



EJTIR

ISSN: 1567-7141
<http://ejtir.tudelft.nl/>

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

Mauro Capurso¹

Institute for Transport Studies and Choice Modelling Centre, University of Leeds, United Kingdom

Stephane Hess²

Institute for Transport Studies and Choice Modelling Centre, University of Leeds, United Kingdom

Thijs Dekker³

Institute for Transport Studies and Choice Modelling Centre, University of Leeds, United Kingdom

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.

Publishing history

Submitted: 18 October 2019
Accepted: 25 September 2021
Published: 29 September 2021

Cite as

Capurso, M., Hess, S., & Dekker, T. (2021). Stated consideration and attribute thresholds in mode choice models: a hierarchical ICLV approach. *European Journal of Transport and Infrastructure Research*, 21(3), 106-123.

© 2021 Mauro Capurso, Stephane Hess, Thijs Dekker

This work is licensed under a Creative Commons Attribution 4.0 International License ([CC BY 4.0](https://creativecommons.org/licenses/by/4.0/))

Keywords: *hierarchical ICLV, stated consideration, stated threshold.*

¹A: via Firenze 16, 70054 Giovinazzo, Italy E: mauro.capurso@live.com

²A: University Road 34-40, LS2 9JT Leeds, United Kingdom T: +44(0)113 343 6611 E: s.hess@leeds.ac.uk

³A: University Road 34-40, LS2 9JT Leeds, United Kingdom T: +44(0)113 343 5334 E: t.dekker@leeds.ac.uk

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 hardly measured (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.

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 the indicators might not correspond to actual levels of consideration. Second, there is scope for endogeneity bias as the indicators might themselves be a function of the utility for the alternatives and/or may be correlated with other unobserved factors (Capurso et al., 2019). Given this, it might be preferable to treat the indicators of consideration as dependent rather than independent variables, using, for example, an Integrated Choice and Latent Variable (ICLV) model (McFadden, 1986; Ben-Akiva et al., 2002; Bolduc et al., 2005), as it is done in the present paper.

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. 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.

The inclusion of the latent variables in the overall framework is made possible by additional information on the decision-making process collected during a SC survey in the form of stated thresholds and stated consideration. The methodology proposed in this paper can be easily used with SC data, since what is needed is additional information on consideration that can be easily collected during SC surveys. Nevertheless, it is also possible to apply the methodology proposed with RP data; for example, one can collect additional information on consideration through a separate survey, as it has been done by Enam and Choudhury (2011), who collected information on consideration of alternatives from a small separate sample of individuals to infer consideration set probabilities on a larger RP dataset.

In the proposed approach latent consideration is used to reduce the utility, and therefore choice probability, of the alternatives. A similar ‘discounting’⁴ 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 time⁵ and travel cost (Table 1). Accordingly, it is reasonable to assume that some travellers might *a priori* disregard alternatives based on self-imposed thresholds for specific attributes.

Table 1. Alternatives’ core characteristics at the time of the SC experiment

Alternatives	Travel time		Travel cost (€)	
	Min	Max	Min	Max
High Speed Rail (HSR)	2h55min	4h28min	19.9	209
Inter-City Rail (IC)	6h27	6h50min	9	79
Full Service Air Carrier (FSC)	2h20min ¹		55.41	244.71
Low Cost Air Carrier (LCC)	2h25min ¹		44.73	267.23
Bus	7h25min	10h45min	1	29
Car-pooling ²	5h41min		25	45
Private car ³	6h22min		99 (41 toll/58 fuel)	

Source: Operators’ websites; Note: 1 - includes an estimate of in-flight (1h10min for FSC and 1h15min for LCC), and boarding time (1h10min) 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

⁴The use of the term “discounting” might induce the reader to think that a discounting factor multiplies the utility of an alternative with the aim of reducing its value; in reality, the (negative) discounting factor enters the utility in an additive way, as it will be better explained in the methodological section.

⁵Total travel time is generally composed of in-vehicle travel time, boarding time (e.g. for the air mode), and access/egress time. In particular, the latter might differ for each individual, depending on the distance between the train station, the airport or the bus terminal and her/his departure/arrival points. Due to software restrictions it was not possible to customise the SC experiment depending on respondents’ specific access/egress time; however, we separately collected this information, and used it as a respondent-specific explanatory variable in the choice model.

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%).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 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). 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. 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 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).

⁶ The sample is not necessarily representative of the population-wide mode shares for the Rome-Milan corridor. However, it should be noted that the latest available official figures with respect to modal shares on this corridor report that, in 2014, 24% of passengers travelled by air, 65% by train, and the remaining 11% by bus and car (Italian Authority for Transport Regulation, ART, 2015).

A1 - If the available alternatives were these, with these characteristics, which one would you choose? (Please choose only one alternative. Total travelling time for air services also includes an estimate of the time needed for security checks and boarding/disembarking)

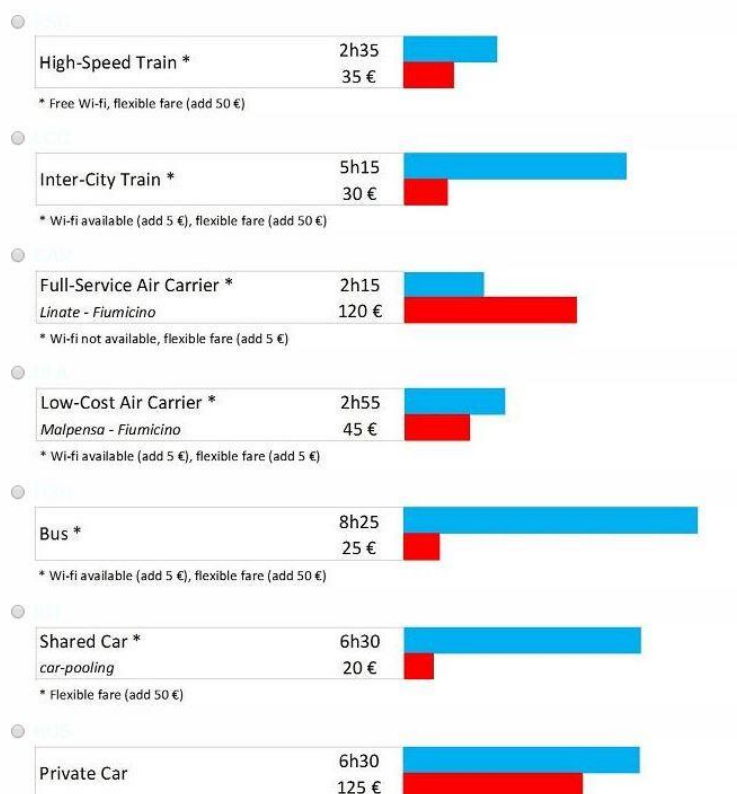


Figure 1. The layout of the choice tasks

Besides choices, information on consideration of the different modes of transport was collected after each choice task.⁷ 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%).⁸ 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, the inter-city train (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.

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

⁷ Information on availability of a private car was also collected and subsequently used in the choice model. The other travel modes were supposed to be always available: several trains, flights and buses were available at any time of the day and overnight (e.g. bus services). Also, at the time of the survey, there was always an oversupply of seats on the car-pooling network "Blablacar" on this route.

⁸ There is no surprise that the percentages of reported consideration follow those of the sample distribution, e.g. with HSR receiving the highest consideration given that the majority of respondents have been recruited while at the train station, and with private car receiving the lowest consideration for a similar reason. As for the sample distribution, the percentages of reported consideration might not be necessarily representative of the population-wide consideration shares.

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). The 'stability' of stated consideration for IC, bus, and car-pooling, and the fact that thresholds on travel time were 'respected' most of the times would indeed suggest that respondents are well aware that the expected average travel time with those alternatives could indeed vary within a range, but it will still be well above/within their thresholds, as it depends on the transport infrastructure in use. In turn, this would suggest that, at least for those alternatives, consideration might not be context-specific and driven by *a priori* beliefs/knowledge for specific journeys. Consideration would be stable over time, or at least until the transport infrastructure in use will not change, with the implication that the information on stated consideration and thresholds could be potentially used also for forecasting. More caution in forecasting would be needed with respect to the thresholds on travel cost, given the unpredictable impact that sales promotions might have on consideration of alternatives that are thought to be not affordable.

3. Methodology

In the proposed hierarchical ICLV model structure - illustrated in Figure 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.

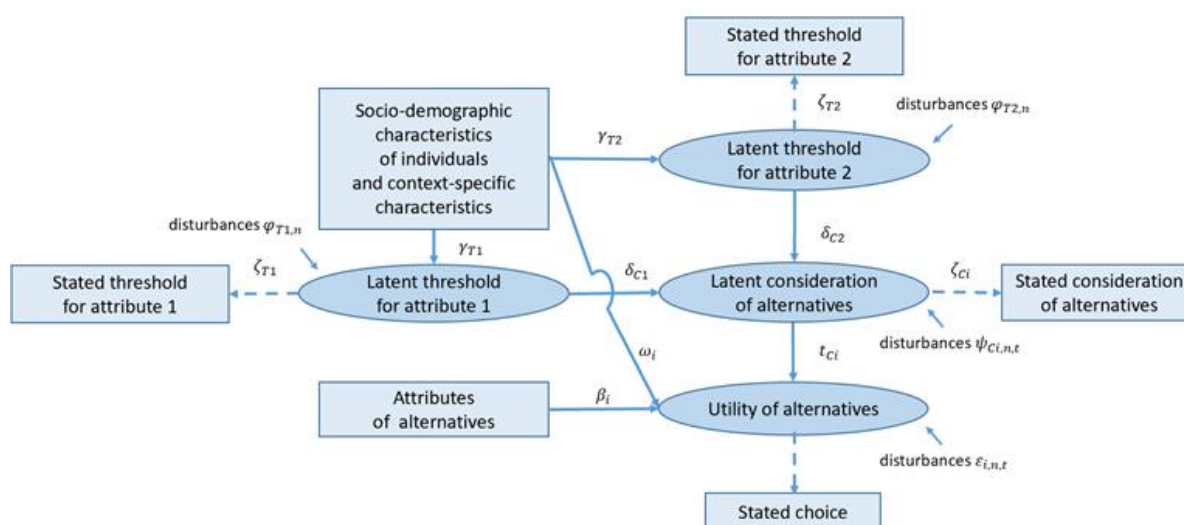


Figure 2. The proposed 'hierarchical ICLV' model

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 1):

$$\alpha_{Tk,n} = \gamma_{Tk} Z_{Tk,n} + \varphi_{Tk,n} \quad (1)$$

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 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.

$$\alpha_{Ci,n} = \sum_{k=1}^K \delta_{Ck} \alpha_{Tk,n} + \gamma_{Ci} Z_{Ci,n} + \psi_{Ci,n} \quad (2)$$

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.⁹

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. Ideally, one would compare the latent thresholds against the presented attribute levels in each equation. A simplified model is however presented here, 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):

$$I_{Tk,n} = \theta_{Tk} + \zeta_{Tk} \alpha_{Tk,n} + \eta_{Tk,n} \quad (3)$$

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.

⁹ In a general framework, correlation parameters to account for potential correlation across the error terms of both latent variables (latent threshold and latent consideration) could be also included. In this particular case the correlation parameters were found to be not statistically significant, hence removed from the model specification.

The density for a stated threshold is therefore given by the normal density function (Equation 4):

$$P(I_{Tk,n}|\alpha_{Tk,n}) = \frac{1}{\sqrt{2\pi\sigma_{I_{Tk,n}}^2}} e^{-\frac{(I_{Tk,n} - (\theta_{Tk} + \zeta_{Tk}\alpha_{Tk,n}))^2}{\sigma_{I_{Tk,n}}^2}} \quad (4)$$

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 5):

$$P(I_{Ci,n}|\alpha_{Ci,n}, \alpha_{Tk,n}) = \prod_{t=1}^T (\lambda_{(I_{Ci,n,t}=0)} \left(\frac{1}{1 + e^{\theta_{Ci} + \zeta_{Ci}\alpha_{Ci,n}}} \right) + \lambda_{(I_{Ci,n,t}=1)} \left(\frac{e^{\theta_{Ci} + \zeta_{Ci}\alpha_{Ci,n}}}{1 + e^{\theta_{Ci} + \zeta_{Ci}\alpha_{Ci,n}}} \right)) \quad (5)$$

where $\lambda_{(I_{Ci,n,t}=1)}$ is a dummy variable which takes value 1 when the alternative is stated to be considered on choice task t , 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 6):

$$U_{i,n,t} = V_{i,n,t} + \varepsilon_{i,n,t} = \varsigma_{i,n} + \beta_i X_{i,n,t} + \omega_i Z_n + \tau_{Ci} \ln(a^*_{Ci,n}) + \varepsilon_{i,n,t} \quad (6)$$

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 Equation 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} .

$$a^*_{Ci,n} = \frac{1}{(1 + \exp(-\alpha_{Ci,n}))} \quad (7)$$

When $a^*_{Ci,n}$ is closer to 0, the utility will be heavily discounted, given that $\ln(a^*_{Ci,n}) \rightarrow -\infty$ as $a^*_{Ci,n} \rightarrow 0$. 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.¹⁰

¹⁰ The use of a penalty function approach (ln of latent consideration) is only applicable in a logit based ICLV structure. The use of the same approach in a probit based ICLV structure will lead to a breakdown of convolution property of normal distribution.

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 $\varsigma_{i,n} = \mu_{\varsigma_i} + \sigma_{\varsigma_i}\xi_{i,n}$, 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 8):

$$P(Y_{i,n,t} | a^*_{Ci,n}, \alpha_{Tk,n}, X_{i,n,t}, Z_n, \varsigma_n) = \prod_{t=1}^T \frac{e^{V_{i,n,t}}}{\sum_{j \in C_n} e^{V_{j,n,t}}} \quad (8)$$

The joint LL function for the proposed ‘hierarchical ICLV’ model is given by (Equation 9):

$$LL = \sum_{n=1}^N \ln \left[\left(\int_{\alpha_{Tk,n}} \int_{a^*_{C,n}} \int_{\varsigma_n} P(I_{Tk,n} | \alpha_{Tk,n}) P(Y_{i,n,t} | a^*_{Ci,n}, \alpha_{Tk,n}, X_{i,n,t}, Z_n, \varsigma_n) P(I_{Ci,n} | \alpha_{Ci,n}, \alpha_{Tk,n}) + \right. \right. \\ \left. \left. f(\varsigma_n | \mu_{\varsigma}, \sigma_{\varsigma}) g(\alpha_{Ci,n} | \alpha_{Tk,n}, Z_{Ci,n}) h(\alpha_{Tk,n} | Z_{Tk,n}) d\varsigma_n da^*_{C,n,t} d\alpha_{Tk,n} \right) \right] \quad (9)$$

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 draws¹¹ (MLHS, Hess et al., 2006).

4. Results and discussion

4.1 Estimation results

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.¹² These alternatives are much slower than HSR, FSC, and LCC, but not necessarily cheaper (see Table 1), and we thereby assume that consideration decisions might be only driven by the presence of thresholds on travel time. This means that latent consideration for the slower alternatives (IC, bus, and car-pooling) 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 data¹³, which suggest that respondents are less likely to *a priori* discard the remaining (faster, but not necessarily more expensive, i.e. HSR, FSC, LCC) alternatives from consideration.

¹¹ 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.

¹² 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.

¹³ 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 2. Estimation results

Structural Models	Model 1		Model 2		Model 3	
	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

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

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(0)		-2319.01		-5142.92		-2319.01
LL(final, complete model):		-1263.23		-2563.49		-1186.60
LL(final, choice model only):		-1263.23		-1198.69		-1186.60

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

Estimation results are summarised in Table 2. Model 1 represents a MMNL model with normally distributed alternative specific constants (ASC) over respondents. This model assumes all available alternatives (e.g. if the respondent does not have a private car, this alternative is not available for her/him) are fully considered and thereby represents the standard practice in the estimation of discrete choice models. 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), as anticipated in the previous paragraph. Model 3 is the reduced-form model of Model 2. This is also a MMNL model in which we do not make direct use of the indicators in the estimation (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. 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.

We estimated four coefficients for travel time, grouping together similar alternatives, 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 according to the functional form in Equation 10. This functional form was selected among others commonly used (e.g. linear, Box-Cox), as the one that best fitted the data. However, given that not all respondents reported their income, we actually estimated two separate cost sensitivities, one for those who disclosed this information (‘Travel cost’), and one for those who did not (‘Travel cost, income na’).

$$\beta_{travel_cost_n} = ((\beta_{travel_cost_income_yes_n} * ((\frac{income_n}{average_income})^{\lambda_{income_n}}) * (income_yes_dummy_n)) + (\beta_{travel_cost_income_na_n} * (1 - income_yes_dummy_n)) + (\beta_{paid_employer_n} * paid_employer_dummy_n)) \quad (10)$$

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.¹⁴ 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.¹⁵ As expected, respondents who are in possession of a loyalty card are more likely to

¹⁴ In particular, the coefficient for travel cost for the former respondents turns out to be less than a half than that for the latter.

¹⁵ 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.

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.

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 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 and 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 5) reflect the fact that the stated consideration rates were on average lower than 50% in the sample (see footnote 3).

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.

Model 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 with respect to its reduced-form (i.e. Model 3), resulting from the use of the additional variables explained by the measurement model components.

4.2 The impact of consideration effects on overall utility and VTT indicators

In order to show what happens to the utility of an alternative when consideration effects are taken into account we plot – just for illustrative purposes – the empirical cumulative distribution functions (ECDFs) for the utility values of the bus alternative according to the aforementioned formulations (Models 1-3, Figure 3). The bus alternative is the one with the largest travel time, hence the alternative for which we could expect lower levels of consideration if this is effectively driven by the presence of thresholds on travel time.

The ECDFs represent the proportions of observations showing specific values of the utility. In Models 2 and 3, variations in utility values are mostly driven by the impact of the discounting factor related to 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.

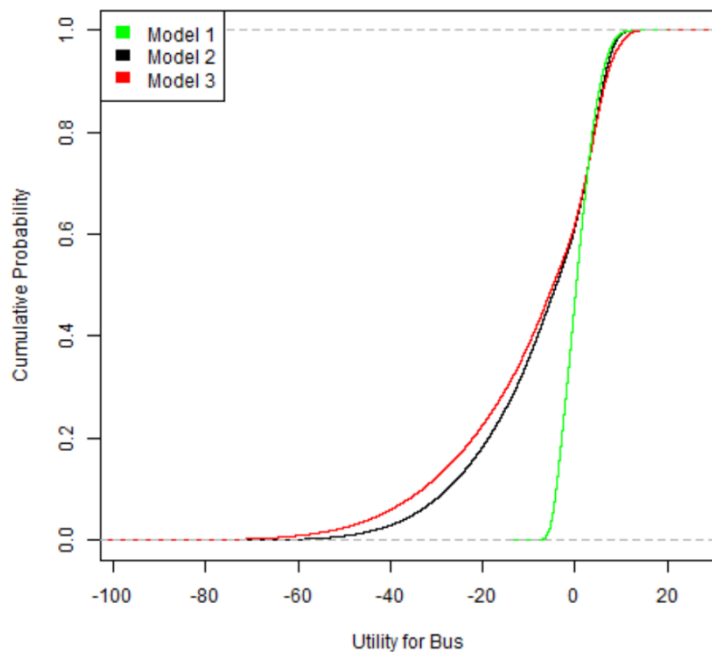


Figure 3. ECDFs for the impact of consideration effects on utility for bus

Considering also the other alternatives, we now discuss the implications of accounting for consideration effects on single 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). VTT indicators are obtained for an individual who pays her/himself for the trip. Standard errors are calculated using the delta method.

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.

Table 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

4.3 *The impact of consideration effects on model fit*

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 reveals 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), as we see here, is not uncommon in the literature ("Author", 2014; Vij and Walker, 2016). 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.

An earlier paper using the same dataset (Capurso et al., 2019) separately employs the two sources of information on respondents' processing strategies collected during the SC experiment, i.e. stated consideration of alternatives and stated thresholds on attributes and finds that consideration of alternatives is better explained by the presence of thresholds of attributes. This was already an interesting finding given that thresholds for attributes minimize the burden on respondents, with respect to stated consideration of alternatives, while also providing the largest informational content. As a step forward, this paper wanted to explore the possibility that there could be a 'hierarchical' nexus between the two sources of indicators, in the sense that (latent) thresholds for attributes could be 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 neither thresholds nor 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 threshold as well as 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 because the reduced form Model 3, which excludes the information provided by the indicators, actually has a slightly better fit than the presented hierarchical ICLV Model 2.

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 wish to acknowledge the 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). Report Autorità Regolazione Trasporti (2015). Turin, Italy.
- 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.

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.

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

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.

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*, 129, 170-184.

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.

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

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.

Fotheringham, A. S. (1988). Note - Consumer store choice and choice set definition. *Marketing Science*, 7(3), 299-310.

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.

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.

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.

Manski, C. F. (1977). The structure of random utility models. *Theory and decision*, 8(3), 229-254.

Mariel, P. and Meyerhoff, J. (2016). Hybrid discrete choice models: gained insights versus increasing effort. *Science of the total environment*, 568, 433-443.

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.

McFadden, D. (1986). The choice theory approach to marketing research. *Marketing Science*, 5(4), 275-297.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behaviour. In Zarembka P. (ed) *Frontiers in econometrics*, 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.

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.

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.

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.

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.

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

Swait, J. (2001). A non-compensatory choice model incorporating attribute cutoffs, *Transportation Research Part B: Methodological*, 35(10), 903-928.

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.

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.

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.

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.

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.