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

Mauro Capurso*, Stephane Hess, and Thijs Dekker
Institute for Transport Studies and Choice Modelling Centre, University of Leeds (UK)

***Corresponding author**
Institute for Transport Studies and Choice Modelling Centre, University of Leeds
34-40 University Road, Leeds, LS2 9JT, United Kingdom
Tel: + 44 1133435325; Email: m.capurso@leeds.ac.uk

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 suggests 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.[1]

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.[2] 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 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

[1] 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).

[2] 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.

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).[3] 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 - still 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 variables. The main difference with Ben-Akiva and Boccara is that we simultaneously control for consideration effects and unobserved preference heterogeneity, to

[3] 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.

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.[4] 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.

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

[4] 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.

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.[5]

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

[5] 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.

online on their official websites[6] and for some operators the Internet is the only available marketplace (car-pooling and bus).

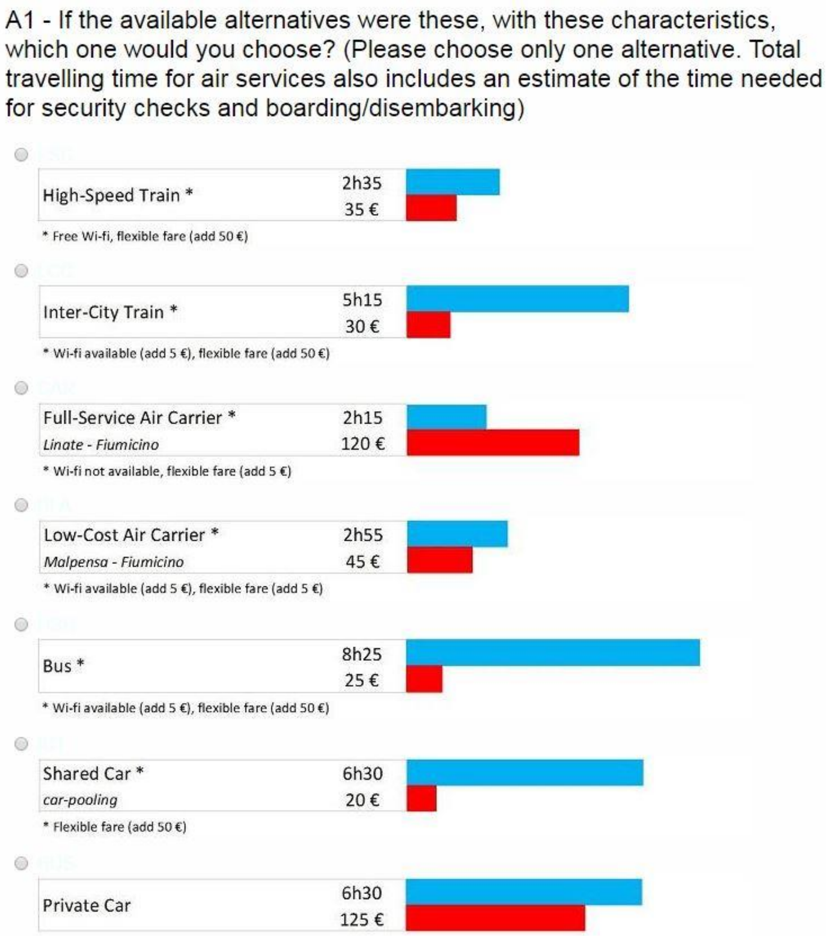
Against this background we designed a SC survey mimicking a real purchasing decision through an online journey planner. Its visual design was comparable to online journey planners as increasingly used by individuals to make travel plans.[7] In online journey planners, however, all ‘objectively’ available (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.[8] 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 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%).[9]

Each respondent completed six choice tasks, and we used a layout similar to the one displayed by the website www.goeuro.com (Figure 1). To avoid possible ordering effects, we randomised the order of the presented alternatives across respondents.

[6] The HSR operator Trenitalia reports that more than 50% of tickets are purchased online (2017).
[7] 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.
[8] 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.
[9] The response rate was higher at bus and train stations (≈ 50%) than at airports and service stations (≈ 20%).

Figure 1. The layout of the choice tasks.



The attributes of the alternatives were travel time[10], 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 1, 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

[10] 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.

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.

Table 1. Current ranges and survey attribute levels.

	Current ranges				**Attribute levels**	
	Travel time		*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.

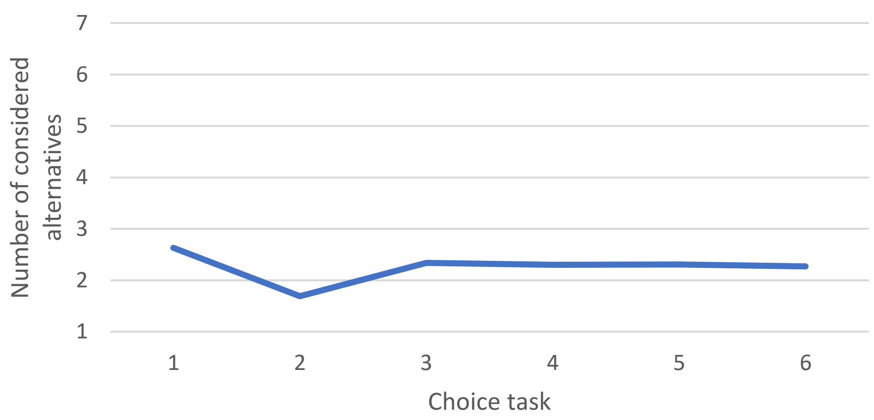
At the end of each choice task we asked respondents to state which non-chosen alternatives they had considered. The following question format was used:

"*Which other alternatives did you consider? (Please select all the other considered alternatives)*"[11]

[11] 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.

In Figure 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.

Figure 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.[12] For example, it could be possible that some respondents did not stated to consider alternatives for which they simply had lower preference. In that case, self-reported information would be associated with

[12] 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.

under report. Given that we are not able to confirm the 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.[13] Although questions about the presence of thresholds are less prone to misunderstanding and over/under reporting issues than those on consideration of alternatives[14], 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, 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.

[13] 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.

[14] 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.

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 1):

$$U_{int} = V_{int} + \varepsilon_{int} \quad (1)$$

V_{int} 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 ε_{int} 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_{nt} by (Equation 2):

$$P_{int} = P\left(U_{int} \geq U_{jnt}, \forall j \neq i \in C_{nt}\right) \quad (2)$$

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 3):

$$P_{int} = \frac{exp(V_{int})}{\sum_{j \in C_{nt}} exp(V_{jnt})} \quad (3)$$

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 $C_{nt}^* \subseteq C_{nt}$. As a result, the choice probability for considered alternatives increases relative to the MNL model in Equation 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 $\tilde{P}_{int}(C_{nt}^*)$ vary across consideration sets due to the changing denominator in (3). The probability of using a particular consideration set $\pi_{nt}(C_{nt}^*)$ is used as a weight in the averaging process to obtain the unconditional choice probability (see Equation 4):

$$P_{int} = \sum_{C_{nt}^* \subseteq C_{nt}} \pi_{nt}(C_{nt}^*) \cdot \tilde{P}_{int}(C_{nt}^*) \tag{4}$$

This two-stage formulation is identical to the expression proposed by Manski (1977). In Equation (5), we follow Swait and Ben-Akiva (1987a)'s *independent availability* model and assume that the consideration probability for alternative *j*, W_{jnt}, is independent across alternatives. This results in the following definition of $\pi_{nt}(C_{nt}^*)$:

$$\pi_{nt}(C_{nt}^*) = \prod_{j \in C_{nt}^*} W_{jnt} \cdot \prod_{j \notin C_{nt}^*} \left(1 - W_{jnt}\right) \tag{5}$$

In this paper, we present two alternative specifications for $\pi_{nt}(C_{nt}^*)$. The first specification makes use of the responses to the stated consideration questions. The second specification makes use of self-reported threshold levels for travel time. Moreover, we assume that probabilistic consideration only applies to a subset of alternatives, rather than to all the available alternatives.

Given that indicators of a different nature are used, the two model specifications therefore differ in the way W_{jnt} is defined. In the first model specification (Equation 6), the predicted consideration probability, $\widehat{W_{jnt}}$, 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.

$$(6) \qquad \widehat{W_{jnt}} = \frac{1}{1 + \exp\left(\hat{\alpha}_j + \hat{\beta}_j Z_{jnt}\right)},$$

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

For the second model, we first estimate a linear regression explaining the stated threshold for travel time as a function of individual socioeconomic and context characteristics X_n. In Equation 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.

$$(7) \qquad \widehat{\max\_TT_n} = \hat{\gamma} + \hat{\delta} X_n$$

Then, we specify a binding function in Equation 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_{jnt} 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.

$$W_{jnt} = \frac{1}{1 + \exp\left(\theta_j + \varphi\,(TT_{jnt} - \widehat{\max\_TT_n})\right)} \quad (8)$$

The use of a predicted consideration probabilities $\widehat{W_{jnt}}$ (Equation 6) and predicted thresholds for the travel time attribute $\widehat{\max\_TT_n}$ (Equations 7-8) 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. 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, we account for the panel nature of the data by introducing the heterogeneity at the level of individual respondents and applying robust standard errors also at the panel level.

5. Results and discussion

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

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 the self-reported consideration data as well as choice data. On average, the self-reported level of consideration for *faster* modes is 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.[15] As a result of modelling consideration probabilistically on only three alternatives, the number of possible consideration sets is reduced to eight.

Table 2a 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 busses (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 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.

[15] 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.

Table 2a. 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 2b, we present the results of a linear regression (OLS) on the logarithm of the stated thresholds for travel time.[16] Results show that, ceteris paribus, male respondents have a higher self-imposed threshold for travel time relative to female respondents. The time threshold is 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 2b. 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.

[16] Given that we estimate a linear regression on the logarithm of the stated threshold, the resulting predicted threshold is exponentiated when implemented in the binding function (see Equation 8 in Section 4).

5.2 Consideration of alternatives and choice

We present the results for three choice models in Table 3. 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 6 and 8, respectively).[17] 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).

[17] 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 3. 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%	
95% Confidence interval	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 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 2a, are directly implemented in the two-stage model (see Equation 6).

In Model 3, we account for consideration effects using information on *stated thresholds*. Consideration probabilities are now calculated within the choice model (Equations 7-8). The predicted thresholds from Table 2b 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). 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 4) 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 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 4. 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.[18] This measure, reported alongside traditional measures of fit in Table 3, 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

[18] 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.

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 former, we present the value of travel time (VTT) for an individual who pays her/himself for the trip (Table 5). [19]

Table 5. 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 5 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

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

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 3a-c) 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 3a), 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 3b 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 3c). Models 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 3a. Predicted aggregate market shares (*status quo*).

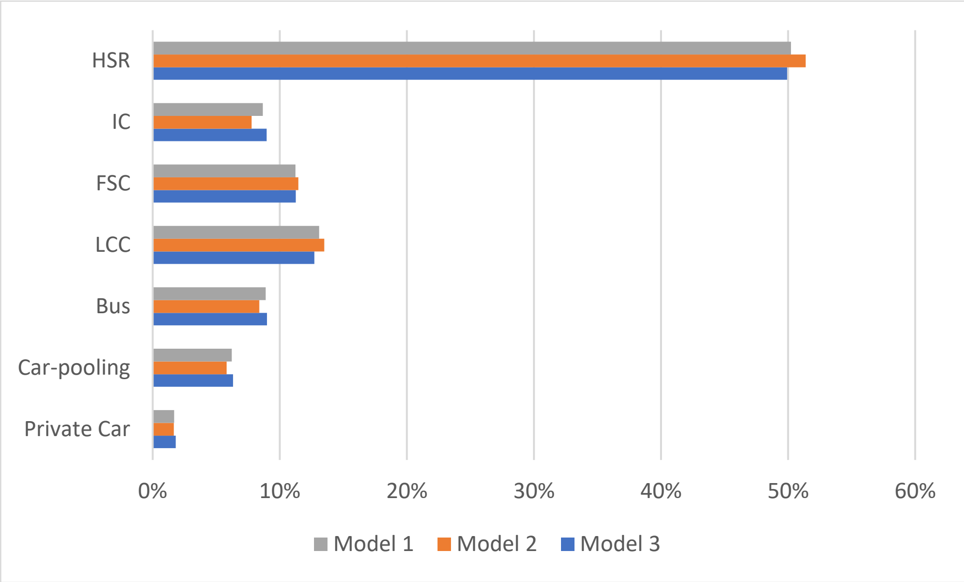


Figure 3b. Predicted aggregate market shares when travel time for HSR is reduced by 20%.

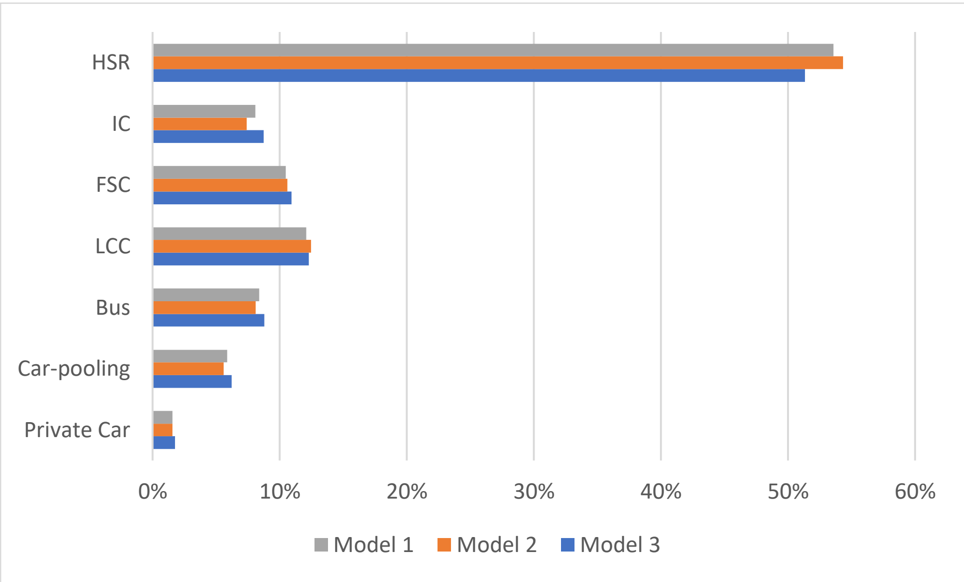
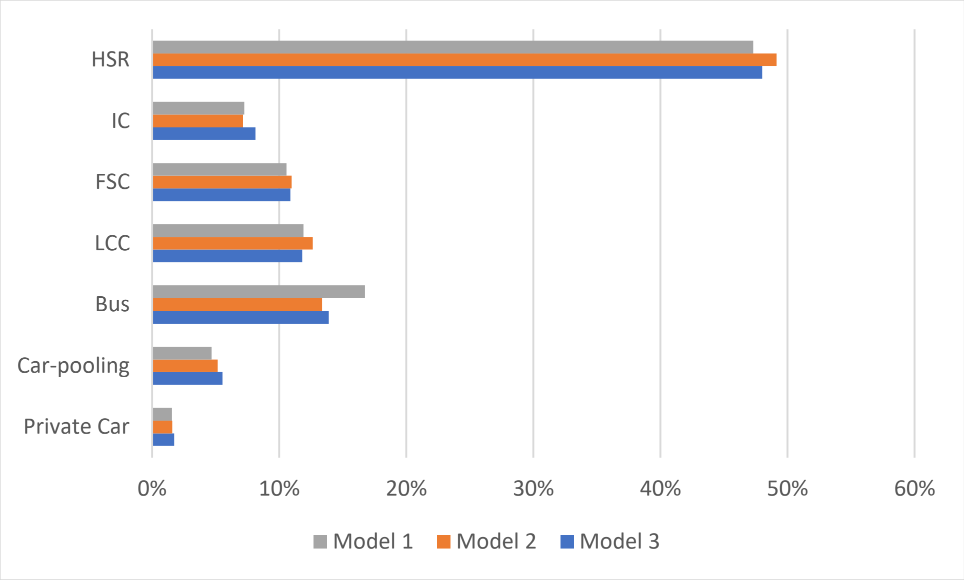


Figure 3c. Predicted aggregate market shares when travel time for bus is reduced by 30%.



6. Conclusion

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 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.

Appendix A: Travel cost

Income information was collected using income classes, therefore 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:

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

References

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

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.

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* ed P. Zarembka, pp. 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. Ph.D. Thesis, Massachusetts Institute of Technology, 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., 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.