility that decision maker n obtains from alternatives j in period t is $U_{njt} = V_{njt} + \varepsilon_{njt}$.

The vector of errors for all alternatives in all times periods is: $\varepsilon_n = (\varepsilon_{n11} \ldots\ldots \varepsilon_{nJ1}, \varepsilon_{n12} \ldots\ldots\ldots \varepsilon_{nJ2}, \ldots\ldots\ldots\ldots, \varepsilon_{n1T} \ldots\ldots \varepsilon_{nJT})$

and the covariance matrix has dimension ***JTxJT***. If the sequence of the choice made by the same decision maker is: ***i***= $\{i_1 \ldots\ldots\ldots i_T\}$, the probability that the decision makers can make this sequence of choice is:

$$
\begin{aligned}
P_{n\mathbf{i}} &= Prob(U_{ni_tt} > U_{njt}\ \forall j \neq i_t,\ \forall t) \\
&= Prob(V_{ni_tt} + \varepsilon_{ni_tt} > V_{njt} + \varepsilon_{njt}\ \forall j \neq i_t,\ \forall t) \\
&= \int_{\varepsilon_n \in B_n} \phi(\varepsilon_n) d\varepsilon_n .
\end{aligned}
$$

In which

$$B_n = \{\varepsilon_n \text{ s.t. } V_{ni_tt} + \varepsilon_{ni_tt} > V_{njt} + \varepsilon_{njt}\ \forall j \neq i_t,\ \forall t\}$$

Compared to the Probit probability the only difference is that in case of panel data the integral is extended to be over *JT* dimensions rather *J*.

B.7. MIXED LOGIT

The mixed logit model has been introduced (Ben –Akiva and Bolduc 1996; Bhat, 1997; Revlet and Train, 1998) to capture the heterogeneity in behavior across respondent.

It also allows for dealing with panel data contexts. Two interpretations are diffused in literature: the first one is defined for *random parameters* the second for *error components*.

B.7.1 Random Parameters interpretation of mixed-logit

In case of interpretation for random parameters the utility for decision makers *n* and for alternative *j* can be expressed by the following function of utility:

$U_j^n = \underline{\beta_n} \cdot \underline{X_j^n} + \varepsilon_j^n$.

In this case β is considerate distributed across respondents in order to capture their random taste variation. The distribution function of β_s can be identified as $f(\beta)$. Consequently (provided that the distribution of ε is considered to be in accordance with a multinomial logit model) for each respondent the choice probability can be calculated as: $P_{ni} = \int \frac{e^{V_{ni}(\beta)}}{\sum_{j=1}^{J} e^{V_{nj}(\beta)}} f(\beta) d\beta$.

Often the $f(\beta)$ distribution function is specified to be a normal (or a lognormal) random variable.

It is worth noting that the computation of the probabilities requires some numerical simulation model in order to deal with the distribution function $f(\beta)$.

B.7.2 Error Components interpretation of mixed logit

In case of specification in terms of error components, the error component of the utility is considerate compose by two parts. The first one is ε_j^n considered to be *i.i.d.* as a *Gumble* random variable. The second one is ϕ^n , supposed be normally distributed:

$$U_j^n = \underline{\beta} \cdot \underline{X_j^n} + \phi^n \cdot z^n{}_j + \varepsilon_j^n$$

The total unobserved portion of the utility $\eta_j^n = \phi^n \cdot z^n{}_j + \varepsilon_j^n$ where $z^n{}_j$ is generally assumed to be a dummy variable with value equal to one for alternatives except one. Standard multinomial logit models are particular cases where all values $z^n{}_j$ are assumed to be null and consequently there is no correlation among the alternatives (leading to the *IIA* propriety).

B.8. MIXED LOGIT FOR PANEL DATA

In cases where the decision maker makes the choice repetitively, for each respondent a sequence of choices can be identified. Such a sequence can be identified as:

i= $\{i_1.........i_T\}$

In this case, both the random parameters and the error component interpretations (previous paragraphs B.7.1 and B.7.2) are adequate to take into account that the sequence of choices is related to the same respondent. In the case of the random parameters interpretation, the model can be formalized as:

$$P_{ni} = \int L_{ni}(\beta) f(\beta) d\beta = \int \prod_{t=1}^{T} \frac{e^{V_{ni}(\beta)}}{\sum_{j=1}^{J} e^{V_{nj}(\beta)}} f(\beta) d\beta$$

In the case of error component interpretation, the model can be formalized as :

$$P_{ni} = \int \prod_{t=t_o}^{T} \frac{\exp\left(\underline{\beta}^T \cdot \underline{X}_{j,t}^n + \phi^n\right)}{\sum_{j'} \exp\left(\underline{\beta}^T \cdot \underline{X}_{j',t}^n + \phi^n\right)} f\left(\phi^n\right) d\phi^n$$

As usual for mixed logit, some numerical simulation techniques are required to solve the model.

APPENDIX C CALIBRATION OF DISCRETE CHOICE MODELS WITH BIOGEME 1.6

In order to calibrate the models proposed in this thesis, the free package Biogeme 1.6 (Bierlaire, M.; 2008) has been adopted. Biogeme allows for the calibration of different types of models and the calibration process is managed by properly specifying the input files by Biogeme. The main files needed to run Biogeme are the input file (*input.mod*) in which the specification of the model to be calibrated is punctually described) and the file in which the sample is described (*sample.dat*). In the sample file the choices made by the respondents of the sample are defined, as well as the values of the attributes for each choice alternative and for each respondent. The sample file depends on th specification of the observable (systeamtic) count-part of utility (all desired attributes have to be evaluated in the sample file) but does not depends on the type of model (multinomial logit, nested logit, probit, mixed logit) that has to be calibrated. The input file (*.mod*) accounts for the type of model to be calibrated and for the actual identification of the observable part of the utility and of the structure of the random distribution function. Several general examples of input files are reported in the following with reference to different types of models to be calibrated.

C.1. LOGIT

The simplest calibrate discrete choice model is the multinomial logit. In the following specification let's suppose that the set of choice is composed by two alternatives Alt1 and Alt 2. For each alternative the expression of the systematic utility to be calibrated is identified. This contains some attributes (X_{ij}) that should corresponds to the ones valuated in the sample file. Each attribute enters the systematic utility togheter with its parameter (β). Parameters β_s are the unknowns of the calibration problem, together with the alternative specific constants (*ASCs*) that can enters the systematic utility expression of all alternatives (except one).

The control variable (*av1* and *av2*, in the following example) are dummy values that allow to control if the alternative are available in the choice set of a decision-maker; they correspond to analogous values listed in the sample file; if the value of the control variable is one (not null) for a given record of the sample file, it means that the corresponding alternative is considered available in the choice set described by the record. In the input file, in all models, the betas and the ASCs have to be declared by defining the initial value (Value), the lower bound, the upper bound and the status (0 if the parameter has to be estimated, 1 in the case in which it has to be considered fixed).

```
[Beta]
// Name Value  LowerBound  UpperBound  status (0=variable, 1=fixed)
BETA1     0      -10000       10000        1
BETA2     0      -10000       10000        0
ASC1      0      -10000       10000        0

[Utilities]
// Id Name  Avail  linear-in-parameter expression (beta1*x1 + beta2*x2 + ... )

  1  Alt1  av1   ASC1 * one + BETA1 * x11 + BETA2 * x12
  2  Alt2  av2                BETA1 * x21 + BETA2 * x22

[Model]
// Currently, only $MNL (multinomial logit), $NL (nested logit), $CNL
// (cross-nested logit) and $NGEV (Network GEV model) are valid keywords
//
$MNL
```

C.2. NESTED LOGIT

In this case the input file is similar to the case of the multinomial logit model but it is enriched with a further section (NLNests) aimed at describing how the nests are structured. Each nest is identified by a name (es. NEST A, NEST B,…) by an intial value of the distribution parameter associated to the nest (those are further unknowns of the calibration problem), by a lower and upper bound for the parameter value, by a status (which controls, if the parameter actually is an unknown of the problem or if it has to be considerd fixed at its initial value) and, finally, by a list of alternatives that belong to the nest.

```
[Beta]
// Name Value      LowerBound UpperBound  status (0=variable, 1=fixed)
ASC1      0.0       -10000          10000      1
ASC2      0.0       -10000          10000      0
ASC3      0.0       -10000          10000      0
ASC4      0.0       -10000          10000      0
ASC5      0.0       -10000          10000      0
ASC6      0.0       -10000          10000      0
BETA1    0.0       -10000          10000      0
BETA2    0.0       -10000          10000      0

[Utilities]
// Id Name  Avail  linear-in-parameter expression (beta1*x1 + beta2*x2 + ... )
  1   Alt1     av1     ASC1 * one + BETA1 * x11 + BETA2 * x12
  2   Alt2     av2     ASC2 * one + BETA1 * x21 + BETA2 * x22
  3   Alt3     av3     ASC3 * one + BETA1 * x31 + BETA2 * x32
  4   Alt4     av4     ASC4 * one + BETA1 * x41 + BETA2 * x42
  5   Alt5     av5     ASC5 * one + BETA1 * x51 + BETA2 * x52
  6   Alt6     av6     ASC6 * one + BETA1 * x61 + BETA2 * x62

[NLNests]
// Name     ParamValue  lb  ub   status   list_of_alt
NESTA      1.01            1   10      0        1   2   3
NESTB      1.01            1   10      0        4   5   6

[Model]
// Currently, only $MNL (multinomial logit), $NL (nested logit), $CNL
// (cross-nested logit) and $NGEV (Network GEV model) are valid keywords
//
$ NL
```

C.3. PROBIT

In this case there are no differences in the input file with respect to the Multinomial Logit case except for the fact that it contains the directive to calibrate probit model (BP instead of MNL).

```
[Beta]
// Name  Value   LowerBound UpperBound  status (0=variable, 1=fixed)
ASC1      0       -10000       10000        1
BETA1   0        -10000       10000         0
BETA2   0        -10000       10000         0

[Utilities]
// Id Name  Avail  linear-in-parameter expression (beta1*x1 + beta2*x2 + ... )
  1   Alt1   av1   ASC1 * one + BETA1 * x11 + BETA2 * x12
  2   Alt2   av2                BETA1 * x21 + BETA2 * x22

[Model]
// Currently, only $BP (Binary probit), $MNL (multinomial logit), $NL (nested
logit), $CNL (cross-nested logit) and $NGEV (Network GEV model) are valid
keywords
//
$BP
```

C.4. MIXED LOGIT

As stated in Appendix B.7, two main approaches can be adopted for the specification of a Mixed Logit model: for *Random Parameters* and for *Error Components.* The input file has to be differently prepared for these two different approaches.

C.5. MIXED LOGIT: RANDOM PARAMETERS

In case of random parameters specification, the betas are considered to be normally distributed. This means that for each beta the associated variance is a further parameter to be estimated. The presence of additional parameters is described in the input file both in the section devoted to the description of the systematic utility.

```
[Beta]
// Name Value  LowerBound  UpperBound  status (0=variable, 1=fixed)
ASC1      0       -10000         10000           1
BETA1    0        -10000        10000           0
BETA2    0        -10000        10000           0
ZERO     0        -10000        10000           1
SIGMA1  0        -10000        10000           0

 [Utilities]
// Id Name  Avail  linear-in-parameter expression (beta1*x1 + beta2*x2 + ... )
  1   Alt1   av1   ASC1 * one + BETA1 [ SIGMA1 ] * x11 + BETA2 * x12
  2   Alt2   av2                        BETA1 [ SIGMA1 ] * x21 + BETA2 * x22

[Draws]
1000

[PanelData]
Id
ZERO_SIGMA

[Model]
// Currently, only $MNL (multinomial logit), $NL (nested logit), $CNL(cross-
nested logit) and $NGEV (Network GEV model) are valid keywords
//
$MNL
```

Differently than the other models, not only the betas and the ASCs, but also the *ZERO* and the *SIGMAj* are introduced in the declaration of parameters.

Provided that the computation of the probabilities (and of the likelihood used to calibrate the model) involves some numerical simulation method, it is also needed to specify in the input file about how many draws of the involved normally distributed random variables have to be employed. Such a value is specified in the section [Draws]. The higher the indicated number of draws the higher the accuracy of the estimates (but the higher also the computation time of the calibration process). Practical evidences suggest that a number of 1000 draws is enough to obtain an accurate estimations.

C.6. MIXED LOGIT: ERROR COMPONENTS/ MIXED LOGIT FOR PANEL DATA

In case of error components the betas are considerate as fixed, but an error component (according to the theory) is introduced. The error component interpretation of mixed logit is also used by Biogeme in order to deal with panel data. In particular the error component for panel data is supposed to be normally distributed with zero mean and sigma variance (which is a further parameter to be estimated). The error component has to be introduced in all alternatives except one. Like the random parameters specification of the mixed-logit model, not only the betas and the ASCs, but also the *ZERO* and the *SIGMAj* are introduced in the declaration of the parameters. Of course, in case of panel data Biogeme requires to be notified about the field (in sample file) that indicates the respondent to whom the records belongs. Such a notification is reported in the session [Id_Respondent].

```
[Beta]
// Name Value  LowerBound  UpperBound  status (0=variable, 1=fixed)
ASC1      0       -10000         10000           1
BETA1    0       -10000         10000           0
BETA2    0       -10000         10000           0
ZERO      0       -10000         10000           1
SIGMA1  0       -10000         10000           0

 [Utilities]
// Id Name  Avail  linear-in-parameter expression (beta1*x1 + beta2*x2 + ... )
  1   Alt1   av1   ASC1 * one +  BETA1 * x11 + BETA2 * x12
  2   Alt2   av2                          BETA1 * x21 + BETA2 * x22 + ZERO [ SIGMA
] *one

[Draws]
1000
```

[PanelData]
Id_ Respondent
ZERO_SIGMA

[Model]
// Currently, only $MNL (multinomial logit), $NL (nested logit), $CNL(cross-nested logit) and $NGEV (Network GEV model) are valid keywords
//
$MNL

APPENDIX D DEFINITION AND COMPUTATION OF THE INACCURACY

Inaccuracies can be defined and computed in several ways[18]. In particular in this research two kinds of inaccuracies are computed:

- Descriptive (In)accuracy;
- Prescriptive (In)accuracy.

D.1. THE DESCRIPTIVE INACCURACY

The descriptive inaccuracy could be measured as the (relative) difference between the travel times dispatched by the ATIS and the actual travel times on the network. With reference to a network with three alternatives routes, assume that ***G*** is the vector of the estimated travel times by the ATIS and that ***C*** is the vector of the actual travel times:

$$G = \begin{bmatrix} g_1 \\ g_2 \\ g_3 \end{bmatrix}; C = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}$$

Three different formulations can be introduced in order to define the descriptive inaccuracy:

- The first one is referred to the computation of the relative difference between the actual travel times and the estimated travel times:

$$Dinc = \sum_{i=1}^{3} \frac{(g_i - c_i)}{c_i}$$

18 These formulations have been introduced and discussed in Bifulco, G.N., Simonelli, F., Di Pace, R. (2007).

- The second one is identical to the previous except that absolute values are:

$$Dinc = \sum_{i=1}^{3} \frac{Abs(g_i - c_i)}{c_i}$$

- The third formulation is obtained by using the classical definition of the Eulerian distance:

$$Dinc = \frac{\sqrt{\sum_{i=1}^{3} (g_i - c_i)}}{\sum_{i=1}^{3} c_i}$$

The calibration of the compliance model carried out during the thesis has shown that the best model performances can be obtained by using the descriptive inaccuracy computed with reference to the Eulerian distance.

Moreover, the inaccuracy can be computed, with reference to the day- to- day dynamics, in three different ways.

The inaccuracy of the ATIS at given day *(t)* can be computed as the:

- *network inaccuracy;* in this case the vectors ***G*** and ***C*** used in the computation at day t are composed by all estimated and actual travel times at day *t* ($\boldsymbol{G^t} = [g^t_1,\ g^t_2,\ g^t_3]$; $\boldsymbol{C^t} = [c^t_1,\ c^t_2,\ c^t_3]$);
- *experienced inaccuracy;* in this case only the estimated and actual travel times of the route actually chosen by the respondent at day *t* are used in the computation; in other terms, if the route actually chosen at day *t* by respondent is route 2, the vectors ***G*** and ***C*** are assumed to be $\boldsymbol{G^t} = [0,\ g^t_2,\ 0]$ and $\boldsymbol{C^t} = [0,\ c^t_2,\ 0]$);
- *induced inaccuracy;* in this case only the estimated and the actual travel times of the minimum estimated travel time route are considered; in other terms, assuming that the route with minimum ATIS estimated travel time is route 1, the vectors ***G*** and ***C*** at day *t* are assumed to be $\boldsymbol{G^t} = [g^t_1,\ 0,$

*0]*and $\boldsymbol{C^t} =[c^t_1,\ 0,\ 0])$; this kind of inaccuracy is said to be *induced* because it coincides with the experienced inaccuracy induced by the system for a compliant respondents.

It is worth noting that the network and induced inaccuracies only depend on the network travel times and on the ATIS estimates, while the experienced inaccuracy also depends on the choices actually made by the respondents. This means that the network and the induced inaccuracies at a given day (and in a given scenario) are the same for all respondents while different experienced inaccuracies could be computed for different respondents.

Provided that, even if calibrated in a disaggregate way, the compliance model of this thesis intended to be framed (in future researches) within a wider traffic assignment model (that is typically applied in aggregate way), the experienced inaccuracy has been excluded from the calibration process. Among the network and induced inaccuracy the one that has shown best performances during the calibration of the compliance model has been the network one.

D.2. THE PRESCRIPTIVE INACCURACY

The prescriptive inaccuracy is defined as the difference between the actual travel time that should have been experienced according to ATIS dispatched information and the actual travel time that would have been experienced if the travellers were aware of the actual travel-times on the network In other terms, assume that:

- $\Phi = \begin{bmatrix} \varphi_1 \\ \varphi_2 \\ \varphi_3 \end{bmatrix}$ is the vector of route choice probabilities induced by the ATIS; it can be differently interpreted in case of prescriptive or descriptive information:
    - for *prescriptive* ATIS, it is the vector of the ATIS prescriptions; each element represents the percentage of travelers that have been

directed by the ATIS on each of the alternative routes; if the ATIS prescription is spread among the travelers, then all the elements of the vector can be positive; if the information is the same for all travelers, then only one element of the vector is not null and assume value 1 (the same route is suggested to all travelers);

- for *descriptive* ATIS, it is the vector of route choice probabilities resulting by applying a route-choice model to the travel time estimates dispatched by the ATIS; a good approximation of such a kind of model choice could be, in order to compute the prescriptive inaccuracy, the deterministic model; in this case only the element of the vector corresponding to the minimum ATIS-estimated travel time assumes value 1, while other elements assume a null value;

- C is the vector of the actual travel-times (as in the case of the descriptive inaccuracy); $C = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}$

- $P = \begin{bmatrix} P_1 \\ P_2 \\ P_3 \end{bmatrix}$ is the vector of route choice probabilities, computed by using the actual travel times of the network; a good approximation, in order to compute the prescriptive inaccuracy, could be a deterministic model, so that only the not null (with value 1) element refers to the actual minimum route. With all the previous, the *prescriptive* inaccuracy can be defined as to be:

$$Pinc = Abs\left(\frac{\Phi^T \cdot C - P^T \cdot C}{P^T \cdot C} \right) = \frac{Abs\left(\sum_{i=1}^{3} \varphi_i C_i - \sum_{i=1}^{3} P_i C_i \right)}{\sum_{i=1}^{3} P_i C_i}$$

It is worth noting that:

- $\Phi^T \cdot C$ is *the total cost spent* on the network if ATIS information is followed;
- $P^T \cdot C$ is the *total cost spent* on the network if ATIS actual travel times are known.

The implemented prescriptive inaccuracy is inherently defined a network (and ATIS) attribute; it does not depend on the respondent being equal (at a given day and in a given scenario) for all travellers.

D.3. DEFINITION AND COMPUTATION OF THE RELIABILITY

A measure of performance for an ATIS could be the reliability. Here the reliability is defined as a measure of the ability of the system to dispatch information that leads at choosing the route that would have been chosen if actual travel times of the network were known to the travellers.

This can be interpreted in two different ways in case of prescriptive or descriptive ATIS:

- *in case of prescriptive ATIS,* if the prescription provided suggests the route that would have been chosen if actual travel time were known, the system has been reliable, otherwise it has been unreliable;
- *in case of descriptive ATIS*, if the supplied travel time estimates induce route choices equals to the ones that would have been observed in case of known actual travel times, the system has been reliable, otherwise it has been unreliable.

In both cases:

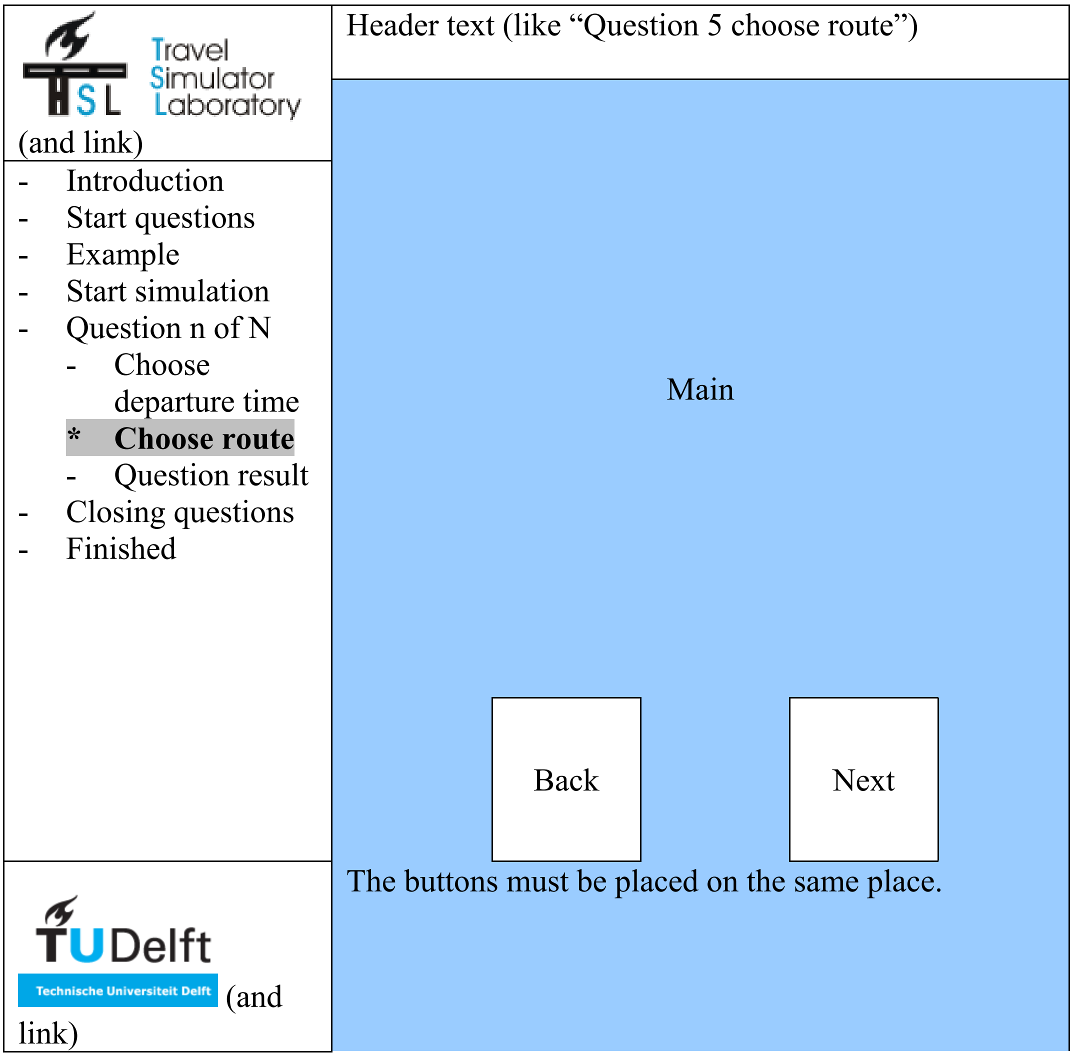
- $\Phi = \begin{bmatrix} \varphi_1 \\ \varphi_2 \\ \varphi_3 \end{bmatrix}$ is the vector of route choice probabilities induced by the ATIS; it can be differently interpreted in case of prescriptive or descriptive information:
    - for *prescriptive* ATIS, it is the vector of the ATIS prescriptions; each element represents the percentage of travelers that have been directed by the ATIS on each of the alternative routes; if the ATIS prescription is spread among the travelers, then all the elements of the vector can be positive; if the information is the same for all travelers, then only one element of the vector is not null and assume value 1 (the same route is suggested to all travelers);
    - for *descriptive* ATIS, it is the vector of route choice probabilities resulting by applying a route-choice model to the travel time estimates dispatched by the ATIS; a good approximation of such a kind of model choice could be, in order to compute the prescriptive inaccuracy, the deterministic model; in this case only the element of the vector corresponding to the minimum ATIS-estimated travel time assumes value 1, while other elements assume a null value;
- $P = \begin{bmatrix} P_1 \\ P_2 \\ P_3 \end{bmatrix}$ is the vector of route choice probabilities, computed by using the actual travel times of the network.

The reliability of the ATIS is here defined as: $REL = 1 - [Abs(\Phi - P)]^T \cdot P = 1 - \sum_{i=1}^{3} Abs[(\Phi_i - P_i)] \cdot P_i$.

The range of values for the reliability is from 0 to 1; 0 means complete unreliability, 1 means complete reliability.
Therefore with this expression we can obtain the difference between ATIS induced route choice probabilities and values of route choice probabilities computed by actual route choice model, and this difference is weighted by the realistic values of probabilities.
To give an example of reliability computation in case in which a deterministic route choice approximation is made for both ATIS-induced and network-induced choices, consider that information is such that $\Phi^T = [0, \quad 1, \quad 0]$ (that means that the ATIS directly or indirectly suggests to follow route 2) and that actual travel times are such that $P^T = [1, \quad 0, \quad 0]$ (that means that actual travel times would have suggested route 1), then the reliability can be computed as $REL = 1 - [1, \quad 1, \quad 0] \cdot [1, \quad 0, \quad 1]$ which confirm that the system has been unreliable.

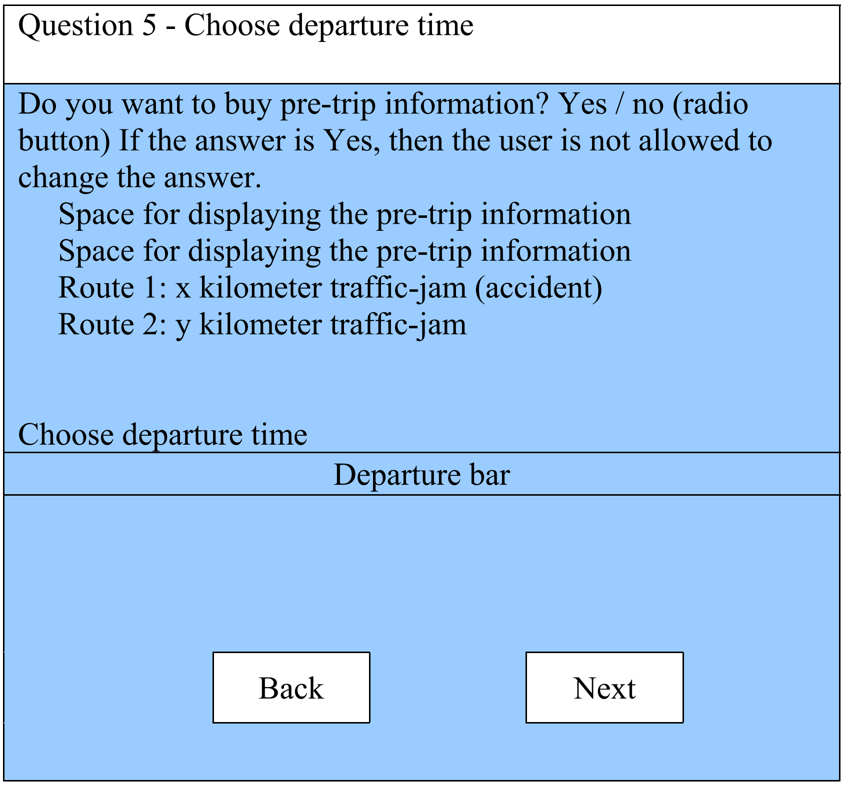
APPENDIX E SCREEN DEFINITION IN TSL

The TSL website is configured like a story board of a film direction. In the following every screen is shown in its general architecture. The most used page of the TSL corresponds to the experiment step where travel decisions are asked to the respondents

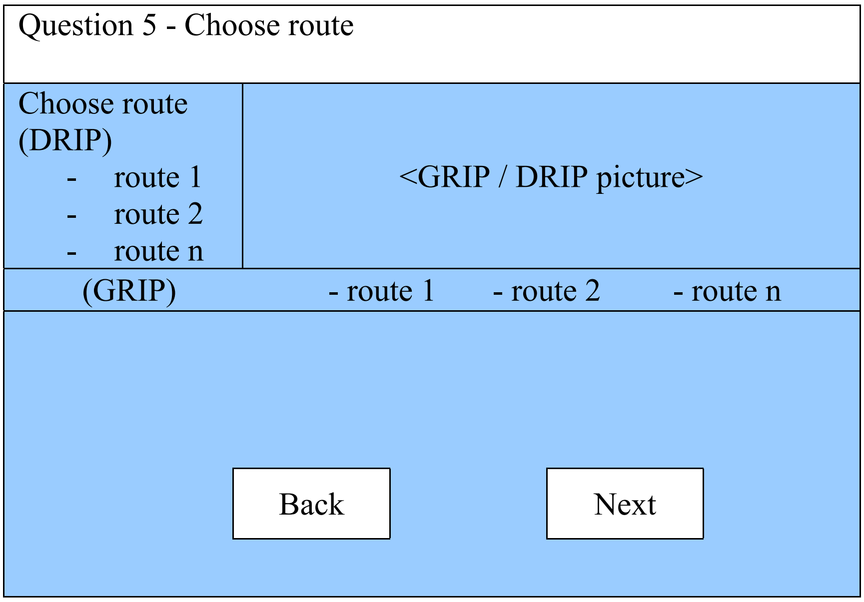


On the left of the page the table of contents is positioned (this canvas of the page is common to all steps of the experiment and is also employed as a bookmark); in the main part it is supposed that the experiment is described and the network shown; both images and text are allowed; at the down two buttons are positioned, they allow the respondent to go back or to continue to the next page.

In the following the departure time screen is described. Typically, a dragging bar is positioned at the centre of the main canvas, as well as some minor controls. The header and the left panel of the page are not displayed and described (here because are common to all pages and are such as described for the route choosing page)



In the next figure the typical configuration of the main step of the providing information and condition for choice formulation is shown. Graphical elements (pictures) are also allowed in order to obtain a more intuitive and friend information for the respondents.



The last screen is referred to the end of the simulation. The traveller is informed on the results of his simulation as shown in next figure.

Question 5 – Result

Text about departure, arrival time, etc.
Text about departure, arrival time, etc.
Text about departure, arrival time, etc.

FEEDBACK (Like, you're too late!)

Table with information about arrival time
Table with information about arrival time
Table with information about arrival time
Table with information about arrival time
Table with information about arrival time
Table with information about arrival time

<busy / quiet picture>

Waiting bar

Back Next

The user should click twice on Next button. The first time is after showing the feedback. The waiting bar is than in progress. The next time is after the waiting bar is completed.

After this last page referred to the end of the simulation a thank you message with a University logo are displayed.

All the pages here briefly displayed and discussed are implemented by using HTML.

APPENDIX F CALIBRATION OF DISCRETE CHOICE MODELS

The aim of this appendix is to provide a basic description of choice model, on how a discrete choice model can be derived.

It's worth noting that a procedure of model estimation substantially is a procedure of *trial and error* (see Figure 38) in which three steps are iterated: specification, estimation and validation.

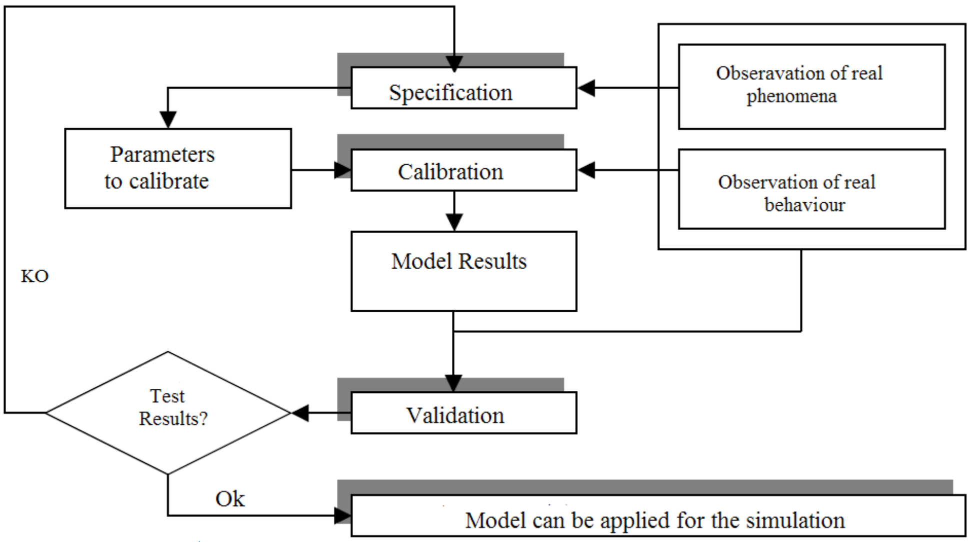


Figure 38: Trial and error in estimation model procedure

F.1. SPECIFICATION

First of all, the structure of the choice model has to be specified. This means that the attributes assumed to influence the perceived utility are framed within the utility expression

$V(s,x)$, where s refers to the considered alternatives and x the candidate explaining attributes.

Generally the common utility expression is a linear, additive form as: $V_{jq} = \beta_{1j} f_1(s_{kjq}) + ... + \beta_{kj} f_k(s_{kjq})$.

In the previous equation it has been explicitly addressed the disaggregate specification of the model, index q in fact refers to the general decision makers, while index k refers to the general candidate attribute.

The model is defined linear because the explaining variables enter in linear form in the utility; from this point of view, explaining variables are not directly the attributes, rather a function ($f_1 \ldots f_k \ldots$) of the attributes. In the linear form could be also, replaced by the others more complex form; however, in this case, some complications arise in the calibration procedure. Of course, on issue to be addressed in the specification phase is to identify the formal structure of the discrete choice model (logit, nested logit, probit, mixed logit, etc.). This implies that also the associated structure of the covariance matrix has to be calculated.

F.2. CALIBRATION: THE STATISTICAL ESTIMATION PROCEDURE

The calibration procedure is finalized to identify the numerical values of the parameters of then model (parameters of the explaining variables), as well as the values of the covariance matrix of the specified model, calibration is performed against all information related to the experiment data survey. Then most diffused procedure is the likelihood method (*Maximum likelihood estimation- MLE*). If n is the total number of observations of some random variables z_j, ($\boldsymbol{Z}$ is denoted as ($z_{1 \ldots\ldots}$ z_n)) and these observations are assumed to derive from a population described by a vector of characteristic parameters $\boldsymbol{\theta}$, it' s possible to associate to $\boldsymbol{Z}$ a probability density function and in particular, if all values can be considerate as independent, the function of probability density can be computed as:

$L(z_{1 \ldots\ldots} z_n / \boldsymbol{\theta}) = f(z_1/ \boldsymbol{\theta}) \ldots. f(z_n/ \boldsymbol{\theta})$, where $f(z_j / \boldsymbol{\theta})$ represents the probability to obtain the observations z_j,, given the vector of parameters $\boldsymbol{\theta}$. If the observations are considered as to be known, maximization of L with respect to $\boldsymbol{\theta}$, leads to the maximization of the probability to observe actual alternatives. This latter estimate is called the maximum likelihood estimate of $\boldsymbol{\theta}$.

Moreover it's mathematically simpler to work with the natural logarithm of the likelihood function, this is allowed by the fact that the *MLEs* of θ are invariant to monotonically increasing transformations of L.

To give an example on how the likelihood method is applied to the calibration of discrete choice mode consider the case in which the specified model is a multinomial logit one. In this case the probability that a decision maker q chooses the alternative i can be written as:

$P_{iq} = \exp(V_{iq}) / \sum \exp(V_{jq})$,

where the observed utility is: $V_{jq} = \sum_{k=1}^{K} \beta_{jk} x_{jk}$ and where x_s and β_s respectively are the explaining variables and the unknown parameters to be estimated. If all the observations are independent, the likelihood function can be written as

$L = \prod_{q=1}^{Q_1} \prod_{j=1}^{J} P_{jq}^{\delta_{jq}}$ where :

Q is the total number of the observed decision makers; J is the total number of the alternatives in the choice set; δ_{ij} is a dummy variable that valuates to 1 if decision-maker q has been observed to choose alternative i and 0 otherwise. The *log-likelihood* expression can be computed as: $L^* = \sum_{q=1}^{Q} \sum_{j=1}^{J} \delta_{iq} \ln(P_{jq})$. The last equation can be formalized in function of the systematic utility (linear combinations of betas) to explicitly show that the maximization of the function can be carried out with respect to the betas. The maximization of the log-likelihood can be carried out by using appropriate algorithms (for instance the gradient one). In case where the discrete choice model to be calibrated is not a multinomial logit one, the computation of the probabilities P_{iq} (and in turn, of the MLE) can't be performed in a closed form and for this reason it's needed to be solved a problem by means of numerical simulation techniques embedded into the standards iterative procedures. The iterative procedure (see Figure 39) are constrained by the tolerance then the iterations continue until the predefined level of tolerance is reached (Goldfeld and Quandt, 1972;Green, 1999). A classical iterative procedure has been described in figure.

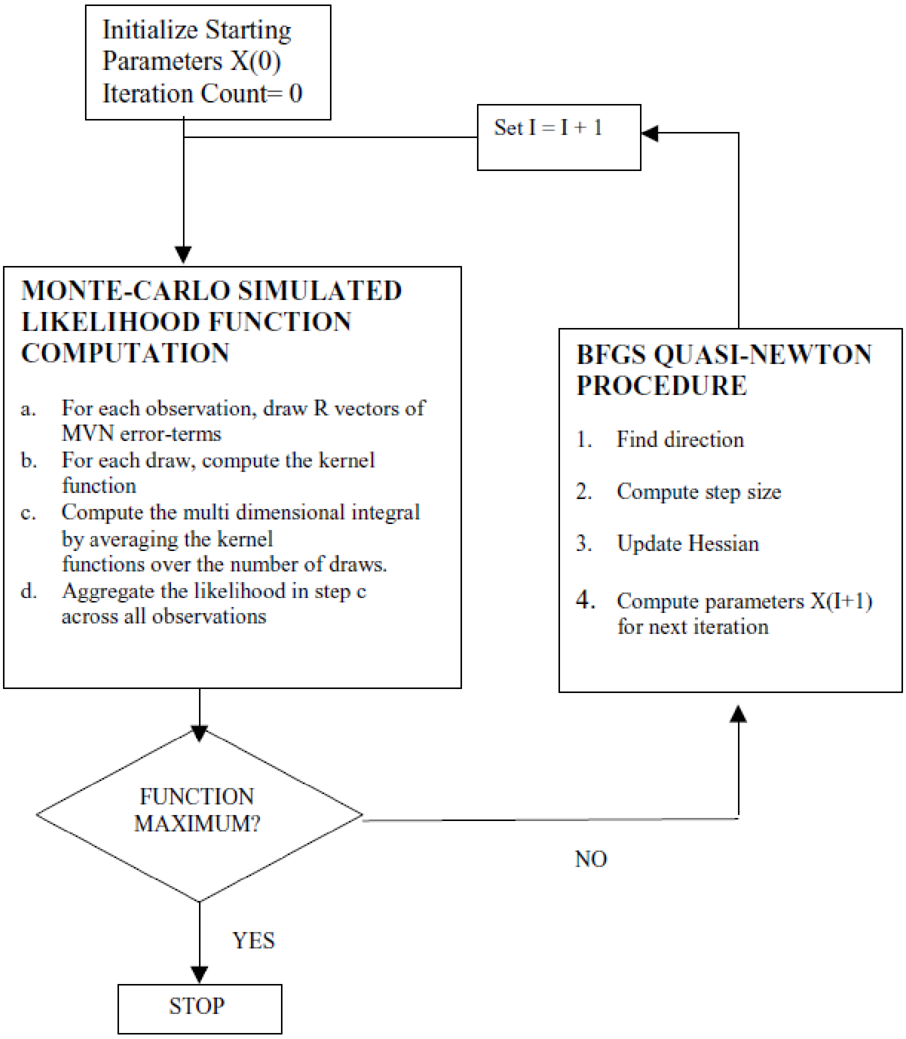


Figure 39: An example of calibration procedure in case of Dynamic Kernel Logit

F.3. VALIDATION

The result obtained by the calibration phase has to be evaluated in order to accept the assumptions (candidate explaining attributes and kind of discrete choice model) of the specification phase. If the validation is not satisfactory, the model has to be specified and calibrated again, until satisfactory results are obtained. The validation phase is generally performed with reference to three kinds of analysis:

i. Based on the significance of the utility parameters (β_s);
ii. Based on the measures of the goodness of fit of the model itself;
iii. Based on the elasticities of choice.

F.4. STATISTICAL SIGNIFICANCE OF UTILITY PARAMETERS

Appropriate standards errors and t-statistics are produced as part of the output of any calibration software. Researchers will seek out mean utility parameters which have sufficiently small standard errors: in this way the estimated mean is a good representation of the attribute influence in explaining the level of relative utility associated with each alternative.

The ratio of the mean parameters to its standard error is the *t-value* . In order to have a *95%* or greater confidence that the estimated value of the β_s has been not only casually obtained different from the null value, it is generally required to have a *t- value* greater or equal than *1.96* (in absolute value).

F.5. GOODNESS-OF-FIT TESTS, THE LIKELIHOOD RATIO TEST

The null hypothesis tested by the likelihood ratio test (H_0) is aimed to test that the true vector ($\boldsymbol{\beta}$) is equal to a given vector ($\boldsymbol{\beta^*}$): $H_0; \boldsymbol{\beta} = \boldsymbol{\beta^*}$.

This hypothesis can be tested by adopting the likelihood ratio test: $LR^* = -2[\ln(L(\beta^*)) - \ln(L(\beta^{MLE}))]$.

The $LR*$ statistic is distributed according to a chi-square variable with a degree of freedom equal to the number of constraints imposed on the betas estimation (as demonstrated by Wilks, 1962).

On the base of test results the interval of confidence of a single vector component (β_k), can be obtained. In particular, the reference vector of betas is supposed to be null one

(β*=0) the previous expression became $LR^* = -2[\ln(L(0)) - \ln(L(\beta^{MLE}))]$. The effect of a null reference betas vector, is that with all coefficients equal to zero all alternatives can be considered as to be equiprobable. This test is based on the difference between the compared values, $\ln(L(0))$ and $\ln(L(\beta^{MLE}))$ and the hypothesis is rejected with increasing confidence as higher is the difference value.

Sometimes the *LR* test is used in order to test if only for chance the calibrated vector βs is different from a vector of null values except than the alternative specific constants. In this case, the reference vector ***(β*)*** is the one where all parameters (but the ASCs) are null.

F.6. GOODNESS-OF-FIT TESTS, RO-SQUARE

Another test able to verify the goodness of fit of the model in reproducing choices is the *ro-square*:

$\rho^2 = 1 - (\frac{L^*(\hat{\beta})}{L^*(0)})$. The ro-square value can be shown to belong in the range from zero to one. The higher the value is, the better the goodness of fit. For transportation models acceptable ro-square values should be higher than 0.2/0.4, even if Domencich and McFadden (1975) have shown that for linear models values in the range from 0.7 to 0.9 can be reached. Representativeness of the ro-square test can be improved by adopting the *adjusted value*

$\overline{\rho}^2 = 1 - \left(\frac{L^*(\hat{\beta}) / [\sum_{q=1}^{Q}(J-1) - k]}{L^*(0) / \sum_{q=1}^{Q}(J-1)}\right)$ where Q is the total number of observed decision makers, J is the number of alternatives and k is the total number of explaining attributes (X_s). The adjusted ro-square is useful to compare models with different number of parameters included in model specification.

F.7. GOODNESS-OF-FIT TESTS, SUCCESS PREDICTION

After the calibration, it is useful to compare the probabilities of choice reproduce by the model with the observed chooses captured by the sample. A prediction model has been described by McFadden (1979). Assume N_{ij} is the number of respondent that are observed to choose the alternative i and are predicted to choose the alternative j. This can computed as: $N_{ij} = \sum_{q=1}^{Q} \delta_{qi} P_{qj}$ where Q is the total number of observed decision makers; δ_{qi} is a dummy variable that valuates to 1 if the observed decision maker q has actually chosen alternative i; P_{qj} is the probability predicted by the model of observing the choice j. From the evaluation of the N_{ij} a square matrix results, has shown in Table 12 for an hypothetical case with three alternatives. Sums can be computed for the totals of rows and columns: $N_{i.} = \sum_j N_{ij}; N_{.j} = \sum_i N_{ij}$.

The sum $N_{i.}$ can be defined as to be the observed number of choices for alternative i. In fact : $N_{i.} = \sum_j \sum_q \delta_{qi} P_{qj} = \sum_q \sum_j \delta_{qi} P_{qj} = \sum_q \delta_{qi} \sum_j P_{qj}$. Provided that the model has to predict choice probabilities that, for a given decision- maker, have to sum up to one ($\sum_j P_{qj} = 1$) over all alternatives, and by indicating A_i the total number of observed choices of alternatives i (that is $A_i = \sum_q \delta_{qi} = 1$), it results that

$N_{i.} = \sum_q \delta_{qi} \sum_j P_{qj} = \sum_q \delta_{qi}.1 = A_i$. Similarly, the sum $N_{.j}$ can be defined as to be the predicted number of choices for alternative j. In fact $N_{.j} = \sum_i \sum_q \delta_{qi} P_{qj} = \sum_q \sum_j \delta_{qi} P_{qj} = \sum_j P_{qj} \sum_q \delta_{qi}$

Provided that each decision- maker is observed to choose one (and only one) alternative ($\sum_q \delta_{qi} = 1$) and by indicating with B_j the total number of choices of alternative j predicted by the model over all decision makers ($\sum_q P_{qj} = B_j$), it results that : $N_{.j} = \sum_q P_{qj} \sum_i \delta_{qi} = \sum_q P_{qj} \cdot 1 = B_j$. The observed share for one alternative (i) can be now defined as: $\omega_i = N_{i.} / T$ where T is the total number of observations ($T = \sum_i P_{qj} N_{i.} = \sum_j N_{j.} = N_{..}$) . Similarly the predicted share for one alternative *(j)* can be defined as: $\pi_j = \dfrac{N_{.i}}{T_{..}}$. It is also possible define the proportion of successful predictions for alternative j. It can be compute as $\dfrac{N_{ii}}{N_{.i}}$. The success index for alternative j can be computed as: $sj = \dfrac{N_{jj}}{N_{.j}} - \pi_j = \dfrac{N_{jj}}{N_{.j}} - \dfrac{N_{.j}}{T}$. The error predicted share can be computed as (for alternative j) as $(N_{.j} - N_{j.}) \cdot T$. The following table (Table 13) shows proportions of successful predictions, success indexes, errors in predicted shares and overall success index for the sample of Table 12.

Table 12

Predicted and observed shares

		Alternatives				
		(1)	(2)	(3)	Row Total	Observed Share %
Alternatives	(1)	100	20	30	150	*45.5=(150/330%)*
	(2)	30	50	20	100	*30.3=(100/330)%*
	(3)	10	20	50	80	*24.2=(80/330)%*
Column Total		140	90	100	330	*100.0*
Predicted Share %		*42.4=(140/330)%*	*27.3=(90/330)%*	*30.3=(100/330)%*	100	

Table 13

Success index and overall prediction success index

Proportion of Successfully predicted	71.4=100/140	55.6=50/90	50.0=50/100
Success index	29.0= (71.4-42.4)	28.3=(55.6-27.3)	19.7= (50-30.3)
Error in predicted Share	-3.03=(140-150)/330	-3.03=(90-100)/330	6.06=(100-80)/330
Overall prediction success index	***0.2599=(0.424*0.290+0.273*0.283+0.303*0.197)***		

F.8. CHOICE ELASTICITIES

Discrete choice models can be considerate as demand model in which the demand (related to the willingness to choose the alternatives) is strictly related to the attributes by which the alternatives one described. The size of change in the level of an attribute is taken into account in elasticties measures. According to the theory two kind of elasticities can be defined: direct and cross elasticities.

The *direct elasticity* measures the (marginal) percentage change in probability of choosing an alternative with respect to the change of value of an attribute which defines the utility of the alternatives itself. For a given alternative *(j)*, and an attribute k referred to the alternative itself (X_{kj}) the direct elasticity in case of marginal change in k^{th} attribute is defined as:

$$E^{p[j]}{}_{kj} = \frac{\partial \ln p[j]}{\partial \ln X_{kj}}$$

The *cross elasticity* measures the percentage change in probability of choosing an alternative with respect to the changing of an attribute defines a *competing* alternative. For a given alternative *(j)*, and an attribute referred to another alternative *h*, (X_{kh}), the cross elasticity in case of marginal change in *k^th* attribute is defined as:

$$E^{p[j]}{}_{h} = \frac{\partial \ln p[j]}{\partial \ln X_{kh}}.$$

Generally the elasticity, for a given alternative *i*, an individual *q* and for the *k^th* attribute of the same alternative (in case of direct elasticity), can be expressed as:

$$E^{P_{iq}}_{X_{ikq}} = \frac{\partial P_{iq}}{\partial X_{ikq}} . \frac{X_{ikq}}{P_{iq}} ; \ \frac{\partial P_{iq}}{\partial X_{ikq}} = P_{iq}\beta_{ik} - P^{2}{}_{iq}\beta_{ik} \Rightarrow E^{P_{iq}}_{X_{ikq}} = P_{iq}(\beta_{ik} - P_{iq}\beta_{ik}) . \frac{X_{ikq}}{P_{iq}}$$

Generally the elasticity, for a given alternative *i*, an individual *q* and for the *k^th* attribute of the competing alternative *j* (in case of cross elasticity), can be expressed as:

$$E^{P_{iq}}_{X_{jkq}} = \frac{\partial P_{iq}}{\partial X_{jkq}} . \frac{X_{jkq}}{P_{iq}} = -P_{iq}P_{jq}\beta_{jk} . \frac{X_{jkq}}{P_{iq}} = -\beta_{jk}X_{jkq}P_{iq}$$

The previous expressions can be combined in following contract relation: $E^{P_{iq}}_{X_{jkq}} = \beta_{jk}X_{jkq}(\delta_{ij} - P_{iq})$, where δ_{ij} =1 in case of direct elasticity otherwise is equal to 0. After these evaluations made for every individual; an aggregation can be made, by adopting the method of the sample enumeration. The elasticity of each individual is weighted by the estimated choice probability:

$$E^{\overline{P}_{iq}}_{X_{jkq}} = \left(\sum_{q=1}^{Q} \hat{P}_{jq} E^{P_{iq}}_{X_{jkq}}\right) / \sum_{q=1}^{Q} \hat{P}_{jq}$$

APPENDIX G THE SOURCE CODES FOR SCENARIOS IMPLEMENTATION IN THE TSL

All the screens briefly introduced in the previous are dynamically managed by appropriate snippet of code. Many of the elements of the HTML pages actually are bookmarks, where some action, animation, data presentation or respondents' action has to be implemented. The implementation of these controls happens by mean of several code behind files.

In particular in the following it's presented the .xml file referred to the experiment configuration. The code is briefly presented, comments allow to have an idea on how the code works (we are referring to the text introduced in the code text and indicated with the symbol *).

It can be noted the presence of way keywords (e.g. "Departure Time") to which several proprieties are associated (e.g."Description", "Min", "txt Min", etc.); these allow the code behind the html pages show customized information and to take paper actions.

```
<?xml version="1.0" encoding="utf-8" ?>
  <ExperimentConfigurationDataset
  xmlns="http://tempuri.org/ExperimentConfigurationDataset.xsd"
  xmlns:xs="http://www.w3.org/2001/XMLSchema">
  <mailnotifier threshold="1" emailaddress="R.diPace@student.tudelft.nl" name="" />
xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
xxxxxxxxxxxxxxxxxxxxxxxx
<!—The dragging bar for the departure time choosing is watched --> (*)

  <DepartureTime Description="Please select your departure time dragging the slider
below" Min="06:30" textMin="Min" Max="09:00" textMax="Max" textValue="Departure
time" Interval="10" ArrivalTime="09:00" TextArrivalTime="You are expected at your
destination at {0} o' clock" />

xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
xxxxxxxxxxxxxxxxxxxxxxxx
<!—Configuration .jgp/panel on which the information is watched --> (*)

<InfoTypes>
<InfoType name="DRIP" pictureURL="Graphics\Roby12.jpg" x1Text="7" y1Text="47"
x2Text="200" y2Text="100" height="100" width="220" />
</InfoTypes>

xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
xxxxxxxxxxxxxxxxxxxxxxxx
```

```
<!—Configuration of the information feedback with which traveler are provided at the end of every simulation-->(*)

<TrafficFeedbackTexts ShowScoreOneQuestion="true" ShowScoreAllQuestions="true">
<FeedbackTexts textInTime="You are in time">
 <FeedbackText number="1" text="You are late" minLate="5" maxLate="10" />
 <FeedbackText number="2" text="You are too late" minLate="11" />
 <FeedbackText number="3" text="You are late again" timesLate="2" />
 <FeedbackText number="4" text="You are early" minEarly="5" maxEarly="10" />
 <FeedbackText number="5" text="You are too early" minEarly="11" />
 <FeedbackText number="6" text="You are early again" timesEarly="2" />
 </FeedbackTexts>
 </TrafficFeedbackTexts>

<ResultFeedbackTexts>
 <ResultTexts Text="Here are the results of your choices">
 <ResultText number="1" name="DepartureTime" display="true" text="Departure time:" />
 <ResultText number="2" name="TravelTime" display="true" text="Travel time:" />
 <ResultText number="3" name="ArrivalTime" display="true" text="Arrival time:" />
 <ResultText number="4" name="Score" display="false" text="Score:" />
 <ResultText number="5" name="ChosenRoute" display="true" text="Chosen route:" />
 </ResultTexts>
 </ResultFeedbackTexts>
xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
xxxxxxxxxxxxxxxxxxxxxxxxxxxx
<!—Configuration  the ex-post information at the end of all simulation--> (*)

<TotalResults>
<TotalResultTexts Text="Here are the total results of your choices">
<TotalResultText number="1" name="MostChosenRoute" display="true" text="Most choosen route:" />
 <TotalResultText number="2" name="AverageTravelTime" display="false" text="Average travel time:" />
 <TotalResultText number="3" name="NrOfTimesLate" display="false" text="Number of times too late:" />
 <TotalResultText number="4" name="NrOfTimesEarly" display="false" text="Number of times too early" />
 <TotalResultText number="5" name="OverallScore" display="false" text="Overall score:" />
 </TotalResultTexts>
 </TotalResults>


<scenarios>
<!—Configuration  of the scenarios under descriptive information--> (*)

xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
xxxxxxxxxxxxxxxxxxxxxxxxxxxx
<!—Configuration  of actual travel times on route 1--> (*)

 <scenario order="1" number="1" name="scenario1" minTestPersons="1" numberOfDays="1" maxBuyTravelInformation="" minFilledQuestions="40">
<realTravelTimes>
 <realTravelTimeManual dayNumber="1" route="1" value="34" />
 <realTravelTimeManual dayNumber="2" route="1" value="38" />
 <realTravelTimeManual dayNumber="3" route="1" value="62" />
 <realTravelTimeManual dayNumber="4" route="1" value="35" />
```

```
  <realTravelTimeManual dayNumber="5" route="1" value="41" />
  <realTravelTimeManual dayNumber="6" route="1" value="40" />
  <realTravelTimeManual dayNumber="7" route="1" value="65" />
  <realTravelTimeManual dayNumber="8" route="1" value="47" />
……… .
<!—Configuration of actual travel times on route 2-->(*)
  <realTravelTimeManual dayNumber="1" route="2" value="35" />
  <realTravelTimeManual dayNumber="2" route="2" value="35" />
  <realTravelTimeManual dayNumber="3" route="2" value="70" />
  <realTravelTimeManual dayNumber="4" route="2" value="35" />
  <realTravelTimeManual dayNumber="5" route="2" value="37" />
  <realTravelTimeManual dayNumber="6" route="2" value="37" />
  <realTravelTimeManual dayNumber="7" route="2" value="72" />
  <realTravelTimeManual dayNumber="8" route="2" value="35" />
  …….
<!—Configuration of actual travel times on route 3-->(*)
  <realTravelTimeManual dayNumber="1" route="3" value="52" />
  <realTravelTimeManual dayNumber="2" route="3" value="52" />
  <realTravelTimeManual dayNumber="3" route="3" value="53" />
  <realTravelTimeManual dayNumber="4" route="3" value="53" />
  <realTravelTimeManual dayNumber="5" route="3" value="52" />
  <realTravelTimeManual dayNumber="6" route="3" value="53" />
  <realTravelTimeManual dayNumber="7" route="3" value="52" />
  <realTravelTimeManual dayNumber="8" route="3" value="52" />
…..
  </realTravelTimes>
xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx

<!—Configuration of en route information on route 1-->(*)
 <enRouteInfoTravelTimes>
  <enRouteInfoTravelTimeManual dayNumber="1" route="2" value="35" />
  <enRouteInfoTravelTimeManual dayNumber="2" route="2" value="37" />
  <enRouteInfoTravelTimeManual dayNumber="3" route="2" value="69" />
  <enRouteInfoTravelTimeManual dayNumber="4" route="2" value="33" />
  <enRouteInfoTravelTimeManual dayNumber="5" route="2" value="34" />
  <enRouteInfoTravelTimeManual dayNumber="6" route="2" value="37" />
  <enRouteInfoTravelTimeManual dayNumber="7" route="2" value="68" />
  <enRouteInfoTravelTimeManual dayNumber="8" route="2" value="35" />

<!—Configuration of en route information on route 2-->(*)
  <enRouteInfoTravelTimeManual dayNumber="1" route="3" value="51" />
  <enRouteInfoTravelTimeManual dayNumber="2" route="3" value="51" />
  <enRouteInfoTravelTimeManual dayNumber="3" route="3" value="54" />
  <enRouteInfoTravelTimeManual dayNumber="4" route="3" value="53" />
  <enRouteInfoTravelTimeManual dayNumber="5" route="3" value="52" />
  <enRouteInfoTravelTimeManual dayNumber="6" route="3" value="54" />
  <enRouteInfoTravelTimeManual dayNumber="7" route="3" value="53" />
  <enRouteInfoTravelTimeManual dayNumber="8" route="3" value="49" />

<!—Configuration of en route information on route 3-->(*)
  <enRouteInfoTravelTimeManual dayNumber="1" route="1" value="34" />
  <enRouteInfoTravelTimeManual dayNumber="2" route="1" value="39" />
  <enRouteInfoTravelTimeManual dayNumber="3" route="1" value="65" />
  <enRouteInfoTravelTimeManual dayNumber="4" route="1" value="34" />
  <enRouteInfoTravelTimeManual dayNumber="5" route="1" value="40" />
  <enRouteInfoTravelTimeManual dayNumber="6" route="1" value="42" />
```

```
  <enRouteInfoTravelTimeManual dayNumber="7" route="1" value="62" />
  <enRouteInfoTravelTimeManual dayNumber="8" route="1" value="45" />
  </enRouteInfoTravelTimes>

<!—Configuration  of ex post information on every day; the ex post information arrives to 5 past days
ago-->
  <postexinformation NrOfPastDays="5" showTravelTimes="true" showOtherRoutes="true"
DayText="Day" TravelTimeText="Travel time" text="You have chosen:"
textOtherRoutes="Here are the results of the other routes:" />
  <QuestionDays Everyday="false" />

<!—Configuration  new scenario under prescriptive information-->(*)
xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
xxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
  -->
<!—Configuration  of  actual travel times on route 1-->(*)
 <scenario order="2" number="2" name="scenario5" minTestPersons="1"
numberOfDays="1" maxBuyTravelInformation="" minFilledQuestions="40">
 <realTravelTimes>
  <realTravelTimeManual dayNumber="1" route="1" value="34" />
  <realTravelTimeManual dayNumber="2" route="1" value="38" />
  <realTravelTimeManual dayNumber="3" route="1" value="62" />
  <realTravelTimeManual dayNumber="4" route="1" value="35" />
  <realTravelTimeManual dayNumber="5" route="1" value="41" />
  <realTravelTimeManual dayNumber="6" route="1" value="40" />
  <realTravelTimeManual dayNumber="7" route="1" value="65" />
  <realTravelTimeManual dayNumber="8" route="1" value="47" />

<!—Configuration  of  actual travel times on route 2-->   (*)
  <realTravelTimeManual dayNumber="1" route="2" value="35" />
  <realTravelTimeManual dayNumber="2" route="2" value="35" />
  <realTravelTimeManual dayNumber="3" route="2" value="70" />
  <realTravelTimeManual dayNumber="4" route="2" value="35" />
  <realTravelTimeManual dayNumber="5" route="2" value="37" />
  <realTravelTimeManual dayNumber="6" route="2" value="37" />
  <realTravelTimeManual dayNumber="7" route="2" value="72" />
  <realTravelTimeManual dayNumber="8" route="2" value="35" />
.........
<!—Configuration  of  actual travel times on route 3-->(*)
  <realTravelTimeManual dayNumber="1" route="3" value="52" />
  <realTravelTimeManual dayNumber="2" route="3" value="52" />
  <realTravelTimeManual dayNumber="3" route="3" value="53" />
  <realTravelTimeManual dayNumber="4" route="3" value="53" />
  <realTravelTimeManual dayNumber="5" route="3" value="52" />
  <realTravelTimeManual dayNumber="6" route="3" value="53" />
  <realTravelTimeManual dayNumber="7" route="3" value="52" />
  <realTravelTimeManual dayNumber="8" route="3" value="52" />
...
  </realTravelTimes>
xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
xxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
<!—Configuration  of  en route prescriptive information-->(*)
 <enRouteInfoTravelTimes>
  <enRouteInfoTravelTimeManual dayNumber="1" route="2" value="1" />
  <enRouteInfoTravelTimeManual dayNumber="2" route="2" value="1" />
  <enRouteInfoTravelTimeManual dayNumber="3" route="2" value="2" />
  <enRouteInfoTravelTimeManual dayNumber="4" route="2" value="1" />
```

```
  <enRouteInfoTravelTimeManual dayNumber="5" route="2" value="1" />
  <enRouteInfoTravelTimeManual dayNumber="6" route="2" value="1" />
  <enRouteInfoTravelTimeManual dayNumber="7" route="2" value="2" />
  <enRouteInfoTravelTimeManual dayNumber="8" route="2" value="1" />
……..
  </enRouteInfoTravelTimes>


  <postexinformation NrOfPastDays="5" showTravelTimes="true" showOtherRoutes="true"
DayText="Day" TravelTimeText="Travel time" text="You have chosen:"
textOtherRoutes="Here are the results of the other routes:" />
  <QuestionDays Everyday="false" />
  </scenario>
  </scenarios>
- <!--
xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
xxxxxxxxxxxxxxxxxxxxxxxxxx
  -->
<!—Configuration  of  the table of contents watched at staring simulation-->(*)
<screens>
 <navigation>
  <navigationstep number="1" text="Welcome to the TSL!" />
  <navigationstep number="2" text="Start questions" />
  <navigationstep number="3" text="Example" />
  <navigationstep number="4" text="Start simulation" />
  <navigationstep number="5" text="Questions" />
  <navigationstep number="6" text="Choose departure time" />
  <navigationstep number="7" text="Choose route" />
  <navigationstep number="8" text="Question result" />
  <navigationstep number="9" text="Closing questions" />
  <navigationstep number="10" text="Finished" />

  <navigationImage name="TSL" urlImage="Graphics\TSL_logo.gif" height="71"
width="175" />
  <navigationImage name="TUDelft" urlImage="Graphics\logo1.gif" height="100"
width="253" />
  </navigation>
xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
xxxxxxxxxxxxxxxxxxxxxxxxxx
<!—Call of the Introduction.html page[19]-->(*)

 <startscreen title="TSL:Introduction" content="TestRoby.html" />
 <startquestionsscreen title="TSL:Start questions">
xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
xxxxxxxxxxxxxxxxxxxxxxxxxx

<!—Preliminary questions before the starting simulation-->(*)
 <startquestions>
  <startquestion number="1" text="Fill in your name, please (*optional!)" type="text"
must="false" />

 <startquestion number="2" text="Fill in your birth year, please" type="listbox"
must="true">
 <startquestionlistitems>
  <startquestionlistitem text="1940-1945" value="1940-1945" />
  <startquestionlistitem text="1946-1950" value="1946-1950" />
```


[19] On this page (.html) is briefly described the experiment (the configuration and aim).

```
  <startquestionlistitem text="1951-1955" value="1951-1955" />
  <startquestionlistitem text="1956-1960" value="1956-1960" />
  <startquestionlistitem text="1961-1965" value="1961-1965" />
  <startquestionlistitem text="1966-1970" value="1966-1970" />
  <startquestionlistitem text="1971-1975" value="1971-1975" />
  <startquestionlistitem text="1976-1980" value="1976-1980" />
  <startquestionlistitem text="1981-1985" value="1981-1985" />
  <startquestionlistitem text="1986-1990" value="1986-1990" />
  </startquestionlistitems>

  </startquestion>
 <startquestion number="3" text="Select your Gender, please:" type="radio"
must="true">
 <startquestionlistitems>
  <startquestionlistitem text="Male" value="M" />
  <startquestionlistitem text="Female" value="F" />
  </startquestionlistitems>
  </startquestion>

</startquestion>
 <startquestion number="4" text="Fill in your Educational Level" type="listbox"
must="true">
<startquestionlistitems>
  <startquestionlistitem text="none" value="none" />
  <startquestionlistitem text="Primary School" value="PS" />
  <startquestionlistitem text="Secondary School" value="SS" />
  <startquestionlistitem text="Bachelor" value="B" />
  <startquestionlistitem text="Master" value="M" />
  <startquestionlistitem text="PhD" value="PhD" />
  </startquestionlistitems>

 </startquestion>
 <startquestion number="5" text="What's your job?" type="listbox" must="true">
<startquestionlistitems>
  <startquestionlistitem text="Architecture" value="Ar" />
  <startquestionlistitem text="Engineering" value="En" />
  <startquestionlistitem text="Finance/Economy/Banking" value="FEB" />
  <startquestionlistitem text="Healt/Biotechnologie/Veterinary" value="HBV" />
  <startquestionlistitem text="Law" value="L" />
  <startquestionlistitem text="Technical/Administration" value="Tc" />
  <startquestionlistitem text="Research" value="RS" />
  <startquestionlistitem text="Other" value="OT" />
  </startquestionlistitems>

  </startquestions>

<standardquestion number="1" text="Assign a value to arrive less than {0} minutes
early" type="listbox" must="true">
<standardquestionlistitems>
  <standardquestionlistitem text="Very pleasant" value="1" />
  <standardquestionlistitem text="Pleasant" value="2" />
  <standardquestionlistitem text="Neutral" value="3" />
  <standardquestionlistitem text="Unpleasant" value="4" />
  <standardquestionlistitem text="Very unpleasant" value="5" />
  </standardquestionlistitems>

  </standardquestion>
```

```
  <standardquestion number="2" text="Assign a value to arrive more than {0} minutes
early" type="listbox" must="true">
  <standardquestionlistitems>
  <standardquestionlistitem text="Very pleasant" value="1" />
  <standardquestionlistitem text="Pleasant" value="2" />
  <standardquestionlistitem text="Neutral" value="3" />
  <standardquestionlistitem text="Unpleasant" value="4" />
  <standardquestionlistitem text="Very unpleasant" value="5" />
  </standardquestionlistitems>

</standardquestion>
<standardquestion number="3" text="Assign a value to arrive less than {0} minutes
late" type="listbox" must="true">
 <standardquestionlistitems>
 <standardquestionlistitem text="Very pleasant" value="1" />
  <standardquestionlistitem text="Pleasant" value="2" />
  <standardquestionlistitem text="Neutral" value="3" />
  <standardquestionlistitem text="Unpleasant" value="4" />
  <standardquestionlistitem text="Very unpleasant" value="5" />
  </standardquestionlistitems>
  </standardquestion>

<standardquestion number="4" text="Assign a value to arrive more than {0} minutes
late" type="listbox" must="true">
 <standardquestionlistitems>
  <standardquestionlistitem text="Very pleasant" value="1" />
  <standardquestionlistitem text="Pleasant" value="2" />
  <standardquestionlistitem text="Neutral" value="3" />
  <standardquestionlistitem text="Unpleasant" value="4" />
  <standardquestionlistitem text="Very unpleasant" value="5" />
  </standardquestionlistitems>


xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
xxxxxxxxxxxxxxxxxxxxxxxxxx
<!—Call of the other pages to which are referred the successive simulation step[20]-->(*)

</standardquestion>
  </standardquestions>
  </startquestionsscreen>
  <examplequestionscreen title="TSL:Example" content="Example.html" />
  <startsimulation title="TSL:Start simulation" content="StartSim.html" />
  <questionscreen title="TSL:Day" usetravelinformation="true" />
  <choosedeparturetime title="TSL:Choose departure time" />
  <chooseroutescreen title="TSL:Choose route" />
  <resultscreen title="TSL:Results" />
 <closingscreen title="TSL:Roundup" content="">
xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
xxxxxxxxxxxxxxxxxxxxxxxxxx

<!—Configuration of the closing questions-->(*)
 <closingquestions>
 <closingquestion number="1" text="Was the Information provided by the System
useful for you?" type="radio" must="false">
```

[20] Each html file is referred to the successive simulation step: *Example.html* (an example is described);*StartSim.html* (the simulation is explained by a text);

```
<closingquestionlistitems>
<closingquestionlistitem text="Yes" value="1" />
  <closingquestionlistitem text="No" value="2" />
  </closingquestionlistitems>
  </closingquestion>

<closingquestion number="2" text="Could you describe the reason of your route choice preference? Route 1 was on average approximately 35 minutes but sometimes was extremely long; travel time on Route 2 was on average 53 minutes, without excessive dispersion over time; travel time on Route 3 was on average 47 minutes." type="text" must="false" maxLenght="20" />

<closingquestion number="3" text="Could you make your considerations about the quality of information provided?" type="listbox" must="false">
<closingquestionlistitems>
  <closingquestionlistitem text="VeryReliable" value="0" />
  <closingquestionlistitem text="Reliable" value="1" />
  <closingquestionlistitem text="Unreliable" value="2" />
  <closingquestionlistitem text="VeryUnreliable" value="3" />
  </closingquestionlistitems>
  </closingquestion>

<closingquestion number="4" text="Would you like be informed of the results of this research?" type="radio" must="false">
<closingquestionlistitems>
  <closingquestionlistitem text="Yes" value="1" />
  <closingquestionlistitem text="No" value="2" />
  </closingquestionlistitems>
  </closingquestion>

<closingquestion number="5" text="Would you like be informed about further research?" type="radio" must="false">
<closingquestionlistitems>
<closingquestionlistitem text="Yes" value="1" />
  <closingquestionlistitem text="No" value="2" />
  </closingquestionlistitems>
  </closingquestion>

  <closingquestion number="6" text="if you want, you may fill in your e-mail address to receive feedback and/ or information about further research (*optional!)" type="text" must="false" />
  </closingquestions>
xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
<!—Call of the last page of the simulation[21]-->
</closingscreen>
  <finishedscreen title="TSL:Finished" content="End.html" />
  </screens>
  </ExperimentConfigurationDataset>
```

[21] *End.html* contains the text referred to the end of the simulation.

REFERENCES

Abdel- Aty, M. A., (1998) *Modeling Incident Related Routing Decisions by Using Nested Logit Structure,* Transportation Research Record, 1645, pp. 103- 110

Abdel- Aty, M., A., Kitamura R. and Jovanis, P.P., (1996) *Using stated preference data for studying the effect of advanced traffic information drivers' route,* Transportation Research C, Vol 5, No.1, pp.389- 405

Adamowicz, W., Louviere, J. and Williams, M. (1994) *Combining stated and revealed preference methods for valuing environmental amenities*, Journal of Environmental Economics and Management 25: pp. 271-92

Adamowicz, W., Swait, J., Boxall, P., Louviere, J. and Williams, M. (1997) *Perceptions versus objective measures of environmental quality in combined revealed and stated preference models of environmental evaluation*, Journal of Environmental Economics and Management 32: pp. 65-84

Avinery, E., Prashker, J. N., *The impact of Travel Time Information on Travellers' Learning under Uncertainty*, 10th International on Travel Behaviour Research, Lucerne 10-15 August 2003

Ben-Akiva, M., E. (1973) *Structure of passenger travel demand models*. PhD Thesis, Department of Civil Engineering, MIT Cambridge

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

Ben-Akiva, M., E. and Morikawa, T. (1990) *Estimation of switching models from revealed preferences and stated intentions*, Transportation Research A 24 A(6): pp.485- 95

Ben-Akiva, M., E. and Morikawa, T. and Shiroshi, F. (1991) *Analysis of the reliability of preference ranking data,* Journal of Business Research 23: pp. 253-68

Bierlaire, M. BIOGEME: a free package for the estimation of discrete choice models, *Proceedings of the 3rd Swiss Transportation Research Conference*, Ascona, Switzerland, 2003

Bierlaire, M. *Users' Manual BIOGEME*: Estimation of discrete choice models with BIOGEME 1.6, March 16 2008

Bierlaire, M., Themans, M., Axhausen, K., (2006): *Analysis of driver's response to real- time information in Switzerland*, European Transport n.34, pp. 21- 41

Bifulco, G.N., Simonelli, F., Di Pace, R. (2007) *The role of the uncertainty in ATIS* applications Proceedings of 12th On-line Conference on Soft Computing in

Industrial Applications, to be published in “Applications of Soft Computing: updating the state of the art”. Advances in Intelligent and Soft Computing vol. 52, Springer

Bogers, E. and van Zuylen, H., (2004) *The influence of reliability on route preferences in freight transport*, Transportation Research Board

Bogers, E., Bierlaire, M., Hoogendoorn, S. (2006) *Modelling Learning in Route Choice,* Transportation Research Record

Bradly, M.A. and Daly, A.J. (1994) *Use of the logit scaling approach to test rank order and fatigue effects in stated preference data*, Transportation 21(2): pp.167-84

Cascetta, E., (2001) Transportation Systems Engineering: Theory and Methods, pp. 331-348

Cascetta, E., Cantarella, G.E. , (1993) *Modelling dynamics in transportation networks: State of the art and future developments*, Elsevier

Cascetta, E., Cantarella, G.E., (1991) *A day- to day and within – day dynamic stochastic assignment model*, Transportation Research Part A 25(5) pp. 277- 291

Chang Gang- Len and Hani S. Mahmassani, (1988) *Travel Time Prediction and Departure Time Adjustment Behaviour Dynamics in Congested Traffic System*, Transportation Reasearch Part B, 22(3): pp. 217-232

Chen, P., Srinivasan, K., K., and Mahmassani, H., S., (1999) *Effect of Information Quality on Compliance Behaviour of Commuters Under Real-Time Traffic Information*, Transportation Research Record 1676

Crittin, F. and Bierlaire, M. (2001) *New Algorithmic Approaches for the Anticipatory Route Guidance Generation Problem*, STRC Conference-1st Swiss Transport Research Conference

Daly, A. (1987) *Estimating “tree” logit models*. Transportation Research Part B: Methodological, 21 (4): pp. 251-268

De Florio, F., P. (2003), *Evaluation of Reactive dynamic route guidance strategy*, Transportation Research C 11 pp. 375-388

De Groot, T., Hellendoorn, M. (2004) *Travel Simulator Laboratory (TSL) Master Thesis* University of Technology Press, Delft The Netherlands

Dell’Orco, M.; Kikuchi, S. (2006) A Mathematical Model for Evaluation of Information Effects in ATIS (Advanced Traveler Information Systems) *Environment in Soft Computing: Methodologies and Applications, Springer Berlin Ed*., pp. 53-70

Dingus, T. A., Hulse H. C., (1993) *Some human factors design issues and* *recommendations for automobile navigation information systems*, Transportation Research 1C(2), pp. 119- 132

Emmerink, R. H. M., Axhausen, K. W., Nijkamp P., Rietveld, P., (1994) *Effects of Information in road transport networks with recurrent congestion*, Transportation 22: pp. 21- 53

Emmerink, R., Nijkamp, P., (1999) *Behavioural Network Impacts of Driver Information Systems*, Ashgate

Erev, I., Y. Bereby-Meyer and A. Roth, (1999) *The effect of Adding a Constant to Al Payoffs: Experiment*: *Experimental Investigation and Implications for Reinforcement Learning Models* Journal of Economic Behaviour and Organization, Vol. 39: pp. 111-128

Ettema, D., Timmermans, H., (2006)*Costs of travel time uncertainty and benefits of travel time information: Conceptual model and numerical examples*, Transportation Research Part C pp. 335- 350

Fujii, S. and Kitamura R., (2000)*Anticipated Travel Time Information Acquisition and Actual Expeience: The case of Hanshin Expressway Route Closure*, paper 00-0712 Presented at 79th Annual Meeting of the Transportation Research Board, Washington, DC

Guo, Z., Srinivasan, K., (2004) *Investigating the effect of User Behaviour Factors and Transportation Control Measures on Day- to – day Network Evolution and Trip- time reliability*, Transportation Research Board

Hall, R., (1996) *Route choice and advanced traveller information systems on a capacitated and dynamic network*, Transportation Research 4C(5), pp. 289-306

Hani S. Mahmassani, Yu- Hsin Liu, (1999) *Dynamics of commuting decision behaviour under advanced traveller information*, Transportation Research Part C

Hato, E., Taniguchi, M., Sugie ,Y. Kuwahara, M., Morita, H., (1999) *Incorporating an information process into a route choice model with multiple information sources,* Transportation Research Part C pp. 109- 129

Heathington, K. W., Worrall, R. D., and Hoff, G. C. (1971) *Attitudes and Behaviour of Drivers Regarding Route Diversion* Highway Research Record, 363, pp. 18-26

Hensher, D. A. and Bradley, M. (1993) *Using stated response choice data to enrich revealed preference discrete choice models*, Marketing Letters 4(2): pp. 139-51

Hensher, D.A. (1998) *Extending valuation to controlled value functions and not uniform scaling with generalized unobserved variances*, in Garling, T. , Laitila, T. and Westin (eds.), Theoretical foundations of travel choice modelling, Oxford:

Pergamon: pp. 75- 102

Horowitz, J., (1984) *The stability of stochastic equilibrium in a two- link transportation network*, Transportation Research B 18(1), pp. 13- 28

Huchingson, R. D., McNees, R. W., and Dudek, C. L. (1977) *Survey of Motorist Route-Selection Criteria* Transportation Research Record, 643, pp. 45-48.

Iida, Y., Akiyama, T. and Unchida, T. (1992) *Experimental Analysis of dynamic route choice behaviour*, Transportation Research Part B 26 pp. 17- 32;

Jane Lappin, Jom Bottom, (2001) *Understanding and predicting Travel Response to Information: Literature Review* pp31- 35

Jayakrishnan, R. Mahmassani, H., Hu, T-Y, (1993) *An evolution tool for advanced traffic information and management systems in urban networks,* Transportation Research 2C(3), pp. 129- 148

Jha, M., Madanat, S., Peeta, S., (1998) *Perception updating and day-to-day travel choice dynamics in traffic networks with information provision*, Transportation Research Part C

Katsikopoulos, K.V., Fisher, D.L., Duse-Anthony, Y, Duffy, S.A., (2002) *Risk attitude reversals in drivers' route choice when range of travel time is provided*, Human Factors 44(3): pp. 466-473

Khattak, A. J., J. L. Scofer and F. S. Koppelman, (1993) *Commuters' Enroute Diversion and Return Decisions: Analysis and Implications for Advanced Traveller Information Systems* Transportation Reasearch A 27A(2): pp101- 111

Khattak, A. J., Y. Yim and Stalker, (1999)*Does Travel Information on Commuter and Non commuter Behaviour? Result from the San Francisco Bay Area Travinfo Project* Transportation Research Record 1694: pp. 48- 58

Khattak, Asad, J. , Koppelman, Frank, S. and Schofer, Joseph, L., (1993) *Stated Preference for investigating commuters' diversion propensity* Transportation20: pp. 107- 127

Kobayashi, K., (1994) *Information, rational expectation and network equilibria- an analytical perspective for route guidance system*s, The Annals of Regional Science 28: pp. 369- 393

Lo, H., Szeto, W. Y. (2002a) *A methodology for sustainable traveller information services*, Transportation Research B 36: pp. 113-130

Lo, H., Szeto, W. Y. (2002b) *Modelling advanced traveller information services: static versus dynamic paradigms*, Transportation Research B 38: pp. 495-515

Louviere, J., J., Hensher, D., A. and Swait, J., D., (2000) *Stated Choice Methods: Analysis and Application*, Cambridge University Press

Mahmassani, H. , Chang, G., (1986) *Experiments with departure time choice dynamics of urban commuters*, Transportation Research B 20(2): pp. 297- 320

Mahmassani, H. S., Jou, R.C. (1998) *Transferring insights into commuter behaviour dynamics from laboratory experiments to field surveys*, Transportation Research Part A: Policy and Practice 34(4): pp. 243-260

Mahmassani, H. S., Liu, Y. H. (1999) *Dynamics of commuting decision behaviour under advanced traveller information systems*, Transportation Research Part C 7: pp. 91- 107

Morikawa, T. (1989) '*Incorporating stated preference data in travel demand analysis', PhD Dissertation*, Department of Civil Engineering, MIT

Muth, J. F., (1961) *Rational expectations and theory of price movements* Econometria, 29: pp. 315- 335

Nolan, R.B., and Polak, J.W, *Travel Time Variability: A Review of Theoretical and Empirical Issues* , Paper Submitted to the 80th Annual Meeting of the Transportation Research Board

Pedersen, D. M. (1998) *Factors in Route Selection*, Perceptual and Motor Skills, 86(3(part 1)): pp. 999-1006

Peter Train, K., (2001) *Discrete Choice Methods with Simulation*, Cambridge University Press, pp. 164- 166

Polak, J. and, Jones, P., (1993) *The acquisition of pre- trip Information: A stated preference approach* Transportation 20: pp. 179-198

Polak, J. W. and Oladeinde, F. , *An empirical model of travellers' day to day learning in the presence of uncertain travel time*s. In M.G.Bell and C.Cassir (ed.), Realiability in Trasport Network. Research Studies Press, Hertfordshire

Polydoropoulou, A., M. Ben- Akiva, (1999) *The effect of Advanced Traveller Information Systems (ATIS) on Travellers' Behaviour*, Behavioural and Network Impacts of Driver Information Systems, R. Emmerink and P. Nijkamp, Ashgate, Chapter 14: pp. 317- 352;

Ratcliffe, E. P. (1972) *A Comparison of Drivers' Route Choice Criteria and Those Used in Current Assignment Processes* Traffic Engineering and Control, 13, pp. 526-529

Schreckenberg, M., Selten, R., *Human Behavior and Traffic Networks* (2004) Springer, pp. 97- 128;

Shen- Te Chen, P., Srinivasan, K. K. and Mahmassani, H. S., (1999) *Effect of Information Quality on Compliance Behavior of Commuters Under Real-Time Traffic Information*, Transportation Research Record, Vol. 1676: pp. 53- 60

Simon, H.A., (1955) A *behavioural model of rational choice, Quarterly Journal of Economics,* 69 SU, pp. 99-118

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

Srinivasan, K., K. and Mahmassani, H., S. (Febr. 2002) *Dynamic Decision and Adjustment Processes in Commuter Behavior under real- time Information*, Research Report SWUTC/02/167204-1 prepared by the Centre for Transportation Research Bureau of Engineering Research University at Austin, pp. 97-112

Swait, L., Louviere, J. and Williams, M. (1994) *A sequential approach to exploiting the combined strengths of SP and RP data: application to freight shipper choice,* Transportation 21: pp. 135- 52

Targa, F. Khattak, A., Yim, Y., (2003) *Understanding access and use of dynamic travel information*, Igert Conference, University of California, Davis

Thakuriah, P. & Sen, A., (1996) *Quality of Information given by advanced traveller information systems,* Transportation Research C 4(5) pp. 249-266

Tindemans, H. D., Van Hofstraeten, A., Verhetsel, A. & Witlox, F. (2003) A spatial approach to the analysis of household activity surveys in Belgium. Working paper presented at the EC- workshop, Behavioural responses to ITS.

TSL homepage, S.P. Hoogendoorn, Delft University of Technology, http://www.tsl.tudelft.nl/, Accessed July 2005

Van der Mede, H., J. and Van Berkum, E., C. (1996) *Modelling Car Drivers' Route Choice Information Environments,* Delft University of Technology, Transportation Planning and Traffic Engineering Section Rudimental Contributions

Viti, F., Bogers, E., A. I., Hoogendoorn, S., P., (2005) *Day- to- Day learning under uncertainty and with information provision: model and data analysis*, ISTT

Wachs, M. (1967) *Relationships Between Drivers' Attitudes Toward Alternate Routes and Driver and Route Characteristics* Highway Research Record, 197, pp. 70-87

Wahle, J., Bazzan, A., L. C., Klugl, F., Schrekenberg, M., (2002) *The impact of real- time information in a two route scenario using agent- based simulation*, Transportation Research Part C 10 pp. 399- 417

Wardman, M. , Bonsall, P.W. and Shires, J., D., (1996) *Driver Response to Variable message signs: a stated preference investigation,* Transportation Research C, Vol 5, No.6, pp. 389- 405

Wilks, S. S. (1962) Mathematical Statistics, New York: Wiley

Yang, H. and Meng, Q., (2001), *Modelling user adoption of advanced traveller information systems: dynamic evolution and stationary equilibrium*, Transportation Research A 35 , pp. 895-912

Yin, Y. and Yang, H. (2003) *Simultaneous determination of the equilibrium market penetration and compliance rate of advanced traveller information systems,* Transportation Research A 37 pp. 165- 181