Consideration of alternatives:

development of hybrid modelling approaches

and applications to transport mode choice

Mauro Capurso

Submitted in accordance with the requirements for the

degree of Doctor of Philosophy

The University of Leeds

Institute for Transport Studies

March 2019

Declaration of contribution

The candidate confirms that the work submitted is his own, except where work which has formed part of jointly authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.

Peer reviewed papers

The following jointly authored papers have been submitted alongside this thesis:

The work in Chapter 2 of this thesis has been accepted for publication:

Capurso, M., Hess, S., and Dekker, T. (2019), *Modelling the role of consideration of alternatives in mode choice: an application on the Rome-Milan corridor*. Transportation Research Part A: Policy and Practice (*forthcoming*).

I developed the main idea for this work, designed the survey, and collected the data under the guidance of Stephane Hess and Thijs Dekker. I performed the data analysis, the modelling work, and wrote most of the manuscript. Stephane Hess and Thijs Dekker contributed to the idea and writing of limited sections of the manuscript, provided recommendations on the modelling and comments on the results. The manuscript was improved by comments from all the co-authors.

The work in Chapter 3 of this thesis is a manuscript ready for submission:

Capurso, M., Hess, S., and Dekker, T. (ready for submission), *Stated consideration and attribute thresholds in mode choice models: a hierarchical ICLV approach*.

I developed the main idea for this work, designed the survey, and collected the data, under the guidance of Stephane Hess and Thijs Dekker. I performed the data analysis, the modelling work and wrote the manuscript. Stephane Hess and Thijs Dekker contributed to the idea and writing of limited sections of the manuscript, provided recommendations on the modelling and comments on the results. The manuscript was improved by comments from all the co-authors.

The work in Chapter 4 of this thesis has appeared in publication as follows:

Bergantino, A.S., **Capurso, M.**, Dekker, T. and Hess S. (2019), *Allowing for heterogeneity in the consideration of airport access modes: the case of Bari airport*. Transportation Research Record (*forthcoming*).

In this paper, authors are listed in alphabetical order. I developed the main idea for this work, under the guidance of Thijs Dekker and Stephane Hess. I designed the survey with Angela Stefania Bergantino, who also coordinated the data collection. I performed the data analysis, the modelling work and wrote the manuscript. Stephane Hess and Thijs Dekker provided recommendations on the modelling and commented on the results. The manuscript was improved by comments from all the co-authors.

Acknowledgments

First and foremost, I would like to thank my supervisors, Stephane Hess and Thijs Dekker for their extremely valuable and generous support. Their feedback have vastly improved the quality of the research in this thesis and their resourcefulness and dedication to research have constantly motivated me to grow as a researcher.

I would also like to thank my co-author and mentor Angela Stefania Bergantino, who has also played an important role in shaping my career as a researcher.

I am also grateful to my fellows and friends of the Institute for Transport Studies, who have made my time in Leeds special.

Finally, a special thanks goes to Liliana for her infinite patience and omnipresent help throughout these years. Thanks also to my family for their boundless support.

Abstract

Transport mode choice models traditionally assume that individuals consider all available alternatives. However, this might not be reasonable, even when the number of alternatives is limited. The biggest challenge for analysts is, however, that consideration sets - defined as the sets of alternatives relevant to the individuals - are unobservable. When only choices are observed, it is impossible to identify whether this is the result of not considering the non-chosen alternatives or preferring the chosen alternative over the other alternatives in the consideration set. Hence, only reduced form models can be estimated to understand the driving factors behind the observed choice. This thesis contributes to the literature on consideration by investigating how direct and indirect indicators of consideration, collected during stated choice surveys, can be used to measure and better understand the role of consideration sets. It aims at answering three research questions. The first is how to measure consideration. By comparing the suitability of a series of indicators of consideration, it emerges that thresholds for attributes are the most informative indicators of consideration. A second research question concerns the driving factors of consideration. Results suggest that thresholds for attributes have a primary role in explaining consideration, and that thresholds themselves are a function of individuals' socio-economic profiles. A third research question relates to the role that consideration effects play in the estimation of the state-of-the-art RUM-based discrete

choice models. Compared to previous studies, it emerges that the identified impact of controlling for consideration effects reduces as soon as the analyst also accounts for unobserved preference heterogeneity. Although calibrated to specific transport choices (i.e. choice of the transport mode for long-distance and airport access trips), the methodology proposed (and, to a limited extent, the findings of this research) can be generalised to other contexts with similar characteristics, e.g. number of alternatives or degree of differentiation amongst the alternatives. These could be either in the transport sector (e.g. mode choice on either systematic or non-systematic urban trips, car purchasing choices, choice of a flight on long-distance trips) or even in other contexts such as purchasing decisions on durable household goods (e.g. TV) or consumer packaged goods (e.g. breakfast cereals or pasta, wine, cosmetics/fragrances).

Contents

List of Tables

List of Figures

Chapter 1 - Introduction

Discrete choice models based on the Random Utility Maximisation (RUM) framework (McFadden, 1974) are widely used to model individuals' decisions in a variety of fields, such as transport, health, marketing, and environment (for a recent overview, see Hess and Daly, 2014). Within this framework, individuals are described as utility-maximisers who select the alternative that provides the highest level of utility. Individuals are thus assumed to be able to assign utility levels to each of the available alternatives and compare those levels. The utility of an alternative is formed by the attributes describing the alternative (Lancaster, 1966). Making a decision requires trade-offs within and across the alternatives (i.e. a good performance on one attribute may compensate for poor performance on another). The analyst can never fully capture the decision process used by individual decision makers, and the use of utility theory in itself is an approximation. To acknowledge this uncertainty, the RUM model treats part of utility as 'random' or unobservable.

Implicit in the use of discrete choice models is the assumption that the analyst can correctly specify the set of alternatives available to a specific individual, i.e. the 'choice set'. In revealed preference (RP) studies, the choice set is unobserved and the analyst is required to make assumptions regarding its composition at the modelling stage. In the context of RP studies, deterministic rules have been used to identify the individual-specific choice set, i.e. the so-called 'expected choice set' (Manski, 1977). For

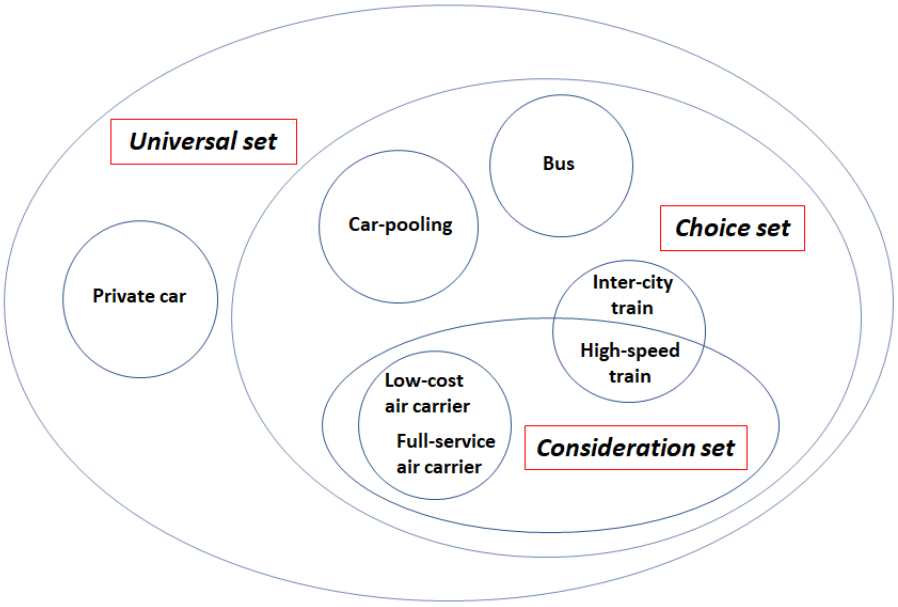
example, in mode choice studies, private car might be removed from the choice set for individuals without a driving licence, or bus for those working/living in areas not served by public transport (Swait and Ben-Akiva, 1987a). In stated preferences (SP) studies, the choice set is defined by the analyst through the experimental design.

In both RP and SP studies, it is however possible that the choice set specified by the analyst excludes relevant alternatives or includes irrelevant alternatives (DeShazo et al., 2009). The risks associated with mis-specification of the choice set in RUM-based choice models have been illustrated by Williams and Ortuzar (1982) and Swait (1984). Namely, mis-specification of the choice set can result in biased parameters estimates and model forecasts. In turn, if the parameter estimates and forecasts are biased, policy and managerial decisions could be also affected (Pancras, 2010; Draganska and Klapper, 2011).

Only in specific circumstances an analyst may obtain consistent estimates of the parameters and choice probabilities when relevant alternatives are excluded from the choice set, i.e. when at least the chosen alternative is included in the choice set (McFadden, 1978). The risks associated with the inclusion of irrelevant (i.e. too many) alternatives, instead, have not been adequately investigated. However, there is no reason that the above risks of mis-specification may not hold since these irrelevant alternatives are assigned a non-zero choice probability which will inherently affect the parameter estimates (Thill, 1992).

The focus of this thesis is on such 'consideration sets' - defined as the set of alternatives relevant to the individual, those that s/he would actually consider, or those 'acceptable' for her/him (Howard and Sheth, 1969; Wright and Barbour, 1977; Hensher and Ho, 2015) – which comprise a subset of the full 'choice set'. For example, in a long-distance transport mode choice context (Figure 1.1), the 'universal set' would comprise all alternatives which are objectively available to at least one individual in the target population (high-speed and inter-city train, full-service and low-cost air carrier, bus, car-pooling and private car). The 'choice set' and the 'consideration set', instead, would be defined at the level of a specific individual. The former subset will contain the alternatives effectively available to the individual, e.g. it will not contain private car if the individual does not have a car. The latter subset, instead, will include alternatives effectively taken into account in the decision-making process, which number might be lower or equal to the number of alternatives in the 'choice set'. For example, the 'consideration set' might only include alternatives which allow the individual to be at destination within a certain amount of time (e.g. high-speed train, full-service and low-cost air carrier), and choice will be made out of these transport modes only.

Figure 1.1 Sets involved in the decision-making process. An example of transport mode choice



Making choices from a subset of the choice set has a strong behavioural intuition. In real life, when individuals are faced with many available alternatives to choose from (e.g. residential choice or consumer goods), it is reasonable to assume that they adopt a (often non-compensatory) screening strategy to reduce the number of alternatives (e.g. *elimination-by-aspect*, Tversky, 1982), and then use a compensatory decision rule (e.g. RUM) to make a final decision. This allows the individuals to rapidly screen

many alternatives whilst reducing cognitive and search costs[1] (Hauser, 2014). The use of consideration sets is therefore 'rational', and consistent with a cost-benefit trade-off (Stigler, 1961). Such behaviour might still be realistic when the number of alternatives is limited, (e.g. transport mode choice), if individuals have *a priori* strong negative preference towards certain alternatives (due, for example, to past experience or to environmental concerns), or depending on the overall context in which the decision is made (e.g. the purpose of the trip).

The biggest challenge is, however, that while consideration sets are known to the individual decision-makers, they cannot be observed by the analyst. Moreover, consideration sets might vary across choice occasions and might depend on specific attribute values of the alternatives. This thesis contributes to the literature on consideration of alternatives by investigating how different indicators of consideration can be used to measure this latent aspect of the decision-making process. Additionally, this research focuses on the impact that accounting for consideration effects has on the estimation of the state-of-the-art RUM-based discrete choice models in the context of transport mode choice.

The remainder of this introduction is structured as follows. It starts by providing the theoretical background behind consideration of the

[1] A similar choice heuristic, *attribute non-attendance*, postulates that certain attributes, rather than alternatives, are irrelevant to (and hence ignored by) the individuals (Hensher, 2006).

alternatives in discrete choice models. Then, it outlines the research questions for the thesis and presents the data used to answer the research questions. Finally, it presents an executive summary of the three papers.

1. The theoretical background

This subsection discusses the approaches proposed in the literature to 'tackle' the consideration set generation problem, defined as the process of establishing the set of alternatives considered by the individual. Even if the majority of the studies reviewed in this Section are empirically grounded in transport decision contexts (and particularly mode choice), the focus will only be on the methodology proposed to model consideration. There will be no specific reference or resume of the insights gained with respect to the specific decision context investigated (e.g. if long-distance vs. urban commuting trips or airport access), since the composition of consideration sets and the drivers of consideration are strongly context dependent.

Although the consideration set may be known to the individual, this aspect has been treated probabilistically in the literature to reflect the fact that the consideration set cannot be directly observed by the analyst. Instead, the analyst observes the final outcome, i.e. the choice.

The literature on this topic originated in the context of RP studies. In these studies, a mis-specification of the choice set can be attributed to lack of awareness and/or availability of the alternatives at the level of the individual, but also to the use of a simplifying choice heuristic. The analyst might also lack information about the context in which the choice was made

(e.g. if the individual was on a business trip). In SP studies, instead, the analyst can relate the presence of consideration sets almost exclusively to the use of simplifying choice heuristics, since individuals are made aware of the available alternatives in the experiment, and information on the choice context is typically collected.[2]

Amongst the first contributions on the topic, Manski (1977) suggested to account for the uncertainty around the composition of the consideration set $C^*_{n,t}$, as a subset of the choice set $C_{n,t}$, expressing the unconditional choice probability for alternative *i* for individual *n* in choice task *t* as (Equation 1.1):

$$P_{i,n,t} = \sum_{C^*_{n,t} \subseteq C_{n,t}} \pi_{n,t}\left(C^*_{n,t}\right) \cdot P_{i,n,t}\left(C^*_{n,t}\right) \tag{1.1}$$

Where $\pi_{n,t}\left(C^*_{n,t}\right)$ is the probability of an individual using a specific consideration set $C^*_{n,t}$, and $P_{i,n,t}(C^*_{n,t})$ is the choice probability conditional on that consideration set. According to this formulation, typically referred to as the *two-stage* model, all possible combinations of alternatives (i.e. $2^J - 1$, where *J* is the number of alternatives in the choice set) have a probability of being the 'true' consideration set. The conditional choice probability above, $P_{i,n,t}(C^*_{n,t})$, can be represented by any choice model, e.g. a multinomial logit (MNL) model. In effect, the *two-stage* Manski expression is similar to a *latent*

[2] Many studies aimed at integrating the uncertainty about the 'available' and 'considered' alternatives in the context of RP. Nevertheless, the two processes are comparable in terms of modelling (Ben-Akiva and Boccara, 1995).

class model with preference parameters kept constant across the classes (Hensher and Green, 2003; Campbell et al. 2014, Calastri et al., 2019).

It is straightforward to notice that this formulation becomes computationally infeasible for even a relatively small number of alternatives. One way to reduce the dimensionality of the consideration set generation problem might be to place *a priori* restrictions on the number of possible consideration sets. For example, the analyst can decide to allow for a minimum or a maximum number of alternatives in the consideration set, or to assume that certain alternatives are always considered. To the same end, Swait and Ben-Akiva (1987a) formulated a model, known as the *independent availability* model, which assumes that the probability of having an alternative in the consideration set is independent on the probability for the other alternatives. Within the Manski expression above, the probability of facing a specific consideration set, $\pi_{n,t}\left(C_{n,t}^{*}\right)$, can therefore be represented as the product of the independent probabilities of the alternatives of being (or not being) in the set.

This assumption might appear counterintuitive, since it implies that consideration of an alternative would be both statistically and behaviourally independent of the consideration of other alternatives. Nevertheless, the *independent availability* formulation allows achieving a good compromise between behavioural realism and computational parsimony for two reasons. First, the independent availability model is actually more consistent with RUM than the Manski model, since it is not context dependent (i.e.

consideration of an alternative does not depend on the attributes and hence consideration of other alternatives). Second, it does not impose any restrictions on the consideration set space (e.g. minimum or maximum number of alternatives in the set); all that it requires is J parameters to be estimated (i.e. one for each alternative) rather than $2^J - 1$.

Swait and Ben-Akiva's study effectively marks a step from the modelling of unobserved consideration sets towards that of the consideration of single alternatives, for which they also propose a behavioural explanation.[3] They hypothesise that constraints of a different nature (e.g. infrastructural, economical, psychological) might act upon the individuals and determine the consideration of alternatives. Basar and Bhat (2004) and Cantillo and Ortúzar (2005) proposed similar constraint-based approaches, hypothesising that individuals might only consider alternatives which provide a certain utility level, or with attributes within some (unobserved) thresholds, respectively. Campbell et al. (2014) also tried to overcome the computational limitations of the *two-stage* Manski formulation, although differently from the *independent availability* model. These authors suggested that consideration sets could include alternatives within exogenously set cost thresholds. Calastri et al. (2019) separately modelled unobserved availability and consideration of the alternatives, using an

[3] This represents another difference with the Manski model, in which consideration set probabilities are random entities.

independent availability formulation for both, i.e. they included two independent probabilistic layers before choice.

The shift towards the modelling of consideration of single alternatives (rather than of combinations of alternatives, i.e. consideration sets) made some researchers suggest that this could be implicitly approximated within the utility function. In the *implicit availability/perception* model, Cascetta and Papola (2001) proposed to add to the utility of each alternative a continuous 'inclusion' function, bounded between 0 and 1, which enters in the logarithmic form (Equation 1.2):

$$U_{i,n,t} = V_{i,n,t} + \log\left(a_{i,n,t}\right) + \varepsilon_{i,n,t} \tag{1.2}$$

According to this formulation, values of $a_{i,n,t}$ close to 0 indicate that the alternative is not considered, i.e. the utility gets heavily discounted. When $a_{i,n,t}$ approaches 1, the alternative is considered, and no discounting of utility is applied.

In a similar vein as the previously reviewed studies, $a_{i,n,t}$ has been expressed as a function of attributes of the alternatives or socio-economic characteristics of the individual. The *constrained multinomial logit* (CMNL) model proposed by Martinez et al. (2009) is a constraint-based version of the model by Cascetta and Papola (2001), in the sense that it recognises the role of attribute thresholds in driving consideration of the alternatives. The CMNL has been shown to be an effective solution to model unobserved

availability of the alternatives in choice context with a large number of alternatives, as for example residential location (see Haque et al. (2019) who proposed an improved version of the CMNL).

The *implicit availability/perception* model and the *independent availability* model share a probabilistic notion of consideration of single alternatives, despite substantially differing in the way they represent it. The former essentially 'simulates' consideration by adding a non-compensatory component to the traditional fully compensatory utility function. The latter instead allows for different behaviours to be separately represented in the two phases of the decision-making process, i.e. consideration and choice. In terms of modelling, this means that the former approach is easier to implement, given that it does not require the enumeration of all possible consideration sets. Here, the analyst only needs to specify one set, to which each alternative will belong with a certain degree of membership $(a_{i,n,t})$.

Bierlaire et al. (2010) empirically compared the *constrained multinomial logit* model with the Manski's *two-stage* model. They showed that these models were providing different distributions for the consideration probabilities, and therefore these had to be considered as completely distinct approaches to model consideration set generation. Their conclusion could be explained by the fact that differently from the Manski's *two-stage* model, consideration probabilities in the CMNL model are a function of selected attributes of the alternatives.

All the aforementioned studies considered the situation that the only information available to the analyst is that on the final outcome, i.e. the observed choice. When this is the case, it is empirically impossible to separate the consideration set generation stage from the choice stage. A number of authors explored the possibility of eliciting (for example, during stated choice experiments or with dedicated surveys) supplementary information on consideration (on the composition of the consideration set or on aspects related to consideration of the single alternatives), which can be used to measure this aspect of the decision-making process.

Ben-Akiva and Boccara (1995) and Swait (2001) used indicators on perceived availability of the alternatives and thresholds for attributes, respectively, to identify consideration set probabilities within the context of a *two-stage* model. Enam and Choudhury (2011) collected information on consideration of alternatives from a small separate sample of individuals to infer consideration set probabilities on a larger RP dataset. Horowitz and Louviere (1995) also use self-reports on consideration to identify consideration set probabilities within the context of a *two-stage* model, but they argue that this information might provide little additional information over that already contained in the observed choices, since these are both reflections of utility. Hensher and Rose (2012), instead, used indicators on acceptability of the alternatives and thresholds for attributes to 'scale' the utility expression for each alternative. Their model formulation is therefore both behaviourally and methodologically closer to the model proposed by Cascetta and Papola

(2001), since the utility for 'unconsidered' alternatives gets discounted, and therefore the choice probabilities for these alternatives reduce.

Finally, Hensher and Ho (2015) treat the stated acceptability of the alternatives in the choice set as a direct measure of the consideration set, and accordingly model the choice of the 'observed' consideration set and the selected alternative (conditional on the respective consideration set).

The use of such supplementary information, however, is not without problems. These indicators cannot be used as *error-free* measures of consideration in choice models for two reasons. First, there is potential for measurement error, since these might not correspond to actual levels of consideration. Second, there is scope for endogeneity bias as these indicators might themselves be a function of the utility for the alternatives. Given this, it might be preferable to treat these indicators of consideration as dependent rather than independent variables, using, for example, a *latent variable* approach (Ben-Akiva and Boccara, 1995; Bolduc et al., 2005; Hess and Hensher, 2013).

2. The research questions

This thesis aims at answering three research questions, empirically grounded in the context of specific transport mode choices:

2.1 How can we measure consideration?

If consideration effects actually play a role in individuals' decision-making processes, one would expect that different indicators used to measure these latent (i.e. unobserved) consideration effects across individuals and choice situations would provide similar results.

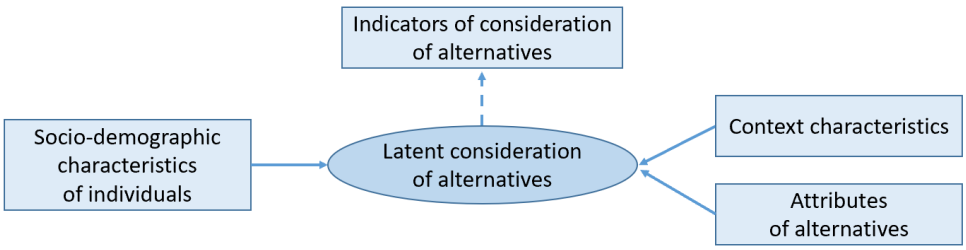
It has been suggested that 'direct' reports of consideration, elicited from the individuals during stated choice experiments, could be used to measure consideration (Horowitz and Louviere, 1995). The use of such statements shows similarities with that of self-reports of 'availability' or 'acceptability' of the alternatives, used in previous studies as proxy for consideration (Ben-Akiva and Boccara, 1995; Hensher and Rose, 2012; Hensher and Ho, 2015). Alternatively, one can measure consideration 'indirectly', using supplementary information on, for example, self-reported thresholds for attributes (Swait, 2001), or on individuals' perceptions towards the alternatives.

The first objective of this thesis is therefore to investigate the 'quality' of those indicators, i.e. to compare them, and ultimately, to provide recommendations on which of them is more appropriate to measure consideration of the alternatives in the context of transport mode choice.

2.2 What drives consideration?

This research question is directly related to the previous one. The use of supplementary information (directly or indirectly) related to consideration (the indicators of consideration) enables the identification of the factors driving consideration and choice, respectively. For example, latent consideration might depend on socio-demographic characteristics of the individuals, on context characteristics, or attributes of alternatives (Figure 1.2).

Figure 1.2 Latent consideration



Note: Items in rectangles can be directly observed by the analyst while Items in the ellipses are unobserved. The broken arrow indicates a measurement component, while plain arrows indicate structural components.

Given that the driving factors of consideration might partially or completely match with those of choice, the separation between the driving factors of these two aspects (i.e. the identification of the individual role of each factor in driving either one or both aspects) is otherwise impossible if the only information available to the analyst is that on choice - as it is typically the

case in many studies. Previous research, however, did not take enough advantage of such supplementary information and primarily measured consideration (i.e. as an additional explanatory variable) without further exploring which factors exclusively drive it.

From an industry perspective, it would be interesting to understand not only if an alternative is considered or not (both because if an alternative is not considered it cannot be chosen, and to identify the set of relevant competitors), but also why an alternative is considered or not, or by which individuals. Even the same individual for a trip on the same O/D might consider different sets of alternatives depending if the trip is for business or for leisure purposes (i.e. the decision context also matters). This could provide new opportunities to develop more effective marketing strategies.

2.3 What role do consideration effects play in the estimation of RUM-based discrete choice models?

Previous studies highlight the benefits that could be obtained from accounting for the consideration of alternatives in the estimation of discrete choice models, particularly in terms of the behavioural representation of individuals' decision-making process, i.e. model fit. These models are also deemed to provide less biased estimates of the parameters and forecasts, despite not knowing the direction of the 'eventual' bias since the true data generating process is unknown.

The role consideration effects play in the estimation of RUM-based discrete choice models needs to be evaluated in light of the substantial econometric advances from the last decade, such as the capacity of these models to accommodate for complex substitution patterns and preference heterogeneity (McFadden and Train, 2000; Hensher and Greene, 2003).

It is possible that a share of what is generally explained as unobserved preference heterogeneity could be due to consideration effects, similarly with what has been unveiled by previous research on other latent effects, such as attribute attendance and variety seeking (Hess and Hensher, 2013; Song et al., 2018). When only choices are observed, all we can do is to estimate reduced form models to understand their driving factors. The utility function will combine all the different stages of the decision-making process, and, if we include random parameters, these will pick up some of these unobserved effects. By adding info on consideration, we might disentangle this from other unobserved effects.

Therefore, the third objective of this research is to investigate the role that accounting for consideration of the alternatives plays alongside preference heterogeneity. This would avoid the risk of putting too much emphasis on the former (or the latter) in the estimation of discrete choice models.

We only decided to focus on RUM-based discrete choice models because only RUM can be properly used for welfare analysis. Moreover, since previous literature analysed consideration in the context of RUM-based models it was easier to make comparisons.

3. The data used in this thesis

To answering the three research questions, data from two stated choice (SC) experiments on transport mode choice have been designed and administered during the PhD. There are two reasons why transport choice contexts have been chosen. First, because the number of alternatives in transport related choice context is generally limited. Given this, it was possible to collect detailed information on consideration of the alternatives with minimal burden on the respondents. Second, because the literature providing the theoretical background for this thesis unveiled the benefits of accounting for this aspect mainly in this field. Previous studies can therefore serve as a basis for comparison of the empirical conclusions.

Amongst a large variety of transport related decision contexts, then, two specific mode choice contexts have been selected in this thesis, namely mode choice on long-distance and on airport access. The rationale for selecting these specific contexts was that the number of alternatives was limited but at the same time sufficiently large to allow for the possibility that respondents were not effectively considering (i.e. making complete trade-offs) all available alternatives when making their decision. Another characteristics of the chosen contexts was the degree of differentiation amongst the alternatives with respect to multiple dimensions (e.g. both travel time and travel cost).

The details of the two datasets used in this thesis are briefly summarised hereafter.

3.1 First dataset: Mode choice on the Rome-Milan corridor (Italy)

This SC experiment was administered in April and May 2016 to a sample of 209 travellers on the Rome-Milan corridor (approximately 600 km). Respondents were mainly recruited in-person in train stations, bus terminals, and airports in Rome and Milan. A small portion of respondents has been also recruited in service stations on the highway, located around half way between Rome and Milan, and online. The experiment has been administered using a tablet.

A layout like that of an online journey planner has been used to increase the realism of the experiment, and thereby increase involvement of the respondents. The choice set comprises seven alternatives: high-speed rail, inter-city rail, full-service air carrier, low-cost air carrier, bus, car-pooling, and private car. Choice tasks (6 for each respondent) were designed using a Bayesian D-efficient experimental design.

Supplementary information on task-level consideration of the alternatives and on the presence of thresholds for attribute were also collected. The first is a 'direct' indicator of consideration, while the second is an 'indirect' one. Additional information on the dataset, on the data collection and on the questionnaire are in Appendix B.

3.2 Second dataset: Airport access mode choice to Bari International Airport (Italy)

This SC experiment has been administered in November and December 2017 to a sample of 746 residents in four cities in a range of 50-100 km from Bari International Airport, who travelled through the airport in the previous three months. Revealed preferences (i.e. actual choice of the access mode) have been also collected, which refer to the respondents' last trip to the airport.

The data was collected under the scientific supervision of Professor Angela Stefania Bergantino, as part of the research project "*An Analysis of demand for the Apulian airport system*" of the Department of Economics, Management and Business Law of the University of Bari (Italy). The project obtained a research grant by Aeroporti di Puglia Spa.

The choice set comprises five alternatives: public transport with at least one change, a direct private bus run by the airport management, car as driver, car as passenger (i.e. the possibility of being dropped-off by someone else), and taxi. Choice tasks (5 for each respondent) were designed using a Bayesian D-efficient experimental design.

Three sets of supplementary information on consideration of the alternatives were also collected. The first set contains perception statements towards the alternatives, and a ranking of the alternatives on the base of respondents' overall preference. The second set refers to respondents' past experience: they were asked to report how many times

they used each of the airport access alternatives in the previous year. The third set comprises choice-task consideration of the alternatives.

Data from a similar SC experiment administered one year earlier (November 2016) on a sample of 300 air travellers from the same study area has been used for models' validation. The validation sample also contains revealed preferences, but not the indicators for consideration. Additional information on the dataset and on the data collection are in Appendix C.

4. Outline of the papers

The first two papers (presented in Chapters 2 and 3, respectively), analyse the mode choice survey on the Rome-Milan corridor. The main difference between the two papers is that the indicators of consideration, i.e. stated consideration for alternatives and stated thresholds for attributes, are used alternatively (in the first paper) rather than jointly (in the second) to measure consideration. We acknowledge that the difference between the two papers might not appear significant; however, each paper can contribute to answering our research questions from a different perspective. In the second paper, the simultaneous use of the indicators to measure consideration do not only increase the robustness of the insights we derive from the investigation, but it also allows to show that the two components of consideration, the one working at the individual level and the one working at the choice level are related. The alternative use of the indicators, instead, can give an idea of which one of the two components is stronger.

In the first paper we start estimating a set of models explaining the indicators of consideration as a function of socio-economic characteristics and attributes of the alternatives. Predicted values for the probability of considering the alternatives and attribute specific threshold levels are then used to obtain consideration probabilities in two distinct two-stage probabilistic choice models (Manski, 1977, Swait and Ben-Akiva, 1987a).

In the second paper, instead, we present an Integrated Choice and Latent Variable (ICLV) model (McFadden, 1986; Ben-Akiva et al., 2002; Bolduc et al., 2005) with inter-related latent variables describing consideration effects. The novelty of the proposed approach, defined as 'hierarchical ICLV', is that latent thresholds for attributes explain latent consideration of the alternatives. The latter are then used to reduce the utility, and therefore choice probability, of the alternatives by means of a 'discounting' factor. Differently from the traditional two-stage probabilistic choice model used in the first paper, the discounting approach does not require enumerating (and modelling) of all possible consideration sets since consideration effects are directly introduced in the utility function of the alternatives.

The third paper (presented in Chapter 4) investigates the role of consideration of the alternatives in the context of airport access mode decisions, using the data collected amongst users of Bari airport. Like in the second paper in this thesis, the indicators of consideration collected during the SC experiment are modelled together with choice within an Integrated Choice and Latent Variable framework (though not in a hierarchical way),

and latent consideration enters the utility of an alternative through a 'discounting' factor. Differently from the first two papers, both revealed and stated preferences have been used in the third paper. The main rationale for conducting this study and for employing a similar methodology as the one employed for the second paper was to provide evidence that the observed results were not an artefact of the methodology employed in conjunction with a specific dataset, but that the resulting *pros* and *cons* of that methodology could be also transferred to other choice contexts.

References

Başar, G. and Bhat, C. (2004) A parameterized consideration set model for airport choice: an application to the San Francisco bay area. *Transportation Research Part B: Methodological*, 38(10), 889-904, https://doi.org/10.1016/j.trb.2004.01.001.

Ben-Akiva, M., and Boccara, B., (1995) Discrete choice models with latent choice sets. *International Journal of Research in Marketing*, 12(1), 9-24, https://doi.org/10.1016/0167-8116(95)00002-J.

Bierlaire, M., Hurtubia, R. and Flötteröd , G. (2010) Analysis of implicit choice set generation using a constrained multinomial logit model. *Transportation Research Record*, 2175, https://doi.org/10.3141/2175-11.

Bolduc, D., Ben-Akiva, M., Walker, J. and Michaud, A. (2005) Hybrid choice models with logit kernel: applicability to large scale models. In *Integrated land-use and transportation models: behavioural foundations*,

edited by M. E. H. Lee-Gosselin and S.T. Doherty, 275-302. Emerald Group Publishing Limited.

Calastri, C., Hess, S. Choudhury, C. F., Daly, A., and Gabrielli, L. (2019) Mode choice with latent availability and consideration: Theory and a case study. *Transportation Research Part B: Methodological* 123, 374-385, https://doi.org/10.1016/j.trb.2017.06.016.

Campbell, D., Hensher, D. A., Scarpa, R. (2014) Bounding WTP distributions to reflect the 'actual' consideration set. *Journal of Choice Modelling* 11, 4-15, https://doi.org/10.1016/j.jocm.2014.02.004.

Cantillo, V., and Ortúzar, J. D. D., (2005) A semi-compensatory discrete choice model with explicit attribute thresholds of perception. *Transportation Research Part B: Methodological*, 39(7), 641-657, https://doi.org/10.1016/j.trb.2004.08.002.

Cascetta, E., and Papola, A. (2001) Random utility models with implicit availability/perception of choice alternatives for the simulation of travel demand. *Transportation Research Part C: Emerging Technologies*, 9(4), 249-263, https://doi.org/10.1016/S0968-090X(00)00036-X.

DeShazo, J. R., Cameron, T.A., Saenz, M. (2009) The effect of consumers' real-world choice sets on inferences from stated preference surveys. *Environmental and Resource Economics*, 42(3), 319-343, https://doi.org/10.1007/s10640-008-9250-8.

Draganska, M., and Klapper, D. (2011) Choice set heterogeneity and the role of advertising: An analysis with micro and macro data. *Journal of Marketing Research*, 48(4), 653-669, https://doi.org/10.1509/jmkr.48.4.653.

Enam, A. and Choudhury, C. F. (2011) Methodological issues in developing mode choice models for Dhaka, Bangladesh. *Transportation Research Record*, 2239(1), 84-92, https://doi.org/10.3141/2239-10.

Haque, M. D., Choudhury, C. F., and Hess, S. (2019) Modelling residential location choices with implicit availability of alternatives. *Journal of transport and land use*, forthcoming.

Hauser, J. R. (2014) Consideration-set heuristics. *Journal of Business Research*, 67(8), 1688-1699, https://doi.org/10.1016/j.jbusres.2014.02.015.

Hensher D. A. (2006) How do respondents process stated choice experiments? Attribute consideration under varying information load. *Journal of Applied Econometrics*, 21(6), 861-878, https://doi.org/10.1002/jae.877.

Hensher, D. A., Greene, W. H. (2003) The mixed logit model: the state of practice. *Transportation*, 30(2), 133–176, https://doi.org/10.1023/A:1022558715350.

Hensher, D. A., and Ho, C. (2015) The role of perceived acceptability of alternatives in identifying and assessing choice set processing strategies in stated choice settings: The case of road pricing reform. *Transportation*

Research Part E: Logistics and Transportation Review, 83, 225-237, https://doi.org/10.1016/j.tre.2015.09.012.

Hensher, D. A., and Rose, J. M. (2012) The influence of alternative acceptability, attribute thresholds and choice response certainty on automobile purchase preferences. *Journal of Transport Economics and Policy,* 46(3), 451-468.

Hess, S. and Daly, A. (2014) *Handbook of choice modelling: the state of the art and the state of practice*. Edward Elgar, Cheltenham

Hess, S., and Hensher, D. A. (2013) Making use of respondent reported processing information to understand attribute importance: a latent variable scaling approach. *Transportation*, 40(2), 397-412.

Horowitz, J. L., and Louviere, J. J. (1995) What is the role of consideration sets in choice modeling? *International Journal of Research in Marketing*, 12(1), 39-54, https://doi.org/10.1016/0167-8116(95)00004-L.

Howard, J.A., Sheth, J.N. (1969) *The theory of buyer behavior*, John Wiley, New York

Kløjgaard, M. E. and Hess, S. (2014) Understanding the formation and influence of attitudes in patients' treatment choices for lower back pain: testing the benefits of a hybrid choice model approach. *Social Science & Medicine*, 114, 138-150, https://doi.org/10.1016/j.socscimed.2014.05.058.

Lancaster, K.J. (1966) A new approach to consumer theory. *Journal of*

political theory, 74(29), 132-157, https://doi.org/10.1086/259131.

Manski, C. F. (1977) The structure of random utility models. *Theory and decision*, 8(3), 229-254, https://doi.org/10.1007/BF00133443.

Martínez, F., Aguila, F., and Hurtubia, R. (2009) The constrained multinomial logit: A semi-compensatory choice model. *Transportation Research Part B: Methodological*, 43(3), 365-377, https://doi.org/10.1016/j.trb.2008.06.006.

McFadden, D. (1974) Conditional logit analysis of qualitative choice behaviour. *Frontiers in Econometrics*, edited by P. Zarembka, 105-142. Academic Press, New York.

McFadden, D. and Train, K. (2000) Mixed MNL models for discrete response*, Journal of Applied Econometrics*, 15(5), 447-470, https://doi.org/10.1002/1099-1255(200009/10)15:5<447::AID-JAE570>3.0.CO;2-1.

Mariel, P. and Meyerhoff, J. (2016) Hybrid discrete choice models: gained insights versus increasing effort, *Science of the total environment*, 568, 433-443, https://doi.org/10.1016/j.scitotenv.2016.06.019.

Pancras, J. (2010) A framework to determine the value of consumer consideration set information for firm pricing strategies. *Computational Economics*, 2010, 35(3), 269-300, https://doi.org/10.1007/s10614-009-9193-3.

Song, F., Hess, S., and Dekker, T. (2018) Accounting for the impact of variety-seeking: theory and application to HSR-air intermodality in China, *Journal of Air Transport Management*, 69, 99-111, https://doi.org/10.1016/j.jairtraman.2018.02.008.

Stigler, G. J. (1961) The economics of information. Journal of Political Economy, 69, 213-225, https://doi.org/10.1086/258464.

Swait, J. (1984) Probabilistic choice set generation in transportation demand models. PhD dissertation, Massachusetts Institute of Technology, Cambridge, USA.

Swait, J. (2001) A non-compensatory choice model incorporating attribute cutoffs. *Transportation Research Part B: Methodological*, 35(10), 903-928, https://doi.org/10.1016/S0191-2615(00)00030-8.

Swait, J. and Ben-Akiva, M. (1987) Incorporating random constraints in discrete models of choice set generation. *Transportation Research Part B: Methodological,* 21(2), 91-102, https://doi.org/10.1016/0191-2615(87)90009-9.

Thill, J.C. (1992) Choice set formation for destination choice modelling. *Progress in Human Geography*, 16(3), 361-382, https://doi.org/10.1177/030913259201600303.

Tversky, A. (1972) Elimination by aspects: a theory of choice. *Psychological Review*, 79, 281–299, https://dx.doi.org/10.1037/h0032955.

Vij, A. and Walker, J. L. (2016) How, when and why integrated choice and latent variable models are latently useful, *Transportation Research Part B: Methodological*, 90, 192-217, https://doi.org/10.1016/j.trb.2016.04.021.

Williams, H. C. W. L. and Ortúzar, J. D. D. (1982) Behavioural theories of dispersion and the mis-specification of travel demand models. *Transportation Research Part B: Methodological*, 16(3), 167-219, https://doi.org/10.1016/0191-2615(82)90024-8.

Wright, P. and Barbour, F. (1977) Phased decision strategies: sequels to an initial screening. In *Multiple criteria decision making*, edited by M. Starr and M. Zeleny. North Holland TIMS Studies in Management Science, Amsterdam.

Chapter 2 - Modelling the role of consideration of alternatives in mode choice: an application on the Rome-Milan corridor

Abstract

In this paper, we investigate the role consideration of the alternatives plays in mode choice models. On the Rome-Milan corridor, in Italy, where seven alternative modes of transport are available, we administered a stated choice (SC) experiment. Responses to supplementary questions on consideration of the different modes of transport and the presence of thresholds for the travel time attribute indicate travellers are less likely to consider the *slower* modes. Two model specifications, in which consideration for the slower alternatives is measured using both sets of supplementary questions, are proposed and contrasted against a model which assumes all alternatives are considered. Our results suggest that some of the unobserved preference heterogeneity could potentially be due to consideration effects. Accounting for consideration of alternatives also has direct impacts on choice probabilities, parameter estimates and willingness-to-pay measures.

Keywords

Consideration of alternatives, mode choice, willingness-to-pay

1. Introduction

The question which of the available alternatives an individual decision-maker considers when making a choice has been a topic of interest in the transportation and marketing literature over the last decades (Manski, 1977; Swait and Ben-Akiva, 1987a-b; Shocker et al., 1991; Roberts and Lattin, 1997; Swait, 2001; Cantillo and Ortúzar, 2005). Behaviourally, considering only a subset of the available alternatives is consistent with the use of task-simplifying heuristics. The latter can be driven by, amongst other things, (self-imposed) thresholds for attributes (e.g. maximum price levels), or searching costs.

Consideration effects are not only relevant in the context of a large number of alternatives (e.g. residential choice and consumer goods), but also when the number of alternatives is limited (e.g. in the case of transport mode choice). Demand models not accounting for consideration have been argued to provide less precise - or even biased - parameter estimates and forecasts of consumer choices (Williams and Ortúzar, 1982; Swait, 1984). From a commercial perspective, a more comprehensive understanding of the role consideration plays in the decision-making process provides new opportunities to develop more effective marketing and pricing strategies (Pancras, 2010; Draganska and Klapper, 2011).

Consideration of the alternatives, as a part of the decision-making process, cannot be directly observed and therefore measured with certainty/without error. A number of authors estimated consideration endogenously, relating

it to some observed attributes of the alternatives (Cascetta and Papola, 2001; Cantillo and Ortúzar, 2005; Martinez et al., 2009). This has been the preferred approach in the presence of knowledge only on the final outcome of the choice process, i.e. on the observed choice.

Other authors instead attempted to directly elicit, i.e. measure, consideration using supplementary questions during surveys. These questions either referred to the perceived 'availability' or 'acceptability' of alternatives (Ben-Akiva and Boccara, 1995; Hensher and Rose, 2012; Hensher and Ho, 2015), or to the presence of thresholds for attributes (Swait, 2001).

The use of supplementary information, however, has its own limitations. These indicators should not be considered as *error-free* measures of consideration. First, there might not be a one-to-one correspondence between stated and actual consideration. That is, there is the potential for measurement error. Second, there is scope for endogeneity bias as these measures may be correlated with other unobserved factors. Third, the indicators might not be suitable (and/or available) for forecasting. This paper serves as an illustration of how to overcome these limitations, allowing the analyst to 'safely' make use of such supplementary information, and thereby aiding identification of consideration effects in the decision-making process. We particularly suggest to explain indicators of consideration as a function of socioeconomic characteristics of the individuals and attributes of the alternatives, i.e. to use the indicators as dependent rather than as

explanatory variables. The 'predicted' values of the indicators are subsequently used as a proxy for consideration in a series of two-stage probabilistic choice models (Manski, 1977; Swait and Ben-Akiva, 1987a). The proposed use of the indicators is therefore similar to the latent variable approach presented in Ben-Akiva and Boccara (1995) in the context of the traditional two-stage approach.

Our indicators for consideration, namely *stated consideration* of the alternatives and *stated thresholds* for attributes have been elicited from the respondents of a stated choice (SC) experiment in the context of transport mode decisions on the Rome-Milan corridor, in Italy. We here contrast two model specifications, which respectively make use of *stated consideration* and *stated thresholds* as indicators of consideration, against a reference model not accounting for consideration effects. In both models, some of the elements conventionally attributed to unobserved preference heterogeneity are now alternatively treated as consideration effects. Parameter estimates and willingness-to-pay measures are affected, particularly when *stated consideration* is used. Compared to previous studies, the identified impact of controlling for consideration effects is however limited. In contrast to Ben-Akiva and Boccara (1995), and comparable studies (e.g. Basar and Bhat, 2004), our models control for consideration effects alongside unobserved preference heterogeneity. Our reference mixed logit model is inherently more flexible than the multinomial logit model adopted in previous studies, and thereby (perhaps incorrectly)

already captures some of the consideration effects, but does not explain them as such. Likewise, solely controlling for consideration effects may put too much emphasis on the role consideration effects play.

The remainder of the paper is structured as follows. We review the relevant literature in Section 2 and describe the case study and the available data in Section 3. Section 4 lays out the empirical strategy and model specifications. In Section 5, we discuss the estimation and forecasting results. Finally, Section 6 concludes.

2. Literature review

Discrete choice models based on the Random Utility Maximisation (RUM) framework are widely used to model individuals' decisions in a variety of fields, particularly transport. Standard discrete choice models treat the choice set, i.e. the set of available alternatives, as given. However, in many circumstances, individuals might not be aware of all available alternatives and (or) employ simplifying choice heuristics. From the perspective of an analyst it is impossible to judge whether the individual has made the decision from a restricted choice set or not when only the final outcome of the decision-making process (i.e. the choice) is observed.

Mis-specifications of the choice set can arise in the context of revealed and stated preference studies. With respect to the latter, despite choices being presented in a controlled experimental setting, i.e. the choice set (potentially of limited size) is designed by the analyst considering the alternatives effectively available, individuals may still apply additional choice

heuristics which further reduce the size of the consideration set (see Hauser, 2014, for a review of such heuristics). In this paper, we treat the composition of the emerging consideration set as probabilistic due to the unobserved nature of this part of the decision-making process.[4]

Choice models making use of a probabilistic consideration set are commonly presented as a variation of the model proposed by Manski (1977). According to this formulation, typically referred to as the *two-stage* model, all $2^J - 1$ (where J is the number of available alternatives) possible combinations of alternatives have a probability of being the *true* consideration set.[5] Conditional on each consideration set, there exists a conditional probability of choosing a given alternative from the consideration set. The expected (or unconditional) choice probability is defined as the sum of weighted conditional (upon the consideration set) choice probabilities.

Although behaviourally appealing, this formulation becomes computationally infeasible for a large number of alternatives. For example, with 5 alternatives, there are already 31 possible consideration sets, and this number increases to 63 with 6 alternatives, 127 with 7 alternatives etc.. Based on Manski's model, several formulations have been proposed in the transportation literature in an attempt to overcome this limitation whilst

[4] The use of restricted choice sets is just one of many process rules or simplifying choice heuristics that individuals can adopt. For example, they might also ignore certain attributes of the alternatives (Hensher et al., 2005).

[5] In the remainder of this Section we will refer to the 'consideration set' regardless of whether the reviewed studies aimed at modelling unobserved availability of the alternatives with revealed preference data or unobserved consideration with either revealed or stated preference data.

providing a behavioural interpretation of the consideration set generation process.

Swait and Ben-Akiva (1987a) assume in the *independent availability logit* model that the probability of an alternative being included in the consideration set is independent of that of the other alternatives. This formulation still requires the enumeration of all possible consideration sets. However, only J independent probabilities need to be estimated (instead of $2^J - 1$). Moreover, these authors hypothesise that random (i.e. unobserved) constraints of a different nature (e.g. physical, psychological, economical) act upon the individuals and determine consideration probabilities. Similar constraint-based approaches have been proposed by Basar and Bhat (2004) and Cantillo and Ortúzar (2005). In the former paper the authors assume in their *probabilistic choice set multinomial logit* model that an alternative is excluded from the consideration set if its consideration utility is lower than a threshold consideration utility level. In the latter paper unobservable threshold levels for attributes of the alternatives are modelled as functions of characteristics of the decision maker and choice conditions (e.g. purpose of the trip).

Gaudry and Dagenais (1979) attempted to reduce the dimensionality of the consideration set generation problem by assuming that individuals either consider all alternatives (i.e. the consideration set coincides with the universal set), or they might be captive to just one alternative (i.e. the consideration set contains only the chosen alternative). In their formulation,

the captivity odds are specified as simple constants; however, these can alternatively be modelled as functions of socioeconomic variables and attributes of the alternatives, as suggested by Swait and Ben-Akiva (1987b). Besides the above two-stage approaches, other conceptual models have been proposed accounting for the consideration set generation process. For example, Cascetta and Papola (2001) assume that unavailability of specific alternatives can be modelled through the use of penalty parameters directly discounting their utility (see also Martinez et al., 2009).

All the aforementioned approaches consider the situation that the only information available to the analyst is that on the final outcome (i.e. the observed choice). It is therefore empirically impossible to separate the consideration set generation stage from the choice stage and thus identify which factors drive each stage respectively. To overcome this limitation, a number of authors explored the possibility of measuring consideration of the alternatives using supplementary information on this stage collected during stated choice experiments.

For example, Ben-Akiva and Boccara (1995) and Swait (2001) use indicators on perceived availability of the alternatives and thresholds for attributes, respectively, to identify consideration set probabilities within the context of the two-stage model. Hensher and Rose (2012) use indicators on alternatives' acceptability and thresholds for attributes to 'scale' the utility expression for each alternative. Hensher and Ho (2015), instead, treat the stated acceptability of the alternatives in the choice set as a direct measure

of the consideration set and accordingly model the choice of the 'observed' consideration set and the selected alternative (conditional on the respective consideration set).

In the remainder of the paper we work in the framework of the two-stage model developed by Manski (1977), under the assumption of independence of consideration of the alternatives proposed by Swait and Ben-Akiva (1987a).[6] We believe that this model is more in line with the notion of consideration of alternatives since an alternative is either considered (included in the choice set) or not (not included), differently from the approaches where this binary inclusion/exclusion is approximated by a smooth function (as, for example, in Cascetta and Papola, 2001). With the aim of separating the consideration stage from the choice stage we use supplementary information on *stated consideration* of the alternatives and *stated thresholds* for attributes. We model these indicators as functions of socioeconomic characteristics of the individuals and attributes of the alternatives, and subsequently use their predicted values as proxy for consideration. Our use of the indicators is similar to the latent variable approach employed - in the context of a two-stage model - by Ben-Akiva and Boccara (1995), in the sense that stated indicators for consideration and thresholds are treated as dependent rather than *error-free* independent

[6] We address the dimensionality problem by further reducing the size of the consideration set by making the simplifying assumption, based on descriptive statistics on stated consideration, that only a subset of alternatives is probabilistically considered.

variables. The main difference with Ben-Akiva and Boccara is that we simultaneously control for consideration effects and unobserved preference heterogeneity, to avoid the risk of putting too much emphasis on the role of the former. With respect to thresholds for attributes, these are compared with the presented attribute levels, as a mechanism for the acceptance or rejection of alternatives. However, differently from other constraint-based approaches, such as those employed by Swait and Ben-Akiva (1987a) or Cantillo and Ortúzar (2005), information on thresholds for attributes is directly available in our study. Unlike Swait (2001), we do not use these thresholds as *error-free* measures of consideration.

3. The case study

3.1 The Rome-Milan Corridor

The Rome-Milan corridor represents an interesting case study to investigate consideration effects among medium-long distance passengers. Individuals can choose amongst seven alternatives (i.e. transport modes): high-speed and inter-city trains, full-service and low-cost flights, bus and car-pooling services, and private car.[7] These alternatives are not homogeneous in terms of core (e.g. travel time and cost), and soft attributes (e.g. Wi-Fi availability and comfort). Hence, it is reasonable to assume that travellers might not consider all the alternatives in their mode choice decision.

[7] Alternatives, such as walking, cycling, or indirect public transport options are also available. However, these were considered infeasible during the design process due to extremely long travel times.

At the time of the data collection (April-May 2016), in the high-speed rail (HSR) market, *Trenitalia* and *Nuovo Trasporto Viaggiatori* were between them offering 65 daily services in both directions, which were taking slightly less than 3 hours. *Trenitalia* was also offering 3 Inter-City (IC) trains. These were slower and could take up to 7 hours. In the air market, the full-service carrier (FSC) *Alitalia* was offering 25 daily services to/from Rome and Milan city airports (Fiumicino and Linate) and 3 to/from Milan Malpensa airport. At the latter airport, *Alitalia* was competing with the low-cost carrier (LCC) *EasyJet* (2 services) and with another FSC, *Meridiana* (2 services). A dozen scheduled coach services were also offered by *Stagecoach-Megabus*, *Flixbus*, and *Baltour*, including over-night services. These coach services were characterised by cheap fares (from €1 with *Stagecoach-Megabus*), and travel times were ranging between 7 and 11 hours. Finally, the car alternative on this corridor was available as a private or a shared mode of transport. The car-pooling network *Bla-bla-car* was connecting riders and passengers willing to share the cost of a 6-hour trip.[8] Finally, the car alternative on this corridor is available as a private or a shared mode of transport. The car-pooling network *Bla-bla-car* connects riders and passengers willing to share the cost of a 6-hour trip.

[8] In the last few years there have been some changes to the competitive environment on the Rome-Milan corridor. In the air market, *EasyJet* decided to abandon its slots on the Rome Fiumicino - Milan Malpensa route (since October 2017), and *Meridiana* was relaunched as *Air Italy* (since March 2018). In the bus market, both the Italian branch of *Stagecoach-Megabus* and *Baltour* joined the *Flixbus* network (in June 2016 and October 2018, respectively), and €1 fares are no longer available.

The Italian Authority for Transport Regulation (ART, 2015) provides the official figures with respect to modal shares on this corridor. In 2014, 24% of passengers travelled by air, 65% by train, and the remaining 11% by bus and car.

3.2 Survey design and descriptive statistics

In the absence of an online journey planner where all alternatives are presented simultaneously, an individual needs to: 1) decide which alternatives to consider from those s/he is aware of, and search on the respective websites; 2) process the information available regarding price and non-price attributes of the considered alternatives; 3) end the process by choosing the preferred alternative or decide to consider more alternatives and repeat steps 1-3 until s/he has made the choice. In this process, some relevant alternatives might be left out due to unawareness or searching costs.

The advent of the Internet has substantially lowered searching costs. Websites such as www.goeuro.com and www.rome2rio.com allow users to compare services for the available modes on a specific route according to travel time, cost etc., and offer the opportunity to purchase tickets. At the same time, alternatives that consumers were previously unaware of might now be chosen. Transport operators report increasing shares of tickets being

purchased online on their official websites[9] and for some operators the Internet is the only available marketplace (car-pooling and bus).

Against this background, we designed a SC survey mimicking a real purchasing decision through an online journey planner. Its visual design was comparable to online journey planners as increasingly used by individuals to make travel plans.[10] In online journey planners, however, all 'objectively' available (i.e. feasible) alternatives are presented, including private transport means (e.g. car) which might not be available to everyone. We acknowledge that the inclusion of all alternatives in the choice set might be questionable and contrasts with typical adjustments made in SC experiments - where choice sets are customised around respondents' personal situation. This might be considered a limitation of the data used in this paper.

The experiment was conducted in Rome and Milan between April and May, 2016.[11] A total of 209 on-site face-to-face TAPI (Tablet Assisted Personal Interview) surveys were administered to travellers going from Rome to Milan (and *vice versa*) while waiting at the platform for their train (57%), at the bus terminals (17%), or in the proximity of the airports (12%). We also

[9] The HSR operator Trenitalia reports that more than 50% of tickets are purchased online (2017).

[10] We indirectly assumed that all respondents were actually familiar with the use of an online journey planner; even if this was not necessarily true for all respondents, this does not necessarily mean that it was harder for them to engage in the experiment. None of the respondents reported difficulties in understanding the layout used.

[11] Prior to final administration to travellers on the corridor, the survey has been individually discussed with international Masters' and PhD students in the transport discipline.

administered a smaller portion of surveys online (8%), and in two service stations on the A1/E35 highway, located around half way between Rome and Milan, in the proximity of Bologna (6%).[12]

Table 2.1 presents descriptive statistics of the sample and of the Italian population. The sample lacks of representativeness of the Italian population by age (particularly amongst those aged 50 and over) and levels of education. Although national-level statistics on wage income are not available, this sample may over-represent individuals with monthly wage income greater than 2,000 €. The extent to which these differences represent a limitation cannot be defined with certainty. On the one hand, younger individuals were more willing to participate in this SC experiment, and this might effectively represent an issue. On the other hand, we should acknowledge that younger individuals tend to travel more compared to their older counterparts (Isfort, 2017). In addition, the Rome-Milan corridor is the most heavily travelled corridor for business purposes in Italy, and thus it is likely that travellers on this route are more educated and wealthier than the Italian population as a whole.

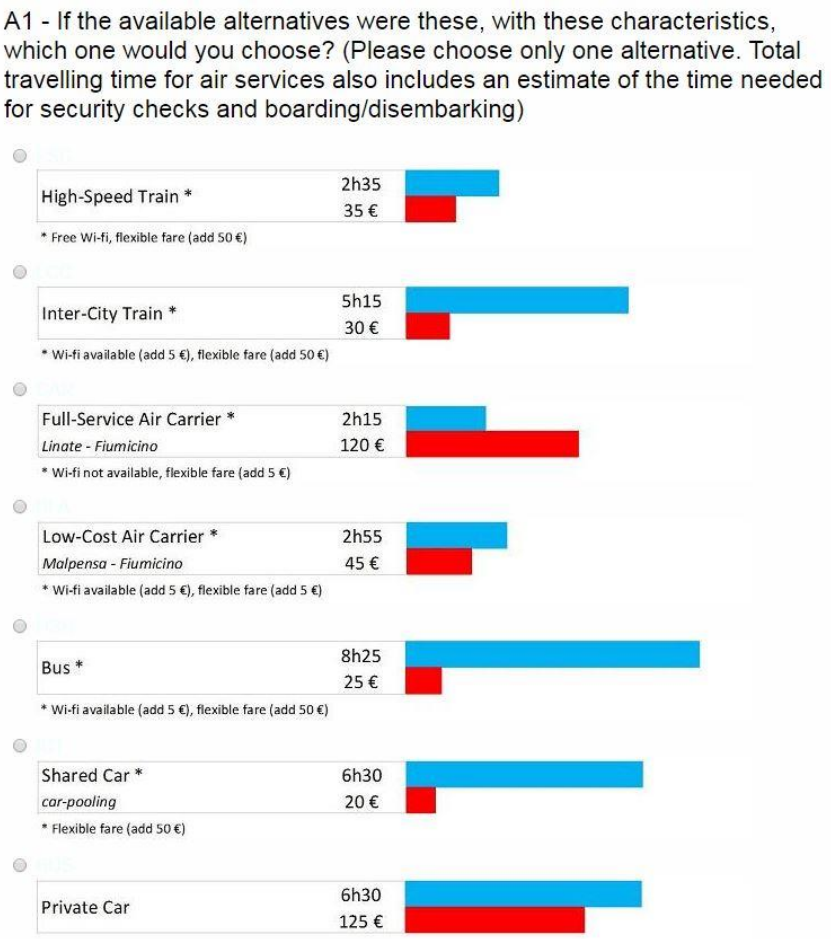
[12] The response rate was higher at bus and train stations (≈ 50%) than at airports and service stations (≈ 20%).

Table 2.1. Descriptive statistics of the sample and of the Italian population (over 18 years old)

Social traits/Year	**Survey**	**Italy (ISTAT, 2018)**
Female	41.6%	51.3%
Age: 18-24	22.0%	6.0%
Age: 25-34	39.2%	13.4%
Age: 35-49	26.3%	26.7%
Age: 50+	12.4%	53.8%
Income: < 500 €	23.9%	-
Income: 500-1,000 €	5.3%	-
Income: 1,000-2,000 €	5.7%	-
Income: 2,000-4,000 €	32.5%	-
Income: >4,000 €	17.2%	-
Income: prefer not to disclose	15.3%	-
Education BSc+	70.0%	15.6%
Business trip	47.8%	-
Student	23.9%	-
Total	209	49,505,718

Each respondent completed six choice tasks, and we used a layout similar to the one displayed by the website www.goeuro.com (Figure 2.1). To avoid possible ordering effects, we randomised the order of the presented alternatives across respondents

Figure 2.1 The layout of the choice tasks



The attributes of the alternatives were travel time[13], travel cost, ticket flexibility, and level of connectivity on-board (Wi-Fi). The attributes all referred to a standard one-way trip between Rome and Milan. In Table 2.2, we report the ranges for travel time and cost for all alternatives on this particular route at the time of the SC survey (*current ranges*, i.e. as displayed on operators' websites), as well as those used in the survey design. The latter were designed around the former, or around values which are expected to be feasible in the near future. For example, the HSR operator *Trenitalia* has already announced it could potentially further reduce travel time between the two cities by increasing speed up to 350km/h. With respect to ticket flexibility, we used three levels, i.e. the possibility of changing the ticket for free, or to do it with a fee of €5 or €50. Wi-Fi availability was also presented in three levels, namely not available, available for free, or available at a fee of €5. We set the choice tasks using a Bayesian D-efficient experimental design, with *priors* drawn from the literature or based on our expectations (Rose et al., 2008). Finally, we decided not to remove strictly dominant alternatives because the independent usage of price discrimination strategies by transport operators sometimes allows for some alternatives to be cheaper and faster than others.

[13] Access/egress time in large cities might play an important role in situations like the one modelled in this experiment. However, due to software restrictions it was not possible to customise the experiment depending on respondents' departure/arrival place. We collected information on respondents' distance (in minutes) from/to departure/arrival place and HSR stations, principal and secondary airports, and bus terminals. This information was accordingly used as a respondent-specific explanatory variable in the choice model.

Table 2.2 Current ranges and survey attribute levels

	Current ranges				Attribute levels	
	Travel time (h/min)		*Travel cost (€)*		*Travel time (h/min)*	*Travel cost (€)*
Alternatives	*min*	*max*	*min*	*max*		
HSR	2h55	4h28	19.9	209	2h15, 2h35, 2h55, 3h20, 3h40	20, 35, 50, 90, 120
IC	6h27	6h50	9	79	5h15, 6h, 6h45, 7h30, 8h15	10, 30, 45, 60, 80
FSC	2h20[1]		55.71	244.71	1h45, 2h, 2h15, 2h30, 2h50	50, 80, 120, 180, 280
LCC	2h25[1]		44.73	267.23	1h50, 2h05, 2h20, 2h35, 2h55	30, 45, 75, 110, 220
Bus	7h25	10h45	1	29	6h15, 7h20, 8h25, 9h30, 10h35	1, 10, 15, 20, 25
Car-pooling[2]	5h41		25	45	5h, 5h45, 6h30, 7h15, 8h	15, 20, 25, 30, 40
Private car[3]	6h22		99 (41 toll/58 fuel)		5h, 6h, 6h30, 7h15, 8h	60, 80, 100, 125, 150

Note: 1 - includes an estimate of in-flight and boarding time as reported by www.goeuro.com; 2 - www.blablacar.it; 3 - www.viamichelin.com

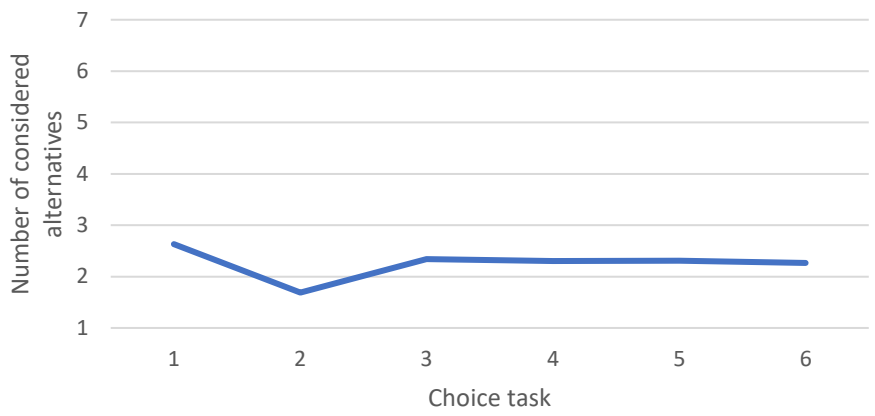
At the end of each choice task we asked respondents to state which non-chosen alternatives they had considered. The following question format was used:

"*Which other alternatives did you consider? (Please select all the other considered alternatives)*"[14]

In Figure 2.2, we show a plot of the average number of considered alternatives (including the chosen one) across choice tasks. Overall, the average number of considered alternatives is 2.26; this number is slightly larger for the first choice task (2.63), and lower for the second choice task (1.69). For the sake of completeness, in around half of the choices (49%), respondents stated to only have considered 2 alternatives. Respondents stated to have considered just one and three alternatives in respectively 22% and 19% of the choices. Only in 2% of the choices respondents stated to have considered all seven alternatives.

[14] These follow-up questions are in line with those on 'availability' or 'acceptability' of the alternatives used in previous studies as a proxy for 'consideration' (Ben-Akiva and Boccara, 1995; Hensher and Rose, 2012; Hensher and Ho, 2015). However, we preferred to ask for 'consideration' since this is more closely related to our objective.

Figure 2.2 The average number of considered alternatives



The fact that a large share of respondents reported to consider only two alternatives might suggest that they felt compelled to say they considered one more alternative with respect to the chosen one to appear cooperative and engaged with the SC experiment. If this is the case, self-reported information on consideration would be associated with over report. However, it is also possible that they reported fewer alternatives than those they actually considered, where this might be due to the fact that the definition of consideration was left vague.[15] For example, it could be possible that some respondents did not state to consider alternatives for which they simply had lower preference. In that case, self-reported information would be associated with under report. Given that we are not able to confirm the

[15] Similar misunderstanding could be observed if one collects information on 'availability', as some respondents might provide answers based on 'objective availability', where others might base their answers based on 'perceived availability', which would be more closely related to consideration.

presence (and the direction) of the measurement error in the number of considered alternatives, this information should be used with caution.

Prior to collecting socioeconomic information, but after presenting the choice tasks, we asked respondents to provide their self-imposed thresholds for total travel time and cost.[16] It could be argued that – particularly for travel time – individuals might have mode-specific rather than generic thresholds. However, in this particular choice context, alternatives could be grouped into two separate groups with comparable travel times within each group, i.e. 'fast' and 'slow' alternatives. Therefore, we expected that individuals would either state to have their travel time thresholds larger than travel times for all alternatives (i.e. to consider alternatives in both groups) or a threshold larger than the travel times for the 'fast' alternatives and lower than that for the 'slow' ones, i.e. that only the former were considered. Although questions on the presence of thresholds are less prone to misunderstanding and over/under reporting issues than those on consideration of alternatives[17], reported thresholds for travel time and travel cost were actually 'respected' in 85% and 91% of choices, respectively. Despite being an indication of the good level of reliability of this information,

[16] Moser and Raffaelli (2014) argue such thresholds should be based on previous experience and not on the information contained in the experiment. This suggests collecting thresholds right at the beginning of the experiment. However, we believe that prior elicitation can similarly condition answers to the choice tasks. Given that there is evidence that the positioning of threshold elicitation questions has no significant influence on parameter estimates (Bush, 2008), we decided to collect this information after the SC tasks.

[17] This would ideally suggest that thresholds for attributes might be better candidates than stated consideration to indirectly infer the size and the composition of consideration sets.

the presence of some 'violations' also recommend attention in using stated thresholds as *error-free* explanatory variables.

With both indicators of consideration, i.e. with *stated consideration* of alternatives and *stated thresholds* for attributes, the presence of possible measurement errors would suggest that there might not be a one-to-one correspondence between stated and actual behaviour. Moreover, there is scope for endogeneity bias as these measures may be correlated with other unobserved factors (Hess and Hensher, 2013); at the same time, the indicators would not be suitable (and/or available) for forecasting if used in a deterministic way (Bergantino et al., 2019). All these reasons motivate the use of the indicators as dependent rather than independent variables, as we explain in the next Section. The proposed approach does not make these measures '*error-free*', but it simply acknowledges the possibility that there might be an error associated with them, and reduces its impact.

4. Methodology

Mode choice is modelled using RUM (McFadden, 1974), where the utility of alternative *i* for individual *n* in choice task *t* is given by (Equation 2.1):

$$U_{i,n,t} = V_{i,n,t} + \varepsilon_{i,n,t} \tag{2.1}$$

$V_{i,n,t}$ is a function of an alternative specific constant, of the attributes of the alternative (e.g. travel time, travel cost, Wi-Fi availability, and ticket flexibility), of individual characteristics in relation to the alternative (e.g.

access/egress time to/from rail and bus stations and airports), and of individual socioeconomic and context-specific characteristics, while $\varepsilon_{i,n,t}$ is the random component. We define the probability of choosing alternative *i* from the *J* available alternatives (i.e. as presented in the SC experiment) comprised in choice set $C_{n,t}$ by (Equation 2.2):

$$P_{i,n,t} = P\left(U_{i,n,t} \geq U_{j,n,t}, \forall\, j \neq i \in C_{n,t}\right) \tag{2.2}$$

For an alternative to be chosen, alternative *i* should provide the highest overall utility over all available alternatives in the choice set. Assuming that the error terms are *i.i.d.* type I extreme value distributed, this probability can be represented by the multinomial logit model (MNL, Equation 2.3):

$$P_{i,n,t} = \frac{exp(V_{i,n,t})}{\sum_{j \in C_{n,t}} exp(V_{j,n,t})} \tag{2.3}$$

Besides accounting for unavailable alternatives from the universal set (e.g. due to not owning a car), we allow individuals to consider only a subset of the available alternatives. Hence, choices are made over $C^*_{n,t} \subseteq C_{n,t}$. As a result, the choice probability for considered alternatives increases relative to the MNL model in Equation 2.3, given that the number of alternatives included in the denominator decreases.

Since the actual consideration set is unobserved, we define a two-stage probabilistic model where the unconditional choice probability is obtained as a weighted average of conditional choice probabilities across all possible consideration sets. The conditional choice probabilities $\tilde{P}_{i,n,t}(C^*_{n,t})$ vary across consideration sets due to the changing denominator in (2.3). The probability of using a particular consideration set $\pi_{n,t}(C^*_{n,t})$ is used as a weight in the averaging process to obtain the unconditional choice probability (see Equation 2.4):

$$P_{i,n,t} = \sum_{C^*_{n,t} \subseteq C_{n,t}} \pi_{n,t}\left(C^*_{n,t}\right) \cdot \tilde{P}_{i,n,t}\left(C^*_{n,t}\right) \tag{2.4}$$

This two-stage formulation is identical to the expression proposed by Manski (1977). In Equation (2.5), we follow Swait and Ben-Akiva (1987a)'s *independent availability* model and assume that the consideration probability for alternative *j*, $W_{j,n,t}$, is independent across alternatives. This results in the following definition of $\pi_{n,t}(C^*_{n,t})$:

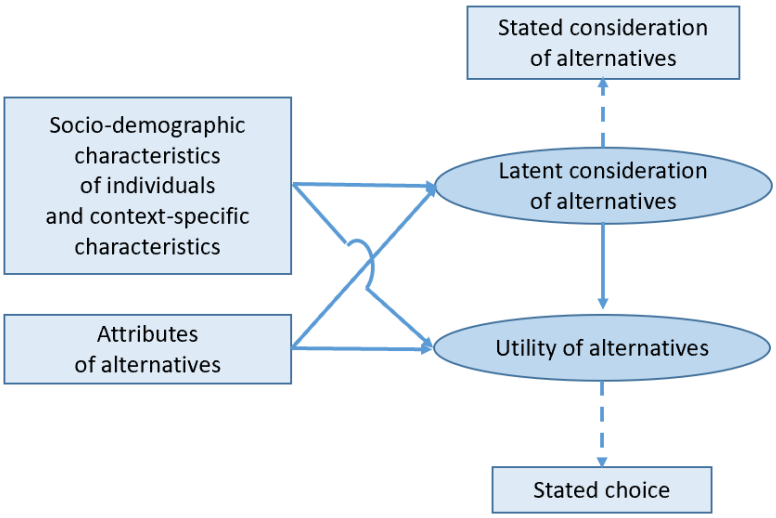
$$\pi_{n,t}\left(C^*_{n,t}\right) = \prod_{j \in C^*_{n,t}} W_{j,n,t} \cdot \prod_{j \notin C^*_{n,t}} \left(1 - W_{j,n,t}\right) \tag{2.5}$$

In this paper, we present two alternative specifications for $\pi_{n,t}(C^*_{n,t})$ Given that indicators of a different nature are used, the two model specifications

therefore differ in the way $W_{j,n,t}$ is defined. In both cases we assume that probabilistic consideration only applies to a subset of alternatives, rather than to all the available alternatives.

The first specification (presented in Figure 2.3) makes use of the responses to the stated consideration questions.

Figure 2.3 The model in which stated consideration is used as indicator for consideration



Note: Items in rectangles can be directly observed by the analyst while Items in the ellipses are unobserved. The broken arrows indicate measurement components, while plain arrows indicate structural components.

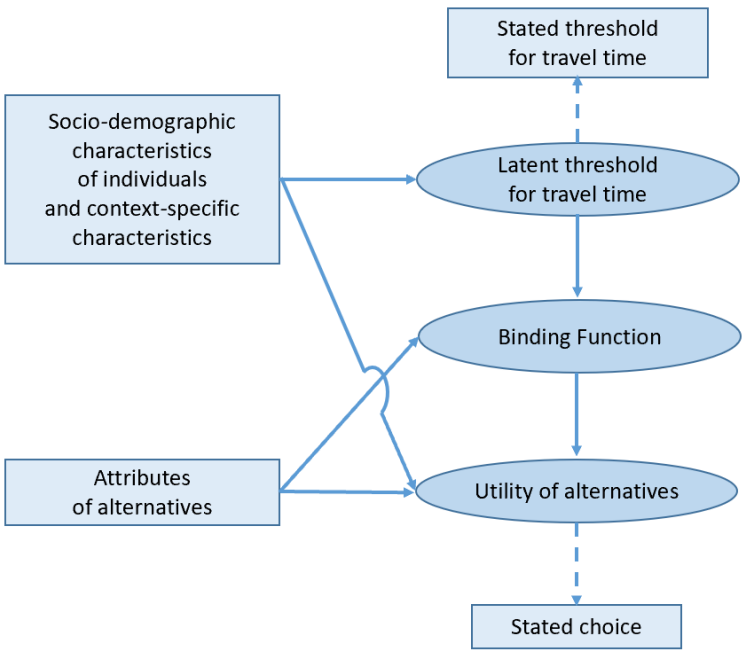
The predicted consideration probability (i.e. latent consideration of alternatives in Figure 2.3), $\widehat{W_{j,n,t}}$, is obtained by making use of the parameters $\hat{\alpha}_j$ and $\hat{\beta}_j$ which are the outcome of a series of alternative-specific binary logit models on the stated consideration data (Equation 2.6).

$$\widehat{W_{j,n,t}} = \frac{1}{1 + \exp\left(\hat{\alpha}_j + \hat{\beta}_j Z_{j,n,t}\right)}, \tag{2.6}$$

$\hat{\alpha}_j$ is an alternative specific consideration constant and $Z_{j,n,t}$ is a vector of attributes of alternatives j, and individual socio-economic and context characteristics, with their impact measured by $\hat{\beta}_j$.

The second specification makes use of self-reported threshold levels for travel time (Figure 2.4).

Figure 2.4 The model in which stated threshold is used as indicator for consideration



Note: Items in rectangles can be directly observed by the analyst while Items in the ellipses are unobserved. The broken arrows indicate measurement components, while plain arrows indicate structural components.

In this second model, we first estimate a linear regression explaining the stated threshold for travel time as a function of individual socio-economic and context-specific characteristics X_n.

In Equation 2.7, $\hat{\gamma}$ represents the estimated constant, and $\hat{\delta}$ the regression coefficients, such that $\widehat{\max\_TT_n}$ becomes the predicted threshold level for travel time for individual n (i.e. latent threshold for travel time in Figure 2.4).

$$\widehat{\max\_TT_n} = \hat{\gamma} + \hat{\delta} X_n \tag{2.7}$$

Then, we specify a binding function in Equation 2.8, which contrasts the predicted threshold for travel time against the presented travel time on mode *j*. It is expected that when the travel time $TT_{j,n,t}$ exceeds the predicted threshold level, the consideration probability reduces. Hence, we expect φ to have a positive sign. θ_j accounts instead for the general level of consideration. Both parameters, i.e. φ and θ_j, are estimated as an integral part of the choice model.

$$W_{j,n,t} = \frac{1}{1 + \exp\left(\theta_j + \varphi\,(TT_{j,n,t} - \widehat{\max\_TT_n})\right)} \tag{2.8}$$

The use of a predicted consideration probabilities $\widehat{W_{j,n,t}}$ (Equation 2.6) and predicted thresholds for the travel time attribute $\widehat{\max\_TT_n}$ (Equations 2.7-2.8) overcomes measurement and endogeneity bias issues arising when

treating these measures as *error-free* indicators for consideration. The proposed approach is similar to the latent variable approach used by Ben-Akiva and Boccara (1995) - or by Hess and Hensher (2013) who model similar indicators on attribute non-attendance - in the sense that we treat supplementary information on consideration as dependent rather than independent variables.

Finally, in the choice model we also account for the presence of unobserved preference heterogeneity. We estimate a mixed logit model (MMNL) with random alternative specific constants. The resulting MMNL models are estimated using maximum simulated likelihood and 500 Halton draws[18]. Accounting for both the role of consideration and unobserved preference heterogeneity introduces additional flexibility in the model specification, but is deemed necessary to avoid putting too much emphasis on the role of consideration. In estimation, the likelihood function accounts for the panel nature of the data. Moreover, the calculation of robust standard errors post-estimation occurs at the level of individuals rather than choice tasks. This ensures that the standard errors are slightly larger by recognising that the observations are not fully independent.

[18] This number of draws resulted in stable models, i.e. by increasing the number of draws we did not observe any improvement in the final LL.

5. Results and discussion

In the models presented in this Section, we assume that respondents always consider the faster but more expensive alternatives, i.e. HSR, FSC, and LCC. For the slower but cheaper modes, i.e. IC, bus, and car-pooling, consideration is modelled probabilistically. These assumptions are supported by preliminary modelling (summarised in Table 2.3), self-reported consideration data as well as choice data. We separately tested the effect of assuming probabilistic consideration of the faster and more expensive alternatives (and the presence of a cost threshold) and of the slower and cheaper alternatives (and the presence of a time threshold). Results from preliminary modelling showed that the former assumptions were less robust than the latter, and that the benefits arising from the inclusion of both were marginal.

Table 2.3 Summary of preliminary modelling

Model tested	Considerations	Action
A. Separately account for probabilistic consideration of slower/cheaper alternatives (IC, bus, car-pooling) and of faster/more expensive alternatives (HSR, FSC, LCC), using stated consideration data. B. Separately account for consideration of slower/cheaper alternatives and of faster/more expensive alternatives, using stated thresholds on travel time and cost, respectively.	The benefits of accounting for consideration for faster alternatives were negligible compared with those of accounting for consideration of slower alternatives.	Proceed with probabilistic consideration for slower/cheaper alternatives only, using stated consideration data and stated thresholds for travel time.
C. Extend A and B to account for probabilistic consideration of all alternatives (but private car).	Increase in estimation time up to 3 (2) times compared to A (B); there were identification issues in the estimation of parameters of the choice model.	Proceed with probabilistic consideration for slower and cheaper alternatives only.
D. Estimate mode-specific parameters for the binding function in model B.	The parameters did not statistically differ from each other.	Use a single parameter for the binding function in model B.
E. Simultaneous estimation of consideration of alternatives and choice.	Improvements in efficiency at the cost of an increase in estimation time compared to sequential estimation. However, there were identification issues in estimation of parameters of both the consideration and choice models.	Proceed with sequential estimation of consideration and choice.

5.1 Stated consideration of alternatives and thresholds for the travel time attribute

On average, the self-reported level of consideration for faster modes is indeed higher than that for slower ones (HSR: 74%; LCC: 37%; FSC: 31%; bus: 25%; IC: 24%; car-pooling: 21%; private car: 14%). Moreover, a large share of respondents (94%) chose at least once (out of 6 choice tasks) one of the faster alternatives. We further assume that the private car alternative is always considered when stated to be available. [19] As a result of modelling consideration probabilistically on only three alternatives, the number of possible consideration sets is reduced to eight.

Table 2.4 presents the results of the three binary logit models explaining stated consideration for the three modes associated with consideration effects. The longer the travel time on a mode, the less likely it is to be considered. Indeed, travel time is found to be a less important driver of consideration for bus (i.e. the slowest mode) compared to IC and car-pooling. Similar effects are found in relation to travel cost. Bus is the cheapest alternative, which might explain why travel cost was found to be insignificant in explaining its stated consideration. Stated consideration for IC increases when Wi-Fi is available on-board; and providing ticket flexibility increases the probability of considering bus. The probability of considering

[19] The information on *stated consideration* for private car was contradictory in several circumstances, i.e. respondents for which car was unavailable stated to consider this alternative during the SC experiment. We tested the implications of this assumption by making the car unavailable for everyone. Results corresponded with those obtained on the full sample, suggesting the very marginal role of this alternative in the choice model, and, therefore, of any assumption related to its consideration.

car-pooling is higher amongst higher educated travellers, but is lower for females. The former result can be explained by the fact that car-pooling has a high ICT component, where seats can only be booked online. The latter result is most likely due to a lower perception of safety. Finally, the probability of consideration for all three slow modes decreases with age, and if the trip is paid by the employer or family members and (or) friends.

Table 2.4 Logistic regressions of stated consideration - Results

	Inter-City Train		Bus		Car-Pooling	
Regressors	*est.*	*t-stat(0)*	*est.*	*t-stat(0)*	*est.*	*t-stat(0)*
Constant	3.178	2.20	4.162	3.17	0.764	0.52
Travel time	-0.008	-2.88	-0.004	-2.77	-0.008	-3.25
Travel cost	-0.040	-5.02			-0.048	-2.18
Wi-fi free (vs not available)	1.442	4.25				
Wi-fi 5 € (vs not available)	0.985	2.45				
Flexible ticket free (vs 50 €)			1.081	2.87		
Flexible ticket 5 € (vs 50 €)			0.980	2.80		
Female	0.881	2.99			-1.487	-4.33
Age (18-24) - base	-	-	-	-	-	-
Age (25-34)			-1.523	-4.00	-1.604	-3.76
Age (35-49)	-1.877	-4.84	-3.895	-7.47	-3.902	-7.05
Age (50+)	-1.562	-2.85	-3.216	-4.72	-5.480	-5.43
Education (years)	-0.112	-1.87	-0.145	-2.20	0.266	3.59
Paid employer (vs paid her/himself)	-0.954	-2.47	-3.749	-6.83	-2.855	-5.61
Paid family/friends (vs paid her/himself)	-0.635	-1.56	-1.390	-3.41	-0.886	-2.04
Predicted consideration (mean)	0.24		0.25		0.21	
Predicted consideration (min)	0.03		0.01		0.00	
Predicted consideration (max)	0.64		0.71		0.77	
Log-Likelihood (null)	-690.21		-703.79		-671.07	
Log-Likelihood (final)	-623.68		-576.41		-544.63	

Note: for all models: observations = 1254, respondents = 209

In Table 2.5, we present the results of a linear regression (OLS) on the logarithm of the stated thresholds for travel time. Results show that, ceteris paribus, male respondents have a higher self-imposed threshold for travel time relative to female respondents. The time threshold decreases with age, and it is lower for those educated at higher level (university), travelling alone for business purposes, whom the trip was for paid by the employer. As expected, the self-imposed threshold is instead higher for respondents travelling with friends on a non-business trip.

Table 2.5 Regressions of stated thresholds for travel time – Results

Regressors	*est.*	*t-stat(0)*
Constant	6.038	189.00
Male	0.102	3.97
Age 25-34	-0.104	-2.90
Age 35-50	-0.210	-5.36
Age 50+	-0.368	-7.48
Higher-education	-0.095	-3.45
Paid employer	-0.118	-3.44
*Travel with friends*non-business trip*	0.090	1.90
*Travel alone*business trip*	-0.235	-7.36
Predicted thresholds (mean)	328	
Predicted thresholds (min)	185	
Predicted thresholds (max)	508	
Adjusted R-squared	0.25	

Note: respondents = 209.

5.2 Consideration of alternatives and choice

We present the results for three choice models in Table 2.6. Model 1 represents a mixed logit model (MMNL) with normally distributed alternative specific constants (ASC). This model does not account for the role of consideration in mode choice, i.e. it assumes that all alternatives are fully considered. Models 2 and 3 probabilistically account for consideration of the slower alternatives (based on Equations 2.6 and 2.8, respectively).[20] The latter two models are compared against Model 1 in terms of parameter estimates and goodness of fit. In addition, we explore the implications of accounting for consideration effects on willingness-to-pay indicators and forecasted market shares.

For Model 1, car-pooling was found to be the minimum variance alternative, and therefore used as baseline alternative to prevent over-identification of the model (Walker et al., 2007). The ASCs reveal a strong preference for FSC over car-pooling, while the opposite occurs for private car which was chosen in only very few occasions (21 out of 1254 choices).

[20] Models 2 and 3 estimated using the indicators for consideration, i.e. stated consideration and stated thresholds, are not presented here but available upon request to the Authors.

Table 2.6 Estimated models - Results

	Model 1		Model 2		Model 3	
Regressors	*est*	*rob t-rat(0)*	*est*	*rob t-rat(0)*	*est*	*rob t-rat(0)*
ASC choice HSR	1.218	2.55	0.251	0.36	1.557	2.44
ASC choice IC	0.655	1.93	0.714	1.51	-0.254	-0.46
ASC choice FSC	3.069	2.91	2.224	1.77	4.768	3.60
ASC choice LCC	1.479	1.41	0.215	0.16	2.858	2.22
ASC choice Bus	0.547	1.58	0.067	0.11	-1.170	-1.73
ASC choice Private Car	-2.951	-1.05	-5.984	-1.94	-2.767	-0.96
Wi-fi free (HRS, IC, FSC, LCC, Bus)	0.246	1.64	0.276	1.43	0.268	1.64
Wi-fi €5 (HRS, IC, FSC, LCC, Bus)	0.107	0.82	0.130	0.78	0.090	0.63
Flexible ticket (up to €5)	0.354	3.24	0.459	3.05	0.385	3.06
Travel time Air	-0.012	-3.17	-0.012	-2.87	-0.014	-3.51
Travel time Train/Bus/Car-pooling	-0.008	-6.42	-0.008	-3.72	-0.004	-1.70
Travel time Private Car	0.003	0.52	0.005	0.8	0.004	0.61
Travel cost	-0.050	-9.45	-0.054	-7.15	-0.055	-7.38
Travel cost, income na	-0.040	-5.92	-0.043	-4.91	-0.044	-5.25
Paid employer or family (travel cost)	0.024	4.69	0.026	3.68	0.027	4.14
Lambda income (elasticity effect on travel cost)	-0.249	-5.85	-0.342	-6.25	-0.294	-5.58
Access/egress time main airports	-0.037	-3.98	-0.036	-4.57	-0.038	-4.30
Access/egress time secondary airports	-0.017	-2.45	-0.017	-2.04	-0.016	-2.47
Fidelity card (FSC)	2.002	3.98	1.828	4.11	1.949	4.23
Female (FSC/LCC)	0.810	2.40	0.828	2.16	0.698	2.05

Age 25+ (HSR)	1.063	2.25	0.298	0.51	0.670	1.27
Age 25+ (FSC/LCC)	1.876	3.76	1.005	1.73	1.407	2.64
Business (HSR)	1.146	3.12	0.922	2.71	0.932	2.69
Higher-education (all but HSR)	-0.830	-3.02	-0.730	-2.07	-1.025	-3.16
Sigma parameters (random coefficients)						
ASC choice HSR (sigma)	1.703	5.95	1.575	5.46	1.616	6.03
ASC choice IC (sigma)	1.034	2.83	-1.201	-1.16	1.661	4.43
ASC choice FSC (sigma)	1.456	3.29	0.582	0.46	-1.275	-3.59
ASC choice LCC (sigma)	-1.420	-4.34	-1.497	-3.92	-1.404	-5.34
ASC choice Bus (sigma)	1.450	4.05	2.298	2.05	2.063	4.86
ASC choice Private Car (sigma)	-2.726	-4.77	-3.155	-4.41	-3.165	-4.54
Consideration component						
Binding function parameter (φ)					0.016	5.44
ASC consideration IC					-1.859	-2.41
ASC consideration Bus					-4.447	-4.46
ASC consideration Car-pooling					-0.231	-0.65
LL(0)	-2319.01		-2083.03		-2118.09	
LL(final)	-1241.73		-1261.36		-1222.67	
AIC	2543.46		2582.72		2513.33	
BIC	2697.48		2736.74		2687.90	
Prob. chosen alternative (100 holdout samples)	41.16%		40.71%		41.48%	
Confidence interval (95%)	40.60%	41.70%	40.20%	41.20%	40.90%	42.00%

Note: for all models: observations = 1254, respondents = 209.

We estimated three travel time coefficients: one for the air alternatives (FSC and LCC), one grouping the train alternatives (HSR and IC), bus, and car-pooling, and one for private car. The first two have the expected (negative) sign and are also statistically significant. The result for the private car travel time coefficient, i.e. that the parameter is not significantly different from zero, can be explained by the fact that this alternative was chosen in very few occasions and our feeling is that those respondents would have chosen to travel by car anyway, regardless of its characteristics and those of the other alternatives.

Travel cost has been interacted with income in a non-linear way (see Appendix A). The negative value for the estimated elasticity ('*Lambda Income'*) implies that the (absolute) sensitivity to travel cost decreases with increases in income. Similarly, accounting for travellers who did not pay for the trip themselves ('*Paid employer or family'*) reveals that these respondents also place a lower importance on the cost attribute. Results also show that respondents are more likely to select a particular mode when they can get a flexible ticket at a reasonable price (i.e. free or up to 5€) instead of having to pay a larger fee of 50€ for this option. The latter value is more in line with current airlines' fees. The presence of Wi-Fi seems, surprisingly, to hardly affect mode choice. We have two possible explanations. First, Wi-Fi connections are currently available only on-board HSR and busses. In the SC experiment, it was also assumed available on-board IC and flights, which will be realistic in the near future. Second,

travellers currently experience low levels of connectivity on this corridor due to the large amount of tunnels.

Coefficients for access/egress time are, as expected, negative and significant for airports. The airports in Rome and Milan are located quite far from the city centres. For train stations and bus terminals, access and egress time were not found to be significant due to being located in more central areas.

Finally, we discuss the influence of socioeconomic and context-specific characteristics on mode choice, and reflect on the degree of random heterogeneity associated with the ASCs. With respect to the former, *ceteris paribus*, car-pooling gains appeal over other modes amongst more educated travellers (university level) and younger travellers. Female and business travellers prefer the air and the HSR alternatives, respectively. Standard deviations (‘*sigma parameters’*) are highly significant.

In Model 2, we account for consideration effects using information on *stated consideration*. We do not estimate any additional parameters relative to Model 1 given that predicted consideration probabilities, derived from Table 2.4, are directly implemented in the two-stage model (see Equation 2.6).

In Model 3, we account for consideration effects using information on *stated thresholds*. Consideration probabilities are now calculated within the choice model (Equations 2.7-2.8). The predicted thresholds from Table 2.5 are included in the binding function and four additional parameters are estimated translating the binding function into consideration probabilities (three alternative specific constants for consideration and one parameter φ

for the binding function[21]). The positive parameter for the non-linear binding function reveals that consideration for the IC, bus, and car-pooling can indeed be explained by the difference between the thresholds for travel time and the actual values for this attribute.

In Model 3, we observe that, whilst the travel time coefficient for the air alternatives increases compared to Model 1, that for the other alternatives is reduced. We offer two possible explanations. First, in Model 3, consideration effects do not act in isolation (whereas in Model 2 these were exogenously introduced), but are integrated within the estimation of the choice model. Second, the implicit consideration probabilities are, on average, larger in Model 3 compared to Model 2 (for IC: 66% vs 23.5%; Bus: 71.2% vs 24.9%; car-pooling: 31% vs 21.3%), thereby reducing the strength of consideration effects on choice.

Finally, we observe that the variance of the utility for the part related to the random ASCs (Table 2.7) is reduced for the fully considered (i.e. *fast*) alternatives in Models 2 and 3. It is possible that consideration models reflect the circumstance that respondents showing a stronger preference towards faster alternatives actually process less information, as their consideration sets are smaller. This, in turn, might imply that elements previously attributed to random heterogeneity can possibly be ascribed to consideration effects. Given that consideration sets are defined at the choice

[21] Mode-specific parameters for the binding function were tested but these did not statistically differ from each other.

task level, while random heterogeneity is added at the individual level, this would be the case particularly when consideration of alternatives is not context dependent, e.g. when consideration is rather dictated by some *a priori* beliefs towards the alternatives. In this case, respondents make decisions about consideration based on the perceived (rather than actual) levels for travel time for these alternatives, i.e. based on their general knowledge of the market. This would be also possible when thresholds for travel time are low enough such that IC, bus, and car-pooling would never be considered.

Table 2.7 Analysis of the variance related to the ASCs

	Model 1	**Model 2**	**Model 3**
HSR	3.42	2.71	2.91
IC	1.21	1.55	2.98
FSC	4.66	2.26	3.96
LCC	3.15	2.92	2.85
Bus	2.25	5.39	4.48
Private car	7.57	10.07	10.24

We now turn our attention to the goodness of fit for the three models. Given that these models are non-nested, the Likelihood ratio tests are not suitable. Similarly, a comparison over the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) would be flawed, because these indicators are based on the final Log-Likelihood. Therefore, model

performance is evaluated using the average probability for the chosen alternative on 100 alternative holdout samples.[22] This measure, reported alongside traditional measures of fit in Table 2.6, indicates that Model 3 is the best-performing model showing a moderate improvement over Model 1 (41.48% vs 41.16%). Previous papers accounting for probabilistic consideration using the two-stage approach have obtained larger improvements in fit (e.g. Swait, 2001; Basar and Bhat, 2004). In contrast to the referred papers, we do not consider supplementary information on consideration as *error-free* measures, and we account for unobserved preference heterogeneity in our choice model. With respect to the use of multinomial logit models inside a two-stage model (or, in general, any model accounting for consideration effects), such a model is in effect a *latent class* model and could thereby erroneously ascribe preference heterogeneity to consideration effects.

We contrast the three models based on marginal *willingness-to-pay* measures and forecasts for the aggregate market shares. With respect to the

[22] The database used in this paper is rather small. For this reason, we randomly split individuals in the sample and their observations in five disjoint subsets, stratified on the base of the mode respondents were travelling with at the time of the survey. Then, in turn, four out of five subsets were used as the training set to estimate the models and we used the other subset as the test set. Therefore, we compared models' forecasting performance on 100 training/test sets (the procedure described has been repeated 20 times, providing 5 different combinations of training/test sets each time), as to make sure that results were robust enough to draw any conclusions from them.

former, we present the value of travel time (VTT, Table 2.8)[23] for an individual who pays her/himself for the trip.

[23] Standard errors are calculated using the delta method for the ratio between travel time and travel cost coefficients.

Table 2.8 VTT (€/hour)

	Model 1		**Model 2**			**Model 3**		
	est.	*t-stat(0)*	*est.*	*t-stat(0)*	*t-stat(Model1-2)*	*est.*	*t-stat(0)*	*t-stat(Model1-3)*
Air	14.874	2.97	13.014	2.70	0.26	15.138	3.26	0.04
Train/Bus/Car-pooling	9.714	5.81	8.550	3.52	0.39	3.822	1.67	-2.08

Table 2.8 reveals differences between Models 2 and 3 on the one hand and Model 1 on the other hand. In Model 2, we particularly observe a reduction in the VTT for the air alternatives. This result is consistent with our expectations: when slower alternatives are hardly (or not) considered, comparisons amongst fastest alternatives, which are therefore more similar in terms of travel time, should result in lower *willingness-to-pay* measures. VTT for the other alternatives also decreases as a result of accounting for consideration effects. In Model 3, instead, we observe that accounting for consideration effects slightly increases the VTT measure for air alternatives and strongly reduces the VTT for the remaining alternatives. The VTT for train, bus, and car-pooling is reduced by 61% compared to Model 1, and this difference is also statistically different from zero. This is due to the fact that measuring consideration using thresholds for the travel time attribute takes away explanatory power from this particular attribute in the utility function. Forecasted aggregate market shares (Figures 2.5-2.7) are also affected by the assumptions we make in the three models about consideration. In general, we observe larger differences in forecasts between Model 2 relative to Models 1, than between the latter and Model 3. This result can be attributed to the average probability of consideration for slower alternatives in Model 3 being higher than in Model 2. In a *status quo* scenario (i.e. applying the model to the choice tasks presented to the respondents, Figure 2.5), Model 2 predicts slightly larger market shares for the fully considered alternatives compared to Model 1 (e.g. for HSR: 51.4% vs 50.2%), and, *vice*

versa, lower market shares for partially considered ones (e.g. for IC: 7.8% vs 8.7%). This is in line with our expectations. When subsequently looking at the effect of a reduction in travel time by 20% for the HSR alternative, Figure 2.6 displays again larger differences in prediction between Model 2 and Model 1 and more comparable predictions between Model 3 and Model 1. If we reduce travel time for the bus by 30%, the difference between Model 1 and the two consideration models (Models 2 and 3) becomes more substantial (Figure 2.7). Model 1 predicts a larger increase over the *status quo* (+88%) and larger market shares for this mode (16.7%) than Model 2 (+59% and 13.4%, respectively) and Model 3 (+55% and 13.9%, respectively) at the expense (mainly) of the HSR alternative.

Overall, this forecasting exercise shows that differences in the average predicted market shares between the traditional mixed logit model and models accounting for consideration effects appear negligible when contrasted against those reported in the previous literature. However, this can be attributable to the fact that we decided to test for more realistic scenarios rather than for more extreme and arguably less realistic ones (e.g. Ben-Akiva and Boccara, 1995, tested a 100% change in travel time).

Figure 2.5 Predicted aggregate market shares (status quo)

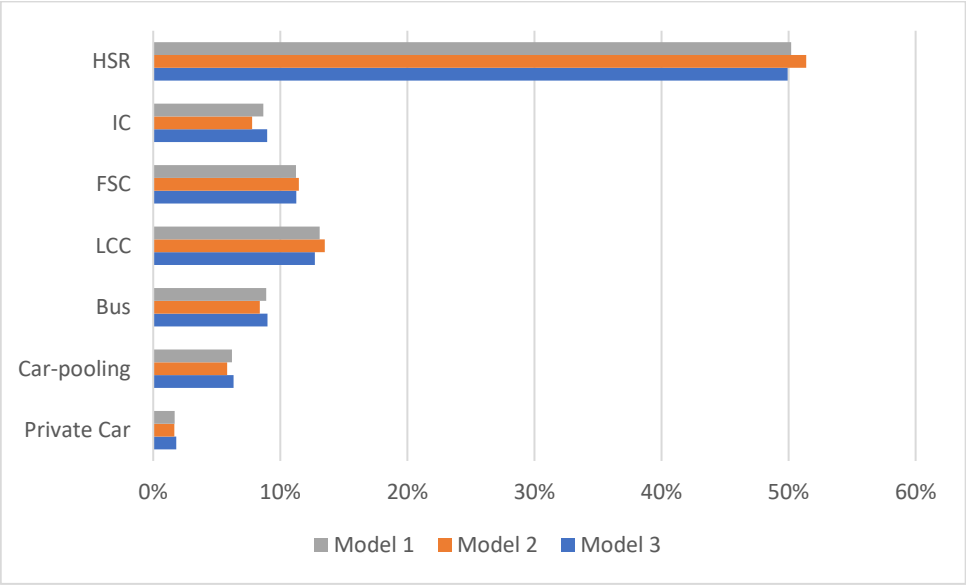


Figure 2.6 Predicted aggregate market shares when travel time for HSR is reduced by 20%

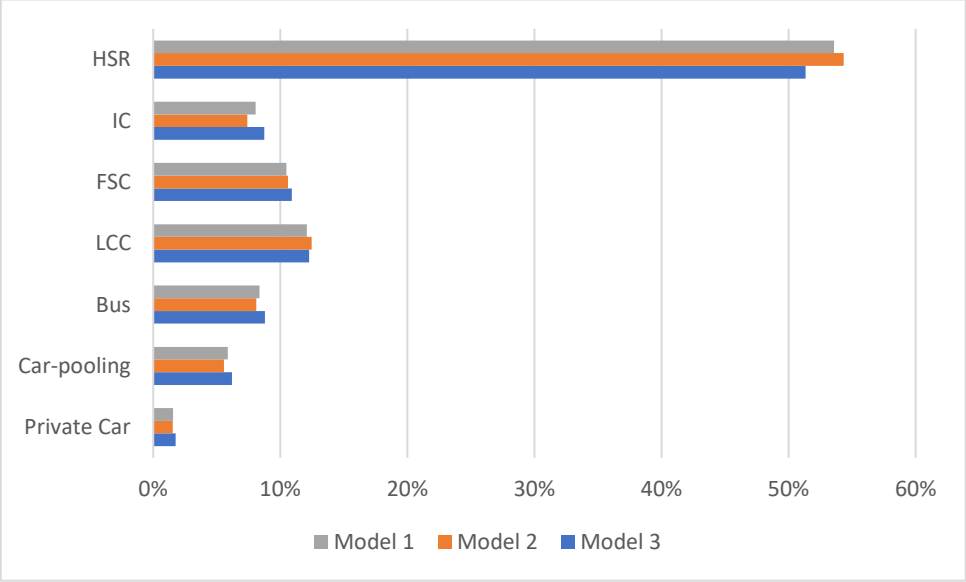
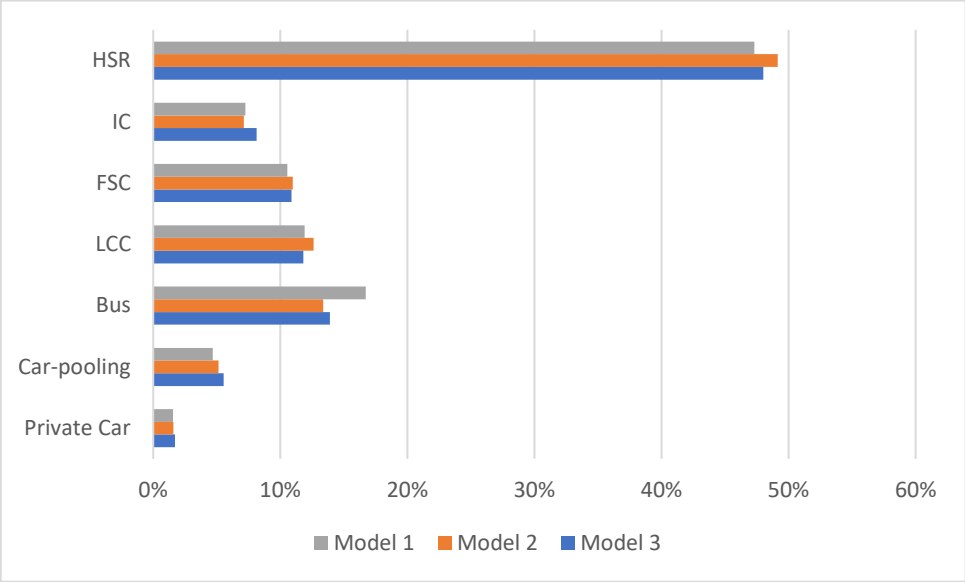


Figure 2.7 Predicted aggregate market shares when travel time for bus is reduced by 30%



6. Conclusions

Within this paper, we have contributed to the ongoing discussion on the role of consideration of the alternatives in the individuals' decision-making process. Consideration of the alternatives cannot be directly observed and therefore measured with certainty, which leads to an empirical identification problem. When the only information available is that on choice, it is impossible to separately identify which factors drive consideration and choice (or both).

It has been argued that consideration and choice cannot (and should not) be separately identified because they represent a unique process. Under that

assumption, estimating a single stage utility function would be sufficient. This would implicitly assume that the majority, if not all, of the choices can be described by a fully compensatory behavioural process where individuals make trade-offs between attributes and across alternatives. However, the presence of many choice heuristics tells us that this is not the case. By not including all alternatives in the choice set, which implies that individuals actually choose from restricted consideration sets, we make the more reasonable assumption that the choice process is non-compensatory to a certain degree.

In this study we propose an extension to the traditional two-stage approach (Manski, 1977; Swait and Ben-Akiva, 1987a), measuring consideration using supplementary information on this stage. This allows us to empirically separate the role (and the driving factors) of both consideration and choice. By assuming that all possible consideration sets have a probability of being the 'true' one, the two-stage model provides the best reflection that consideration sets are unobserved.

In particular, we study the role of consideration of the alternatives in a transport mode choice context, using data from a SC survey administered to a sample of 209 travellers on the Rome-Milan corridor. The SC experiment was designed to mimic a real purchasing occasion through an online journey planner, which implied a strong limitation that all 'objectively' available (i.e. feasible) alternatives - not those effectively available (e.g. private car) - were presented to the respondents in the experiment. The use of such

experimental data (rather than stated preference data pivoted around individual's actual choice sets and/or of revealed preference data) in combination with the small sample size and the lack of representativeness with the Italian population limits the generalisability of our results on travellers' preferences on the Rome-Milan corridor. Indeed, rather than suggesting policy measures, the aim of this paper was to propose a methodology with respect to the measurement and modelling of consideration of the alternatives.

In addition to choices, during the experiment we also collected additional information on consideration of the alternatives at the task level, and on self-imposed thresholds for the travel time attribute at the respondent level. This additional information is used to measure consideration of the alternatives within two distinct model specifications, which are in turn compared with a choice model where all alternatives are assumed to be considered.

The use of exogenous information related to consideration is not new in the literature. Similarly to Ben-Akiva and Boccara (1995) we treat these indicators as dependent rather than independent and error-free variables, and the resulting functional forms are then combined with the data to derive the consideration probabilities required in a two-stage model. Moreover, we also account for additional unobserved preference heterogeneity in the choice model to avoid the risk of putting too much emphasis on the role of consideration effects.

In the first model, a series of binary logit models are estimated on stated consideration and used to predict consideration probabilities. In the second model, consideration probabilities are instead calculated within the choice model. We use a binding function which compares the values for the travel time attribute with the predicted value for the threshold for the travel time attribute. The latter are the outcomes of a standard regression model.

Consideration probabilities differ substantially depending on which supplementary information is used to obtain them. In particular, those obtained using stated consideration are, on average, lower than those obtained using the thresholds. As a result, differences with respect to the reference model – particularly in terms of parameter estimates and forecasted market shares – are more evident (and more in line with expectations) in the first model than in the second. On the other hand, only the second model shows an improvement in fit with respect to the reference model, which is most likely due to the estimation of additional parameters relative to the consideration stage. In both models, elements conventionally attributed to unobserved preference heterogeneity could alternatively be attributed to consideration effects. To conclude, we acknowledge that collecting additional information on consideration of alternatives and thresholds for attributes might be burdensome, and not always feasible. However, it can convey additional insights into the consumer's decision-making process, including preferences. Its usage within the proposed approaches does not completely overcome the limitations common to the

other consideration models, but it moves towards a more precise identification of the two stages, i.e. consideration and choice, and of their respective drivers.

Despite our findings not being as strong as those found in previous studies – most likely due to simultaneously accounting for unobserved preference heterogeneity in the choice model – we still recommend the inclusion of consideration effects to get a more realistic representation of individuals' behaviour. Consideration of alternatives does influence willingness-to-pay measures and forecasted market shares, and can thereby influence transport planning investment decisions. However, this more likely happen when the market share of not considered alternatives is anything but marginal.

Acknowledgements

The authors wish to acknowledge three anonymous reviewers, whose suggestions at different stages on the development of the manuscript greatly improved the quality of this paper, as did final suggestions by the editor. The authors also acknowledge the Italian Railway Infrastructure Manager, Rete Ferroviaria Italiana, and the ownership of two service stations located on the A1/E35 Highway, for gathering special authorisations to administer surveys at platforms, and inside the service stations, respectively. Moreover, Stephane Hess acknowledges financial support by the European Research Council through the consolidator grant 615596-DECISIONS.

References

ART - Autorità Regolazione Trasporti (2015) Secondo rapporto annuale al parlamento (in Italian).

Ben-Akiva, M., and Boccara, B., (1995) Discrete choice models with latent choice sets. *International Journal of Research in Marketing*, 12(1), 9-24, https://doi.org/10.1016/0167-8116(95)00002-J.

Bergantino, A.S., Capurso, M., Dekker, T., and Hess, S. (2019) Allowing for heterogeneity in the consideration of airport access modes: the case of Bari airport. *Transportation Research Record*, forthcoming, https://doi.org/10.1177/0361198118825126.

Cantillo, V., and Ortúzar, J. D. D., (2005) A semi-compensatory discrete choice model with explicit attribute thresholds of perception. *Transportation Research Part B: Methodological*, 39(7), 641-657, https://doi.org/10.1016/j.trb.2004.08.002.

Cascetta, E., and Papola, A. (2001) Random utility models with implicit availability/perception of choice alternatives for the simulation of travel demand. *Transportation Research Part C: Emerging Technologies*, 9(4), 249-263, https://doi.org/10.1016/S0968-090X(00)00036-X.

Draganska, M., and Klapper, D. (2011) Choice set heterogeneity and the role of advertising: An analysis with micro and macro data. *Journal of Marketing Research*, 48(4), 653-669, https://doi.org/10.1509/jmkr.48.4.653.

Gaundry, M. J., and Dagenais, M. G. (1979) The dogit model. *Transportation Research Part B: Methodological*, 13(2), 105-111, https://doi.org/10.1016/0191-2615(79)90028-6.

Hauser, J. R. (2014) Consideration-set heuristics. *Journal of Business Research*, 67(8), 1688-1699, https://doi.org/10.1016/j.jbusres.2014.02.015.

Hensher, D. A., Rose, J. M. and Greene, W. H. (2005) The implications on willingness to pay of respondents ignoring specific attributes**.** *Transportation***,** 32(3), 203-222, https://doi.org/10.1007/s11116-004-7613-8.

Hensher, D. A., and Ho, C. (2015) The role of perceived acceptability of alternatives in identifying and assessing choice set processing strategies in stated choice settings: The case of road pricing reform. *Transportation Research Part E: Logistics and Transportation Review*, 83, 225-237, https://doi.org/10.1016/j.tre.2015.09.012.

Hensher, D. A., and Rose, J. M. (2012) The influence of alternative acceptability, attribute thresholds and choice response certainty on automobile purchase preferences. *Journal of Transport Economics and Policy,* 46(3), 451-468.

Hess, S., and Hensher, D. A. (2013) Making use of respondent reported processing information to understand attribute importance: a latent variable scaling approach. *Transportation*, 40(2), 397-412.

Horowitz, J. L., and Louviere, J. J. (1995) What is the role of

consideration sets in choice modeling? *International Journal of Research in Marketing*, 12(1), 39-54, https://doi.org/10.1016/0167-8116(95)00004-L.

Isfort (2017) Annual report of the Audimob survey on the mobility in Italy (In Italian).

Manrai, A. K., and Andrews, R. L. (1998) Two-stage discrete choice models for scanner panel data: An assessment of process and assumptions. *European Journal of Operational Research*, 111(2), 193-215, https://doi.org/10.1016/S0377-2217(98)00145-3.

Manski, C. F. (1977) The structure of random utility models. *Theory and decision*, 8(3), 229-254, https://doi.org/10.1007/BF00133443.

Martínez, F., Aguila, F., and Hurtubia, R. (2009) The constrained multinomial logit: A semi-compensatory choice model. *Transportation Research Part B: Methodological*, 43(3), 365-377, https://doi.org/10.1016/j.trb.2008.06.006.

McFadden, D. (1974) Conditional logit analysis of qualitative choice behaviour. *Frontiers in Econometrics*, edited by P. Zarembka, 105-142. Academic Press, New York.

Moser, R. and Raffaelli, R. (2014) Does attribute cut-off elicitation affect choice consistency? Contrasting hypothetical and real-money choice experiments. *Journal of choice modelling*, 11, 16-29, https://doi.org/10.1016/j.jocm.2014.02.003.

Narayana, C. L. and Markin, R. J. (1975) Consumer behavior and product performance: an alternative conceptualization. *The Journal of Marketing*, 39(4), 1-6, https://doi.org/10.2307/1250589.

Pancras, J. (2010) A framework to determine the value of consumer consideration set information for firm pricing strategies. *Computational Economics*, 2010, 35(3), 269-300, https://doi.org/10.1007/s10614-009-9193-3.

Roberts, J. H. and Lattin, J. M. (1997) Consideration: review of research and prospects for future insights. *Journal of Marketing Research*, 34(3), 406-410, https://doi.org/10.2307/3151902.

Rose, J. M., Bliemer, M. C., Hensher, D. A. and Collins, A. T. (2008) Designing efficient stated choice experiments in the presence of reference alternatives. *Transportation Research Part B: Methodological*, 42(4), 395-406, https://doi.org/10.1016/j.trb.2007.09.002.

Shocker, A. D., Ben-Akiva, M., Boccara, B. and Nedungadi, P. (1991) Consideration set influences on consumer decision-making and choice: issues, models, and suggestions. *Marketing letters*, 2(3), 181-197, https://doi.org/10.1007/BF02404071.

Swait, J. (1984) Probabilistic choice set generation in transportation demand models. PhD dissertation, Massachusetts Institute of Technology, Cambridge, USA.

Swait, J. (2001) A non-compensatory choice model incorporating attribute cutoffs. *Transportation Research Part B: Methodological*, 35(10), 903-928, https://doi.org/10.1016/S0191-2615(00)00030-8.

Swait, J. and Ben-Akiva, M. (1987a) Empirical test of a constrained choice discrete model: mode choice in Sao Paulo, Brazil. *Transportation Research Part B: Methodological*, 21(2), 103-115, https://doi.org/10.1016/0191-2615(87)90010-5.

Swait, J. and Ben-Akiva, M. (1987b) Incorporating random constraints in discrete models of choice set generation. *Transportation Research Part B: Methodological,* 21(2), 91-102, https://doi.org/10.1016/0191-2615(87)90009-9.

Trenitalia (2017) Oltre il 50% di biglietti delle Frecce venduto su canali digitali (in Italian). Retrieved from http://www.fsnews.it/fsn/Gruppo-FS-Italiane/Trenitalia/Trenitalia-biglietti-Frecce-canali-digitali.

Walker, J. L., Ben-Akiva, M. and Bolduc, D. (2007) Identification of parameters in normal error component logit-mixture (NECLM) models. *Journal of Applied Econometrics*, 22(6), 1095-1125, https://doi.org/10.1002/jae.971.

Williams, H. C. W. L. and Ortúzar, J. D. D. (1982) Behavioural theories of dispersion and the mis-specification of travel demand models. *Transportation Research Part B: Methodological*, 16(3), 167-219, https://doi.org/10.1016/0191-2615(82)90024-8.

Chapter 3 - Stated consideration and attribute thresholds in mode choice models: a hierarchical ICLV approach

Abstract

Consideration of alternatives, as many other aspects related to the decision-making process, is not observable and challenging to measure. Even when supplementary information is collected during stated choice experiments, its use as an additional explanatory variable is discouraged due to potential endogeneity issues, measurement error and limited suitability for forecasting. To overcome these limitations, we propose an Integrated Choice and Latent Variable model where consideration of an alternative is treated as a latent variable. The novelty of the presented model is that the latent variable for consideration of an alternative itself is a function of another set of latent variables that represent thresholds applied by the decision maker to individual attributes (such as travel time and cost). The proposed hierarchical relationship between *latent thresholds* and *latent consideration* enables us to explain a share of otherwise purely random heterogeneity, and identify the structural drivers of consideration. The latter is of interest to policymakers and private operators.

Keywords

Hierarchical ICLV, stated consideration, stated threshold

1. Introduction

One of the strongest assumptions underlying mode choice studies is that all available alternatives are considered. This might not be a reasonable assumption because individuals are often not aware of all alternatives and/or employ simplifying choice heuristics. Past work suggests that ignoring consideration effects can have severe implications on parameter estimates and forecasting (Williams and Ortúzar, 1982; Swait, 1984). Namely, biased parameter estimates and forecasts may lead to incorrect policy and managerial decisions (Pancras, 2010; Draganska and Klapper, 2011).

The challenge with consideration of alternatives, as part of the decision-making process, is that it cannot be observed and challenging to measure (i.e. at least not directly or without error). Previous studies mainly inferred consideration solely on the basis of the observed choice behaviour (Gaudry and Dagenais, 1979; Swait and Ben-Akiva 1987a, 1987b; Basar and Bhat, 2004), or related consideration to some observed attributes of the alternatives (Cascetta and Papola, 2001; Cantillo and Ortúzar, 2005; Martinez et al., 2009).

A handful of scholars, generally when using stated choice (SC) surveys, have collected additional information covering aspects related to consideration, such as availability (Ben-Akiva and Boccara, 1995) and acceptability (Hensher and Ho, 2015) of alternatives, or self-imposed thresholds for individual attributes (Swait, 2001). Indeed, the answers to these questions

do not give an exact or error free measure of the underlying behavioural processes. In the present paper, we use an Integrated Choice and Latent Variable (ICLV) model (McFadden, 1986; Ben-Akiva et al., 2002; Bolduc et al., 2005), which recognises this property of the data.

The ICLV approach has been extensively used in many fields (e.g. transport, health, and environment) to incorporate either psychological factors such as attitudes and perceptions (see, e.g., Soto et al., 2018; Kløjgaard and Hess, 2014; Mariel and Meyerhoff, 2016) or respondents' processing strategies (Hess and Hensher, 2013) into choice models. In this paper, we provide a novel use of the ICLV framework by incorporating consideration effects through inter-related latent variables. In particular, *latent thresholds* for attributes are used to explain *latent consideration* of the alternatives. These latent variables are in turn used to help explain mode choice behaviour. The inclusion of the latent variables in the overall framework is made possible by additional information collected during a SC survey on the decision-making process in the form of *stated thresholds* and *stated consideration*. We adopt the term 'hierarchical ICLV' model as introduced by Paulssen et al. (2014), because the *latent threshold* only affects individual choices indirectly through *latent consideration*. There is a strong behavioural mechanism supporting such a hierarchical relationship since the consideration of alternatives is likely to be driven by the presence of thresholds for individual attributes.

In the proposed approach *latent consideration* is used to reduce the utility, and therefore choice probability, of the alternatives. A similar 'discounting' approach has been proposed by Fotheringham (1988) in the context of consumer store choice, and by Cascetta and Papola (2001) and Martinez et al. (2009) in transport contexts, even though these authors related consideration to observable (as opposed to latent) characteristics of the alternatives. This discounting approach represents a convenient alternative to the traditional *two-stage* modelling of consideration and choice (Manski, 1977), given that it does not require enumerating (and modelling) of all possible consideration sets (i.e. combinations of alternatives).

Our work unveils the strong behavioural link between consideration of alternatives and thresholds for attributes, and their role in the decision making-making process. We illustrate a mechanism through which these links can be captured with the use of additional information collected during standard surveys. The remainder of the paper is structured as follows. In Section 2, we describe the available data, coming from a SC experiment on transport mode choice on the Rome-Milan corridor, in Italy. Section 3 lays out the empirical strategy and explains the proposed models. In Section 4, we report and discuss the estimation results. Finally, in Section 5, we draw conclusions from our study.

2. Data

We use data from a SC experiment that was administered in April and May 2016 to a sample of travellers on the Rome-Milan corridor (approximately

600 km). Here, seven alternatives (i.e. modes of transport) are available to travellers. The alternatives vary significantly in terms of travel time and travel cost (Table 3.1). Accordingly, it is reasonable to assume that some travellers might *a priori* disregard alternatives based on self-imposed thresholds for specific attributes.

Table 3.1 Alternatives' core characteristics at the time of the SC experiment

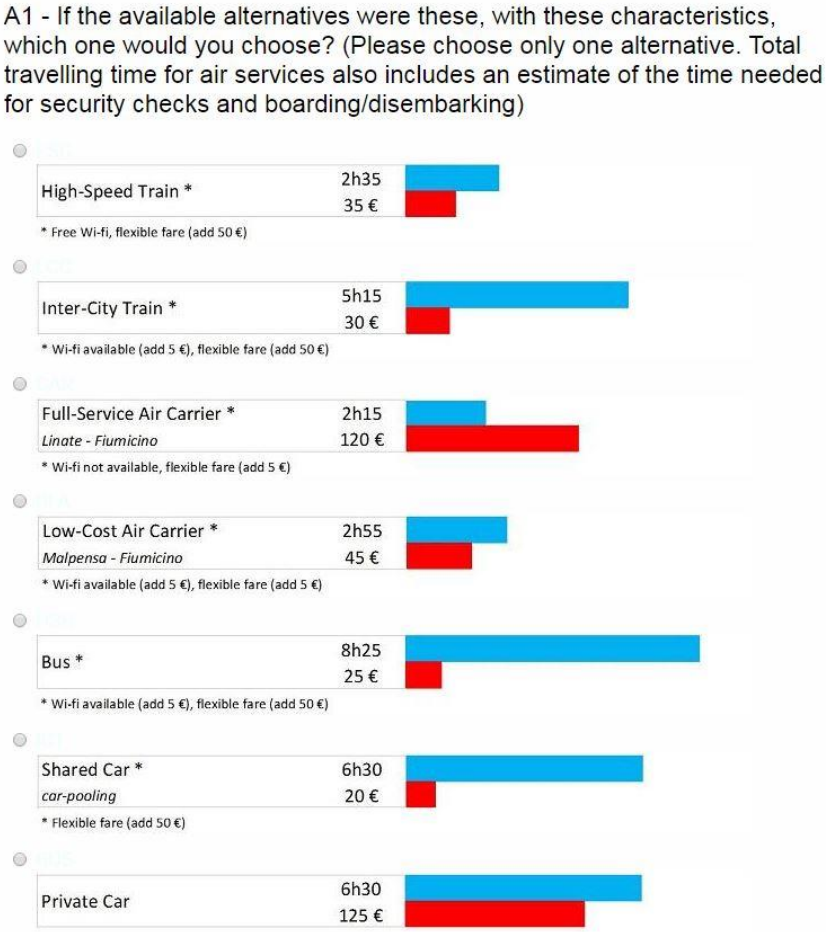
	Travel time (h/min)		Travel cost (€)	
Alternatives	*min*	*max*	*min*	*max*
High Speed Rail	2h55	4h28	19.9	209
Inter-City Rail	6h27	6h50	9	79
Full Service Air Carrier	2h20[1]		55.71	244.71
Low Cost Air Carrier	2h25[1]		44.73	267.23
Bus	7h25	10h45	1	29
Car-pooling[2]	5h41		25	45
Private car[3]	6h22		99 (41 toll/58 fuel)	

Source: Operators' websites; Note: 1 - includes an estimate of in-flight and boarding time as reported by www.goeuro.com; 2 - www.blablacar.it; 3 - www.viamichelin.com.

A total of 209 on-site face-to-face *Tablet Assisted* surveys were administered to travellers going from Rome to Milan (or *vice versa*) while waiting on the platform for their train (57%), at the bus stations for their bus (17%), or in the proximity of the airports (12%). A smaller portion of surveys was administered online (8%), and in two service stations on the A1/E35 highway, located around halfway between Rome and Milan, in the proximity of Bologna (6%). Each respondent completed 6 choice tasks, which were designed to mimic a real purchasing decision through an online journey planner. To this end a similar layout to the one displayed by the website www.goeuro.com (Figure 3.1) was used. To avoid possible ordering effects, we randomised the order of the presented alternatives across respondents. The attributes of the alternatives selected for the SC experiment were travel time, travel cost, ticket flexibility, and the level of connectivity on-board (Wi-Fi). The attributes all referred to a standard one-way trip between Rome and Milan.

Due to software limitations it was not possible to customise the design around respondents' most recent trip. The attribute levels presented in Table 3.1 were therefore designed around the *current ranges* (as displayed on operators' websites) and values which are expected to be feasible in the near future. The use of generic values is justifiable by the use of the same origin and destination across all respondents. We generated the choice tasks using a Bayesian D-efficient experimental design, with priors drawn from the literature or based on our expectations (Rose et al., 2008).

Figure 3.1 The layout of the choice tasks



Besides choices, information on consideration of the different modes of transport was collected after each choice task. High-speed rail (HSR) obtained the highest average reported consideration (74%), followed by low-cost air carrier (LCC, 37%), and full-service air carrier (FSC, 31%). Private car obtained the lowest level of reported consideration (14%). Across respondents, the average number of considered alternatives in the 6 choice tasks is 2.26 (with an average standard deviation of 0.56). Moreover, there is little variation in the alternatives considered across choice tasks. For

example, IC, bus, or car-pooling, are found to be either considered or not considered in at least 4 out of 6 choice tasks by 70% of respondents. This suggests that, for some alternatives, consideration is not context-specific and driven by *a priori* beliefs/knowledge for specific journeys.

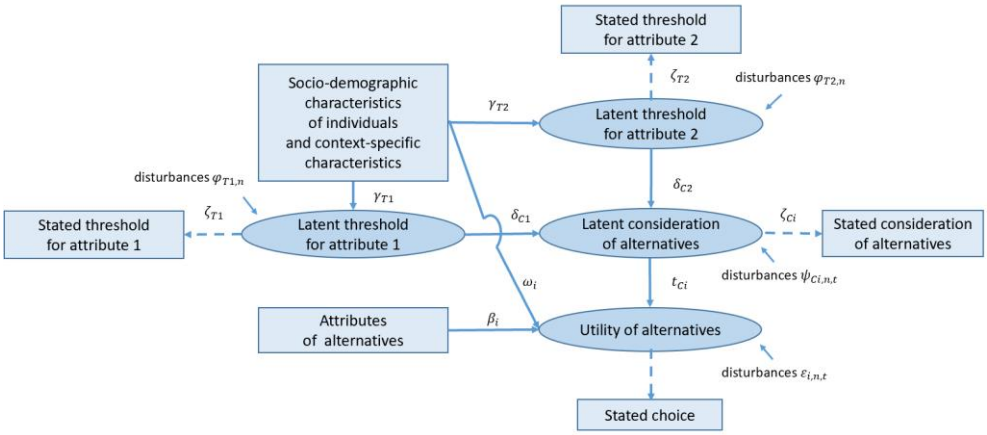
In addition to stated consideration, the existence of thresholds on travel time and cost was also collected for each individual. The average reported value for the threshold on travel time was close to 6 hours (5h57min), while that on travel cost was 123€. Across respondents and choice tasks, the reported thresholds for travel time and travel cost were 'respected' in 85% and 91% of choices, respectively. This gives some measure of the reliability of this information, but the presence of some 'violations' supports the treatment of the thresholds as latent (i.e. acknowledging error in the stated thresholds) as well as the use of a probabilistic approach (rather than deterministically excluding alternatives that exceed thresholds).

3. Methodology

In the proposed hierarchical ICLV model structure - illustrated in Figure 3.2 - *latent thresholds* for attributes are used to explain the *latent consideration* of each alternative, which is then in turn used in the choice model. The model structure contains latent variables for thresholds, for example, one for time and one for cost, where these are explained on the basis of socio-demographic and context-specific characteristics. At the second layer, there are then mode-specific latent variables for consideration, where these again are a function of observable characteristics but are also informed by the

latent threshold variables. Latent consideration then enters into the choice model via a discount factor on the utilities. We will now explain the individual model components in turn.

Figure 3.2 The proposed 'hierarchical ICLV' model



Note: Items in rectangles can be directly observed by the analyst while Items in the ellipses are unobserved. The broken arrows indicate measurement components, while plain arrows indicate structural components.

3.1 Structural model for latent variables

The structural equation for the *latent threshold* $\alpha_{Tk,n}$ for attribute *k* (where, for example, *k=1* for time and *k*=2 for cost) and respondent *n*, which is assumed to be constant across choice situations, is defined by (Equation 3.1):

$$\alpha_{Tk,n} = \gamma_{Tk} Z_{Tk,n} + \varphi_{Tk,n} \tag{3.1}$$

where $Z_{Tk,n}$ denotes a vector of socio-demographic (e.g. gender/income/age of the respondent) or context specific characteristics (e.g. nature of the trip), γ_{Tk} measures their impact on the *latent threshold* for attribute *k*, and $\varphi_{Tk,n}$ represents the error term. The latter is assumed to follow a standard normal distribution across attributes and respondents.

Latent consideration of an alternative *i* and respondent *n* is assumed to be a function of the relevant *latent thresholds* $\alpha_{Tk,n}$, as well as of socio-economic and trip characteristics, $Z_{Ci,n}$, (Equation 3.2). This allows for the possibility that, besides the role of thresholds (i.e. of its structural drivers) in explaining consideration of similar alternatives (for example, in terms of travel time or cost), there are further characteristics of the individuals which are able to explain why specific alternatives are considered or not.

$$\alpha_{Ci,n} = \sum_{k=1}^{K} \delta_{Ck} \alpha_{Tk,n} + \gamma_{Ci} Z_{Ci,n} + \psi_{Ci,n} \tag{3.2}$$

In this equation, δ_{Ck} and γ_{Ci} measure the impact of the *latent thresholds* and of the socio-economic characteristics, respectively, and $\psi_{Ci,n}$ represents a standard normally distributed error term across alternatives and respondents.

Latent consideration is specified at the person level because responses to the stated consideration questions suggest consideration is not context-

specific and driven by a priori beliefs/knowledge for specific journeys (see Section 2). Ideally, one would compare the *latent thresholds* against the presented attribute levels in each equation. A simplified model is however presented, where the *latent thresholds* are implicitly contrasted against (constant) a priori beliefs. On the Rome-Milan corridor, the available alternatives can be categorised in two groups with respect to travel time or cost (e.g. 'fast' and 'slow', 'cheap' and 'expensive'), and this is assumed to guide consideration of the alternatives.

3.2 Measurement model

The stated threshold for attribute *k* and respondent *n,* $I_{Tk,n}$, is used as indicator for the *latent threshold*. When the indicator for the threshold takes the form of a continuous variable (as it would be the case with thresholds for travel time and travel cost), it can be modelled by the following measurement equation (Equation 3.3):

$$I_{Tk,n} = \theta_{Tk} + \zeta_{Tk}\alpha_{Tk,n} + \eta_{Tk,n} \tag{3.3}$$

where θ_{Tk} is a constant, $\alpha_{Tk,n}$ is the latent variable for the threshold for attribute *k*, ζ_{Tk} measures its impact on the value of the corresponding stated threshold. $\eta_{Tk,n}$ is the error term, which follows a zero-mean normal density with a standard deviation of $\sigma_{I_{Tk}}$, which is to be estimated. Using zero-centered thresholds and latent variables obviates the need to estimate the constant θ_{Tk}.

The probability of having a threshold is therefore given by the normal density function (Equation 3.4):

$$P_{(I_{Tk,n}|\alpha_{Tk,n})} = \frac{1}{\sqrt{2\pi\sigma_{I_{Tk,n}}{}^2}}\, e^{-\frac{(I_{Tk,n}-(\theta_{Tk}+\zeta_{Tk}\alpha_{Tk,n}))^2}{\sigma_{I_{Tk,n}}{}^2}} \tag{3.4}$$

Stated consideration for alternative *i,* respondent *n*, and choice situation *t*, $I_{Ci,n,t}$, is used as the indicator for *latent consideration*. This indicator is a binary variable, and the probability of consideration over the sequence of choice tasks takes the form of a binary logit (Equation 3.5):

$$P\left(I_{Ci,n}|\alpha_{Ci,n},\alpha_{Tk,n}\right) = \prod_{t=1}^{T}\left(\lambda_{(I_{Ci,n,t}=0)}\left(\frac{1}{1+\ e^{\theta_{Ci}+\zeta_{Ci}\alpha_{Ci,n}}}\right) + \lambda_{(I_{Ci,n,t}=1)}\left(\frac{e^{\ \theta_{Ci}+\zeta_{Ci}\alpha_{Ci,n}}}{1+\ e^{\theta_{Ci}+\zeta_{Ci}\alpha_{Ci,n}}}\right)\right) \tag{3.5}$$

where $\lambda_{(I_{Ci,n,t}=0)}$ is a dummy variable which takes value 1 when the alternative is stated to be considered, and 0 otherwise, θ_{Ci} is a constant, $\alpha_{Ci,n}$ is the latent variable for consideration, and ζ_{Ci} measures its impact on the value of stated consideration. Even though indicators for stated consideration were collected at the choice-level, these have been modelled

using *latent consideration* specified at the respondent level as explained before.

3.3 Choice model

The choice component is consistent with the Random Utility Maximisation (RUM) theory (McFadden, 1974). In the proposed approach, the modelled component of utility of alternative *i,* for respondent *n* in choice occasion *t*, $V_{i,n,t}$, depends on both observed and latent characteristics, where the latter are deemed to account for the consideration stage in respondents' decision-making process (Equation 3.6):

$$U_{i,n,t} = V_{i,n,t} + \varepsilon_{i,n,t} = \varsigma_{i,n} + \beta_i X_{i,n,t} + \omega_i Z_n + \tau_{Ci} \log\left(a^*{}_{Ci,n}\right) + \varepsilon_{i,n,t} \qquad (3.6)$$

where $X_{i,n,t}$ is a vector of attributes of alternative *i* for respondent *n* and choice situation *t*, whose impact on utility is measured by a vector of estimated parameters β_i, Z_n is a vector of socio-demographic characteristics of respondent *n*, whose impact on utility (which differs across alternatives) is measured by a vector of estimated parameters ω_i, and $\varepsilon_{i,n,t}$ is the error. $a^*{}_{Ci,n}$ is the transformed *latent consideration* variable $\alpha_{Ci,n}$. The transformation in 3.7 is required to bound the variable between 0 and 1 and thereby discount the utility of unconsidered alternatives through the use of a log-transform. The impact of this discount factor on utility is measured by τ_{Ci}.

$$a^{*}{}_{Ci,n} = \frac{1}{(1+exp(-\alpha_{Ci,n}))} \tag{3.7}$$

When $a^{*}{}_{Ci,n}$ is closer to 0, the utility will be heavily discounted, given that $\log(a^{*}{}_{Ci,n}) \rightarrow -\infty$ as $\rightarrow 0$. When the alternative is very likely to be considered, and therefore $a^{*}{}_{Ci,n}$ approaches 1, no discounting of utility is enforced. Therefore, *latent consideration* effectively accounts for the role of consideration by giving a lower choice probability to alternatives that are unlikely to be considered.

We also introduce random alternative-specific constants for all but one alternatives, $\varsigma_{i,n}$, with mean μ_{ς_i} and standard deviation σ_{ς_i}, such that $\varsigma_{i,n}$ = μ_{ς_i}+ $\sigma_{\varsigma_i}\xi_{i,n}$, where $\xi_{i,n}$ follows a standard normal distribution over respondents. Assuming that the error terms for all alternatives are i.i.d. type I extreme value distributed, the probability that alternative *i* is chosen by respondent n – amongst the *J* available alternatives in the set C_n – over the sequence of choice situation *t* can be represented by the standard logit probability (Equation 3.8):

$$P\left(Y_{i,n,t} \mid a^{*}{}_{Ci,n}, \alpha_{Tk,n}, X_{i,n,t}, Z_n, \varsigma_n\right) = \prod_{t=1}^{T} \frac{e^{V_{i,n,t}}}{\sum_{j \in C_n} e^{V_{j,n,t}}} \tag{3.8}$$

The joint LL function for the proposed 'hierarchical ICLV' model is given by (Equation 3.9):

$$LL = \sum_{n=1}^{N} \ln\left[\left(\int_{\alpha_{Tk,n}} \int_{a^*_{C,n}} \int_{\varsigma_n} P(I_{Tk,n}|\alpha_{Tk,n})\, P(Y_{i,n,t}|a^*_{Ci,n}, \alpha_{Tk,n}, X_{i,n,t}, Z_n, \varsigma_n) P(I_{Ci,n}|\alpha_{Ci,n}, \alpha_{Tk,n})\right.\right.$$
$$\left.\left. \mathrm{f}(\varsigma_n|\mu_\varsigma, \sigma_\varsigma)\mathrm{g}(\alpha_{Ci,n}|\alpha_{Tk,n}, Z_{Ci,n})\mathrm{h}(\alpha_{Tk,n}|\, Z_{Tk,n}) d\varsigma_n da^*_{C,n,t} d\alpha_{Tk,n}\right)\right] \quad (3.9)$$

The repeated choice nature of both consideration and choice data is taken into account through the use of a panel mixed multinomial logit (MMNL) model and the estimation of robust standard errors (cf. Daly and Hess, 2011). The models are all estimated using maximum simulated likelihood and 1000 Modified Latin Hypercube Sampling draws[24] (MLHS, Hess et al., 2006).

4. Results and discussion

The present paper serves as proof of concept of accounting for consideration of the alternatives using a hierarchical ICLV framework. In this study we assume that only a subset of the alternatives, namely IC, bus, and car-pooling, are 'partially' considered by the respondents.[25] These alternatives are much slower and cheaper than HSR, FSC, and LCC and we thereby assume consideration decisions are only driven by the presence of

[24] This number of draws resulted in stable models, i.e. by increasing the number of draws we did not observe any improvement in the final LL.

[25] This therefore means that latent consideration is not included in the utility function for the remaining alternatives, meaning that these are always 'fully' considered, whereas the other alternatives are discounted, but still receive a positive choice probability.

thresholds on travel time. This means that latent consideration for the slower alternatives are here explained by only one latent threshold, i.e. that for the travel time attribute. Our assumptions are supported by both the stated consideration and choice data[26], which suggest that respondents are less likely to a priori discard the remaining (faster and more expensive) alternatives from consideration.

[26] On the one hand, the average self-reported levels of consideration for HSR, FSC, and LCC are larger than those for IC, bus, and car-pooling (HSR: 74%; FSC: 31; LCC: 37%; IC: 24%; Bus: 25%; Car-pooling: 21%). On the other hand, HSR, FSC, and LCC have been chosen at least once by 94% of respondents, which would suggest that these alternatives were not a priori discarded, while the remaining alternatives were chosen at least once only by 52% of respondents. Private car deserves a separate discussion. The information on stated consideration for this alternative was contradictory in several circumstances, i.e. a share of respondents stated to consider private car even when this was unavailable for them. For this reason, we decided not to use this information; nevertheless, we took information on car availability into account in the modelling, making car deterministically available/unavailable accordingly.

Table 3.2 Estimation results

	Model 1		Model 2		Model 3	
STRUCTURAL MODELS	*est*	*rob t-rat(0)*	*est*	*rob t-rat(0)*	*est*	*rob t-rat(0)*
Latent threshold travel time						
γ Business			-0.147	-1.97	-0.052	-0.40
γ Age 35+			-0.423	-5.96	-0.464	-3.20
Latent consideration IC						
δ Latent Threshold IC			1.778	6.02	1.898	3.95
Latent consideration BUS						
δ Latent Threshold Bus			2.177	4.96	2.470	5.76
γ Paid myself (vs employer and relatives)			0.458	2.99	0.426	1.53
γ Income 2000+ € or na			-0.539	-3.58	-0.440	-1.79
Latent consideration CAR POOLING						
δ Latent Threshold CAR POOLING			0.908	3.89	2.104	2.47
γ Female			-0.347	-2.96	-0.295	-1.45
γ High-education (university level)			0.273	1.93	0.237	1.11
γ Income 2000+ € or na			-0.449	-4.03	-0.507	-2.05
MEASUREMENT MODELS						
Stated threshold						
ζ Latent threshold travel time			0.256	10.44		
σ Stated threshold travel time			0.375	16.84		
Stated consideration IC						
ζ Latent threshold travel time IC			1.149	5.72		

θ *Stated consideration IC*			-2.263	-10.75		
Stated consideration BUS						
ζ *Latent threshold travel time BUS*			1.368	5.71		
θ *Stated consideration BUS*			-3.122	-9.67		
Stated consideration CAR POOLING						
ζ *Latent threshold travel time CAR POOLING*			2.307	6.14		
θ *Stated consideration CAR POOLING*			-3.149	-8.98		
CHOICE MODELS						
ASC choice IC	-1.072	-1.85	1.080	1.88	1.714	2.80
ASC choice FSC	1.569	1.91	1.715	2.06	1.658	1.87
ASC choice LCC	0.133	0.14	0.434	0.43	-0.047	-0.04
ASC choice Bus	-1.103	-1.34	1.871	2.28	2.136	2.36
ASC choice Car-pooling	-1.443	-1.91	1.179	1.68	0.428	0.50
ASC choice Private Car	-6.239	-1.90	-8.736	-1.88	-7.015	-1.99
Wi-fi free (HRS, IC, FSC, LCC, Bus)	0.141	0.93	0.244	1.66	0.210	1.34
Wi-fi €5 (HRS, IC, FSC, LCC, Bus)	0.071	0.58	0.089	0.70	0.055	0.41
Flexible ticket (free)	0.364	2.81	0.390	3.00	0.457	3.31
Flexible ticket (€5)	0.345	2.92	0.389	3.27	0.395	3.27
Travel time train (HSR, IC)	-0.008	-3.76	-0.008	-3.69	-0.009	-4.18
Travel time air (FSC, LCC)	-0.012	-3.32	-0.012	-3.27	-0.012	-3.05
Travel time Bus/Car-pooling	-0.009	-5.64	-0.008	-5.32	-0.009	-5.87
Travel time Private Car	0.003	0.39	0.004	0.59	0.002	0.31
Travel cost	-0.045	-9.16	-0.047	-8.88	-0.051	-7.85
Travel cost, income na	-0.037	-5.75	-0.043	-5.68	-0.043	-5.13

Paid employer (travel cost)	0.023	4.32	0.026	4.47	0.027	3.91
Income elasticity (travel cost)	-0.221	-3.92	-0.181	-3.22	-0.241	-3.73
Access/egress time main airports	-0.030	-4.37	-0.031	-4.60	-0.032	-4.40
Access/egress time secondary airports	-0.013	-2.25	-0.016	-2.25	-0.014	-2.01
Fidelity card (FSC)	1.897	4.81	2.045	5.06	2.059	4.92
Age 25-34 (IC, Bus, Car-Pooling)	-0.830	-1.75	-0.511	-1.39	-0.533	-1.02
Age 35+ (IC, Bus, Car-Pooling)	-1.602	-2.89	0.183	0.37	0.504	0.62
Business (FSC, LCC)	-0.468	-1.70	-0.507	-1.73	-0.501	-1.66
Business (IC, Bus, Car-pooling)	-1.344	-2.79	-0.637	-1.67	-1.195	-1.86
High-education (all but HSR)	-0.477	-1.88	-0.471	-1.94	-0.672	-2.22
Female (FSC, LCC)	0.570	2.03	0.594	2.00	0.561	1.78
τ Latent Consideration IC			5.859	5.84	6.474	3.14
τ Latent Consideration BUS			14.150	4.38	15.937	2.97
τ Latent Consideration CAR POOLING			6.868	5.62	5.743	2.83
Random coefficients standard deviations						
ASC choice IC sd	-2.037	-6.01	0.491	1.48	-0.050	-0.71
ASC choice FSC sd	1.187	3.80	1.277	4.08	1.289	3.97
ASC choice LCC sd	-1.389	-6.48	1.430	6.06	-1.648	-6.53
ASC choice Bus sd	2.452	8.87	1.073	2.64	0.392	1.39
ASC choice Car-pooling sd	1.547	4.96	0.538	1.82	0.214	0.48
ASC choice Private Car sd	-4.414	-3.18	4.873	2.41	-3.647	-5.05
LL(final, complete model):	-1263.23		-2563.49		-1186.60	
LL(final, choice model only):	-1263.23		-1198.69		-1186.60	

Estimation results are summarised in Table 3.2. Model 1 represents a MMNL model with normally distributed alternative specific constants (ASC) over respondents. This model assumes all alternatives are fully considered and thereby represents the standard practice. Model 2 is the proposed 'hierarchical ICLV' model which accounts for latent consideration effects by 'discounting' the utility of a subset of alternatives (i.e. IC, bus, and car-pooling). Model 3 is the reduced-form model of Model 2. This is also a MMNL model in which we do not make use of the indicators (i.e. stated threshold and stated consideration), but still include the discounting factor (unlike Model 1). This discounting factor is a function of the same set of observed explanatory variables used in the structural equations for latent threshold and latent consideration in Model 2. The estimation of this reduced-form model is aimed at unveiling the actual benefits of using supplementary information (Vij and Walker, 2016).

In Model 1, the estimates for the ASCs reveal a strong preference for FSC over HSR, which was used as the reference alternative in our models.[27] The opposite occurs for IC, car-pooling and private car. Standard deviations, which reflect the degree of heterogeneity for the ASCs at the respondent level, are all significant; in particular, we notice that those for IC, LCC, bus, and car-pooling are larger than the respective mean values.

[27] This alternative has been chosen as baseline even though car-pooling was found to be the minimum variance alternative (Walker et al., 2007). We opted for this inferior solution given that, in the proposed model formulations, consideration effects are directly included in the utility level though latent considerations.

We estimated four coefficients for travel time, namely one for the rail alternatives (HSR and IC), one for the air alternatives (FSC, LCC), one for the slow and low-cost alternatives (bus and car-pooling), and one for the private car alternative. These coefficients show the right (negative) sign and are all statistically significant, except for private car. This result can be explained by the fact that this alternative was chosen in very few occasions (21 out of 1254 choices). Coefficients for access/egress time for airports are also negative and significant, while those for train and bus stations were found to be in-significant.

We interacted travel cost with income in a non-linear way and estimated the income elasticity (3.10). Given that not all respondents reported their income, we estimated separate cost sensitivities (without an income effect) , one for those who disclosed this information (‘*Travel cost*’), and one for those who did not (‘*Travel cost, income na*’).[28]

$$\beta_{travel\_cost_n} = ((\beta_{travel\_cost\_income\_yes_n} * (\left(\frac{income_n}{average_{income}}\right)^{lambda\_income_n}) * (income\_yes\_dummy_n)) + (\beta_{travel\_cost\_income\_na_n} * (1 - income\_yes\_dummy_n)) + (\beta_{paid\_employer_n} * paid\_employer\_dummy_n)) \quad (3.10)$$

[28] Income information was collected using income classes, and we used class-midpoints to compute both income and average income.

The travel cost coefficients have the expected (negative) sign and are both significant; the negative, and significant value for the income elasticity implies that the (absolute) sensitivity to travel cost decreases with increases in income. The shift on the cost sensitivity for those respondents whose trip was paid by the employer ($\beta_{paid\_employer_n}$) is positive and statistically significant, implying that they care less about travel cost (i.e. the travel cost coefficient is less negative) than those who paid the trip themselves or whose trip was paid by some family members.[29] Model 1 also shows that respondents are more likely to select alternatives for which they can get a flexible ticket at a reasonable price (i.e. free or up to 5€). Surprisingly, the presence of Wi-Fi onboard was found to be insignificant.[30] As expected, respondents who are in possession of a loyalty card are more likely to choose the FSC alternative. Those aged 25+ are less likely to choose IC, bus, and car-pooling over HSR compared to their younger counterparts. The HSR alternative is the most likely alternative to be chosen by respondents on a business trip and those educated to at least the university level. We additionally observe a strong preference for the air alternatives over HSR by female travellers.

[29] In particular, the coefficient for travel cost for the former respondents turns out to be less than a half than that for the latter.

[30] Although insignificant, coefficients for Wi-Fi on board were retained in final estimation given that this attribute was modelled in the SC experiment design, differently from access/egress time, for which information was collected afterwards.

Model 2, the 'hierarchical ICLV' model, has three separate components. First, in the structural models, the latent threshold for the travel time is described as a function of observable exogenous variables. The latent consideration for IC, bus, and car-pooling is described as a function of the latent threshold for the travel time and an additional set of observable exogenous variables. Second, in the measurement models, the aforementioned latent variables are linked to the stated threshold for travel time and to stated consideration for IC, bus, and car-pooling (i.e. indicators) respectively. Third, in the choice model, the utility for the alternatives is specified on the basis of attributes of observable exogenous variables and latent consideration.

In the structural model for the latent travel time threshold (see Equation 3.1), it can be seen that the latent threshold on the travel time attribute is lower for those on a business trip and aged at least 35. Consistent with our expectation, the δ parameters indicate that latent consideration for IC, bus, and car-pooling is larger for those respondent with a higher latent threshold for travel time. Latent consideration for bus is lower for those who declared an income of at least 2,000 € per month, for those who did not declare their income, and for those who did not pay the trip themselves. Latent consideration for car-pooling is also lower for those who declared an income of at least 2,000 € per month and those who did not declare their income, but also for female, and for less educated travellers.

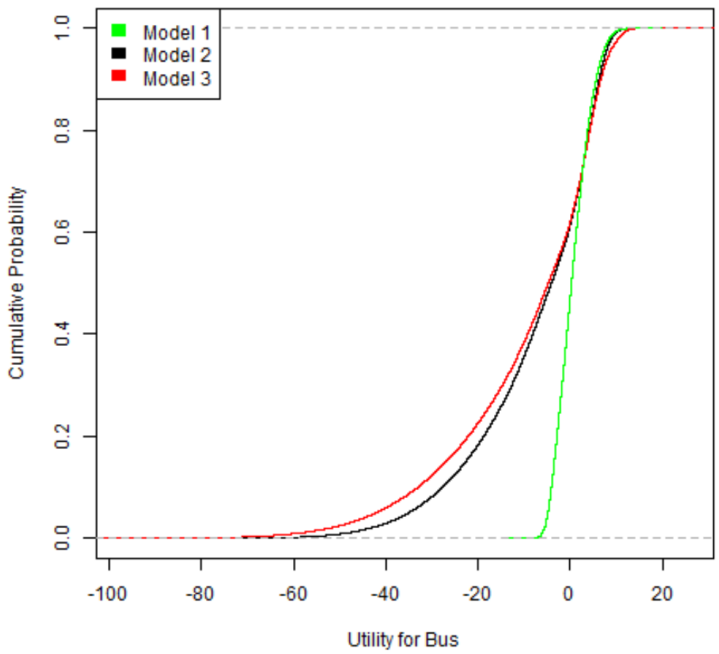
In the measurement models, the ζ parameters (see Equations 3.3 and 3.4) suggest that as our latent variables increase, the probability of respondents stating a higher threshold, or to consider either IC, bus, or car-pooling, increases. In the measurement models for consideration of IC, bus, and car-pooling, negative values for the θ parameters (see Equation 3.5) reflect the fact that the stated consideration rates were on average lower than 50% in the sample (see footnote 26).

The τ parameters in the choice model measure the marginal impact of latent consideration on the utility for the supposed unconsidered alternatives, and their magnitude is simply an outcome of the functional form used; our results show that a value for the transformed latent consideration closer to unity (zero) leads to higher (lower) utility, i.e. less (more) discounting.

Models 3 is the reduced-form MMNL model of Model 2. It has the same structure of Model 2 but it does not make use of the indicators and therefore we do not estimate the measurement models. Looking at Model 3, we notice that many observable exogenous variables in the structural models are no longer significant. This result can be explained by the circumstance that these characteristics now only explain choice (while in Model 2 these also explain the indicators via the latent threshold and latent consideration); this is particularly relevant when the same variable is also included as a free parameter in the choice model (e.g. '*Business*'). This confirms an efficiency gain by Model 2, resulting from the use of the additional variables explained by the measurement model components.

In Figure 3.3 we plot empirical cumulative distribution functions (ECDFs) for the utility values of the bus alternative according to the aforementioned formulations (Models 1-3). The ECDFs represent the proportions of observations showing specific values of the utility. In Models 2 and 3, variations in utility are mostly driven by the impact of latent consideration. Therefore, the distribution of the bus utility in these two models differs with that in Model 1, where consideration effects are not taken into account. Interestingly, for around 60% of the sample the utility of the bus alternative is strongly discounted, thus assigning a lower choice probability for this alternative.

Figure 3.3 ECDFs for the impact of consideration effects on utility for bus



We now discuss the implications of accounting for consideration effects on the parameters of the utility function. Relative to Model 1, the standard deviations of the ASC strongly reduce for IC, bus, and car-pooling when consideration effects are introduced. In particular, in Model 2, the standard deviation parameter becomes insignificant for IC. Differently from previous studies employing the ICLV approach (e.g. Kløjgaard and Hess, 2014; Mariel and Meyerhoff, 2016; Song et al., 2018) we are not able to quantify which share of preference heterogeneity is explained by the latent variables. This is due to the hierarchical nature of our latent variables and the transformations these are subjected to before including them in the utility function. The reduction in size of the standard deviations for the ASCs' for the 'discounted' alternatives, however, indicates that also in this case at least a share of preference heterogeneity is explained by the introduction of latent constructs.

The impact on the travel time and travel cost coefficients can be more effectively analysed in terms of Value of Travel Time (VTT) indicators (Table 3.3). VTT indicators are obtained for an individual who pays her/himself for the trip. Standard errors are calculated using the delta method.

Table 3.3 VTT (€/hour)

	Model 1		Model 2				Model 3			
	est.	*t-stat(0)*	*est.*	*t-stat(0)*	*change (2vs1)*	*t-stat (2vs1)*	*est.*	*t-stat(0)*	*change (3vs1)*	*t-stat (3vs1)*
Train	10.33	3.66	9.58	3.56	-7%	0.19	10.35	3.97	0%	-0.01
Air	16.26	3.03	15.71	3.03	-3%	0.07	13.98	2.81	-12%	0.31
Bus/Car-pooling	11.43	5.31	10.63	4.86	-7%	0.26	10.02	5.14	-11%	0.48

Overall, we observe a reduction in the VTT for all the alternatives in Models 2 and 3 relative to Model 1. The differences are, however, not significant and therefore suggest that consideration effects actually have a rather limited impact on VTT estimates.

Turning our attention to model fit, the final Log-Likelihood of the traditional MMNL model (Model 1) and the reduced-form MMNL model (Model 3) cannot be compared with that of the 'hierarchical ICLV' model (Model 2). This is due to the fact that whilst Models 1 and 3 are estimated on the choice data alone, the ICLV structure also explains respondents' stated thresholds on travel time and stated consideration for the IC, bus, and car-pooling alternatives. It is however possible to derive the final Log-Likelihood for the choice model component separately from the other components. A comparison of these measures reveal that Models 2 and 3 outperform Model 1. Vij and Walker (2016) suggest that model fit for the ICLV model and its MMNL reduced form model should be similar. A worse fit for the ICLV model (Model 2) with respect to its reduced form (Model 3) is not uncommon in the literature. In this case, the difference in fit between Model 2 and Model 3 is, however, not negligible but can be explained by the fact that the ICLV model evaluates a joint likelihood function.

5. Conclusions

The latent nature of the consideration stage, as a part of the decision-making process, implies that variations in consideration of the alternatives across individuals are not observable. Reports of consideration – or of aspects

related to this stage – might indeed be collected during SC experiments. Their direct use as additional explanatory variables, to account for consideration of alternatives in the estimation of discrete choice models is, however, highly discouraged. In this paper we propose an Integrated Choice and Latent Variable (ICLV) model to account for consideration of the alternatives, with an application to transport mode choice. The ICLV approach helps circumventing the aforementioned drawbacks by treating information on respondents' processing strategies as dependent variables rather than as explanatory variables.

Our approach is 'hierarchical', in the sense that latent thresholds for attributes are used to explain latent consideration of the alternatives, reflecting what might happen in the individuals' decision-making process. These inter-related latent variables are in turn used to explain both choice outcomes and self-reported information on the decision-making process in the form of stated thresholds for attributes and stated consideration of the alternatives. Latent consideration enters the utility of the alternatives through a 'discounting' factor, which accounts for the role of consideration lowering choice probability for the supposed unconsidered alternatives.

In this study, we incorporate consideration effects only on a subset of alternative transport modes which are deemed to be hardly considered by the respondents' of a mode choice experiment on the Rome-Milan corridor. Here, seven alternatives are available, which vary substantially in terms of their characterising attributes, particularly travel time. We assume slower

(but also less expensive) alternatives are not always considered, most likely due to the presence of thresholds for the travel time attribute. The assumption is supported by both stated consideration and stated choice data.

The proposed 'hierarchical ICLV' model is compared against two reference models. The first is a traditional MMNL model where consideration effects are not taken into account and all alternatives are assumed to be 'fully' considered. The second is a reduced-form MMNL model of the proposed 'hierarchical ICLV' model in which we keep the structural equations for the latent variables, but we do not make use of the respective indicators. The first reference model represents the current practice in most mode choice studies and we estimate the second in order to unveil the actual benefits of the proposed ICLV model.

Consistent with our expectations, results suggest that the latent threshold on travel time is lower for respondents on a business trip and for those aged at least 35. Latent consideration for IC, bus, and car-pooling is larger for those respondents with a higher latent threshold for travel time. Latent consideration for bus is also lower for richer respondents, and for those who did not pay for the trip themselves. Latent consideration for car-pooling is instead lower not only for richer travellers, but also for female and less educated travellers. The latter results could potentially be explained by safety concerns and by the fact that this mode has a very high ICT component. Latent consideration for IC, bus, and car-pooling has a

significant (and positive) marginal effect on the overall utility of these alternatives; conversely, the utility for those respondents with predicted lower levels of latent consideration gets highly discounted, and choice probability for these alternatives approaches zero.

Interestingly, willingness-to-pay indicators are hardly affected by the introduction of consideration effects. Previous studies found more tangible differences in these metrics with respect to models assuming that all alternatives are 'fully' considered (Ben-Akiva and Boccara, 1995; Basar and Bhat, 2004). We believe that this is due to the fact that we simultaneously account for additional random heterogeneity and that we appropriately account for measurement errors in the indicators for consideration and thresholds in the ICLV model. In terms of model fit, we observe an improvement with respect to a traditional MMNL as a result of explicitly account for consideration effects. However, consistent with Vij and Walker (2016), who discuss pros and cons of any latent variable approach, we find that such improvement in fit cannot be fully ascribed to the use of the indicators.

The ICLV model shows benefits when compared with traditional RUM-based choice models. First, it enables us to explain a share of otherwise completely random heterogeneity, which can therefore be associated to latent thresholds for attributes and latent consideration of the alternatives. Second, thanks to the indicators we are able to obtain more insights into the

structural drivers of consideration. This might be of interest for policymakers and private operators, and useful when applying the model.

Acknowledgements

The authors acknowledge the Italian Railway Infrastructure Manager, Rete Ferroviaria Italiana, and the ownership of two service stations located on the A1/E35 Highway, for gathering special authorisations to administer surveys at platforms, and inside the service stations, respectively. The second author acknowledges financial support by the European Research Council through the consolidator grant 615596-DECISIONS.

References

Başar, G. and Bhat, C. (2004) A parameterized consideration set model for airport choice: an application to the San Francisco bay area. *Transportation Research Part B: Methodological*, 38(10), 889-904, https://doi.org/10.1016/j.trb.2004.01.001.

Ben-Akiva, M., McFadden, D., Train, K., Walker, J., Bhat, C., Bierlaire, M., Bolduc, D., Börsch-Supan, A., Brownstone, D., Bunch, D. S., Daly, A., de Palma, A., Gopinath, D., Karlstrom, A. and Munizaga, M. A. (2002) Hybrid choice models: progress and challenges. *Marketing Letters*, 13(3), 163-175, https://doi.org/10.1023/A:1020254301302.

Bolduc, D., Ben-Akiva, M., Walker, J. and Michaud, A. (2005) Hybrid choice models with logit kernel: applicability to large scale models. In *Integrated land-use and transportation models: behavioural foundations*,

edited by M. E. H. Lee-Gosselin and S.T. Doherty, 275-302. Emerald Group Publishing Limited.

Cantillo, V. and Ortúzar, J. D. D. (2005) A semi-compensatory discrete choice model with explicit attribute thresholds of perception. *Transportation Research Part B: Methodological*, 39(7), 641-657, https://doi.org/10.1016/j.trb.2004.08.002.

Cascetta, E. and Papola, A. (2001) Random utility models with implicit availability/perception of choice alternatives for the simulation of travel demand. *Transportation Research Part C: Emerging Technologies*, 9(4), 249-263, https://doi.org/10.1016/S0968-090X(00)00036-X.

Daly, A. and Hess, S. (2011) Simple methods for panel data analysis. Paper presented at the 90th annual meeting of the Transportation Research Board, Washington, D.C., January.

Daly, A., Hess, S., Patruni, B., Potoglou, D. and Rohr, C. (2012) Using ordered attitudinal indicators in a latent variable choice model: a study of the impact of security on rail travel behaviour. *Transportation*, 39(2), 267-297, https://doi.org/10.1007/s11116-011-9351-z.

Fotheringham, A. S. (1988) Note - Consumer store choice and choice set definition. *Marketing Science*, 7(3), 299-310, https://doi.org/10.1287/mksc.7.3.299.

Hess, S., Train, K. E. and Polak, J. W. (2006) On the use of a modified

latin hypercube sampling (MLHS) method in the estimation of a mixed logit model for vehicle choice. *Transportation Research Part B: Methodological*, 40(2), 147-163, https://doi.org/10.1016/j.trb.2004.10.005.

Hess, S., and Hensher, D. A. (2013) Making use of respondent reported processing information to understand attribute importance: a latent variable scaling approach. *Transportation*, 40(2), 397-412, https://doi.org/10.1007/s11116-012-9420-y.

Kløjgaard, M. E. and Hess, S. (2014) Understanding the formation and influence of attitudes in patients' treatment choices for lower back pain: testing the benefits of a hybrid choice model approach. *Social Science & Medicine*, 114, 138-150, https://doi.org/10.1016/j.socscimed.2014.05.058.

Manski, C. F. (1977) The structure of random utility models. *Theory and decision*, 8(3), 229-254, https://doi.org/10.1007/BF00133443.

Mariel, P. and Meyerhoff, J. (2016) Hybrid discrete choice models: gained insights versus increasing effort, *Science of the total environment*, 568, 433-443, https://doi.org/10.1016/j.scitotenv.2016.06.019.

Martínez, F., Aguila, F. and Hurtubia, R. (2009) The constrained multinomial logit: a semi-compensatory choice model, *Transportation Research Part B: Methodological*, 43(3), 365-377, https://doi.org/10.1016/j.trb.2008.06.006.

McFadden, D. (1986) The choice theory approach to marketing

research, *Marketing Science*, 5(4), 275-297, https://doi.org/10.1287/mksc.5.4.275.

McFadden, D. (1974) Conditional logit analysis of qualitative choice behaviour. In *Frontiers in econometrics*, edited by P. Zarembka, 105-142. Academic Press, New York.

Paulssen, M., Temme, D., Vij, A. and Walker, J. (2014) Values, attitudes and travel behavior: a hierarchical latent variable mixed logit model of travel mode choice, *Transportation*, 41(4), 873-888, https:/doi.org/10.1007/s11116-013-9504-3.

Roberts, J. H. and Lattin, J. M. (1997) Consideration: review of research and prospects for future insights, *Journal of Marketing Research*, 34(3), 406-410, https://doi.org/10.2307/3151902.

Rose, J. M., Bliemer, M. C.J., Hensher, D. A. and Collins, A. T. (2008) Designing efficient stated choice experiments in the presence of reference alternatives, *Transportation Research Part B: Methodological*, 42(4), 395-406, https://doi.org/10.1016/j.trb.2007.09.002.

Song, F., Hess, S., and Dekker, T. (2018) Accounting for the impact of variety-seeking: theory and application to HSR-air intermodality in China, *Journal of Air Transport Management*, 69, 99-111, https://doi.org/10.1016/j.jairtraman.2018.02.008.

Soto, J.J., Cantillo, V., and Arellana J. (2018) Incentivizing alternative fuel

vehicles: the influence of transport policies, attitudes and perceptions, *Transportation*, 45(6), 1721-1753, https://doi.org/10.1007/s11116-018-9869-4.

Swait, J. (1984) Probabilistic choice set generation in transportation demand models. PhD dissertation, Massachusetts Institute of Technology, Cambridge, USA.

Swait, J. (2001) A non-compensatory choice model incorporating attribute cutoffs, *Transportation Research Part B: Methodological*, 35(10), 903-928, https://doi.org/10.1016/S0191-2615(00)00030-8.

Swait, J. and Ben-Akiva, M. (1987a) Empirical test of a constrained choice discrete model: mode choice in Sao Paulo, Brazil, *Transportation Research Part B: Methodological*, 21(2), 103-115, https://doi.org/10.1016/0191-2615(87)90010-5.

Swait, J. and Ben-Akiva, M. (1987b) Incorporating random constraints in discrete models of choice set generation *Transportation Research Part B: Methodological*, 21(2), 91-102, https://doi.org/10.1016/0191-2615(87)90009-9.

Vij, A. and Walker, J. L. (2016) How, when and why integrated choice and latent variable models are latently useful, *Transportation Research Part B: Methodological*, 90, 192-217, https://doi.org/10.1016/j.trb.2016.04.021.

Walker, J. L., Ben-Akiva, M. and Bolduc, D. (2007) Identification of parameters in normal error component logit-mixture (NECLM) models, *Journal of Applied Econometrics*, 22(6), 1095-1125, https://doi.org/10.1002/jae.971.

Williams, H. C. W. L. and Ortúzar, J. D. D. (1982) Behavioural theories of dispersion and the mis-specification of travel demand models, *Transportation Research Part B: Methodological*, 16(3), 167-219, https://doi.org/10.1016/0191-2615(82)90024-8.

Chapter 4 - Allowing for heterogeneity in the consideration of airport access modes: the case of Bari airport

Abstract

Mode choice models traditionally assume that all objectively available alternatives are considered. This might not always be a reasonable assumption, even when the number of alternatives is limited. Consideration of alternatives, like many other aspects of the decision-making process, cannot be observed by the analyst, and can only be imperfectly measured. As part of a stated choice survey aimed at unveiling air passengers' preferences towards access modes to Bari International Airport, in Italy, we collected a wide set of indicators that either directly or indirectly measure respondents' consideration for the public transport alternative. In our access mode choice model, consideration for public transport services is treated as a latent variable, and enters the utility function for this mode through a 'discounting' factor. The proposed *Integrated Choice and Latent Variable* approach allows the analyst not only to overcome potential endogeneity and measurement error issues associated with the indicators, but also makes the model suitable for forecasting. As a result of accounting for consideration effects, we observe an improvement in fit which also holds in a validation sample; moreover, the effects of policy changes aimed at improving modal share for public transport are considerably reduced.

Keywords

Consideration of alternatives, latent variables, airport access

1. Introduction

The number of air travellers in the European Union has significantly increased in recent years (Eurostat, 2018). This growth was largely driven by low-cost carriers, which made air transport economically affordable to a larger share of the population. This expansion continuously imposes a challenge for airport managers and regional mobility planners, who have to deal with the increasing number of (infrequent) travellers, but also additional staff and accompanying persons needing to access the airport. There is no generic solution to this challenge which is valid everywhere; in addition to this, each user segment (e.g. resident vs. non-resident, business vs. non-business, or airport employees) has its own needs and preferences towards airport access services (Leigh Fisher Associates et al., 2000; 2002).

Most studies investigating the drivers of airport access mode decisions have relied on revealed preference (RP) and (or) stated preference (SP) data in combination with discrete choice models. These studies were aimed at understanding the choice between existing access modes (Harvey, 1986; Tam and Lam, 2008), or focused on the implications of introducing a new access mode (Monteiro and Hansen, 1996; Jou et al., 2011). In some cases, access mode decisions have been modelled jointly with airport and/or airline decisions (Pels et al., 2003; Hess and Polak, 2006; Gupta et al., 2008).

The underlying assumption in all these studies is that all objectively available airport access modes are effectively considered by each airport user. However, this assumption might be questioned since some access modes might be discarded a priori, i.e. regardless of their characteristics. For example, in the case of air travellers, trips to the airport are only the first 'leg' of a longer trip and are associated with a hard constraint, i.e. the departure time of the flight. Hence, the possible consequences of a delay in arriving at the airport may be severe. Even though unexpected delays might occur with all modes, air travellers might consider as feasible only those alternatives that they 'perceive' to have a sufficiently low risk of getting to the airport late. Other factors that might influence which alternatives are considered or not are concerns for personal safety, or the need to access a train station/bus stop which is inconveniently located with respect to their location of departure. Comfort also matters, particularly because passengers perceive the need to transfer and wait (e.g. with public transport) as a significant 'discomfort' (Coogan, 2008).

The assumption that individuals might consider only a subset of the available alternatives has been tested in several transport contexts, particularly route and mode choice (Frejinger et al., 2009; Swait and Ben-Akiva, 1987). However, to the best of our knowledge, this assumption has never been tested in the specific context of airport accessibility, which is the focus of this paper. The biggest challenge with consideration of alternatives is that this aspect of the decision-making process is not observable to the analyst.

Some researchers have tried to incorporate consideration effects into probabilistic models only on the base of the observed choices (Cascetta and Papola, 2001; Basar and Bhat, 2004; Cantillo and Ortúzar, 2005). Others have explored the possibility of using supplementary information as direct (but imperfect) measures of consideration, including for example perceived availability (Ben-Akiva and Boccara, 1995) or acceptability (Hensher and Ho, 2015) of the alternatives and thresholds for attributes (Swait, 2001), elicited using ad-hoc questions in travel surveys. These indicators, however, might not correspond to actual levels of consideration, i.e. there is potential for measurement error, and they may be correlated with other unobserved factors, i.e. there is scope for endogeneity bias (Hess and Hensher, 2013). Given this, rather than using them as 'error-free' measures of consideration, it might be preferable to recognise that these are a function of latent consideration, and treat them as dependent rather than independent variables using an Integrated Choice and Latent Variable (ICLV) model (Bolduc et al., 2005). The ICLV approach has been extensively used in many fields, not only transport, to incorporate either psychological factors such as attitudes and perceptions (Kløjgaard and Hess, 2014) or respondents' processing strategies (Hess and Hensher, 2013) into models based on random utility maximisation (McFadden, 1974). Besides allowing the analyst to overcome potential endogeneity and measurement error issues with the indicators, the ICLV approach also allows us to make the indicators suitable for forecasting.

In this paper, we adopt the ICLV framework to measure consideration of airport access modes using three distinct sets of indicators collected as part of a stated choice (SC) survey. The first set consists of the level of agreement with various perception statements and of a preference-based ranking of the alternatives; the second refers to thresholds for attributes inferred from respondents' previous choices; the third set comprises direct reports of consideration of the alternatives. These indicators have been chosen because they represent additional sources of information which are generally collected during travel surveys (the first two), or because they have been used in previous studies to measure consideration of the alternatives (the third).

In our proposed formulation, latent consideration explains the indicators and enters the utility of an alternative through a discounting factor. The discounting factor accounts for consideration lowering the utility, and therefore choice probability, of a supposed unconsidered alternative.

Data for this study comes from a SC experiment on a sample of air travellers of Bari International Airport, in Italy. This airport recently experienced a substantial increase in travellers (Eurostat, 2018) as a result of the increase in the number of low-cost connections available. A direct train connects the airport with the city centre in 15 minutes; however, more peripheral areas within the Metropolitan City of Bari and the Apulian region are not as easily accessible, since the railway link to the airport is not interconnected with the main regional railway networks. Other public transport means are available

(e.g. local buses), but these involve at least one interchange, are even less frequent, and their timetables are not coordinated. As a result, travellers from these areas mainly access the airport by car.

Given these premises, in this paper we estimate mode choice models in which we allow for the possibility that some air travellers might not consider public transport as a feasible alternative. Both RP and SP data is used in the estimation, and the proposed ICLV models are compared with two reference models: the first is a traditional Mixed Multinomial Logit (MMNL) model in which all alternatives are assumed to be considered. The second is a reduced-form MMNL model of the proposed ICLV models, which only infers the latent consideration for public transport through the observed choice data.

The remainder of the paper is structured as follows. We describe the data in Section 2. Section 3 explains the proposed model. In Section 4, we report and discuss the estimation results, and in Section 5 we present the validation exercise. Finally, in Section 6 we draw conclusions from our study.

2. Data

The data used in this paper were gathered through pen-and-paper personal interviews (PAPI) conducted in autumn 2016 and autumn 2017. A total of 1,046 randomly selected residents in four cities in a range of 50-100 km from Bari International Airport were interviewed at their homes. Our sample comprises only air travellers, i.e. individuals who had flown through the airport at least once in the previous three months. A preliminary screening

question was used in the survey to ensure that respondents were within scope. Official statistics on the actual profiles of the airport users are not available, and thus we are not able to assess the representativeness of our sample with respect to the target population. However, our sample is balanced across key socio-demographic characteristics (e.g. sex, age, and level of education, Table 4.1). Although more than 50% of the Apulian population is aged 50 years and over, this may still hold because individuals in this age group tend to travel less than their younger counterparts (Isfort, 2017). One worrisome sample characteristic is that over 30% of the respondents declared to not have any wage income. Because around 80% of them are under 30 and/or students, they most likely rely on their parents' financial support. Therefore, caution is needed when interpreting the estimation results. We decided to only focus on residents for three reasons. First, because they are more likely to have a private car, and therefore to use it to access the airport, whether in the '*kiss-and-ride*' or the '*park-and-ride*' mode (i.e. 'as passengers' or 'as drivers'). Second, because they are more likely to have better knowledge of all available alternatives. Third, because they are more familiar with regional traffic patterns. Given these premises, residents represent a major potential market for public transport services (Leigh Fisher Associates et al., 2002).

The catchment area for this airport goes far beyond the city of Bari. It comprises the geographical boundaries of the Apulian region and the adjacent county of Matera in the Basilicata region. It has been estimated

that approximately 3,150,000 individuals can access the airport within 90 minutes (ENAC, 2010). Only 9% of these potential passengers live in the city of Bari (ISTAT, 2018), and this explains why this paper focuses on regional rather than urban mobility patterns towards the airport.

Table 4.1 Descriptive statistics of the sample and of the population in the Apulian region

Social traits/Year	Survey 2016	Survey 2017	Region (ISTAT, 2018)
Female	49.7%	51.6%	51.7%
Age: 18-24	27.0%	37.3%	9.2%
Age: 25-34	30.3%	26.7%	13.9%
Age: 35-49	24.0%	20.1%	25.7%
Age: 50+	18.7%	16.0%	51.2%
Do not have monthly income	37.0%	31.7%	-
Income: < 500 €	2.0%	5.0%	-
Income: 500-1,000 €	10.0%	10.7%	-
Income: 1,000-2,000 €	34.0%	26.6%	-
Income: 2,000-4,000 €	13.7%	9.1%	-
Income: >4,000 €	3.3%	2.1%	-
Income: prefer not to disclose	-	14.7%	-
Education BSc+	64.0%	37.7%	18.8%
Business trip	30.0%	18.8%	-
Student	29.0%	48.4%	-
City: Matera	12.7%	19.4%	-
City: Altamura	64.3%	24.7%	-
City: Gravina	26.3%	25.5%	-
City: Corato	-	30.4%	-
Total	300	746	3,381,008

Both revealed and stated preferences were collected during the survey. The former refer to the respondents' last trip to the airport. In the SC experiment, respondents were asked to choose their preferred access mode amongst i) public transport with at least one change, ii) a direct private bus run by the airport management in cooperation with private operators, iii) car 'as driver', iv) car 'as passenger' (i.e. the possibility of being dropped-off by someone else), and v) taxi. The attributes of the alternatives modelled in the experiment were in-vehicle travel time, out-of-vehicle travel time (i.e. the waiting time between connecting services for public transport), travel cost, and headway (i.e. the time until the next public transport service to the airport). When the departure place is located within 50-100 km from the airport, it is reasonable to assume that passengers will use a timetable to schedule their arrival at the train station/bus stop. The headway might still be an important factor in their decision-process because if there is a reliability issue with the scheduled public transport journey, a high headway could lead to passengers missing their flight.

Each respondent was shown 5 choice tasks which were generated using Bayesian D-efficient experimental designs (Rose et al., 2008), with priors inferred from a pilot study. Experimental designs with different attribute levels but the same priors were generated for each city in the study area. The attribute levels were designed around the current ranges (as reported by the transport operators and www.viamichelin.com), and the order of the presented alternatives was randomised across respondents to avoid

possible *left-to-right* effects (i.e. always choose the first alternative on the left).

The 2017 survey (i.e. the second wave) also collected three sets of supplementary information which could be used as indicators of *latent consideration* of the available alternatives, particularly public transport.

The first set of indicators consists of the level of agreement with perception statements towards the alternatives and of a ranking of these on the base of their overall preference. With respect to public transport, we collected responses about agreement with the following statements on a 5-points Likert scale:

1) *If I had to use public transport to get to the airport, I would take an earlier bus/train to make sure I will not get there late*.
2) *I do not consider the possibility of getting to the airport by public transport, because I can only be there too early or too late due to the reduced frequency*.

The second set of measures refers to respondents' previous choices. Respondents were asked to report how many times they had used each of the airport access alternatives in the previous year. This information is in turn employed to infer respondents' threshold for the travel time attribute, i.e. travel time for the (used) alternative with the longest travel time is assumed to be their threshold. This indicator would give an idea of the

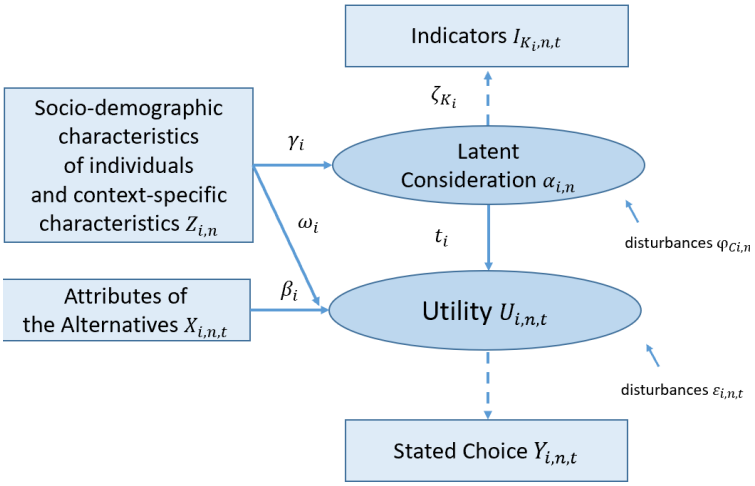
maximum travel time the respondents are willing to accept outside the SC experiment, i.e. in real situations, and its robustness (as an indicator of a potential self-imposed threshold on travel time) would certainly increase with the number of trips made. Of course, this presents a lower limit on this threshold; just because a respondent has never chosen a mode taking longer than the slowest mode chosen in the past does not mean that the travel time for these modes exceeds that traveller's threshold. Moreover, there might be a potential correlation between the threshold, i.e. with the chosen mode(s), and the number of airport trips made, as well as with the haul of the flight(s) taken. For example, an individual might have done only one trip to the airport and for an intercontinental flight and asked a family member to drive her/him to the airport; however s/he might be actually willing to accept a longer travel time than that by car if taking a domestic flight. These reasons make the treatment of these values as indicators rather than direct measures of threshold even more important.

The third set of indicators comprises self-reports of consideration of the alternatives. Respondents were asked to reveal which alternatives they actually considered at the end of each choice task. Despite being directly related to consideration, self-reports of consideration are still imperfect measures of consideration, and therefore these might not necessarily correspond to an individual's 'actual' behaviour. An additional limitation resides in the possibility that, if collected after each choice task, these follow-up questions might influence subsequent choices.

3. Methodology

In Figure 4.1 we illustrate the general ICLV model formulation, consisting of three sub-models: a structural model, where *latent consideration* is described as a function of socio-demographic characteristics of the respondent; a measurement model, which links *latent consideration* to the values of the indicators; and a choice model, where the utility for the alternatives and hence the choice is specified on the basis of attributes of observable exogenous variables and *latent consideration*.

Figure 4.1 The proposed ICLV model



Note: Items in rectangles can be directly observed by the analyst. Items in the ellipses are unobserved: an error is added to take account of this. The broken arrows indicate measurement equations, while plain arrows indicate structural equations.

3.1 Structural model

In the structural equation, *latent consideration* for alternative *i* (e.g. public transport) and respondent *n*, α_{in}, is defined by (4.1)

$$\alpha_{i,n} = \gamma_i Z_{i,n} + \varphi_{i,n} \tag{4.1}$$

where $Z_{i,n}$ denotes a vector of socio-demographic characteristics of the respondent whose impact on *latent consideration* is measured by γ_i, and $\varphi_{i,n}$ represents a normally distributed error term (Bolduc et al., 2005). Changes in the structural equation impact both the measurement model and the choice model components, given that *latent consideration* is an explanatory variable in both.

3.2 Measurement model

In the measurement model, the indicator (dependent variable) is explained by *latent consideration* (independent variable, as defined by Equation 4.1). Depending on the nature of the selected indicator, distinct measurement models can be specified. In this paper we test for the use of ordinal, continuous, and binary indicators. Therefore, we specify the corresponding measurement models as an ordinal logit, a probability distribution function, and a binary logit, respectively.

Ordinal indicators

The level of agreement with statements such as those related to public transport reported in the previous Section can be recorded, for example, on a 5-point scale, ranging from 1 being 'completely disagree' to 5 referring to 'completely agree'. The ranking of a single alternative, e.g. public transport,

amongst the five alternatives is also treated as ordinal, with the value ranging from 1 if the alternative is the 'most preferred' to 5 if the alternative is the 'least preferred'. Of course, if the ranking of multiple alternatives were to be included in the model, an exploded logit would be more appropriate than the ordered model used here. Both level of agreement to the statement and the ranking of the alternatives can be used as indicators for *latent consideration*. The probability of observing a specific response to these ordinal indicators K relative to alternative *i* and respondent *n*, can be modelled using an ordered logit form (Equation 4.2):

$$P_{K_i,n}\left(I_{K_i,n} = s|\alpha_{i,n}\right) = \frac{e^{\mu_{K_i,s}-\zeta_{K_i}\alpha_{i,n}+\psi_{K_i,n}}}{1+\ e^{\mu_{K_i,s}-\zeta_{K_i}\alpha_{i,n}+\psi_{K_i,n}}} - \frac{e^{\mu_{K_i,s-1}-\zeta_{K_i}\alpha_{i,n}+\psi_{K_i,n}}}{1+\ e^{\mu_{K_i,s-1}-\zeta_{K_i}\alpha_{i,n}+\psi_{K_i,n}}} \quad (4.2)$$

Where $\mu_{K_i,s}$ are estimated threshold parameters, s ∈ (1,2,3,4,5) if a 5-point scale is used, $\alpha_{i,n}$ is the *latent consideration*, ζ_{Ki} measures its impact on the value of the indicator, and $\psi_{K_i,n}$ is the error term. For normalisation purposes, we set $\mu_{K_i,0}$ to - ∞ and $\mu_{K_i,5}$ to + ∞; therefore, only the intermediate four threshold values can be estimated for each indicator.

The likelihood of the observed value $I_{K_i,n}$ is then given by (Equation 4.3):

$$L_{I_{K_i,n}} = \sum_{s=1}^{S} \lambda_{\left(I_{K_i,n}=s\right)} \left(\frac{e^{\mu_{K_i,s}-\zeta_{K_i}\alpha_{i,n}}}{1+\ e^{\mu_{K_i,s}-\zeta_{K_i}\alpha_{i,n}}} - \frac{e^{\mu_{K_i,s-1}-\zeta_{K_i}\alpha_{i,n}}}{1+\ e^{\mu_{K_i,s-1}-\zeta_{K_i}\alpha_{i,n}}} \right) \quad (4.3)$$

where λ is a dummy variable which takes value 1 when the value for the indicator equals s, and 0 otherwise.

Continuous indicators

The threshold for an attribute *d* and respondent *n,* $I_{Td,n}$ *,* can also be used as an indicator for *latent consideration*. Assuming the indicator takes the form of a continuous variable, it can be modelled by the following measurement equation (4.4):

$$I_{Td,n} = \theta_{Td} + \zeta_{Td}\alpha_{i,n} + \eta_{Td,n} \tag{4.4}$$

where θ_{Td} is a constant, $\alpha_{i,n}$ is the *latent consideration*, ζ_{Td} measures its impact on the value of the threshold, and $\eta_{Td,n}$ is the error term, which follows a zero-mean normal density and standard deviation of $\sigma_{I_{Td}}$. By centering the indicators on zero, i.e. subtracting the sample mean from each indicator, we obviate the need to estimate the constant θ_{Td} .

The likelihood for observing a particular threshold is given by the normal density function (Equation 4.5):

$$L_{I_{Td,n}} = \frac{1}{\sqrt{2\pi\sigma_{I_{Td,n}}{}^2}}\, e^{-\frac{(I_{Td,n}-(\theta_{Td}+\zeta_{Td}\alpha_{i,n}))^2}{\sigma_{I_{Td,n}}{}^2}} \tag{4.5}$$

Binary indicators

Stated consideration for alternative *i,* respondent *n*, and choice situation *t*, $I_{C_i,n,t}$, is our third candidate indicator for *latent consideration*. This is a binary variable, and probability of consideration takes the form of a binary logit (Equation 4.6):

$$P_{C_i,n,t}\left(I_{C_i,n,t}|\alpha_{i,n}\right) = \frac{e^{\theta_{C_i}+\zeta_{C_i}\alpha_{i,n}+\nu_{C_i,n,t}}}{1+e^{\theta_{C_i}+\zeta_{C_i}\alpha_{i,n}+\nu_{C_i,n,t}}} \tag{4.6}$$

where θ_{C_i} is a constant, $\alpha_{i,n}$ is the *latent consideration*, ζ_{C_i} measures its impact on the value of stated consideration, and $\nu_{C_i,n,t}$ is the error term. Although indicators for stated consideration are collected at the choice-level, we decided to model them using *latent consideration* specified at the respondent level, with choice specific measurement equations, to make them comparable with the other two sets of indicators. The likelihood function for this part of the model is (Equation 4.7):

$$L_{I_{C_i,n,t}} = \lambda_{\left(I_{C_i,n,t}=0\right)}\left(1 - P_{C_i,n,t}\left(I_{C_i,n,t}|\alpha_{i,n}\right)\right) + \lambda_{(I_{C_i,n,t}=1)}P_{C_i,n,t}\left(I_{C_i,n,t}|\ \alpha_{i,n}\right) \tag{4.7}$$

where λ is a dummy variable which takes a value of 1 when the alternative is stated to be considered, and 0 otherwise.

3.3 Choice model

The mode choice model uses a random utility specification, where the utility of alternative *i,* for respondent *n* in choice occasion *t* depends on both observable explanatory variables and *latent consideration* (Equation 4.8):

$$U_{i,n,t} = \varsigma_{i,n} + \beta_i X_{i,n,t} + \omega_i Z_n + \tau_{Ci} \log\left(a^*{}_{i,n}\right) + \varepsilon_{i,n,t}\,, \tag{4.8}$$

where $X_{i,n,t}$ is a vector of attributes of alternative *i* for respondent *n* and choice situation *t* whose impact on utility is measured by β_i, and Z_n is a vector of socio-demographic characteristics of respondent *n* whose impact on utility (which differs across alternatives) is measured by ω_i.

$a^*{}_{i,n}$ is the transformed *latent consideration* variable (which has been bounded between 0 and 1 through a logit transformation to enable the use of a log-transform, Equation 4.9), and its impact on utility is measured by τ_{Ci}. Finally, $\varepsilon_{i,n,t}$ is the typical type I extreme value error term.

$$a^*{}_{Ci,n} = \frac{1}{(1+\exp(-\alpha_{Ci,n}))} \tag{4.9}$$

According to the proposed formulation, when $a^*{}_{i,n}$ is closer to 0, the utility will be heavily discounted, since $\log(a^*{}_{i,n}) \to -\infty$ as $a^*{}_{i,n} \to 0$, and the alternative will be also given lower choice probability. When the alternative is very likely to be considered, and therefore $a^*{}_{i,n}$ approaches 1, no

discounting of utility is enforced. A similar utility-discounting approach has been used by Cascetta and Papola (2001).

We specify a Mixed Multinomial Logit (MMNL) choice model introducing random alternative-specific constants (ASCs) for all but one alternative (i.e. *Car passenger*), $\varsigma_{i,n}$, with mean μ_{ς_i} and standard deviation σ_{ς_i}, such that $\varsigma_{i,n}$ = μ_{ς_i}+ $\sigma_{\varsigma_i}\xi_{i,n}$, where $\xi_{i,n}$follows a standard normal distribution over respondents. The choice probability of the sequence of choices for individual *n* is then defined by (Equation 4.10):

$$P_{i,n}\left(Y_{i,n}|\ a^*{}_{i,n}, X_{i,n,t}, Z_n\ , \varsigma_n\right) = \int_{\varsigma_n} \prod_{t=1}^{T} \frac{e^{U_{i,n,t}}}{\sum_{j\ \in C_n} e^{U_{j,n,t}}}\ f\left(\varsigma_n \middle| \mu_\varsigma, \sigma_\varsigma\right) d\varsigma_n \quad (4.10)$$

where $Y_{i,n}$is the vector of *stated choices*, C_n is the set of available alternatives, and t represents the sequence of observations.

Assuming that the first set of indicators (ordinal) is used, the final LL function for the proposed ICLV model is given by (Equation 4.11):

$$LL = \sum_{n=1}^{N} \ln\left[\left(\int_{a^*{}_{i,n}} \int_{\varsigma_n} \prod_{t=1}^{T} \left(P_{n,t}\left(Y_{n,t}|a^*{}_{i,n}, X_{n,t}, Z_n, \varsigma_n\right) P_{K_{i,n}}\left(I_{K_{i,n}}|\alpha_{i,n}\right)\right) \mathrm{f}\left(\varsigma_n|\mu_\varsigma, \sigma_\varsigma\right) \mathrm{g}\left(\alpha_{i,n}\middle|Z_{Ci,n}\right) d\varsigma_n da^*{}_{i,n}\right)\right] \quad (4.11)$$

The repeated choice nature of the data is taken into account through the use of a panel MMNL and the estimation of robust standard errors. The

models are all estimated using maximum simulated likelihood and 500 Modified Latin Hypercube Sampling draws.[31]

4. Results and discussion

We account for latent consideration for the public transport alternative when accessing Bari International airport. This might be due to possible negative judgements about its reliability, safety concerns, lack of convenience (with respect to the departure location), or comfort, since it involves at least one interchange. Other modes are assumed to be fully considered.

The choices from the RP (i.e. the access mode used by the respondents during their last trip to the airport) and SP data have been jointly estimated. Table 2 presents the results for five alternative model specifications. Model 1 is a MMNL model where all alternatives are fully considered, representing standard practice in the mode choice literature. In the following columns we report the estimation results for the proposed ICLV models, where the indicators vary across models. We first use responses to perception statements and a preference-based ranking of the alternatives (Model 2), followed by an inferred travel time threshold (Model 3), and finally stated consideration (Model 4) as indicators for latent consideration, respectively. Model 5 is a reduced-form MMNL model of the ICLV models in which we do not make use of any indicators. This latter model still includes the

[31] This number of draws resulted in stable models, i.e. by increasing the number of draws we did not observe any improvement in the final LL.

discounting factor (unlike Model 1), which is defined as a function of the same observed explanatory variable used in the structural equation for latent consideration in Models 2-4. The estimation of this reduced-form model is consistent with the discussion on the role of latent variables in Vij and Walker (2016), and it aims at unveiling the actual benefits of using the indicators.

Table 4.2 Estimation results

	Model 1		Model 2		Model 3		Model 4[a]		Model 5	
STRUCTURAL MODEL	Est.	t-stat(0)	Est.	t-stat(0)	Est.	t-stat(0)	Est.	t-stat(0)	Est.	t-stat(0)
γ Student			0.204	3.95	0.213	3.13	*(0.217)*	*(3.40)*	0.368	2.33
MEASUREMENT MODEL										
Preference Ranking										
ζ Latent Consideration PT			1.909	5.55						
μ1 Threshold Ranking			-1.125	-6.84						
μ2 Threshold Ranking			1.419	7.84						
μ3 Threshold Ranking			3.030	9.51						
Perception Statement Frequency										
ζ Latent Consideration PT			1.094	8.28						
μ1 Threshold Statement Frequency			-1.773	-13.61						
μ2 Threshold Statement Frequency			-0.261	-2.84						
μ3 Threshold Statement Frequency			0.990	9.71						
μ4 Threshold Statement Frequency			3.665	16.02						
Perception Statement Reliability										
ζ Latent Consideration PT			0.191	2.19						
μ1 Threshold Statement Reliability			-3.662	-15.83						
μ2 Threshold Statement Reliability			-2.037	-17.85						
μ3 Threshold Statement Reliability			-1.180	-13.71						

μ4 Threshold Statement Reliability			0.817	10.06						
Travel Time Threshold										
ζ Latent Consideration PT					3.018	7.42				
σ Travel Time Threshold					4.952	15.24				
Stated Consideration										
ζ Latent Consideration PT							*(2.795)*	*(11.63)*		
θ Stated Cconsideration PT							*(0.787)*	*(4.19)*		
CHOICE MODEL										
ASC PT	-2.096	-5.57	-0.887	-2.99	-0.858	-2.98	-1.142	-3.60	-1.221	-3.76
ASC Direct Bus	-1.146	-3.08	-0.681	-2.03	-0.721	-2.17	-0.758	-2.25	-0.734	-2.18
ASC Car Driver	-1.168	-3.73	-0.845	-2.97	-0.890	-3.12	-0.938	-3.27	-0.921	-3.27
ASC Taxi	-1.758	-3.41	-1.340	-2.90	-1.272	-2.84	-1.334	-2.97	-1.346	-2.95
ASC PT, sd	0.822	10.33	0.602	8.61	0.566	8.40	0.637	8.55	0.662	8.90
ASC Direct Bus, sd	0.977	9.50	-0.900	-9.16	-0.888	-9.01	-0.904	-9.23	-0.905	-9.07
ASC Car Driver, sd	-1.150	-11.33	-1.013	-10.82	-1.022	-10.77	-1.037	-10.97	-1.022	-11.02
ASC Taxi, sd	-1.455	-7.63	1.287	7.38	1.316	7.52	1.310	7.84	-1.276	-8.74
β In-Vehicle Travel Time PT	-0.011	-5.17	-0.009	-4.48	-0.009	-4.53	-0.009	-4.51	-0.009	-4.52
β In-Vehicle Travel Time Direct Bus	-0.018	-3.51	-0.016	-3.49	-0.015	-3.46	-0.016	-3.61	-0.017	-3.64
β In-Vehicle Travel Time Car Driver	-0.017	-4.15	-0.013	-3.57	-0.013	-3.48	-0.014	-3.65	-0.014	-3.75
β In-Vehicle Travel Time Car Passenger	-0.033	-7.12	-0.025	-6.11	-0.025	-6.25	-0.026	-6.37	-0.026	-6.40
β In-Vehicle Travel Time Taxi	-0.037	-4.12	-0.030	-3.78	-0.033	-3.93	-0.034	-4.11	-0.033	-3.97
β Out-Vehicle Travel Time PT	-0.017	-5.94	-0.014	-5.24	-0.014	-5.22	-0.015	-5.29	-0.015	-5.35
β Headway time	-0.009	-8.71	-0.008	-8.12	-0.008	-8.10	-0.008	-8.17	-0.008	-8.16
β Travel Cost, income yes	-0.080	-11.25	-0.074	-11.39	-0.074	-11.25	-0.074	-11.35	-0.073	-11.31
β Travel Cost, income na	-0.090	-7.95	-0.081	-7.98	-0.083	-7.82	-0.082	-7.86	-0.081	-8.14
β Income Elasticity (Travel Cost)	-0.077	-2.97	-0.063	-2.51	-0.057	-2.23	-0.064	-2.56	-0.062	-2.36

β Travel cost paid employer (multiplier)	0.765	1.49[b]	0.753	-1.78[b]	0.776	-1.61[b]	0.766	-1.64[b]	0.772	-1.64[b]
scale SP	2.152	10.78[b]	2.422	10.24[b]	2.411	10.01[b]	2.399	10.29[b]	2.428	10.42[b]
β Female (Car Driver)	-0.526	-4.42	-0.467	-4.30	-0.456	-4.16	-0.473	-4.26	-0.467	-4.35
β Change Ruvo	0.595	4.89	0.389	3.71	0.605	5.03	0.518	4.65	0.498	4.56
β Business Trip (PT)	-0.211	-1.84	-0.093	-0.94	-0.105	-0.93	-0.142	-1.33	-0.124	-1.22
τ Latent Consideration PT			2.235	8.53	2.483	8.47	1.811	5.76	1.575	4.64
LL(0)	-6630.290		-10070.35		-24969.170		-6630.290 *(-1063.288)*		-6630.290	
LL(final, complete model)	-4676.765		-7595.827		-7003.418		-4660.365 *(-791.121)*		-4662.519	
LL(final, choice model only)	-4676.765		-4662.557		-4663.984		-4660.365 *(-791.121)*		-4662.519	

Note: a) The structural and measurement models in Model 4 have been estimated separately from the choice model, on a sub-sample of 307 respondents. The estimated parameters and the LL values for these models are in parenthesis; the γ parameter in the structural model has been fixed in the choice model, which has been estimated on the full sample of 746 respondents; b) t-stat against 1. Abbreviation: PT = public transport; sd = standard deviation.

In Model 1, the estimates for the normally distributed ASCs reveal a strong preference for the car passenger alternative over all other alternatives, particularly public transport. Standard deviations, which reflect the degree of heterogeneity for the ASCs, are all significant.

Alternative-specific in-vehicle travel time coefficients show the right (negative) sign and are all statistically significant. Similarly, parameters for the out-of-vehicle travel time (which only refers to the public transport alternative) and for the headway time (which refers to both the public transport and the private direct bus alternatives) also show the right (negative) sign and are statistically significant.

Travel cost has been interacted with income in a non-linear way, and we estimated the respective income elasticity. Not all respondents disclosed their income; therefore, we estimated two coefficients for travel cost, one for those who reported this information ('*Travel cost, income yes*'), and one for those who did not ('*Travel cost, income na*'). Both travel cost coefficients have the expected (negative) sign and are statistically significant, where the negative, and significant value for the income elasticity implies that the (absolute) sensitivity to travel cost decreases with increases in income. Respondents whose trip was paid by the employer show a lower sensitivity to travel cost, although the '*Travel cost paid employer*' coefficient - estimated as a multiplier of the overall travel cost coefficient - is not statically different from unity in this model.

Travellers on a business trip are less likely to choose public transport, while female respondents show a negative preference for the car driver alternative. '*Change Ruvo*' accounts for fact that passengers coming from (directed to) Corato (i.e. one of the four cities under investigation) need to transfer in Ruvo railway station rather than in Bari railway station.

Finally, given that we employed both RP and SP data, we also estimated a scale parameter for the SP observations to allow for difference in the variance of the error terms between SP and RP. The utility function can be re-written as (Equation 4.12):

$$U_{i,n,t}{}^{*} = \left(RP_{dummy} + scale_{SP} * \left(1 - RP_{dummy}\right)\right) * U_{i,n,t} \qquad (4.12)$$

Where *RP_dummy* equals 1 for RP observations, and 0 otherwise (i.e. for SP observations). As expected, the scale parameter for SP is greater and statically different from one (which is the RP case).

We now move towards the discussion of the results of Models 2-4, where latent consideration for public transport has been included in the utility. These are all ICLV models which differ by the indicators used to measure consideration. The three components (the structural, the measurement, and the choice sub-models) have been estimated simultaneously in Models 2 and 3, and sequentially in Model 4, since the indicators for stated consideration were available only for approximately 40% of respondents.

In the structural sub-models for Models 2-4, we parametrised the latent consideration as a function of a dummy variable taking the value of one if the respondent was a student, and zero otherwise.[32] Consistent with our expectation, the γ parameters (see Equation 4.1) indicates that the latent consideration for the public transport alternative is higher for students.

In Model 2, three distinct measurement sub-models have been estimated, given that three indicators have been used. Preference ranking was re-scaled on a 4-points scale, since the direct private bus alternative was not available for all routes. Therefore, for this indicators, we only estimated three thresholds (see Equation 4.2). In Model 4, the positive θ parameter (see Equation 4.6) reflects the fact that the stated consideration rates for public transport were larger than 50% in the sub-sample. The response-order for the preference ranking has been shifted, such that the general assumption in all cases is that more positive responses to the indicators are observed when latent consideration increases. As expected, the ζ parameters are all positive.

In the choice sub-models for Models 2-4, the τ parameters measure the marginal impact of latent consideration on the utility for the public transport alternative, which is found in all cases to be statistically significant. This implies that a value for the latent consideration closer to unity (zero) would

[32] The inclusion of other variables such as the frequency of the airport trips, the haul of the flight taken and the nature of the trip has been also tested. However, these have not been retained in the final models because the corresponding parameter was found to be not statistically different from zero.

lead to higher (lower) utility for this alternative. We also observe that the parameters accounting for the likelihood of choosing public transport for respondents on a business trip is no longer significant. This might indicate a possible (negative) correlation with the γ parameters in the structural models, since students are less likely to travel for business purposes.

Model 5 is the reduced-form MMNL model of Models 2-4. In this model we do not estimate any measurement models since we do not make use of any indicators. The latent construct now only explains choices, and, as a result of this, we observe a larger standard error for the γ parameter (structural model) compared to Models 2-4, i.e. there is an efficiency loss.

Interestingly, all parameters in Models 2-5 (except for '*scale SP*') are reduced in size with respect to Model 1. This might be due to the fact that in the former models there is an additional explanatory variable (i.e. the log-discounting factor accounting for consideration of public transport), which takes away explanatory power from the other variables. In turn, the estimated Value of Travel Time (VTT) indicators are also lower than - although not statistically different from - those obtained for Model 1 (Table 4.3).

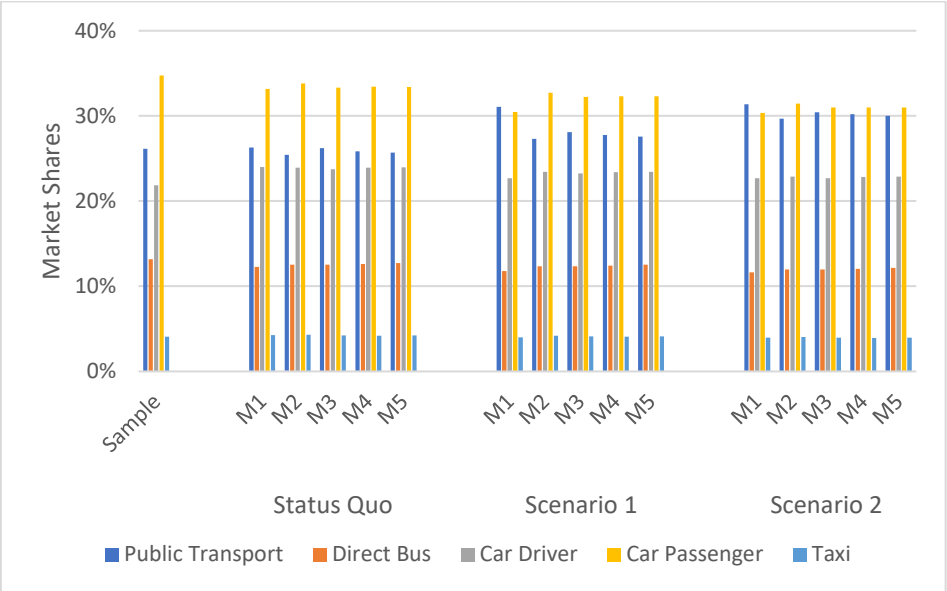
Table 4.3 VTT (€/hour)

	Model 1	**Model 2**		**Model 3**		**Model 4**		**Model 5**	
	Est.	*Est.*	*Change (%)*	*Est.*	*Change (%)*	*Est.*	*Change (%)*	*Est.*	*Change (%)*
Public Transport	8.6	7.2	-16%	7.1	-17%	7.3	-15%	7.4	-14%
Direct Private Bus	13.2	12.8	-3%	12.5	-5%	13.3	1%	13.7	4%
Car Driver	12.8	10.9	-15%	10.5	-18%	11.2	-13%	11.4	-11%
Car Passenger	24.4	20.0	-18%	20.0	-18%	21.3	-13%	21.4	-12%
Taxi	27.5	24.8	-10%	26.6	-4%	27.4	-1%	26.6	-3%

Note: VTT indicators for an individual whose trip was not paid by employer.

Forecasted aggregate market shares are represented in Figure 4.2. As expected, differences between Model 1 and Models 2-5 are negligible in a *status quo* scenario, i.e. when the model is applied to the attribute levels actually faced by the respondents. This is because any model including a full set of alternative-specific constants would perfectly recover the market shares in the data. More pronounced differences in forecasts between Model 1 and Models 2-5 can be observed when instead looking at the effect of a reduction in headway time by 30% for the public transport alternative (Scenario 1). For example, Model 1 predicts a larger increase over the Status Quo for the public transport alternative (+18.2%) and a larger decrease for the other alternatives (e.g. -8.2% for Car passenger) with respect to Models 2 (+7.5% and -3.2%, respectively). A similar pattern is observed when we reduce travel time for the public transport by 30% (Scenario 2). This means that a more traditional MMNL model - which assumes that public transport is considered by everyone in the sample - might overestimate the gains of policy actions aimed at improving modal share for this alternative.

Figure 4.2 Forecasted aggregate market shares



Turning our attention to model fit, we acknowledge that the final Log-Likelihood across Models 1-5 cannot be compared, given that in Models 2 and 3 we actually estimate a joint Likelihood function for the choices and for the indicators. It is however possible to derive final Log-Likelihood measures for the choice model components separately from the other components. A comparison of these measures reveals that Models 2-4 outperform Model 1. However, such improvement in fit cannot be completely ascribed to the use of the indicators (Vij and Walker, 2016), given that it could be also attained by a properly specified reduced-form MMNL model (Model 5).

5. Validation exercise

Model validation on a different sample allows for a more rigorous comparison with respect to final Log-Likelihood measures, ensuring that the

estimation results are not due to overfitting. We estimated the models using the data collected in 2017 (richer in terms of supplementary information used to measure consideration) and to keep the data collected in 2016 for validation purposes. Indeed, the use of the ICLV approach allows for *latent consideration* to be directly predicted from Equation 4.1 without relying on the availability of the indicators in the validation sample. This is because *latent consideration* rather than the indicators is used as explanatory variable in the choice model.

Table 4.4 Probability for the chosen alternative in the validation sample

	Model 1	**Model 2**		**Model 3**		**Model 4**		**Model 5**	
	Prob.	*Prob.*	*Change*	*Prob.*	*Change*	*Prob.*	*Change*	*Prob.*	*Change*
RP+SP	25.2%	25.6%	+0.4	25.7%	+0.5	25.6%	+0.4	25.5%	+0.3
Only RP	25.8%	27.6%	+1.8	28.0%	+2.2	27.5%	+1.7	27.0%	+1.2

The average probability for the chosen alternative is used as measure of fit on the validation sample. As we can see from Table 4.4, the ICLV models (Models 2-4) produce slightly better predictions with respect to both their reduced-form MMNL (Model 5) and the more traditional MMNL model (Model 1) form. Such an improvement is however almost negligible when using both RP and SP observations, meaning that *latent consideration* brings little additional correlation when multiple observations for each respondent are available. Nevertheless, when only RP observations are used we observe a more substantial increase in the probability for the chosen alternative up to 2.2 percentage points.

6. Conclusions

The challenge with consideration of alternatives is that this aspect of an individual's decision-making process is unobservable. When the only information available is that on the final outcome of the process, i.e. the choices individuals make, it is impossible to separately identify what drives 'consideration' and what drives 'choice'.

Supplementary information on aspects related to consideration can be collected during SC experiments; however, the direct use of such indicators as '*error-free*' explanatory variables in the estimation of discrete choice models is highly discouraged due to potential measurement errors, endogeneity bias, and unsuitability of the resulting model for forecasting.

In this paper we overcome these drawbacks by treating indicators for consideration as dependent rather than independent variables, and modelling these together with choice within an ICLV framework. *Latent consideration*, rather than the indicators, enters the utility of an alternative through a 'discounting' factor, which accounts for consideration lowering the utility, and therefore choice probability of that alternative.

The proposed approach is tested in the context of airport access mode decisions for journeys to Bari International Airport, in Italy, using data from a SC experiment on a sample of air travellers resident within the catchment area of the airport. Despite being always available, we assume in this paper that the public transport alternative might not be always considered by those travellers. Three sets of supplementary information directly or indirectly related to consideration of this alternative have been collected during the SC experiment, which are tested as potential indicators for *latent consideration*.

Our results suggest that *latent consideration* has a significant (and positive) marginal effect on the overall utility of public transport; this means that the utility for those respondents with predicted lower levels of *latent consideration* gets highly discounted, and their choice probability for this alternative approaches zero. However, since we also use revealed preferences data in the estimation, it is possible that, for those observations, a share of what we identify as the effect of *latent consideration* might also capture unawareness or unavailability effects.

We additionally observe a decrease in the size of the estimates for key parameters (travel time, travel cost, headway time) relative to a more traditional MMNL model which assumes that all alternatives are considered; in turn, this affects willingness-to-pay indicators and most importantly forecasts for aggregate market shares. Interesting is the case of headway time. If this is reduced by 30% for public transport, a traditional MMNL would predict an increase in the modal share for this mode by 18.2%, while the proposed models accounting for consideration would still predict an increase, but only by 7.5%. This result would suggest that not accounting for consideration of the alternatives might have serious implications in predicting the effect of planned or expected changes in the airport ground transportation system such as the introduction or removal of a key access mode, or a change in the quality of its services. Of course, given that the true data generating process is unknown, it is impossible to identify the size and direction of a 'possible' bias. All that we can observe is the difference with a more traditional MMNL model.

In general, accounting for consideration of public transport seems to provide a more realistic representation of airport access mode decisions with respect to a more traditional MMNL model, where all alternatives are assumed to be considered: this is shown through an improvement in model fit which also holds on a separate validation sample. However, consistent with the discussion in Vij and Walker (2016), we acknowledge that this improvement cannot be completely ascribed to the use of the indicators,

since a properly specified reduced-form MMNL model is able to attain very similar results. Nevertheless, the availability of indicators - similar to those used in this paper, which can be easily included when designing an air passenger survey - allows us to identify the structural drivers of consideration, in this particular case that students are more likely to consider public transport as a feasible access mode.

As with any paper, there are many areas for future research. In the specific case of airport access mode decisions, it would be interesting to test, also in the case of SC surveys, whether the consideration for public transport alternatives varies by time of day (e.g. becoming less likely in the early mornings and late evenings) as well as to attain a better understanding the drivers of consideration in general, e.g. as a result of luggage, who people were travelling with, or even destination of the flight.

Author contribution statement

The authors confirm contribution to the paper as follows: study conception and design: Capurso, M., Dekker, T., Hess, S.; data collection: Bergantino, A.S., Capurso, M.; analysis and interpretation of results: Capurso, M, Dekker, T., Hess, S.; draft manuscript preparation: Capurso, M., Dekker, T., Hess, S. All authors reviewed the results and approved the final version of the manuscript.

Acknowledgements

Stephane Hess acknowledges financial support by the European Research Council through the consolidator grant 615596-DECISIONS. All authors acknowledge Angela Stefania Bergantino for making the data available for the purpose of this study.

References

Başar, G. and Bhat, C. (2004) A parameterized consideration set model for airport choice: an application to the San Francisco bay area. *Transportation Research Part B: Methodological*, 38(10), 889-904, https://doi.org/10.1016/j.trb.2004.01.001.

Ben-Akiva, M. and Boccara, B. (1995) Discrete choice models with latent choice sets. *International Journal of Research in Marketing*, 12(1), 9-24, https://doi.org/10.1016/0167-8116(95)00002-J.

Bolduc, D., Ben-Akiva, M., Walker, J. and Michaud, A. (2005) Hybrid choice models with logit kernel: applicability to large scale models. In *Integrated land-use and transportation models: behavioural foundations*, edited by M. E. H. Lee-Gosselin and S. T. Doherty, Emerald Group Publishing Limited.

Cantillo, V. and Ortúzar, J. D. D. (2005) A semi-compensatory discrete choice model with explicit attribute thresholds of perception. *Transportation Research Part B: Methodological*, 39(7), 641-657, https://doi.org/10.1016/j.trb.2004.08.002.

Cascetta, E. and Papola, A. (2001) Random utility models with implicit availability/perception of choice alternatives for the simulation of travel demand. *Transportation Research Part C: Emerging Technologies*, 9(4), 249-263, https://doi.org/10.1016/S0968-090X(00)00036-X.

Coogan, M. A. (2008) ACRP Report 4: Ground access to major airports by public transportation. Transportation Research Board, Washington, D.C., https://dx.doi.org/10.17226/13918.

ENAC - National Civil Aviation Authority (2010) Studio sullo sviluppo degli aeroporti italiani (in Italian), available at https://www.enac.gov.it/pubblicazioni/studio-sugli-aeroporti-nazionali/studio-sullo-sviluppo-degli-aeroporti-italiani. Accessed Nov. 14 2018.

Eurostat. Transport statistics. https://ec.europa.eu/eurostat/web/transport/data/main-tables. Accessed Jul. 27 2018.

Frejinger, E., Bierlaire, M., and Ben-Akiva, M. (2009) Sampling of alternatives for route choice modelling. *Transportation Research Part B: Methodological*, 43(10), 984-994, https://doi.org/10.1016/j.trb.2009.03.001.

Gupta, S., Vovsha, P., and Donnelly, R. (2008) Air passenger preferences for choice of airport and ground access mode in the New York City metropolitan region. *Transportation Research Record*, 2042, 3-11, https://doi.org/10.3141/2042-01.

Harvey, G. (1986) Study of airport access mode choice. *Journal of transportation Engineering*, 112(5), 525-545, https://doi.org/10.1061/(ASCE)0733-947X(1986)112:5(525) .

Hensher, D. A. and Ho, C. (2015) The role of perceived acceptability of alternatives in identifying and assessing choice set processing strategies in stated choice settings: the case of road pricing reform. *Transportation Research Part E: Logistics and Transportation Review*, 83, 225-237, https://doi.org/10.1016/j.tre.2015.09.012.

Hess, S. and Hensher, D. A. (2013) Making use of respondent reported processing information to understand attribute importance: a latent variable scaling approach. *Transportation*, 40(2), 397-412, https://doi.org/10.1007/s11116-012-9420-y.

Hess, S. and Polak, J. W. (2006) Exploring the potential for cross-nesting structures in airport-choice analysis: a case-study of the greater london area. *Transportation Research Part E: Logistics and Transportation Review*, 42(2), 63-81, https://doi.org/10.1016/j.tre.2005.09.001.

Isfort (2017) Annual report of the Audimob survey on the mobility in Italy (In Italian).

ISTAT - National Institute of Statistics, Popolazione Residente (in Italian), available at https://demo.istat.it/pop2017/index.html. Accessed Nov. 14 2018.

Kløjgaard, M. E. and Hess, S. (2014) Understanding the formation and influence of attitudes in patients' treatment choices for lower back pain:

testing the benefits of a hybrid choice model approach. *Social Science & Medicine*, 114, 138-150, https://doi.org/10.1016/j.socscimed.2014.05.058.

Jou, R. C., Hensher, D. A. and Hsu, T. L. (2011) Airport ground access mode choice behaviour after the introduction of a new mode: a case study of Taoyuan international airport in Taiwan. *Transportation Research Part E: Logistics and Transportation Review*, 47(3), 371-381, https://doi.org/10.1016/j.tre.2010.11.008.

Leigh Fisher Associates, Coogan, M.A. and MarketSense (2000) TCRP Report 62: Improving public transportation access to large airports. Transportation Research Board, Washington, D.C., https://onlinepubs.trb.org/onlinepubs/tcrp/tcrp_rpt_62-a.pdf.

Leigh Fisher Associates, Coogan, M.A. and MarketSense (2002) TCRP Report 83: Strategies for improving public transportation access to large airports. Transportation Research Board, Washington, D.C., https://onlinepubs.trb.org/onlinepubs/tcrp/tcrp_rpt_83a.pdfs.

McFadden, D. (1974) Conditional logit analysis of qualitative choice behaviour. In *Frontiers in econometrics*, edited by P. Zarembka. Academic Press, New York.

Monteiro, A. B. and Hansen, M. (1996) Improvements to airport ground access and behaviour of multiple airport system: BART extension to San Francisco international airport. *Transportation Research Record*, 1562, 38-47, https://doi.org/10.1177/0361198196156200105.

Pels, E., Nijkamp, P. and Rietveld, P. (2003) Access to and competition

between airports: a case study for the San Francisco bay area. *Transportation Research Part A: Policy and Practice*, 37(1), 71-83, https://doi.org/10.1016/S0965-8564(02)00007-1.

Rose, J. M., Bliemer, M. C., Hensher, D. A. and Collins, A. T. (2008) Designing efficient stated choice experiments in the presence of reference alternatives. *Transportation Research Part B: Methodological*, 42(4), 395-406, https://doi.org/10.1016/j.trb.2007.09.002.

Swait, J. A (2001) Non-compensatory choice model incorporating attribute cutoffs. *Transportation Research Part B: Methodological*. 35(10), 903-928, https://doi.org/10.1016/S0191-2615(00)00030-8.

Swait, J. and Ben-Akiva, M. (1987) Empirical test of a constrained choice discrete model: mode choice in Sao Paulo, Brazil. *Transportation Research Part B: Methodological*, 21(2), 103-115, https://doi.org/10.1016/0191-2615(87)90010-5.

Tam, M. L., Lam, W. H. and Lo, H. P. (2008) Modeling air passenger travel behaviour on airport ground access mode choices. *Transportmetrica*, 4(2), 135-153, https://doi.org/10.1080/18128600808685685.

Vij, A. and Walker, J. (2016) How, when and why integrated choice and latent variable models are latently useful. *Transportation Research Part B: Methodological*, 90, 192-217, https://doi.org/10.1016/j.trb.2016.04.021.

Chapter 5 - Discussion and conclusions

1. Discussion

This thesis contributes to the ongoing discussion on the role of consideration of alternatives in the decision-making process of individuals, and on the impact of accounting for this aspect in the estimation of RUM-based discrete choice models. Although calibrated to specific choice contexts (i.e. choice of the transport mode for long-distance and airport access trips), the methodology proposed (and, to a limited extent, the findings of this research) can be applied (generalised) to other choice contexts with similar characteristics such as number of alternatives and degree of differentiation amongst the alternatives.

Three papers have been developed as 'stand-alone' contributions to the literature; nevertheless, each contributes in its own way to answer the three research questions set out in the first chapter of this thesis. The first two research questions, namely '*how to measure consideration*' and '*what drives consideration*' are directly related, as are therefore their answers. Since consideration of alternatives is unobserved by the analyst, the suitability of either direct (stated consideration) and indirect (stated thresholds for attributes, perceptions towards the alternatives, preference-based ranking of the alternatives, previous choices) indicators has been tested in the three papers. In particular, since similar ICLV models have been employed in the papers presented in Chapters 3 and 4, it was possible to test the robustness of the proposed approach to a wide range of indicators of a different nature.

Ceteris paribus, it emerges that thresholds for the travel time attribute (either elicited from the individuals or inferred from their previous choices) are the most informative indicators of consideration[33]: thresholds for travel time are able to summarise into a single unit of merit most of the information required to measure consideration for the *slower* alternatives (i.e. bus, IC, and car-pooling in the first paper and public transport in the third paper). In particular, while the third paper makes the assumption that consideration of an alternative is driven by the thresholds *per se*, i.e. that individuals might exclude alternatives from consideration on the basis of their *a priori* believes, the first paper assumes that individuals always compare their self-imposed threshold for an attribute with the actual level of that attribute, i.e. that consideration of an alternative is context (i.e. choice-task) specific. Both assumptions can contribute to further explaining heterogeneity in the individual behaviour, but sometimes data does not allow comparing latent thresholds against the attribute levels. Even when this would be possible, the fact that attribute levels vary at the task level makes strict comparisons more complicated, leading to empirical identification problems.

Thresholds for attributes can be considered a valid indicator for consideration particularly when the alternatives can be easily categorised based on the fulfilment (or not) of observable requirements (e.g. having

[33] The other indicators tested in the thesis were: stated consideration, the ranking of the alternatives based on overall preference, the level of agreement with perception statement towards the alternatives and thresholds inferred from previous choices.

travel time below or above 4 hours, as in the Rome-Milan example). As the alternatives can be grouped into separate categories, it can be reasonable to have separate groups of individuals with similar consideration sets.

However, thresholds alone might not be able to explain, for example, why, for two alternatives with the same travel time, one is considered and the other not. This is the type of information that is instead provided by direct indicators of consideration, i.e. stated consideration. Nevertheless, self-reports of consideration may not contain a different set of information on individuals' preferences with respect to that contained in the observed choice (Horowitz and Louviere, 1995). Despite being the most straightforward indicators of consideration, this feature makes stated consideration (but also the remaining indirect indicators tested in this research, i.e. perceptions towards the alternatives and preference-based ranking of the alternatives) to contribute less in terms of additional informational content with respect to thresholds. This appears even more evident when these indicators are further explained using the same (or a similar) set of explanatory variables in the utility functions for the corresponding alternatives.

All these considerations motivated the simultaneous use of both direct and indirect indicators, rather than just one or the other. In the second paper, a hierarchical relation between the stated threshold for travel time and stated consideration of alternatives has been proposed, i.e. the former is used to 'explain' the latter. By doing this, the analyst can fully exploit the

informational content of both indicators, and hence obtain a more robust (i.e. precise) measurement of consideration.

2. Main contributions of the research

To sum up with respect to the first research question presented in the first chapter of the thesis, i.e. "*how can we measure consideration"*, we can conclude that each indicator provides a different set of information. Thresholds for attributes are certainly more informative than stated consideration of alternatives, and their collection "costs" are smaller in terms of additional burden on respondents; however, these are not able to tell us which alternatives are considered or not amongst those with similar attribute levels. Therefore, it goes without saying that collecting both direct and indirect indicators of consideration would be a "first best" solution; however, this might not be always possible, since all these questions might exponentially increase the "costs" in terms additional burden to the respondents of SC surveys.

To sum up with respect to the second research question, i.e. *"what drives consideration",* we show that thresholds have a primary role in explaining consideration, and that thresholds themselves are a function of respondents' socio-economic profiles. Over the three papers we see that the availability of good (although still imperfect) indicators of consideration makes it 'easier' to empirically separate consideration of alternatives and choice. Being able to identify what exclusively drive consideration does not mean that 'consideration of alternatives' and 'choice' are independent

processes, since their drivers are likely to be highly correlated. However, this would provide efficiency gains in modelling, allowing the analyst to use the same variables to explain both processes, rather than only in one or the other. For example, we observe this in the second paper with respect to the dummy accounting for 'business' trips, which is simultaneously included in the structural model for the latent threshold for travel time and in the choice model. Ideally, there might be also variables which can explain consideration but not choice, such as 'being a student' in the third paper. From an industry perspective, being able to identify those characteristics can provide useful insights at the early product development stage and/or, later on, in the evaluation of a more appropriate product positioning.

With respect to the third research question, i.e. '*what role do consideration effects play in the estimation of RUM-based discrete choice models*', this research mainly focused on four aspects, namely i) sources of heterogeneity, ii) model fit, iii) willingness-to-pay (WTP) indicators, and iv) forecasts. The first and second papers suggest that as soon as the analyst incorporates in the models elements of the decision-making process (in the form of consideration effects) alongside preference heterogeneity, the role of the latter reduces. Having said that, if the analyst only accounts for unobserved preference heterogeneity, s/he would neglect the fact that at least a share of the observed heterogeneity in the data could be actually explained by the fact that individuals do not process the available information in the same way, e.g. that some of them might actually ignore certain alternatives, and

therefore have different consideration sets. Accounting for consideration of alternatives therefore provides a deeper understanding of the decision process by opening the 'black box'. On the contrary, if the analyst only accounts for consideration effects and not preference heterogeneity (i.e. s/he estimates a MNL choice model) there is an actual risk of overstating the role of former and vice versa.

The latter reflection could also explain the substantial improvements in model fit seen in previous research as a result of accounting exclusively for consideration effects. This research shows that there may be an improvement in fit, but this could be a rather 'mild' one. Interestingly, a validation exercise on a different sample in the third paper suggests that when the validation sample is made of RP rather than SP observations, the improvement in fit could be slightly larger. This could be explained by the fact that a share of what was identified as consideration effect, could be ascribed to unawareness or unavailability effects. The latter effects are supposed to have a role in the estimation of discrete choice models only when RP observations are used; here the analyst does not typically have information on which alternatives were effectively available to the individuals, and which alternatives the individuals were aware of when they made their choice. This lack of information is generally reduced with SP data, since only objectively available alternatives are typically presented during SC experiments, and individuals are somehow made aware of these alternatives through their presentation in the choice cards.

Nevertheless, the second and third paper also show that when the analyst estimates a MMNL that does not make use of the indicators of consideration, but still includes all variables (e.g. socio-economic variables) used to explain the indicators of consideration (through the so-called 'structural' models), even the aforementioned additional mild improvement in fit vanishes. Results from these papers reinforce the findings of Vij and Walker (2016) on the actual econometric benefits of the ICLV model, making this research also contribute to the literature on latent variables. As a result, if an exhaustive set of explanatory variables is used, one might account for consideration effects even without indicators. However, it would be possible to disentangle the heterogeneity which is due to consideration effects from that which is due to other unobserved effects only when supplementary information in the form of indicators is available.

With respect to possible impacts on WTP, results from the all three papers confirmed that when the analyst accounts for consideration effects, the size of WTP measures is affected for the unconsidered alternatives. When consideration of alternatives is measured using the difference between the threshold for travel time and the actual level for this attribute (e.g. as in Model 3 in the first paper), the WTP indicator is statistically different from that obtained from a fully compensatory MMNL model. This is due to the fact that accounting for consideration effects in this way would affect the size of the parameters which are directly involved, such as travel time in this case.

Finally, the first and third paper also investigated the impact of accounting for consideration effects in forecasting. When the model is applied to the attribute levels faced by the respondents (i.e. in a *status quo* scenario) treating consideration of certain alternatives probabilistically increases the choice probabilities of the remaining alternatives. This was an expected result, even though the differences with a MMNL model not accounting for consideration effects appear negligible. This would suggest that, in terms of modelling, a fully compensatory MMNL model might be flexible enough to accommodate any source of heterogeneity in the data, in line with McFadden and Train (2000). This does not rule out other potential benefits of accounting for consideration effects. Models accounting for these effects would keep providing a more realistic description of individuals' behaviour. More noticeable differences with traditional MMNL models can be observed when the models are used to forecast different scenarios; however, this only applies when testing for changes on attributes which explain both consideration and choice (e.g. travel time). Indeed, in the first paper we observe differences only when testing for the impact of a reduction in travel time for the bus - which is deemed to be alternatively considered - and not when a similar change applies to the high-speed rail - which is instead assumed to be always considered by all individuals in the sample.

3. Practical implications of the research

This research does not come out with a "yes/no" answer to the question "*do we need to account for consideration effects?*", but rather with some stylised facts.

As generally reported by the previous literature on consideration, demand models not accounting for consideration effects may provide biased parameter estimates and forecasts. However, it is not possible to identify the direction and size of the 'eventual' bias with a model not accounting for consideration effects since the true data generating process is unknown. In this research we indeed observe differences between models accounting and not accounting for consideration effect. In particular, results from the papers presented in Chapter 2 and 4 suggest that WTP estimates might be statistically different under the two approaches and we might obtain improved out-of-sample predictions if we account for consideration effects. We also observe that the importance of accounting for consideration effects might depend on the nature of the data used for the analysis. With RP data, not only consideration, but also awareness and availability of the alternatives are in most cases unobserved to the analyst. However, we should recognise that it is not possible to separately identify the effect of all those stages, and their confounding with individuals heterogeneous preferences.

When forecasting future scenarios, differences between models that do and do not account for consideration effects seem larger when testing for

changes in the attributes' levels of alternatives deemed to be probabilistically considered, compared with other alternatives.

Finally, it seems worth to recognise that the identification of the driving factors of consideration might provide managers with useful insights on which markets to target to increase the probability of specific alternatives (e.g. transport modes or other products/services) to be first considered and then eventually chosen.

4. Limitations of the research

Limitations of this research could be grouped in two main categories: the first refers to the data available and choice contexts investigated, hence to the temporal and spatial transferability of the results; the second refers to the models presented and their estimation.

With respect to the first category, since SP - but also RP in this case - are reflections of preferences at a given moment in time, it is not possible to study temporal stability of preferences. This would have been possible only with RP panel data.

Then, the decision of focusing on specific decision contexts (namely mode choice for long distance and airport access) does not even allow the spatial transferability of the results and of the insights about the measurement of consideration, but only of the methodology proposed to model it. Consideration sets are strongly context dependent, and in other decision contexts (or even when the nature of the trip over the same O/D is different), the role of each of the drivers of consideration we investigated

(i.e. the degree with which each of these drivers is likely to lead to the creation of consideration sets) might be completely different. For example, in the case of more frequent/habitual trips (e.g. commuting to work), it might be possible that, rather than by the presence of thresholds for attributes, consideration of alternatives could be more closely measured by *a priori* judgements with respect to the alternatives, including any previous (bad) experience with certain transport modes. Inertia might also lead individuals to only consider one alternative, i.e. the one that they have always chosen.

Another limitation with the data regards the size of the samples and the possible lack of representativeness. The first dataset is particularly small (6 SP choice tasks for 209 respondents) and it has been collected for the exclusive purpose of this research without any financial support. The second dataset is larger in size (5 SP choice tasks and 1 RP for 746 respondents) and I once again acknowledge Angela Stefania Bergantino and the Department of Economics, Management, and Business Law of the University of Bari for making it available. In both datasets, less than 20% of respondents is aged 50 and over, compared with the Italian and Apulian populations in which this percentage is larger than 50%. However, the extent to which this represents a limitation cannot be defined with certainty, since official statistics on the exact target populations are not available. In any case, while this aspect might potentially affect some of our modelling results (e.g. those related to the value of travel time, which therefore must be interpreted with caution),

it is not supposed to invalidate the answers to the research questions presented in this thesis. Indeed, this thesis aims at suggesting methodologies with respect to the measurement and modelling of consideration of the alternatives and not at providing policy measures related to specific choice contexts.

With respect to the second category, which includes the limitations of this research in terms of the models presented and their estimation, it should be acknowledged that the estimation time for models accounting for consideration effects is substantially larger (up to 5 times on standard PCs, up to 3 on super computer) than that of more traditional models. This explains why the presented models do not include additional error components to capture correlation among alternatives and/or among consideration of alternatives, which would have increased the estimation time even more. It also explains why the number of alternatives assumed to be probabilistically considered was limited, and why we only decided to make consideration of alternatives to be driven by thresholds for a single attribute. The decision of which alternatives to make probabilistically considered and on which attribute to take into account when modelling consideration was taken on the basis of preliminary modelling in which we separately evaluated from which of these it was possible to exploit the largest informational content.

5. Guidelines based on this research

The guidelines for other researchers based on the experience gained with this research can be grouped into three categories: the first refers to the data used; the second refers to the measurement of consideration; the third category refers to the estimation of complex models.

With respect to the first category, it would be preferable, whether possible, to replicate the analysis, i.e. to apply the proposed methodology, to more than one dataset. The availability of more than one dataset, containing not only data of different nature (e.g. RP vs. SP), but also on different decision contexts (e.g. frequent vs. infrequent choices, long-distance vs. urban commuting) certainly increases the robustness of the results and of the insights shared with the research community.

With respect to the second category, it is not possible to *a priori* establish which indicator is the most suited to measure unobserved attributes such as consideration of alternative. It would be preferable to compare different indicators and then choose the one that minimizes the burden on respondents while providing the largest informational content. In the particular case of consideration of alternatives in relatively infrequent decision contexts – such as those investigated in this research – we compared a series of direct (stated consideration) and indirect (level of agreement with various perception statements and of a preference-based ranking of the alternatives; thresholds for attributes inferred from respondents' previous choices or stated by respondents) indicators of

consideration. From this comparison, it emerged that it might be preferable to collect information on threshold for attributes, which have the above mentioned characteristics.

With respect to the third category, my recommendation would be to first try and test simpler models before estimating more complicated models. The cost of estimating complicated models (e.g. in terms of additional information needed or in terms of estimation time) must be always compared with the benefits obtainable (e.g. increase in model fit).

6. Conclusions

To sum up, two main messages emerge from this research. First, with respect to the measurement of consideration of alternatives, as soon as individuals effectively ignore certain alternatives amongst those available (i.e. presented to them during SC experiments), either direct and/or indirect indicators can be used to measure this unobserved aspect of their decision-making process. The choice between direct or indirect indicators (or both) will be dictated by the characteristics of the alternatives and/or by the available data, i.e. by how much additional burden can be put on the respondents of the surveys. Second, consideration effects should be considered alongside unobserved preference heterogeneity: by doing this, the analyst will avoid the risk of (wrongly) putting too much emphasis on the role of the former or the latter. Moreover, this will allow us to explain a share of otherwise purely random heterogeneity as consideration effects. Accounting for consideration of alternatives might also affect the estimates

of the parameters, and therefore WTP indicators. This, however, does not mean that traditional models not accounting for consideration effects would provide 'biased' parameters (as suggested by previous research), given that the true data generating process is unobserved.

7. Avenues for future research

To conclude, consideration of alternatives requires further investigation. There might be at least five main potential avenues for future research on this topic. The first one could be related to the design of SC experiments. Although SC experiments are designed to be optimal in a statistical sense, this does not necessarily mean they are optimal also in a behavioural sense; indeed, these do not account for consideration effects. The most naïve approach would be to run an explorative wave of surveys to understand, for a given set of objectively available alternatives and of potential respondents, which alternatives are considered (and by whom) and which not. Then, a traditional SC experiment could be produced where only a subset of alternatives are presented to respondents with certain characteristics (as unveiled in the explorative survey). Although tempting, such an approach - which would incorporate some of the uncertainty on the composition of the consideration set at the design stage - is however at risk of introducing endogeneity bias (c.f. a similar discussion in the context of adaptive SP, see Cherchi and Hensher, 2015). Alternatively, one might still present to respondents the most complete choice set and allow them to 'filter' a subset of alternatives for further consideration. Such an approach would closely

mimic respondents' behaviour in the real world when facing many alternatives, e.g. when booking flights online (Collins et al., 2012). The by-product information on the filtered-unfiltered alternatives could be then probabilistically (rather than deterministically) used to model - and possibly predict - consideration of alternatives.

A second avenue for future research could be an investigation of the role that consideration of alternatives would have in the estimation of non-RUM models, such as elimination-by-aspects (EBA) or random regret minimisation (RRM), even though we acknowledge that only RUM can be properly used for welfare analysis. Despite EBA representing a process fundamentally different from RUM, it looks consistent with the existence of a consideration stage in which selection/grouping of alternatives is driven by the presence (or not) of certain characteristics, and (or) even on the importance of attribute thresholds (Manrai, 1995). It would be interesting to understand if EBA is a reduced form model that needs to make use of indicators - similar to those used in this thesis - to better understand the whole decision process. With respect to RRM, the fact that it considers non-linearities and relative performance of attributes might make the identification of consideration effects challenging. This is best illustrated by the approach employed in Chapters 3 and 4, where consideration of alternatives is directly included in the utility function through non-linear discounting functions.

A third avenue for future research could be an investigation of the role of consideration of alternatives in contexts with similar characteristics such as number of alternatives and degree of differentiation amongst the alternatives. With respect to the number of alternatives, while the methodology proposed in Chapter 2 would require it to be fairly limited – since the analyst has to enumerate all possible consideration sets – the one proposed in Chapters 3 and 4 would be less restrictive – and potentially allow an extension to choice contexts with a large number of alternatives (e.g. residential choices). With respect to the degree of differentiation amongst the alternatives, instead, the proposed methods would be more easily applied when it is possible to group the alternatives according to a well-defined observable "core" characteristic. Examples of possible applications in the transport sector are mode choice on either systematic or non-systematic urban trips (for which consideration of alternatives might be driven by travel time or distance), car purchasing decisions (where consideration might be driven by size or by car manufacturer), and choice of a flight on a long-distance trip (a classical situation in which we apply filters to narrow down the number of alternatives). The proposed models could be potentially applied also in other fields. In our everyday life, we often face the choice amongst a wide range of goods and services that could be grouped into "high-end" and "low-end" solutions, suggesting that the monetary cost could in many cases be a driver for consideration due to differences in personal budget constraints. However, depending on the nature of the

goods/services, there might be other characteristics, not necessarily correlated to the monetary cost, that might contribute to narrow down the set of alternatives to consider. For example, in food purchasing decisions consideration of alternatives might be driven by the nature of the ingredients (e.g. organic, traditional, GMO), while for wine one discriminant might be the country of origin. In the case of decisions related to cosmetics or fragrances, individuals might only consider those exclusively made with natural ingredients. In all those cases, an additional driver for consideration would be the brand, with individuals can be ideally grouped into those who are loyal to a specific brand and those who are open to consider different ones at each choice occasion, e.g. depending on promotions.

A fourth avenue for future research could be related to the use of innovative data sources to measure this unobserved aspect of individuals' decision-making process. For example, it might be interesting to use eye-tracking information - which are also collected as a part of SC experiments in laboratory environments - to understand if additional insights on consideration and choice could be inferred from such "revealed" rather than "stated" indicators. Also, it would be interesting to understand if the so-called "big data" sources could be also used to measure consideration. These could be in the form of a continuous flow of information on our behaviours (e.g. the places we visit, the products we buy, the websites we surf, the social interactions we maintain) rather than information on the outside world (e.g. real time information on traffic flows, train/bus delays, geo-referenced data,

weather conditions). While the latter set of information could of some help in identifying actual consideration sets (see, for example, Calastri et al., 2019), the risk with the information on our behaviours is that these might be too much related with preferences, and therefore that it would not be possible to separately identify consideration effects.

Finally, a fifth avenue for future research could be related to the incorporation in consideration models of other *latent* attributes, such as attitudes specifically related to the choice context under investigation. For example, with respect with transport mode choices, it would be interesting to see if attitudes towards the environment drive consideration of the alternatives. Individuals, unfortunately still only in few countries, are increasingly conscious that they can have an active "role" in slowing down the effects of climate changes, for example by choosing more environmentally-friendly transport modes. In Sweden, as a consequence of the protests run by Greta Thunberg, demand for air trips has drastically reduced while that for train has increased (Open Online, 2019). In order to measure and then model environmental attitudes, however, it would be necessary to include specific additional questions in the surveys, in the form, for example, of the level of agreement/disagreement with environmental statements.

References

Calastri, C., Hess, S. Choudhury, C. F., Daly, A., and Gabrielli, L. (2019) Mode choice with latent availability and consideration: Theory and a case

study. *Transportation Research Part B: Methodological* 123, 374-385, https://doi.org/10.1016/j.trb.2017.06.016.

Cherchi, E., and Hensher, D.A. (2015) Stated preference surveys and experimental design: an audit of the journey so far and future research perspectives, *Transportation Research Procedia*, 11, 154-164.

Collins, A.T., Rose, J.M., and Hess, S. (2012) Interactive stated choice surveys: s study of air travel behaviour, *Transportation*, 39(1), 55-79, https://doi.org/10.1007/s11116-011-9327-z.

Horowitz, J. L., and Louviere, J. J. (1995) What is the role of consideration sets in choice modeling? *International Journal of Research in Marketing*, 12(1), 39-54, https://doi.org/10.1016/0167-8116(95)00004-L.

McFadden, D. and Train, K. (2000) Mixed MNL models for discrete response*, Journal of Applied Econometrics*, 15(5), 447-470, https://doi.org/10.1002/1099-1255(200009/10)15:5<447::AID-JAE570>3.0.CO;2-1.

Open Online (2019) Gli svedesi si vergognano di prendere l'aereo (in Italian), https://www.open.online/2019/04/23/effetto-thunberg-gli-svedesi-si-vergognano-di-prendere-laereo/.

Vij, A. and Walker, J. (2016) How, when and why integrated choice and latent variable models are latently useful. *Transportation Research Part B: Methodological*, 90, 192-217, https://doi.org/10.1016/j.trb.2016.04.021.

Appendix A

Income information in the first dataset was collected using income classes. In the paper presented in the second chapter, we used class-midpoints to compute both income and average income for those respondents who stated the income class they belonged to. A separate travel cost coefficient was estimated for those respondents who preferred not to disclose this information.

We also accounted for who paid the trip, choosing those who paid themselves as baseline.

The specification for the travel cost coefficient is the following:

$$\beta_{travel\_cost_n} = \left[\left(\beta_{travel\_cost\_n} \cdot \left(\left(\frac{income_n}{average\_income} \right)^{lambda\_income_n} \right) \cdot income\_yes\_dummy_n + \beta_{travel\_cost\_income\_na_n} \cdot (1 - income\_yes\_dummy_n) \right) + \beta_{paid\_employer\_or\_family_n} \cdot paid\_employer\_family\_dummy_n \right]$$

Appendix B

This first SC survey was administered in April and May 2016 to a sample of 209 travellers on the Rome-Milan corridor. Respondents were recruited in-person in train stations in Rome and Milan (Rome Termini, Rome Tiburtina, Milan Centrale), bus terminals (Rome Tiburtina, Milan Lampugnano), and airports (Rome Fiumicino, Milan Linate). A small portion of respondents was recruited in service stations on the highway, located around half way between Rome and Milan (in the proximity of Bologna), and online. The experiment was administered by myself using a tablet and the online survey platform Qualtrics.

Prior to designing the survey, a Monte Carlo analysis was has been conducted, in order to better understand the data requirements and to explore modelling opportunities. In particular, a dataset containing 1000 cross-section observations was created. The number of alternatives, and the attributes of the alternatives were chosen according to the current situation on the corridor.

Choice tasks (6 for each respondent) were designed using a Bayesian D-efficient experimental design and the software NGene. To this extent, priors have been imputed, based on expectations on the possible outcomes of the estimates relative to travel time and travel cost. Soft attributes (Wi-Fi and ticket flexibility) have instead been modelled as "*effects*". Constraints were also included as to avoid clearly dominated choices using the command "*reject*". In particular, scenarios in which the travel time for the bus

alternative was lower than the one for the traditional train (IC) and, those in which the travel cost for the former was higher than for the latter were rejected. Moreover, travel cost for faster alternatives (HSR, FSC, and LCC) could not be higher than for slower ones (IC, bus, car-pooling, and private car). Several designs were evaluated aiming at minimizing the D-error criterion (Rose et al., 2008).

Unfortunately, it was not possible to pivot the attribute levels in the SC experiment around those of the respondents' last trip due to limitations with the online survey platform used to administer the survey. Similarly, information on access-egress time (and cost) was also collected but not used to represent individual specific scenarios.

A copy of the survey was submitted to the University of Leeds Ethical Review Committee, and it received the ethics clearance. Prior to start the data collection, a legal authorization to conduct surveys at the platforms in the Rome and Milan rail stations and in two service stations on along the A1 highway was granted by the Italian Railway Infrastructure Manager (RFI) and by the ownership of the service stations, respectively.

Questionnaire

You are being invited to participate in a research study titled "A mode choice study on the Rome - Milan corridor". This study is being done by Mauro Capurso, prof. Stephane Hess, and Dr. Thijs Dekker from the University of Leeds.

The purpose of this research study is to analyse the drivers of the choice of the mean of transport for passengers on this corridor, and will take you approximately 10 minutes to complete. Your participation in this study is entirely voluntary and you will be able to withdraw from the survey at any time. You do not have to answer any questions you do not want to.

We believe there are no known risks associated with this research study, which does not have any commercial purpose; however, as with any online related activity the risk of a breach is always possible. To the best of our ability your answers in this study will remain confidential.

We will minimise any risks by storing the data collected in secure places, and we will use it only for statistical purposes.

1 - Location of the interview

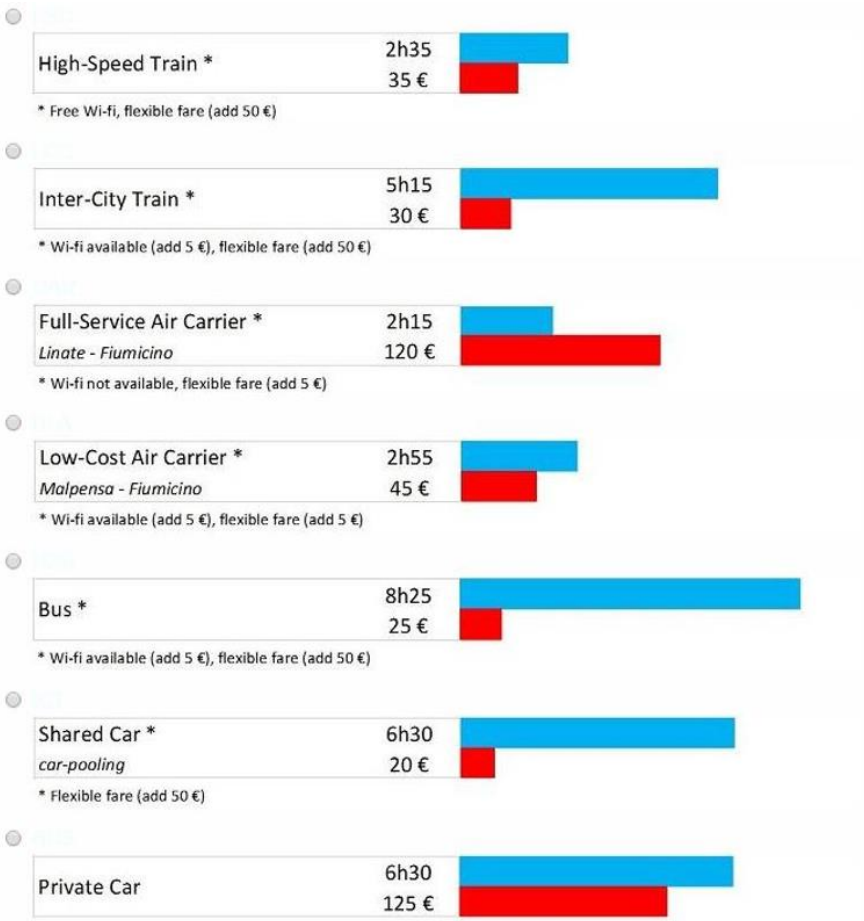
- Rome
- Milan
- Online

2 - Which mode are you using for your current journey on this corridor?

- High-Speed Train
- InterCity Train
- Full-Service Air Carrier
- Low-Cost Air Carrier
- Bus
- Car Pooling
- Private Car

3 - If the available alternatives were these, with these characteristics, which one would you choose? (please choose only one alternative. Total travelling time for the air services also includes an estimate of the time needed for security check and boarding/disembarkation).

(Six choice tasks like this)



4 - Which other alternatives did you really consider? (Please select all the other considered alternatives)

(this question repeated after each choice task)

- High-Speed Train
- InterCity Train
- Full-Service Air Carrier
- Low-Cost Air Carrier
- Bus
- Car Pooling
- Private Car
- I only considered the chosen alternative

5 - According to which criteria you decided not to consider one or more alternatives? (you can choose more than one)

- Total travelling time
- Travel cost
- Fear of flying
- Car disease
- Security issues deriving from travelling with unknowns
- Other

6 - Which is your maximum admissible travelling time (in hours)?

7 - Which is your maximum admissible travelling cost (in euros)?

8 - In which city did your current journey start?

- Rome
- Milan

9 - In which place did your current journey start? (city/neighborhood)

- Rome/Milan, city centre
- Rome/Milan, other areas
- Another city in Rome/Milan Metropolitan Areas
- Another city outside Rome/Milan Metropolitan Areas
- Another city in another region

10 - With reference to your typical journeys, please make an estimate of the time which is necessary for you to reach the following places:

Rome Fiumicino Airport:

Rome Termini Train Station:

Rome Tiburtina Train and Bus Station:

Milan Linate Airport:

Milan Malpensa Airport:

Milan Centrale Train Station:

Milan Lampugnate Bus Station:

11 - Which is the reason of your trip?

- Study
- Leisure
- Health
- Work
- Other

12 - Who paid for your trip?

- The company I work for
- Myself
- Parents or relatives
- Partner
- Other

13 - Did you book this trip by yourself?

- Yes
- No, someone else booked it for me
- I do not remember

14 - How much did it cost? (in Euros)

15 - Are you travelling alone?

- Yes
- No

16 - How many people are you travelling with?

- 1
- 2
- 3-5
- More than 5

17 - Who are you travelling with?

- Friends
- Parents or relatives
- Work colleagues
- Partner
- People I never met before
- Children

18 - Do you remember which mode did you use during your last trip on the corridor?

- High-Speed Train
- InterCity Train
- Full-Service Air Carrier
- Low-Cost Air Carrier
- Bus
- Car Pooling
- Private Car
- I do not remember

19 - Do you usually travel with the same mode on this corridor? (you can select more than one answer)

- Yes
- No, depending from the nature of the journey (e.g. for work or for leisure)
- No, depending on the price
- No, depending on the availability
- No, depending on who I am travelling with (e.g. if I am travelling with children)
- No, depending on the exact place of departure/arrival
- No, for other reasons

20 - Are you member of any fidelity program?

- Fidelity program train company
- Fidelity program airline
- Experience level (car-pooling)
- I am not member of any fidelity program

21 - How far in advance do you usually plan to arrive at the airport (in minutes)?

22 - Do you have a car?

- Yes
- No

23 - Would this car be available for you to use to travel from Rome to Milan?

- Yes
- No

24 - Do you have a driving licence?

- Yes
- No

25 - Please define your sex

- Male
- Female

26 - Please define your age band

- 18 – 24
- 25 – 34
- 35 – 49
- 50 – 69
- 70

27 - Please define your educational level

- Primary school
- After primary school
- High-school
- Professional education
- Bachelor
- Master
- PhD

28 - Which is your net monthly salary? (This survey is completely anonymous. We ensure that this information will not be transmitted to any financial authority, but only used for statistical purposes)

- I do not have a monthly salary
- < 500 €
- Between 500 - 1000 €
- Between 1000 - 2000 €
- Between 2000 - 4000 €
- > 4000 €
- I prefer not to say

29 - Please, choose the category that better describes your status

- ○ Student
- ○ Worker
- ○ White collar (public sector)
- ○ White collar (private sector)
- ○ Entrepreneur
- ○ Unemployed
- ○ Manager/Businessman
- ○ Artisan
- ○ Teacher
- ○ Housewife
- ○ Other

Appendix C

This second SC survey has been administered in November and December 2017 to a sample of 746 residents in four cities in a range of 50-100 km from Bari International Airport, who travelled through the airport in the previous three months. Revealed preferences (i.e. actual choice of the access mode) have been also collected, which refer to the respondents' last trip to the airport.

The data was collected under the scientific supervision of Professor Angela Stefania Bergantino, as part of the research project "*An Analysis of demand for the Apulian airport system*" of the Department of Economics, Management and Business Law of the University of Bari (Italy). The project obtained a research grant by Aeroporti di Puglia Spa.

Choice tasks (5 for each respondent) were designed using a Bayesian D-efficient experimental design and the software NGene. To this extent, priors were imputed based on a preliminary modelling run on a dataset collected in November 2016 in the same study area as a part of the same research project of the University of Bari. The latter dataset - which in turn was created using an orthogonal fractional factorial design - has also been used to validate the outcome of the models presented in Chapter 5.

For both datasets (i.e. for the 2016 and the 2017 waves), constraints were included as to avoid clearly dominated choices using the command "*reject*". In particular, scenarios in which travel time and travel cost for the public transport alternative were lower than those for the direct bus and train were

rejected. Similarly, those in which travel cost for the taxi option was lower than travel cost for the car alternatives (as a driver or as a passenger).