ABOUT ATTITUDES AND PERCEPTIONS – FINDING THE PROPER WAY TO CONSIDER LATENT VARIABLES IN DISCRETE CHOICE MODELS

Francisco J. Bahamonde-Birke*
Energy, Transportation and Environment Department
Deutsches Institut für Wirtschaftsforschung – Berlin
and
Technische Universität Berlin
e-mail: bahamondebirke@gmail.com

Uwe Kunert, Heike Link
Energy, Transportation and Environment Department
Deutsches Institut für Wirtschaftsforschung – Berlin
e-mails: ukunert@diw.de; hlink@diw.de

Juan de Dios Ortúzar
Department of Transport Engineering and Logistics,
Centre for Urban Sustainable Development (CEDEUS)
Pontificia Universidad Católica de Chile
e-mail: jos@ing.puc.cl

ABSTRACT

In this work, we provide an in-depth theoretical discussion about the differences between attitudes and perceptions, as well as an empirical experience to analyse its effects. This discussion is of importance, as the large majority of papers considering attitudinal latent variables, just consider those as attributes affecting directly the utility of a certain alternative and systematic taste variations are rarely taken into account; perceptions are normally completely ignored.

The results of our case study show that perceptions may affect indeed the decision making process and that they are able to capture of lot of the variability that is normally explained by the alternative specific constants. In the same line, our results indicate that attitudes may be the reason for systematic taste variations, and that a proper categorization of the latent variables, in accordance with the underlying theory, may outperform the customary assumption of linearity.

Keywords: *Hybrid Discrete Choice Modelling, Latent Variables, Attitudes, Perceptions*

1. INTRODUCTION

The last decades have seen discrete choice models (DCM) become a key element in travel demand modelling and forecasting (Ortúzar and Willumsen, 2011). Their current state-of-practice considers objective characteristics of the alternatives and the individuals as explanatory variables and yield as output individual probabilities of choice between different alternatives. It is also is well known that attitudes and perceptions play a role in the decision making process, and the usual approach to take these into account considers the estimation of a Multiple Indicator MultIple Cause (MIMIC) model, as suggested by Bollen (1989). The joint use of MIMIC models and DCM leads to state-of-the-art hybrid discrete choice (HDC) models (Ben-Akiva *et al.*, 2002; Bolduc and Alvarez-Daziano, 2010; Bahamonde-Birke and Ortúzar, 2014a).

In the last years, the literature has provided abundant empirical and theoretical evidence about the advantages of this approach and the use of HDC models has gained popularity (v. Acker, 2011; Ashok et al., 2002; Bahamonde-Birke *et al.*, 2010; Raveau *et al.,* 2012; Alvarez-Daziano and Bolduc, 2013; among others). Notwithstanding, attitudes and perceptions are usually addressed as a whole, not considering that both are expressions of different value judgments. This way, attitudes express a characteristic of the individuals toward life, society, etc. and are intrinsically related to them, playing a role in every decision made. Perceptions, instead, are exclusively related to the way certain alternatives are perceived. This being the case, an attitude resembles a socio-economic characteristic of the individual, while a perception is intrinsically associated with an alternative attribute.

This difference has important implications and the way in which both should be treated in a discrete choice model is completely different. In turn, this issue affects not only the fashion in which latent variables are estimated but almost all hypotheses concerning them. Hence, different assumptions will have an effect on both the way the latent variables are constructed through the MIMIC model as well as on the manner in which these constructs are reflected in the utility function of the DCM.

The rest of the paper is organized as follows. Section 2 offers a theoretical overview of HDC models, while Section 3 presents an extensive discussion about the different ways to consider latent variables in DCMs. Section 4 describes an experiment carried out to test the hypotheses of the previous section, and its results are discussed in section 5. Finally, section 6 summarises our conclusions.

2. THEORETICAL BACKGROUND

Under the assumption that individuals are rational decision makers, it can be postulated that individuals q facing a set of available alternatives $A(q)$, will choose the alternative i that maximizes their perceived utility. In accordance with Random Utility Theory (Thurstone, 1927; McFadden, 1974), it is possible to depict this utility as the sum of a representative component (V_{iq}) and an error term (ε_{iq}), which leads to the following expression (Ortúzar and Willumsen, 2011):

$$U_{iq} = V_{iq} + \varepsilon_{iq} \tag{2.1}$$

The representative utility (V_{iq}), considering all attributes that can be quantified by an observer, is usually characterized through concrete and measurable properties of the alternatives and the individuals; the error term, in turn, is considered to take into account all unknown or abstract elements affecting the decision.

When considering a Hybrid Discrete Choice (HDC) modelling framework (Ben-Akiva *et al.*, 2002), the modeller attempts to depict abstract attributes as measurable variables in order to include them as part of the systematic utility. Hereby, immaterial constructs, known as latent variables (η_{liq}), are also included into the modelling. These variables are supposed to represent attitudes and/or perceptions of the individuals and, as they cannot be directly observed, they must be constructed as a function of positively observed variables. The usual approach to construct these latent variables relies on a MIMIC structure (Zellner, 1970; Bollen, 1989). Here, the latent variables are explained by a set of characteristics of the individuals and the alternatives (s_{iqr}), through so called *structural equations*, while explaining, at the same time, a set of attitudinal and/or perception indicators (y_{ziq}), previously gathered from the individuals, through so called *measurement equations*. This framework can be represented through the following equations:

$$\eta_{liq} = \sum_r \alpha_{lri} \cdot s_{riq} + \upsilon_{liq} \tag{2.2}$$

$$y_{ziq} = \sum_{l} \gamma_{lzi} \cdot \eta_{liq} + \varsigma_{ziq} \tag{2.3}$$

where the indices i, q, r, l and z refer to alternatives, individuals, exogenous variables, latent variables and indicators, respectively. The error terms ν_{liq} and ζ_{ziq} can follow any distribution, but they are typically considered to be normally distributed with mean zero and a certain covariance matrix. Finally, α_{lri} and γ_{lzi} are parameters to be jointly estimated.

If we assume a linear specification in V_{iq}, the utility function can be expressed as (2.4). This specification can be understood as a first-order Taylor expansion of any multi-variable complex function (and therefore it is always valid in the neighbourhood of the estimation point); further, if the attributes are also assumed to be linear, the estimated parameters θ_{ik} and β_{il} (related to the tangible attributes and latent variables, respectively) can be directly interpreted as marginal utilities:

$$U_{iq} = \sum_{k} \theta_{ki} \cdot X_{kiq} + \sum_{l} \beta_{li} \cdot \eta_{liq} + \varepsilon_{iq} \tag{2.4}$$

Under the assumption that the error terms ε_{iq} in (2.1) are independent and identically distributed (IID) Extreme Value Type 1 (EV1) with the same variance σ^2, the differences between the utilities associated with the alternatives follow a Logistic distribution with mean zero and scale factor λ, leading to the well-known Multinomial Logit (MNL) model (Domencich and McFadden, 1975); in this case, the probability of choosing alternative i is given by:

$$P_{iq} = \frac{e^{\lambda \cdot V_{iq}}}{\sum_{j} e^{\lambda \cdot V_{jq}}} \tag{2.5}$$

and λ is inversely related to the standard deviation of the error terms:

$$\lambda = \frac{\pi}{\sigma\sqrt{6}} \tag{2.6}$$

However, as the scale factor cannot be estimated (assuming a linear function as usual), it is customary to *normalize* it to one (Walker, 2002).

The estimation of both parts of the model should be performed simultaneously, as a sequential estimation considering first the MIMIC part as an isolated system and evaluating afterwards the expected values for the latent variables cannot guarantee consistent and unbiased estimators (Train *et al.*, 1987; Ben-Akiva *et al.*, 2002). However, empirical evidence sustains the thesis that the sequential estimation produces no major discrepancies regarding the ratios between the estimated parameters and, therefore, the marginal rates of substitution (Raveau *et al.*, 2010; Bahamonde-Birke *et al.*, 2010). Nevertheless Bahamonde-Birke and Ortúzar (2014a) prove that the estimators may indeed be affected by a significant deflation bias (affecting all estimated parameters), while Bahamonde-Birke and Ortúzar (2014b) propose the following expression to correct this deflation:

$$\lambda_{HDC} = \frac{\lambda_{DC}}{\sqrt{1 + \frac{6 \cdot \lambda_{DC}^2 \cdot \sum_l \beta_l^2 \cdot \sigma_l^2}{\pi^2}}} \tag{2.7}$$

where $\sum_l \beta_l^2 \cdot \sigma_l^2$ stands for the variability induced into the model through the latent variables. This correction term performs in an acceptable manner as long as the ratio between the induced variability and the model's own variability is sufficiently small. If that is not the case, the sequential estimation must be disregarded.

3. ABOUT ATTITUDES AND PERCEPTIONS

Prior to discuss the different ways, in which latent variables may be considered into a HDC-model it is necessary to understand the difference between attitudinal and perception indicators. The former depend only on the individuals and are therefore constant for all alternatives. Thus, one set of attitudinal indicators will be enough to represent all decisions taken by the individual in question. Contrarywise, the latter depend on both the individuals and the alternatives, and for every different alternative a new set of indicators must be gathered. Even more, every variation in the alternatives will lead to a different valuation of them, as every detail affects the way in which the population perceives the various alternatives. Therefore, in order to work with perception indicators, it is necessary to gather more than one set of indicators per individual.

This issue can lead to a significant increase in the information required, as normally the alternatives would consist of different attributes that are subject to variations. Therefore, it is mandatory to make certain simplifying assumptions. Firstly, it may be assumed that certain attributes will not affect the way in which a given alternative is perceived and, consequently, this dimension may be excluded from the design (e.g. price and accessibility indicators). In the same line, it can be assumed that the model is valid across individuals (avoiding the need that every one states his/her perceptions for every combination of possible attributes), as long as other individuals are faced with the remaining combinations.

Once the indicators are gathered it is possible to construct the latent variables[1]. Obviously, attitudinal variables will be related to attitudinal indicators and *vice versa*. Thus, while attitudinal variables must be solely explained by characteristics of the individuals (as no variation across alternatives will be observed), perceptions should be also explained by the attributes of the different alternatives considered in the design of the experiment.

For estimation purposes, different sets of perception indicators associated with different alternatives (in terms of the attributes that constitute them) may be treated jointly or separately. If the sets of indicators are considered jointly, the estimated MIMIC model will hold for any variation of the alternatives and the estimated latent variables should be comparable among each other. An isolated estimation would offer a better goodness-of-fit, but the estimated latent variables would not be longer comparable. This must be taken into account when estimating the utility function as it can lead to considerable trouble if the alternatives cannot be isolated in the DCM. Note that using common estimators (and a common structure) for the measurement equations overcomes the difference between both approaches. When more than one set of indicators per person is considered, correlation among individuals must be taken into account.

[1] Most of the points discussed will still hold in case of latent classes models, but in this paper we will focus on the latent variables approach.

Further, the treatment of both kinds of variables (attitudes and perceptions) in the DCM should not be equal and some attitudes, just as socio-economic variables, should be considered through systematic taste variations and not directly in the utility function, as they affect the way in which the attributes of the alternatives are perceived.

That being the case, it is possible to identify three kinds of latent variables:

a) Non-alternative related attitudes: Most researchers working with HDC models consider this kind of variables (Bolduc and Alvarez-Daziano, 2010; v. Acker *et al.,* 2011, among many others). They represent general attitudes of the individuals toward their social and physical environment, such as a more ecological mind-set or a higher valuation of her social status. Even when using variables that may be understood as perceptions, such as comfort, security, etc., the modeller is in fact dealing with a non-alternative related attitude, as in this case these variables stand for the importance assigned by the individuals to these aspects and not for a perception of the alternative itself. This way, inferences such as “Alternative A is perceived as more comfortable” would not be accurate but rather “individuals caring for comfort favour alternative A”, which is not equivalent. On the other hand, as these variables resemble socio-economic characteristics, for the DCM to be identified they must be considered together with alternative specific attributes in the utility function. However, in most reported cases they are just considered in conjunction with alternative specific constants (Vredin-Johansson *et al.*, 2006; Bolduc *et. al*, 2008); this restriction may neglect important aspects of the decision, as it can be expected that individuals with different attitudes toward life exhibit a different valuation of the attributes of the alternatives, and therefore systematic taste variation should be allowed for (Ortúzar and Willumsen, 2011, page 279). Further, as in the case of socio-economic variables it is not clear that the latent variables should have a linear impact over the utility, a categorization should be considered (resembling in part a latent classes approach).
b) Alternative related attitudes: These variables are similar to the abovementioned, variables with the exception that attitudes are unequivocally related to a given alternative. Thus, these variables must be considered in conjunction with the alternative specific constant. For instance, Alvarez-Daziano and Barla (2012) considered the effect of a favourable predisposition toward automobiles or transit

systems in this way. As in the previous case, systematic taste variations (within the same alternative) and a possible categorization should also be analysed.

c) Perceptions: These variables are alternative related, i.e. they exhibit a different valuation depending on the alternative considered and as such they resemble observed attributes of the alternatives such as price or travel time; hence, both kind of variables should be treated in the same fashion. Alternative specific and generic estimators may be considered.

4. STUDY CASE

Departing from usual practice (which considers, typically, only attitudes), we developed an experiment considering both attitudinal and perception indicators in a transport choice framework. This allows testing for the more appropriate manner to consider both kinds of latent variables in the DCM, regarding the underlying theoretical concerns.

We conducted a stated choice (SC) experiment where respondents were asked to choose between different interurban public transport alternatives (regional[2] and intercity trains, and interurban coaches). The experiment was carried out in three waves (January 2014, March 2014 and April/May 2014), contacting both students and employees of the two universities in Berlin (the Technische Universität Berlin and the Humboldt-Universität zu Berlin), as well as employees of member institutions of the Leibniz-Gemeinschaft (WGL)[3]. After data cleaning, the survey yielded a total of 1,832 responses.

The questionnaire had four parts. In the first part, respondents were asked to describe the main characteristics (fare, travel time, number of transfers, etc.) of their last trips with the regional and intercity trains of Deutsche Bahn. At the end of this module and based on their experience travelling with Deutsche Bahn (i.e. considering the same kind of trains and the same number of transfers of the journey described), participants were required to state their level of agreement with the following statements:

---

[2] Regional trains should not be confused with commuter rail. Regional trains operate over long interurban distances, stopping more and over shorter distances than intercity trains. It is possible to travel across the country using only regional trains.

[3] The Leibniz-Gemeinschaft is the shelter association of publicly funded research institutes in Germany.

I was able to relax during the trip (y_{11})	***Relax***
I felt secure from thefts and losses (y_{12})	***Security***
Traveling with heavy luggage was (would have been) uncomplicated (y_{13})	***Luggage***
The departure time was reliable (y_{14})	***Departure***
The arrival time was reliable (y_{15})	***Arrival***
It was possible to use the travel time productively (y_{16})	***Productivity***
The station was easily accessible (y_{17})	***Station***
Purchasing the ticket was uncomplicated (y_{18})	***Tickets***

In the same line, respondents were also asked to state their level of agreement with these statements under the assumption that a bus carrier with no transfers would offer the service. The level of agreement was stated on a scale which ranged from strongly disagree (1) to strongly agree (10).

The second part of the survey gathered travel behaviour data as well as indicators related to the travellers' attitudes toward current political issues discussed in Germany. Hereby, the respondents had to state the level of agreement with the following sentences:

I agree with the nuclear power phase-out (y_{21})	***NuclearPhaseOut***
Environment protection is more important than economic growth (y_{22})	***Environment***
I am willing to pay a 25% surcharge on my electric bill to reduce CO_2 *emissions from coal power plants* (y_{23})	***ElectricSurcharge***
Highway tolls should be introduced to compensate CO_2 *emissions* (y_{24})	***HighwayTolls***
Automobiles with higher engine power should pay more taxes (y_{25})	***CarTax***
Investing on the development of high-speed trains should be encouraged (y_{26})	***HSTrains***
New highways or additional lanes to the existing ones should be built (y_{27})	***Highways***
New high-speed rail lines should be built (y_{28})	***RailLines***
I agree with the introduction of speed limits on highways (y_{29})	***SpeedLimits***

The third part of the questionnaire was the SC experiment itself. Here, respondents were required to choose between a first pivotal alternative, representing the trip previously described, and a new travel alternative. Altogether, respondents were confronted with 12 choice situations, where the first six used a pivotal alternative based on the trip with the Deutsche Bahn regional trains and the last six considered a trip with Deutsche Bahn intercity trains as the base situation. Alternatives were described in terms of their travel time, fare, number of transfers, mode of transport - regional trains (RE), intercity trains (FVZ) and coaches (LB) - and a safety level (represented through the number of severely injured passengers and the number of fatalities in the overall network over a year). Finally the fourth part of questionnaire gathered socioeconomic information about the respondents.

5. MODEL ESTIMATION

5.1 Model Structure

Before starting with the estimation of HDC models, it was necessary to establish the structure of the MIMIC-model considered. For this, the indicators were analysed using factor analysis to guarantee a correct specification of the latent variables (LV). This way, it was possible to identify three components explaining 69.9% of the variance of the perception indicators (y_{11} to y_{18}). In the same way, it was possible to establish that two variables captured 53.9% of the variability associated with the attitudinal indicators (y_{21} to y_{29}). Table 1 presents the rotated component matrices for both types of indicators. On the basis of these results, we constructed five latent variables, as highlighted in Table 1. The first was identified as "Comfort", as it was exclusively related to comfort indicators. The second component was called "Stress-free", as it was associated with situations causing tension during the trip. Finally, the third component was identified as "Reliability".

Table 1 – Rotated Component Matrix of Perception and Attitudinal Indicators

Indicator	Comfort	StressFree	Reliability
Relax	0.548	0.591	0.171
Security	0.144	0.782	0.132
Luggage	0.061	0.810	0.178
Departure	0.117	0.245	0.892
Arrival	0.280	0.125	0.867
Productivity	0.663	0.432	0.099
Station	0.810	0.064	0.177
Tickets	0.711	0.059	0.153

Indicator	Green	TrainFan
NuclearPhaseOut	0.688	-0.029
Environment	0.726	-0.074
ElectricSurcharge	0.704	0.030
HighwayTolls	0.658	0.214
CarTax	0.686	0.192
HSTrains	0.114	0.860
Highways	-0.546	0.365
RailLines	0.046	0.891
SpeedLimits	0.610	0.082

Regarding the attitudinal indicators, the first component was associated with a "Green" attitude, including a negative predisposition toward automobiles (y_{24}, y_{25}, y_{27} and y_{29}). The second component was related to individuals who have great appreciation for the development of trains and rail lanes (for this reason, this LV was called "TrainFan"). These results are interesting for our analysis as it was possible to identify a non-alternative ("Green") and an alternative ("TrainFan") related attitude.

5.2 MIMIC models

Given the complex structure and size of the data set (1,832 individuals; 3,900 sets of perception indicators; five latent variables and 13,138 choice situations), it was not computationally possible to perform a simultaneous estimation of the HDC model. In addition, we wanted to analyse the effect of attitudinal latent variables both as continuous and as categorized variables, which complicated the structure of the model even more. Therefore, a sequential estimation was attempted and it was possible to establish that the bias caused by this second-best estimation technique was manageable. So, the MIMIC model was estimated first and the latent variables considered in the DCM component were built according to these estimates.

It was necessary to estimate two different MIMIC models. First, a model for attitudinal variables, which only considered individual characteristics as explanatory variables. Figure 1 presents the final structure of the selected model (several specifications were considered).

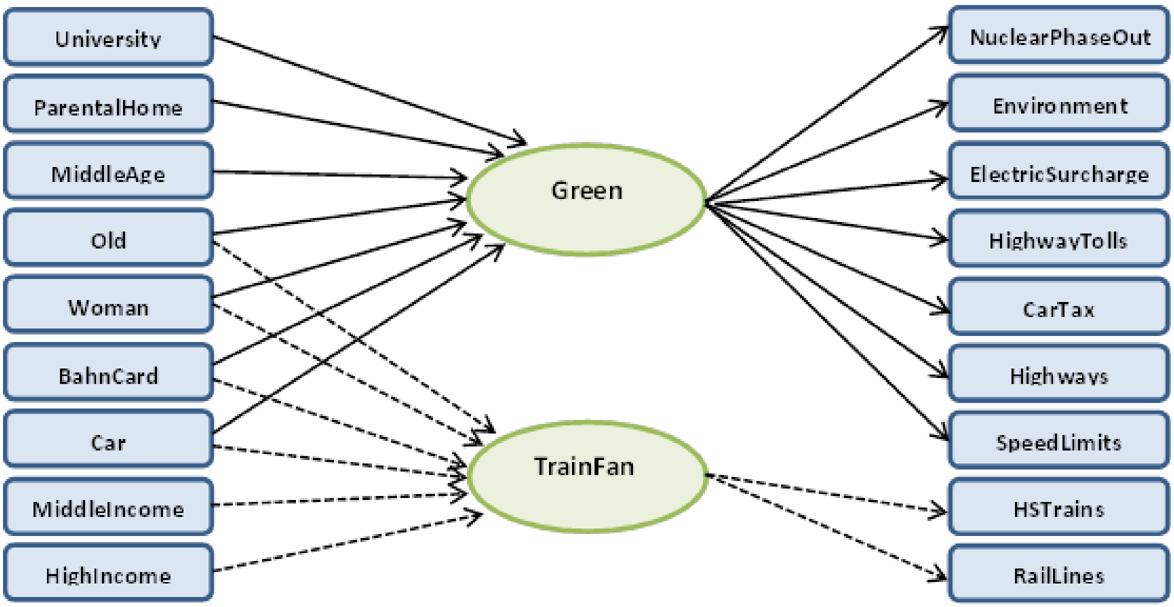

Figure 1 – Structure of the Attitudinal MIMIC model[4]

[4] The different line types are only used to ease the observation of the figure.

In this case, “University” is associated with working individuals holding this educational degree. “Parental Home” only applies to students and it indicates that the individual still lives at the parental home. “BahnCard” indicates that the individual holds a Deutsche Bahn yearly discount card (which is common in Germany due to the price discrimination policies adopted by Deutsche Bahn AG), while “Car” indicates automobile ownership. The remaining variables are self-explanatory. Table 2 presents the estimated parameters.

Table 2 – Estimated Parameters for the Attitudinal MIMIC model

Explanatory Variable	**Estimate**	**t-test**	**Attitudinal Indicator**	**Estimate**	**t-test**
Green Attitude			***Green Attitude***		
University	0.258	4.134	***NuclearPhaseOut***	1.463	44.416
ParentalHome	-0.181	-2.815	***Environment***	1.178	49.917
MiddleAge	0.298	6.124	***ElectricSurcharge***	1.666	51.02
Old	0.497	3.713	***HighwayTolls***	2.14	51.02
Woman	0.287	5.917	***CarTax***	1.628	46.489
BahnCard	0.334	6.565	***Highways***	-1.053	-37.228
Car	-0.524	-10.075	***SpeedLimits***	2.282	51.902
TrainFan			***TrainFan***		
Old	0.282	2.048	***HSTrains***	2.088	49.695
Woman	-0.282	-5.639	***RailLines***	2.12	49.937
BahnCard	0.333	6.356			
Car	-0.058	-1.116[5]			
MiddleIncome	0.141	2.807			
HighIncome	0.128	1.66[6]			

As stated above, a second MIMIC model was estimated for the perception indicators. In this case, not only the characteristics of the individuals but also the attributes of the transport modes were considered as explanatory variables. It is also important to consider interactions between these two kinds of variables, as different population groups perceive differently the attributes of the alternatives (i.e. systematic taste variations). The structure of the estimated model is shown in Figure 2. Here, “Losses” and “Accidents” indicate that the individual had suffered losses during a trip in the past or had been involved in a train accident, respectively. The number of transfers is represented by a discrete variable ranging between zero and four, while “BusUser” indicates whether the individual had

[5] The variable was kept in the model as it is considered a policy variable and has the proper sign.
[6] As the signs of the estimators were known *a priori*, a one-tailed test was performed ($\alpha_{5\%}$ =1.645).

undertaken at least one trip with coach services during the last three years. The estimation results are presented in Table 3.

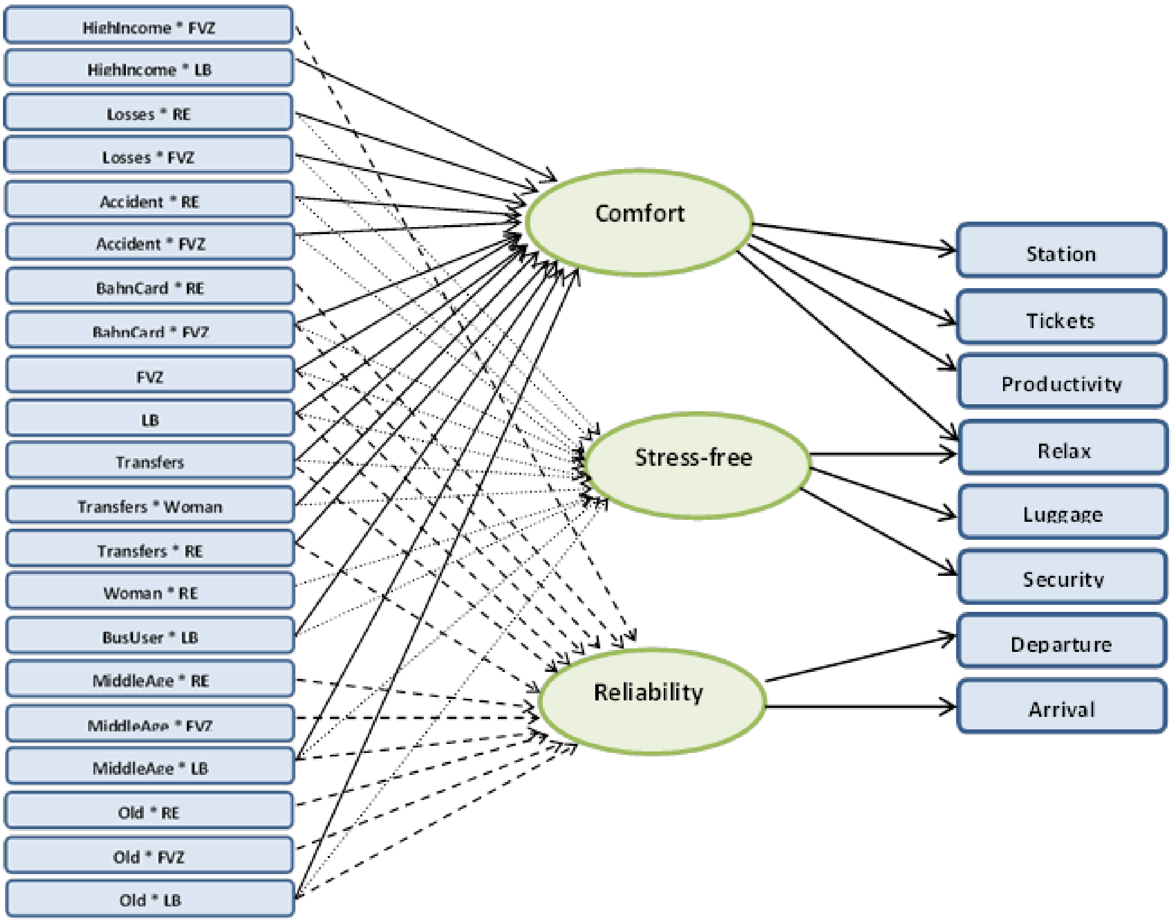

Figure 2 – Structure of the Perception MIMIC model

In line with our hypotheses, all explanatory variables affecting perceptions are directly related with the specific alternatives for which they were calculated, whether considering the attributes directly or through systematic taste variations (e.g. *Transfers * Woman*). It is important to note that not considering the latter works as well, but provides a worse goodness-of-fit and therefore, a worse representation of the way in which the alternatives are apparently perceived by the individuals.

5.3 Discrete Choice

This section reports the results of the estimation of the discrete choice component of the model. In addition to the previously described latent variables, socioeconomic characteristics of the individuals and attributes of the alternatives (price, travel time, number of transfers, transport mode and safety level) were considered. In addition, an inertia variable taking the value of one when individuals chose their revealed preference option in spite of the advantages of new alternatives was introduced.

Table 3 – Estimated Parameters for the Perception MIMIC model

Explanatory Variable	Estimate	t-test	Attitudinal Indicator	Estimate	t-test
Comfort			***Comfort***		
HighIncome * LB	-0.249	-2.958	***Station***	1.403	64.12
Losses * RE	-0.28	-3.577	***Ticket***	1.136	57.733
Losses * FVZ	-0.173	-2.028	***Productivity***	1.935	72.199
Accident * RE	-0.243	-3.527	***Relax***	1.36	62.262
Accident * FVZ	-0.217	-2.816			
BahnCard * FVZ	0.341	7.199			
FVZ	0.471	12.704			
LB	-0.907	-23.639			
Transfers	-0.161	-8.05			
Transfers * Woman	-0.06	-1.765[7]			
Transfers * RE	0.059	2.448			
BusUser * LB	0.338	8.011			
MiddleAge * LB	-0.282	-5.917			
Old * LB	-0.522	-3.796			
Stress-free			***Stress-free***		
Losses * RE	-0.539	-6.808	***Relax***	0.719	36.669
Losses * FVZ	-0.468	-5.403	***Luggage***	2.119	72.975
Accident * RE	-0.188	-2.704	***Security***	1.546	64.834
Accident * FVZ	-0.246	-3.162			
BahnCard * FVZ	0.283	5.928			
FVZ	0.311	8.35			
LB	0.399	10.674			
Transfers	-0.12	-5.965			
Transfers * Woman	-0.144	-4.193			
Woman * RE	-0.207	-4.56			
BusUser * LB	0.104	2.438			
MiddleAge * LB	-0.128	-2.67			
Old * LB	-0.462	-3.327			
Reliability			***Reliability***		
LB	-0.227	-6.209	***Departure***	2.115	72.855
FVZ	0.063	1.724[7]	***Arrival***	2.284	74.573
Transfers	-0.158	-7.921			
Transfers * Woman	-0.099	-2.934			
HighIncome * FVZ	-0.173	-2.07			
BahnCard * RE	0123	4.155			
BahnCard * FVZ	0.083	2.583			
BusUser * LB	0.185	1.773[7]			
MiddleAge * RE	0.133	2.974			
MiddleAge * FVZ	0.083	1.754[7]			
MiddleAge * LB	-0.103	2.185			
Old * RE	0.133	0.974[8]			
Old * FVZ	0.166	1.212[8]			
Old * LB	-0.347	-2.536			

[7] As the signs of the estimators were known *a priori*, a one-tailed test was performed ($\alpha_{5\%}$ =1.645).
[8] The variables ware kept in the model, despite their low significance, as the sign as well as the magnitude of the estimated parameters were consistent with the values obtained for the other age related estimators.

Altogether, there were 13,138 observations available for estimation and the potential correlation (panel effect) between the responses of a given individual was taken into account. The latent variables "Green" and "TrainFan" were considered both linearly as well as categorized in three levels. Estimation was performed sequentially using BIOGEME (Bierlaire, 2003), so it was necessary to correct the estimates accounting for the bias described by Bahamonde-Birke and Ortúzar (2014b). Table 4 presents the estimation results for five different specifications.

Table 4 – Estimated Parameters for the Discrete Choice model

Variable	Model 1	Model 2	Model 3	Model 4	Model 5
Inertia	0.332 (13.33)	0.332 (13.34)	0.331 (13.32)	0.331 (13.31)	0.33 (13.31)
LB	-0.205 (-1.03)	-0.0538 (-0.28)	-0.25 (-1.26)	-0.0663 (-0.25)	-1.37 (-12.42)
RE	-0.029 (-0.32)	0.0553 (0.63)	-0.0585 (-0.64)	-0.0378 (-0.41)	-0.376 (-9.17)
Travel Time	-0.0159 (-25.15)	-0.0159 (-25.14)	-0.0154 (-23.8)	-0.0159 (-25.15)	-0.0156 (-24.85)
*Travel Time * LV Green*	-	-	0.000995 (0.99)	-	-
*Travel Time * LV Green (+67%)*	0.00248 (3.03)	0.00253 (3.07)	-	0.00243 (2.96)	0.00206 (2.57)
*Ln(Price) * Very Low Income*	-5.19 (-31.56)	-5.21 (-31.69)	-5.17 (-31.42)	-5.19 (-31.54)	-5.26 (-32.04)
*Ln(Price) * Low Income*	-4.66 (-26.38)	-4.64 (-26.32)	-4.67 (-26.4)	-4.66 (-26.39)	-4.66 (-26.48)
*Ln(Price) * Middle Income*	-3.73 (-15.36)	-3.7 (-15.27)	-3.74 (-15.39)	-3.74 (-15.37)	-3.63 (-15.18)
*Ln(Price) * High Income*	-2.78 (-8.5)	-2.72 (-8.32)	-2.76 (-8.43)	-2.79 (-8.5)	-2.66 (-8.26)
Safety Level	-0.00441 (-5.38)	-0.00445 (-5.42)	-0.00437 (-5.33)	-0.00444 (-5.4)	-0.00374 (-4.61)
Transfers	-0.207 (-4.28)	-0.215 (-4.47)	-0.201 (-4.16)	-0.216 (-4.34)	-0.442 (-18.34)
LV Comfort	0.564 (3.66)	0.756 (5.19)	0.511 (3.31)	0.661 (3.36)	-
LV Reliability	0.715 (2.77)	0.523 (2.07)	0.78 (3.01)	0.708 (2.74)	-
LV Stress-free	-	-	-	-0.124 (-0.79)	-
*FVZ * LV TrainFan*	0.426 (2.89)	-	0.439 (2.98)	0.409 (2.74)	0.758 (6.14)
*FVZ * LV TrainFan (+67%)*	-	0.0539 (0.74)	-	-	-
Log-Likelihood	-7425.05	-7428.952	-7429.123	-7424.737	-7454.932
ρ^2	0.185	0.184	0.184	0.185	0.181
Induced Variability	41,8%	39.9%	41,4.7%	43.6%	7.7%

As can be observed, two of the three perception indicators were found to be statistically significant. This way, both the perception of reliability and comfort affect positively the utility ascribed to a certain alternative. On the contrary, the perception of a stress-free travel appears not to be important in the decision making process (Model 4).

Note that when the perceptions attributes are omitted, the alternative specific constants (ASC) are highly significant (Model 5), but when perceptions are taken into account, they are able to capture almost the whole variability previously described by the ASC and the latter become statistically insignificant (Model 1). Also, everything is accompanied by a significant improvement in goodness-of-fit.

Regarding the attitudinal latent variables, it was found that our non-alternative related attitude (“Green”) affects the way in which travel time is perceived (i.e. it was possible to identify a systematic taste variation related to this attitude). As can be observed, when the variable was linearly considered, the interaction was not statistically significant (Model 3), but when only the individuals with a higher environmental concern (*LV Green +67%)* were taken into account, the systematic taste variation became important. This finding is in line with the perception that shorter travel times imply higher speeds and, therefore, more CO_2 emissions and a larger damage to the environment. Also, the fact that the effect of this variable is not linear, is in agreement with the notion that only highly environmentally concerned individuals are willing to accept larger travel times in order to reduce the ecological harm.

Finally, as expected, our alternative related attitude (“TrainFan”) is statically significant in conjunction with the intercity trains. It was possible to detect a social group of train enthusiasts willing to favour the railways in spite of the apparent advantages of other alternatives. However, this favouritism does not extend to regional trains. In this case, the categorized variable does not outperform the linear specification (Model 2). It was also not possible to identify a systematic taste variation within the alternative intercity trains.

6. CONCLUSIONS

Despite the significant technical and methodological improvements in the estimation of HDC models during the last decade, this has not led to a significantly better understanding of the way in which perceptions affect the decision making process, as these aspects are usually ignored by modellers. Even in the case of attitudes, which have been largely studied, the specification of latent variables has tended to be rather simplistic and rarely depart from the linearity assumptions (fortunately latent class models have been an alternative in this regard), while the analysis of systematic taste variations in association with attitudes appears to be practically inexistent.

This reticence may be related to deeper concerns about artificial constructs, such as latent variables and the information that can be acquired from them. Nevertheless, it should not be forgotten that we as modellers aim to depict reality in the best way possible, and therefore if we decide to work with latent variables, we should guide our efforts to represent as accurately as possible the decision making process and the fashion in which the different variables take part on it.

This paper gives an overview of the different ways in which attitudes and perceptions may affect the decision making process as well as providing practical recommendation about data collection and estimation issues.

Our empirical analysis provides evidence sustaining the fact that perceptions affect the way in which individuals ascribe a utility to a certain alternative. In the same line, our evidence shows that perceptions may explain a lot of the variability that is normally captured by the ASC, offering significant improvements in goodness-of-fit for the whole model. Also, our results sustains our hypotheses, in the sense that attitudes may be indeed related with systematic taste variations and that attitudinal latent variables should be treated in the same way as socio-economic variables.

Although we were able to identify systematic taste variations as well as a categorization for latent variables that outperform the linearity assumption, this does not imply that every attitudinal latent variable should be considered in this way. Prior to estimation, or even better, prior to constructing the experiment, the analyst should study which variables take part on the decision making process and decide the way in which they are considered in accordance with the underlying theory.

ACKNOWLEDGMENTS

We gratefully acknowledge the support of Becas Chile given by the Chilean Council for Scientific and Technological Research (CONYCIT), the Millennium Institute in Complex Engineering Systems (ICM: P05-004F; FONDECYT: FB016), the All Latitudes and Cultures BRT Centre of Excellence funded by the Volvo Research and Educational Foundations, the Alexander von Humboldt Foundation and the Centre for Sustainable Urban Development, CEDEUS (Conicyt/Fondap/15110020).

REFERENCES

van Acker, V., Mokhtarian, P.L. and Witlox, F. (2011). Going soft: on how subjective variables explain modal choices for leisure travel. *European Journal of Transport and Infrastructure Research* **11**, 115-147.

Alvarez-Daziano, R. and Barla, P. (2012). On the way to sustainable commuting to campus: stated, revealed, and attitudinal response using a mixed logit model. *Kuhmo Nectar Conference of ITEA 2012,* Berlin, 21-22, June, 2012.

Alvarez-Daziano, R. and Bolduc, D. (2013). Incorporating pro-environmental preferences towards green automobile technologies through a Bayesian hybrid choice model. *Transportmetrica* **9A**, 74-106.

Ashok, K., Dillon, W. and Yuan, S. (2002). Extending discrete choice models to incorporate attitudinal and other latent variables. *Journal of Marketing Research* **39**, 31-46.

Bahamonde-Birke, F.J., Raveau, S., Yáñez, M.F. and Ortúzar, J. de D. (2010). The role of tangible attributes in hybrid discrete choice models. *European Transport Conference 2010*, Glasgow, 11-13, October, 2010.

Bahamonde-Birke, F.J. and Ortúzar, J. de D. (2014a). On the variability of hybrid discrete choice models. *Transportmetrica* **10**, 74-88.

Bahamonde-Birke, F.J. and Ortúzar, J. de D. (2014b). Is sequential estimation a suitable second best for estimation of hybrid choice models? *93rd Annual Meeting of the Transportation Research Board 2014*, Washington, D.C., 12-16, January 2014.

Ben-Akiva, M.E., Walker, J.L., Bernardino, A.T., Gopinath, D.A., Morikawa, T. and Polydoropoulou, A. (2002). Integration of choice and latent variable models. In H.S. Mahmassani (ed.), *In Perpetual Motion: Travel Behaviour Research Opportunities and Challenges*. Pergamon, Amsterdam.

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

Bolduc, D., Boucher, N. and Alvarez-Daziano, R. (2008). Hybrid choice modeling of new technologies for car choice in Canada. *Transportation Research Record: Journal of the Transportation Research Board*, ***2082*****(1)**, 63-71.

Bolduc, D. and Alvarez-Daziano, R. (2010). On the estimation of hybrid choice models. In S. Hess and A. Daly (eds.), *Choice Modelling: The State-of-the-Art and the State-of-Practice,* 259–288. Emerald, Bingley.

Bollen, K.A. (1989). *Structural Equations with Latent Variables*. John Wiley and Sons, Chichester.

Domencich, T. and McFadden, D. (1975). *Urban Travel Demand – A Behavioural Analysis.* North Holland, Amsterdam.

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

Ortúzar, J. de D. and Willumsen, L.G (2011). *Modelling Transport*. Fourth Edition, John Wiley and Sons, Chichester.

Raveau, S., Alvarez-Daziano, R., Yáñez, M.F., Bolduc, D. and Ortúzar, J. de D. (2010). Sequential and simultaneous estimation of hybrid discrete choice models: some new findings. *Transportation Research Record* **2156**, 131-139.

Raveau, S., Yáñez, M.F. and Ortúzar, J. de D. (2012) Practical and empirical identifiability of hybrid discrete choice models. *Transportation Research* **46B**, 1374-1383.

Thurstone, L.L. (1927). A law of comparative judgment. *Psychological Review* **34**, 273 286.

Train, K.E., McFadden, D.L. and Goett, A.A. (1987). Consumer attitudes and voluntary rate schedules for public utilities. *Review of Economics and Statistics* **64**, 383-91.

Vredin-Johansson, M., Heldt, T. and Johansson, P. (2006). The effects of attitudes and personality traits on mode choice. *Transportation Research* **40A**, 507-525.

Walker, J.L. (2002). The mixed logit (or logit kernel) model: dispelling misconceptions of identification. *Transportation Research Record* **1805**, 86-98.

Zellner A. (1970). Estimation of regression relationships containing unobservable variables. *International Economic Review* **11**, 441-454