



## How to categorize individuals on the basis of underlying attitudes? A discussion on latent variables, latent classes and hybrid choice models

Francisco J. Bahamonde-Birke & Juan de Dios Ortúzar

To cite this article: Francisco J. Bahamonde-Birke & Juan de Dios Ortúzar (2021) How to categorize individuals on the basis of underlying attitudes? A discussion on latent variables, latent classes and hybrid choice models, *Transportmetrica A: Transport Science*, 17:4, 856-877, DOI: [10.1080/23249935.2020.1817171](https://doi.org/10.1080/23249935.2020.1817171)

To link to this article: <https://doi.org/10.1080/23249935.2020.1817171>



© 2020 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 22 Sep 2020.



Submit your article to this journal [↗](#)



Article views: 1436



View related articles [↗](#)



View Crossmark data [↗](#)



Citing articles: 1 View citing articles [↗](#)

# How to categorize individuals on the basis of underlying attitudes? A discussion on latent variables, latent classes and hybrid choice models

Francisco J. Bahamonde-Birke<sup>a</sup> and Juan de Dios Ortúzar <sup>b</sup>

<sup>a</sup>Sociale Geografie en Planologie, Universiteit Utrecht, Utrecht, the Netherlands; <sup>b</sup>Department of Transport Engineering and Logistics, Instituto Sistemas Complejos de Ingeniería, BRT+ Centre of Excellence, Pontificia Universidad Católica de Chile, Santiago, Chile

## ABSTRACT

Classifying individuals into distinctive groups in discrete choice modelling work has become a common procedure since latent class models were introduced in the field. However, latent classes have certain shortcomings regarding the interpretation of the identified classes. On the other hand, hybrid choice models incorporating latent variables and psychometric data, are a powerful tool to treat and identify some underlying attitudes affecting behaviour; however, the treatment of the latent variables into the utility function has not been analysed in sufficient depth. Latent variables accounting for attitudes resemble socio-economic characteristics and, therefore, both systematic-taste-variations and categorizations may also be considered.

We examine different ways to categorize individuals based on latent characteristics, and explain why this may be convenient. In particular, we propose a direct categorization of individuals based on underlying latent variables and conduct theoretical analyses contrasting this method with existing approaches. Based on this analysis, we conclude that some of the methods used in the past exhibit certain shortcomings that can be overcome by relying on a direct categorization. Then, we show some of the analytical advantages of the approach with the aid of two illustrative examples. Furthermore, the proposed approach allows bridging the gap between latent classes and latent variable models.

## ARTICLE HISTORY



Received 17 March 2020  
Accepted 22 August 2020

## KEYWORDS

Hybrid discrete choice models; latent variables; latent classes; categorization

## 1. Introduction

Since the introduction of latent class (LC) models (Kamakura and Russell 1989), classifying individuals into distinctive behavioural groups has become a standard procedure in discrete choice modelling work. LC models consider a discrete mixture distribution, allowing to associate every step of it with a distinctive behavioural class, establishing a probability that an individual belongs to each class. Initially, these models were developed to form groups with clear potentially-identifiable behavioural patterns and differently

**CONTACT** Francisco J. Bahamonde-Birke  bahamondebirke@gmail.com  Sociale Geografie en Planologie, Universiteit Utrecht, Princetonlaan 8a, 3584 CB, Utrecht, the Netherlands

© 2020 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group  
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

specified utility functions (e.g. individuals without all alternatives available or ignoring a given attribute in the decision). Now, in practice, LC models have been extended to represent preference heterogeneity by simply allowing different estimators in otherwise equally specified classes following a somewhat exploratory approach. Then, after careful observation of the estimated parameters, the different classes are associated with underlying unobservable characteristics of decision-makers, such as attitudes or perceptions (e.g. DeSarbo, Ramaswamy, and Cohen 1995; Bujosa, Riera, and Hicks 2010). However, this means that the latent classes identified by the model (as their identification improves the adjustment) may be associated with preconceived ideas in a rather *ad-hoc* fashion.

An alternative approach to treat the heterogeneity caused by individuals' underlying characteristics is the so-called *hybrid discrete choice* (HDC) models (McFadden 1986; Train, McFadden, and Goett 1987). Starting from the usual hypotheses of discrete choice models (McFadden 1974), HDC models aim at enriching the functional forms by incorporating psychometric data, collected in addition to the individual choices in the form of indicators (usually stated on a Likert-scale; Likert 1932<sup>1</sup>). In an HDC framework, the psychometric information is not used directly in the behavioural model, as indicators are not deemed to be causal. Rather, they are an expression of underlying unobservable/unobserved characteristics of the individuals (Ben-Akiva et al. 2002), which are, in turn, modelled through latent variables (LV).

Then, the indicators are considered a function of these underlying LV, via so-called *measurement equations*. The LV are, then incorporated into the HDC model as a continuous mixing distribution (Ortúzar and Willumsen 2011, Chapter 8). Hence, and opposite to the LC-framework, the representation of unobserved attributes is based on additional psychometric data in HDC models. Further, the LV are typically interpreted based on a previously defined Multiple Indicator Multiple Causes (MIMIC) model (Zellner 1970), and not on *ad-hoc* hypotheses (as a causal relationship is assumed before the estimation of the model, given the directional nature of the MIMIC framework).

The HDC modelling approach has become popular and used by different disciplines ranging from marketing (Ashok, Dillon, and Yuan 2002; Luo, Kannan, and Ratchford 2008; Lam et al. 2010) to labour economics (Humlum, Kleinjans, and Nielsen 2012), as well as spatial economics (Hurtubia and Bierlaire 2014), tourism (Chia-Jung and Pei-Chun 2014), environmental economics (Hess and Beharry-Borg 2012) and transportation (van Acker, Mokhtarian, and Witlox 2011; Kim, Rasouli, and Timmermans 2016). However, the HDC framework has been subject to the criticism that the model's adjustment should be worse than or equal to that provided by a similarly specified Mixed Logit (ML) model (Vij and Walker 2016). By using the same number of explanatory variables at the level of the reduced utility functions, the HDC model aims at maximizing the choices and the indicators simultaneously. This fact, however, should not be considered a shortcoming, as satisfying both the choices and relevant behavioural indicators increases the model reliability (avoiding an eventual spurious over adjustment).

Furthermore, being able to disentangle which proportion of a given estimator (including the alternative specific constants, ASCs) can be attributed to the valuation of a variable *per se* and which can be ascribed to attitudes and/or perceptions, offers significant insights to understand the actual behavioural mechanisms behind decision-making (these advantages add to the mathematical advantages described by Vij and Walker 2016).

However, despite the extensive use of HDC models, modellers have not extensively analysed how to consider different types of LV in the utility function; and generally, introduce them in fairly simplistic ways. Bahamonde-Birke et al. (2017) argued that as attitudes resemble socioeconomic characteristics of individuals, one should not expect a linear impact in the utility function. That being the case, systematic taste variations (Ortúzar and Willumsen 2011, 279) or a discrete treatment of LV should be considered, rather than just including variables accounting for attitudes in an additive fashion.

If we do not expect LV to have a linear impact on utility, and it is not straightforward to decide on a parametric functional form, it may be preferable to describe their effects using a categorical non-parametric treatment. Note that such treatment (or categorizing individuals based on socioeconomic variables), does not imply that the LV's intensity is being ignored. We can use several categories to represent different intensities in a non-parametric fashion (e.g. most analysts would favour categorizing individuals by age or income group over testing age or income directly in the utility function). For instance, when using LV to describe underlying environmental attitudes or unreported information (i.e. lacking income data), it may be preferable to identify green-minded individuals or to classify them by income level (Bahamonde-Birke and Hanappi 2016). It is well-established that such attributes do not have a linear-continuous impact on decision-making. Furthermore, a categorical representation allows redefining the original variable piecewise, allowing for a discrete-continuous representation of the phenomenon.

Categorizing individuals based on LV depicts *de facto* a discrete mixture model, akin to LC models (beyond theoretical issues, related to directionality assumptions), where every category represents a different LC. This procedure has the added value that classes are described by behavioural variables, which, in turn, are based on psychometric data. In fact, by starting from a LC model and attempting to enrich class identification using indicators (Ben-Akiva and Boccara 1995), the modeller faces a similar problem. Hence, for all policy implications, models with discretized LV should be treated in the same way as LC models. Thus, a categorical treatment of LV allows bridging the gap between the latent variable and latent class components in a generalized choice model (Walker and Ben-Akiva 2002<sup>2</sup>).

In this paper, we conduct a theoretical analysis, contrasting different approaches to enrich the categorization of individuals based on psychometric data, as well as to link latent variables and latent classes highlighting their advantages and shortcomings. From this analysis, we propose a new method to construct latent classes using a direct categorization based on LV, establishing a direct and clear link between them. We conclude that some of the previous methods exhibit certain shortcomings that may be overcome by relying on a new direct categorization, as proposed in this paper. Then, we show some of the analytical advantages of the approach with the help of two illustrative examples. Our discussion offers valuable insights into how attitudes and other unobserved variables could be treated in HDC models.

## 2. Methodological framework

In an HDC framework (Walker and Ben-Akiva 2002), individuals are assumed to exhibit utility functions with the following form (under the assumption of additive linearity):

$$U_q = X_q \cdot \beta_X + \eta_q \cdot \beta_\eta + \varepsilon_q \quad (1)$$

where  $U_q$  is a vector of utilities associated with the  $J$  alternatives in the choice-set of each individual  $q$ ;  $X_q$  is a matrix of observed attributes of alternatives  $j \in J$  and characteristics of the individuals;  $\eta_q$  is a matrix representing unknown latent variables (associated with individual  $q$  and the various alternatives  $j$ ). Finally,  $\beta_x$  and  $\beta_\eta$  are matrices of parameters to be estimated ( $\beta = [\beta_x' | \beta_\eta']'$ ), and  $\epsilon_q$  an error term.

If  $\epsilon_q$  follows an EV1 distribution with mean zero and a diagonal covariance matrix  $\Sigma_U$ , the choice probabilities (conditional on  $\eta_q$ ) are given by a Multinomial Logit (MNL) model (Domencich and McFadden 1975). The LV accounting for unobserved characteristics of the individuals (e.g. attitudes) or alternatives (e.g. perceptions) are then constructed in accordance with a MIMIC model and are a function of positively observed explanatory variables and, eventually, other latent constructs (Link 2015). This way, assuming a linear additive specification, these *structural equations* may be written as follow:

$$\eta_q = X_q \cdot \alpha_\gamma + \eta_q^* \cdot \alpha_\eta + \nu_q \tag{2}$$

where  $X_q$ , once again, is a set of observed explanatory variables;  $\eta_q$  and  $\eta_q^*$  are vectors of latent constructs (which may be alternative specific or not) and  $\nu$  an error term following any distribution (usually Normal with mean zero and a given covariance matrix  $\Sigma_\eta$ ). Finally,  $\alpha_\gamma$  and  $\alpha_\eta$  are matrices of parameters to be estimated (and  $\alpha = [\alpha_\gamma' | \alpha_\eta']'$ <sup>3</sup>).

In this framework, the *structural equations'* set is unidentified and it is mandatory to consider it jointly with a set of *measurement equations*. Notwithstanding, even when considering both sets of equations jointly, to achieve identification it is necessary to fix certain parameters without loss of generality – typically the variances of the structural equations (Vij and Walker 2014). Assuming a linear specification, the latter may be represented as follows:

$$I_q = X_q \cdot \gamma_x + \eta_q \cdot \gamma_\eta + \eta_q^* \cdot \gamma_{\eta^*} + \varsigma_q \tag{3}$$

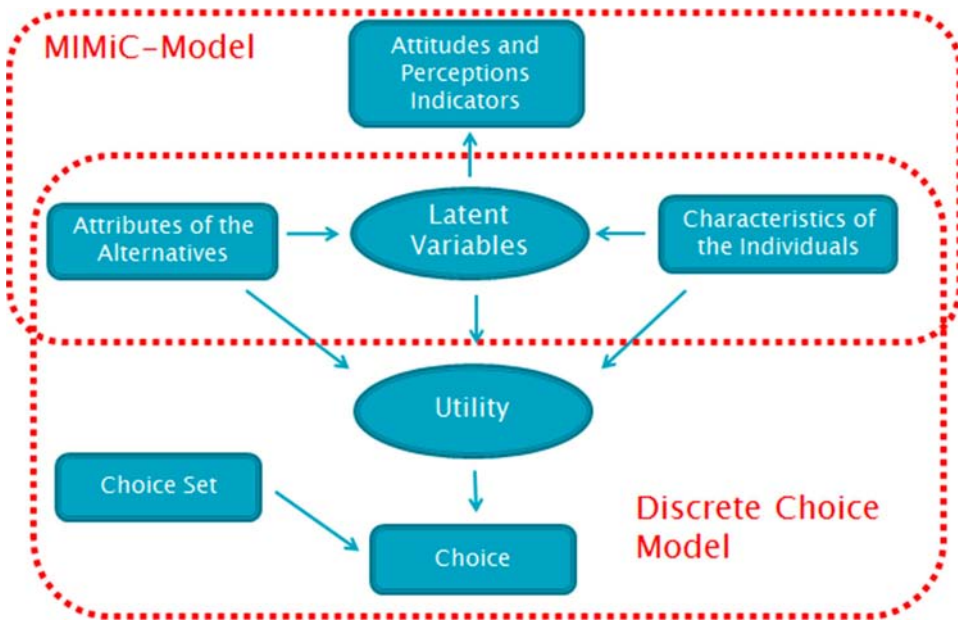
where  $I_q$  is a vector of indicators (considered to be an expression of underlying LV),  $X_q$  is the matrix of observed explanatory variables,  $\varsigma_q$  is an error term, the distribution of which depends on the assumptions regarding the indicators (i.e. when the indicators are continuous outputs,  $\varsigma_q$  is usually assumed to be Normal; in turn, when they have a discrete nature,  $\varsigma_q$  may be assumed to follow a Logistic distribution); in any case, the distributions have zero mean and a diagonal covariance matrix  $\Sigma_I$ ; finally,  $\gamma_x$  and  $\gamma_\eta$  are matrices of parameters to be estimated ( $\gamma = [\gamma_x' | \gamma_\eta' | \gamma_{\eta^*}]'$ ).

Figure 1 describes the typical structure of a hybrid discrete choice (HDC) model.

It is convenient to estimate the model simultaneously (i.e. MIMIC and discrete choice models jointly) as a sequential estimation may lead to biased results (Bahamonde-Birke and Ortúzar 2014a, 2014b). In this case, the likelihood function takes the following form:

$$L_q = \int_U P(y_q | X_q, \eta_q; \alpha, \beta, \Sigma_U, \Sigma_\eta) \cdot P(I_q | X_q, \eta, \eta_q^*; \alpha, \gamma, \Sigma_I, \Sigma_\eta) \cdot f(\eta_q | X_q, \eta_q^*; \alpha, \Sigma_\eta) \cdot d\nu_q \tag{4}$$

where the first part corresponds to the likelihood of the discrete choice component, the second to that of observing the indicators and the third to the distribution of the LV over which the likelihood function must be integrated.



**Figure 1.** Hybrid discrete choice model with latent variables.

## 2.1. Categorizing individuals on the basis of unobserved variables

This is not an easy subject (in contrast with observed socio-economic characteristics, such as income or age), as attitudinal variables are not observed and consequently exhibit intrinsic variability (Walker and Ben-Akiva 2002). Under these circumstances, it is not possible to associate individuals with specific groups, but only to establish a probability that they belong to them. Moreover, as unobserved variables do not exhibit an unequivocal scale, it is not easy to establish thresholds for the categorization and, as such, the process may appear to be somewhat arbitrary. Below we discuss some methods reported in the literature for this purpose, and propose a new method (DCLV) that allows overcoming some of their shortcomings.

### 2.1.1. Latent classes with psychometric indicators approach (LCPI)

This approach (Hurtubia et al. 2014) does not attempt to categorize individuals on the basis of an HDC and unobserved latent variables *per se*, but rather to include psychometric indicators into a LC framework, enriching the categorization process. In this structure, class-membership is assumed to depend exclusively on positively observed individual characteristics ( $X_q$ ), while the probability of belonging to a certain class is given by a Multinomial Logit (MNL) model. However, as the authors only consider two different classes, their model can also be expressed as an Ordinal Logit model (McCullagh 1980). For the purposes of this work (for comparison reasons), when classes are an expression of an underlying latent variable, the ordinal interpretation seems more appropriate. Then, assuming that the error term follows a Logistic distribution with zero mean and scale parameter  $\lambda$  (fixed without loss of

generality), the probability of an individual belonging to a certain class  $k$  is given by:

$$P_{qk} = \frac{1}{e^{\lambda(\alpha_X \cdot X_q - \psi_T)} + 1} - \frac{1}{e^{\lambda(\alpha_X \cdot X_q - \psi_B)} + 1} \quad (5)$$

where  $\psi_B$  and  $\psi_T$  represent the lower (bottom) and upper (top) thresholds describing the class, respectively (and must be estimated). The main difference between this approach and the HDC framework, is that here the LV impact only indirectly on the measurement equations. This way, these are just a function of positively observed individual characteristics ( $X_q$ ) and the estimated parameters depend on the individual's LC. Thus, under this approach equation (3) has the following structure:

$$I_q = X_q \cdot \gamma_k + \varsigma_q \quad (6)$$

where  $\gamma_k$  is a LC specific matrix of parameters. As in this case no LV are being considered, equation (1) simply takes the following form:

$$U_q = X_q \cdot \beta_k + \varepsilon_q \quad (7)$$

leading to the likelihood function (8):

$$L_q = \sum_k P(y_q | X_q; \alpha, \beta_k, \Sigma_U) \cdot P(I_q | X_q; \alpha, \gamma_k, \Sigma_I) \cdot P_{qk} \quad (8)$$

where  $\beta_k$  is a LC specific matrix of parameters. Note that while (6) and (7) do not have common parameters, they are connected with each other via (5) as  $\gamma_k$  and  $\beta_k$  are LC-specific. This approach offers computational advantages, as it requires no integration over the domain of the LV.

Nevertheless, despite enriching the categorization process with additional information, the absence of causality assumptions between classes and outcomes in the specification of the model, which is something inherent to all LC models, remains in the LCPI approach.

In contrast with the HDC framework, where it is assumed that the stated indicators are an expression of underlying attitudes or perceptions (and the analyst is forced to make assumptions, specifying *a priori*, which indicators are going to be explained by which LV), in the LCPI no *ex-ante* causality assumptions are required (note that these causality assumptions refer to the specification of the model and not to causality *per se*, which cannot be directly established). Moreover, in most cases the analyst cannot explicitly specify different structures to be tested for the different classes, and the classes are simply defined by the optimization process to improve the likelihood of observing the outcomes.<sup>4</sup> As a consequence, in an LCPI structure, the indicators are mainly a tool to improve class-identification (individuals are more or less likely to belong to a certain LC given their stated indicators) and it is not easy for the analyst to guide the model towards identifying a specific unobserved/unobservable trait. Given this inherent difficulty, associating a given latent class with a specific characteristic is mostly done *a posteriori* after considering the estimated parameters; but this often appears to be a somewhat *ad-hoc* approach to classify individuals.

Furthermore, as the complex structure of the LCPI framework (requiring the joint consideration of several sets of LC specific measurements equations) is very demanding in terms

of data variability, empirical identification problems (Cherchi and Ortúzar 2008) may also arise.

However, it is important to acknowledge that the LCPI framework will offer the best possible adjustment to the data, as LC represent the optimal classification given input and output variables. Thus, it carries the potential to replicate the data more accurately than any other classification approach. But, as previously mentioned, it is challenging for the analyst to elucidate which are the actual underlying behavioural triggers. Nevertheless, when using LC in a confirmatory fashion – rather than in an exploratory fashion – it becomes easier to ascribe a meaning to the identified classes (El Zarwi, Vij, and Walker 2018), although the difficulties to guide the model towards the identification of a specific group remain given the usual impossibility of specifying LC differently (especially when dealing with attitudes or perceptions resulting in systematic taste variations).

Finally, as the approach aims at identifying behavioural groups only (which are of discrete nature), it is not possible to extend it to consider a discrete-continuous treatment of unobserved variables (which may be interesting for the analyst in several instances, for example, when considering missing income data).

**2.1.2. Latent variable latent class approach (LVLC)**

An alternative approach to classify individuals into distinct behavioural groups, while allowing for the analyst to explicitly specify the links between indicators and latent constructs, is the LVLC method (Hess, Shires, and Jopson 2013; Motoaki and Daziano 2015). It relies on a latent variable  $\eta^c$  (modelled within an HDC framework – (2), (3) and (4)) for the categorization. This way, the likelihood of an individual belonging to a certain class (i.e. being categorized in a certain way) is given by the probability of  $\eta^c$  being smaller (or larger) than a set of thresholds (to be estimated), under a probabilistic categorizing function.<sup>5</sup> Basically, it means including  $\eta_c$  into a typical LC framework.

If we assume that the class probability is given by a Logistic function, the likelihood of an individual belonging to a given class can be written as follows (Hess, Shires, and Jopson 2013<sup>6</sup>):

$$P_{qk} = \frac{1}{e^{\lambda(\eta_q^c - \psi_T)} + 1} - \frac{1}{e^{\lambda(\eta_q^c - \psi_B)} + 1} \tag{9}$$

$$P_{qk} = \frac{1}{e^{\lambda(X_q \cdot \alpha_{Xc} + \eta_q^* \cdot \alpha_{\eta c} + v_c - \psi_T)} + 1} - \frac{1}{e^{\lambda(X_q \cdot \alpha_{Xc} + \eta_q^* \cdot \alpha_{\eta c} + v_c - \psi_B)} + 1}$$

Then, the utility functions are given by:

$$U_q = X_q \cdot \beta_{Xk} \cdot + \eta_q \cdot \beta_{\eta} + \varepsilon_q \tag{10}$$

where  $\beta_{Xk}$  is a LC specific matrix of parameters and  $\eta_q$  is a vector of continuously treated LV. Then, the overall likelihood function for the LVLC framework is given by:

$$L_q = \sum_k \int_v \frac{P(y_q | X_q, \eta_q; \alpha, \beta_k, \Sigma_U, \Sigma_{\eta}) \cdot P(I_q | X_q, \eta; \alpha, \gamma, \Sigma_I, \Sigma_{\eta})}{f(\eta_q | X_q, \eta_q^*, \alpha, \Sigma_{\eta})} \cdot P_{qk} \tag{11}$$

Unfortunately, this approach implies adding the variability associated with the distribution of the categorizing function (or of the LC model, if we approach modelling from this perspective) to the LV's own error term  $v$  (see [2]). In fact, if just for explanatory purposes we



wrote equation (9) in terms of utility functions (which is not entirely appropriate in this case) rather than cumulative probabilities (and considering just the first class, where  $\psi_B = -\infty$ ), we would obtain:

$$\begin{aligned} U_{\neq k} &= \psi_T \\ U_{\in k} &= X \cdot \alpha_{Xc} + \eta^* \cdot \alpha_{\eta c} + v_c + \varepsilon \end{aligned} \quad (12)$$

where  $\varepsilon$  follows a Logistic distribution (alternatively  $\varepsilon$  could be the difference of two independent and identically distributed EV1 error terms – one of them associated with the upper equation in [12]). From here, it is obvious that by including this stochastic unobserved variable into the LC framework, the analyst is (artificially) imposing that the magnitude of the error term associated with the LV has to be larger when considering it in the classification function than in the measurement equations. The reasons for this relate to the structure of the model: when considering random error terms, their effects are necessarily additive and cannot cancel out (as variances are additive); thus, when considering random disturbances in a model, we must allow for their impact to be zero. However, in the LVLC specification, the modeller considers two additive error terms, the variability of which is, by construction, non-zero.

### 2.1.3. Direct categorization of individuals on the basis of latent variables approach (DCLV)

To overcome the aforementioned problem in the treatment of the error terms, this paper introduces a new method based on a direct categorization on the basis of LV. It basically consists in associating an individual with a latent class  $k$  if a given latent variable  $\eta^c$  falls between a set of thresholds  $\underline{\psi}$  (without using an additional categorizing function). This way, it is possible to establish a probability that a certain individual should be categorized in a given way. Then, it should be possible to estimate the thresholds, and the probability of an individual being categorized in certain fashion ( $P_{qk}$ ) would be given by (13):

$$\begin{aligned} P_{qk} &= P(\psi_B < \eta_q^c < \psi_T | s, \alpha, \Sigma_\eta) \\ P_{qk} &= P(X_q \cdot \alpha_{Xc} + \eta_q^* \cdot \alpha_{\eta c} + v_q^c < \psi_T) - P(X_q \cdot \alpha_{Xc} + \eta_q^* \cdot \alpha_{\eta c} + v_q^c < \psi_B) \end{aligned} \quad (13)$$

Note that this approach avoids inducing extra variability into the categorization process (as the LVLC does), because no additional error term is added. Unfortunately, HDC models do not have a closed form and, therefore, they are normally estimated using simulation methods (Ben-Akiva et al. 2002; Bierlaire 2003). This way, we do not observe a continuous distribution for the error term  $v$  of the latent variable  $\eta^c$ , but a set of stochastic (or pseudo-stochastic) draws describing a probability function. As a consequence, discontinuity issues arise, the threshold cannot be estimated and perfect convergence of the algorithm is not theoretically warranted.<sup>7</sup>

An alternative to overcome this problem would be to rely on auxiliary variables. First, we define an auxiliary latent variable to be categorized  $\eta_q^a$ , following the exact same specification (i.e. same specification and same parameters) as  $\eta_q^c$ , such that:

$$\eta_q^a = X_q \cdot \alpha_Y + \eta_q \cdot \alpha_\eta + v_q^a \quad (14)$$

Here, the error term  $v_q^a$  may be considered via simulation and, consequentially,  $\eta_q^a$  can be used to compute the probabilities every time the actual numeric and continuous value

of  $\eta^c_q$  is required. In turn, to compute the likelihood of being categorized in a given fashion, the actual latent variable  $\eta^c_q$  (i.e. not the auxiliary one  $\eta^a_q$  the error term of which is computed via simulation) is used. If we then impose that the distribution of the error terms  $v^a_q$  and  $v^c_q$  (which is the same by construction) is such that their cumulative distribution function can be represented through a smooth closed-form expression (defined for all real numbers<sup>8</sup>), the probability that  $\eta^c_q$  be larger than a certain threshold, depicted in (13), is given by a closed-form expression (which allows estimating the thresholds). A convenient assumption satisfying this condition, is to assume an i.i.d. Logistic distribution with zero mean and a given scale parameter (fixed without loss of generality), so that (13) may be as an Ordinal Logit model:

$$P_{qk} = P(\psi_B < \eta^c_q < \psi_T | s, \alpha, \nu) = \frac{1}{e^{\lambda(X_q \alpha_{Xc} + \eta^* \alpha_{\eta c} - \psi_T)} + 1} - \frac{1}{e^{\lambda(X_q \alpha_{Xc} + \eta^* \alpha_{\eta c} - \psi_B)} + 1} \quad (15)$$

This formulation leads to a likelihood function of the form:

$$L_q = \sum_k \int_{\nu} P(y_q | X_q, \eta_q, \eta^a_q; \alpha, \beta_k, \Sigma_U, \Sigma_{\eta}) \cdot P(I_q | X_q, \eta_q, \eta^a_q; \alpha, \gamma, \Sigma_I, \Sigma_{\eta}) \cdot P_{qk} \cdot f(\eta_q | X_q, \eta^a_q; \alpha, \Sigma_{\eta}) \cdot d\nu \quad (16)$$

Note that the formulation of (16) is exactly the same as (11) except that the integral is only affected by  $\eta^a_q$ , while  $P_{qk}$  is only a function of  $\eta^c_q$ .

Finally, it is important to note that assuming i.i.d. terms for  $v^c_q$  and  $v^a_q$  implies that all observations are independent. This may appear to be a rather innocuous limitation, but has major implications when working with panel data (or pseudo-panel data as in stated choice experiments), as, in this case, integration over the domain of the latent variables should be performed at an individual level (Ortúzar and Willumsen 2011, Chapter 8). This problem can be reduced by including random panel effects into both  $\eta^c_q$  and  $\eta^a_q$ , but at least a proportion of variability would still be independent.

Another way to overcome the problem, while using the DCLV approach (but not auxiliary variables), would be to assume a distribution for the error terms such that a non-simulated-based estimation could be performed (e.g. as in the MACML approach of Bhat 2011). Another inconvenient is related with the non-monotonicity of the categorizing function. This leads to the existence of local optima and therefore the categorization's threshold will depend on the starting value. Nevertheless, this problem is common to all latent class approaches.

## 2.2. Empirical issues

As previously mentioned, the modeller can rely on a different number of categories to represent the intensity of the LV (or even on the combination of different categorized LV, akin to constructing classes combining different categorical socioeconomic variables, such as income or education level). To establish the optimal number of categories, the analyst may estimate models considering different numbers of classes. To judge if including an additional category improves the fit, the analyst can rely on the likelihood-ratio (LR) test (Ortúzar and Willumsen 2011), bootstrapping, or information criteria such as the AIC or BIC (Nyland et al., 2007).

Note that when working with several categories, the first and/or last thresholds may tend to diverge, so individuals categorized into the extreme categories would be associated with

the tails of the distribution. Even though, this could improve the model's goodness-of-fit, it may not translate into a better model, as these categories would just characterize outliers (i.e. this is equivalent to selecting a category with two or three individuals when categorizing on the basis of observed socioeconomic variables). In this case it may be advisable to fix exogenously the first and/or last thresholds.

Even though fixing a threshold *a priori* is an arbitrary decision (which again, must be in accordance with underlying theory), relying on the following heuristic may be useful: (i) estimate a model considering the desired number of categories plus one. (ii) ignore the first (or last) calibrated threshold and estimate a new model considering the desired number of categories; in this case, the first (or last) threshold is fixed at the level of the second (or second from last) estimated threshold in the first model.

Further, and even though in our derivation we considered only the categorization of a single latent variable, it is possible to extend this approach to categorize more variables. That being the case, the categories would be given by the combination of the categorized latent variables. For instance, when considering two categorized latent variables, we would have:

$$P_k = P(\psi_{B1} < \eta_{c1} < \psi_{T1} \wedge \psi_{B2} < \eta_{c2} < \psi_{T2} | s, \alpha, \Sigma_\eta) \quad (17)$$

And if we consider the error terms of both structural equations to be independent, then equation (17) can be written as:

$$P_k = P_{k1} \cdot P_{k2} = P(\psi_{B1} < \eta_{c1} < \psi_{T1} | s, \alpha, \Sigma_\eta) \cdot P(\psi_{B2} < \eta_{c2} < \psi_{T2} | s, \alpha, \Sigma_\eta) \quad (18)$$

It must be pointed out, however, that establishing categories on the basis of several latent variables may lead to a significant increase in estimation costs.

### 3. Illustrative examples

To illustrate the analytical considerations made in Section 2, we applied the LVLC and DCLV approaches to empirical data. A fair comparison with the LCPI approach was not possible due to its completely different structure and aims, so our conclusions relating to this approach are based exclusively on the previous theoretical arguments. As classes in the LCPI are defined exclusively to maximize goodness-of-fit, the approach has to be more likely than models estimated according to LV specifications. In fact, a simple LC model (without indicators) would always outperform all competing approaches as, in this case, the LC are defined with the single purpose of maximizing the goodness-of-fit of the observed choices (i.e. having to satisfy jointly both the observed choices and the psychometric indicators, can be understood as a restriction in the definition of the latent classes). This is related with the discussion raised by Vij and Walker (2016): a simple LC model represents the reduced form of both the LVLC and the DCLV (as a LC model is a ML with a discrete mixing distribution). Hence, a LV approach cannot outperform a LC model in terms of fit.

Similarly, the LCPI represents the reduced form of a discrete mixed model maximizing the likelihood associated with the representation of both choices and indicators simultaneously. Imposing any functional form on the way both types of variables are related, necessarily implies a worsening of fit (as it imposes a restriction in the categorization). However, if these functional forms are driven by theory, the model still can offer a better representation of the phenomenon (despite the worst fit). Furthermore, and also akin to

the discussion between the HDC models and the reduced ML, the representation based on the LVLC or the DCLV offers significant advantages in terms of behavioural insights.

Both case studies are only intended to serve as illustrative examples of the methodological issues discussed in the previous section. For a deeper discussion on their findings the reader is referred to Bahamonde-Birke and Hanappi (2016; Case 1) or to Bahamonde-Birke et al. (2014; Case 2).

### 3.1. Case 1 – electromobility in Austria

Our first example comes from the DEFINE project. A web-based survey conducted in Austria during February 2013, considered a stated preference (SP) vehicle purchase experiment set in the context of different electromobility options (Bahamonde-Birke and Hanappi 2016). The discrete choice experiment considered 787 respondents, and used a labelled experimental design, including four alternatives referring to different propulsion technologies: conventional vehicles (CV), plug-in hybrid-electric vehicles (PHEV), hybrid-electric vehicles (HEV) and electric vehicles (EV). Each alternative was described by the following attributes: purchase price (PP), power (PS), fuel costs (FC) and maintenance costs (MC). In addition, the EV were further characterized by the following attributes: full driving range (RA), availability of charging stations (LS) and policy incentives (IM). Charging station availability varied across three categories (low, intermediate and high) and was described qualitatively using a separate pop-up box. Policy incentives included a Park & Ride subscription for one year (IM2), investment subsidies to support private charging stations (IM3), and a one-year-ticket for public transport (IM4).

Additionally, attitudinal indicators related to the degree of agreement with eight different sentences were collected. Bahamonde-Birke and Hanappi (2016) considered the following five of them to construct a latent variable related to the environmental attitude of the individuals: *'I am an ecologically aware person'* (EcAwareness), *'I pay attention to regional origins when shopping foods and groceries'* (LocalFood), *'I buy ecologically friendly products'* (EcoFriendly), *'Environmental protection measures should be enacted even if they result in job losses'* (Protection) and *'I pay attention to the CO2 footprint of the products I buy'* (CO2Footprint). The level of agreement was stated using a six-point Likert scale. Table 1 presents an overview of the relevant variables for this study.<sup>9</sup>

Using PythonBiogeme (Bierlaire 2003), we estimated two models following both approaches presented in the previous section. Taking advantage of having just one latent variable, the log-likelihood was computed using numerical integration.

The latent variable (*LV Green*) accounts for environmental awareness and was discretized into two levels. Hence, the dichotomous latent variable *DLV Green* (taking a value of one if *LV Green* is larger than a threshold to be estimated and zero otherwise) defines the two different latent classes (note that the inclusion of a single dichotomous variable allows estimating different parameters for the entire class). The inclusion of the aforementioned dummy variable, allows us to differentiate between a latent class of more environmentally concerned individuals and a latent class of less environmentally concerned individuals.

The specification of the utility functions of both latent classes is basically the same, apart from the inclusion of *DLV Green* in association with the variables HEV, PHEV and BEV, which allows for estimating class-specific parameters for the latter. While usually LC models consider the estimation of a different sets of parameters for each class, here we imposed the

**Table 1.** Definition of the variables considered in the model. Case 1.

Variable	Definition
<i>MidSkill</i>	Dummy variable indicating a career and technical education.
<i>HighSkill</i>	Dummy variable indicating a college education or higher.
<i>Vienna</i>	Dummy variable indicating a residence in Vienna.
<i>Male</i>	Dummy variable indicating masculine gender.
<i>Old</i>	Dummy variable indicating individuals older than 60 years
<i>MidAge</i>	Dummy variable indicating individuals older than 35 years, but no older than 60 year.
<i>Carsharing</i>	Dummy variable indicating that the individual relies on Car Sharing on a regular basis.
<i>CarUser</i>	Dummy variable indicating that the individual drives to their main occupational activity on a regular basis.
PP	Purchase price in €.
FC, MC	Fuel and maintenance cost in € / 100 km., respectively.
PS	Power of the engine in hp.
RA	Driving range in km.
IM2, IM3, IM4	Dummy variables indicating the execution of the respective policy incentive.
<i>LSMid</i> , <i>LSHigh</i>	Dummy variables indicating medium or high availability of loading stations for EV.
<i>LV Green</i>	Latent variable Green accounting for environmental awareness.
<i>DLV Green</i>	Discretized latent variable Green accounting for individuals with high environmental awareness.
<i>EcAwareness</i>	Attitudinal Indicator for 'I am an ecologically aware person'.
<i>LocalFood</i>	Attitudinal Indicator for 'I pay attention to regional origins when shopping foods and groceries'.
<i>EcoFriendly</i>	Attitudinal Indicator for 'I buy ecologically friendly products'.
<i>Protection</i>	Attitudinal Indicator for 'Environmental protection measures should be enacted even if they result in job losses'.
<i>CO2Footprint</i>	Attitudinal Indicator for 'I pay attention to the CO2 footprint of the products I buy'.

restriction that most parameters were generic across classes and only one parameter was different. The reason was that the empirical exercise had an illustrative purpose and having both classes differing in only one parameter facilitated the illustration of the differences. However, note that considering a full-set of class-specific parameters is not only possible, but potentially one of the main advantages of the approach compared with HDC models with only continuous LV, as it easily allows addressing systematic taste variations among classes.

In this application, it was only possible to identify two statistically different classes and the model results are presented in Table 2. The t-test for statistical significance are presented in parenthesis and the log-likelihood for the overall model, as well as for the discrete choice (DC) component only, are also reported (the results for the measurement equations are shown in the appendix). The indicators were considered discrete outputs following an Ordered Logit model approach.

The differences between the parameter values obtained following both approaches are not large and many are not statistically different. This applies to all parameters of the structural and measurement equations as well as the utility functions, with the exception of the threshold parameter of the categorizing function. This last parameter evidently exhibits a different value, as the LVLC approach is associated, by construction, with a larger error and a wider distribution, implying that the threshold must be located further away from the expected value to capture the behaviour of the same individuals.

The DCLV approach exhibits a better goodness-of-fit than the LVLC, despite the fact that the latter has one extra degree-of-freedom (providing further evidence about the incorrect error treatment in the LVLC). This result is in accordance with our expectations, as the LVLC

**Table 2.** Parameter estimates. Case 1.

Variable	Equation	LVLC		DCLV	
Vienna	S.E. LV Green	<b>-0.333</b>	(-2.56)	<b>-0.182</b>	(-1.71)
Male	S.E. LV Green	<b>-0.597</b>	(-5.58)	<b>-0.587</b>	(-5.74)
HighSkill	S.E. LV Green	<b>1.05</b>	(6.16)	<b>0.879</b>	(5.89)
MidSkill	S.E. LV Green	<b>0.536</b>	(3.85)	<b>0.477</b>	(3.72)
Old	S.E. LV Green	<b>1.09</b>	(7.18)	<b>0.955</b>	(6.67)
MidAge	S.E. LV Green	<b>0.657</b>	(5.15)	<b>0.615</b>	(5.14)
Carsharing	S.E. LV Green	<b>1.33</b>	(4.9)	<b>0.988</b>	(4.99)
CarUser	S.E. LV Green	<b>-0.567</b>	(-5.53)	<b>-0.474</b>	(-5.7)
Threshold	C.F	<b>3.76</b>	(5.27)	<b>2.23</b>	(7.18)
Scale Parameter	C.F	<b>0.538</b>	(5.06)	<b>1</b>	(fixed)
ASC_CV	Utility CV	<b>0</b>	(fixed)	<b>0</b>	(fixed)
ASC_HEV	Utility HEV	<b>-0.0165</b>	(-0.07)	<b>0.0945</b>	(0.41)
ASC_PHEV	Utility PHEV	<b>-0.235</b>	(-0.94)	<b>-0.207</b>	(-0.88)
ASC_EV	Utility EV	<b>-1.09</b>	(-2.65)	<b>-0.982</b>	(-2.36)
PP	Utility CV	<b>-1.21</b>	(-7.92)	<b>-1.27</b>	(-9.07)
PP	Utility HEV	<b>-1.8</b>	(-22.39)	<b>-1.79</b>	(-22.85)
PP	Utility PHEV	<b>-1.76</b>	(-19.94)	<b>-1.77</b>	(-20.48)
PP	Utility EV	<b>-1.66</b>	(-9.79)	<b>-1.73</b>	(-10.69)
MC	Utility CV, HEV, PHEV, EV	<b>-21.1</b>	(-9.32)	<b>-20.6</b>	(-9.6)
FC	Utility CV, HEV, PHEV, EV	<b>-21.4</b>	(-14.96)	<b>-20.7</b>	(-15.87)
PS	Utility CV	<b>0.0327</b>	(5.22)	<b>0.0341</b>	(5.74)
PS	Utility HEV	<b>0.0353</b>	(7.18)	<b>0.0348</b>	(7.37)
PS	Utility PHEV	<b>0.0354</b>	(7.2)	<b>0.0367</b>	(7.5)
PS	Utility EV	<b>-0.000269</b>	(-0.05)	<b>-0.000232</b>	(-0.05)
PS * Male	Utility CV	<b>-0.019</b>	(-4.14)	<b>-0.0195</b>	(-4.21)
PS * Male	Utility HEV	<b>-0.0164</b>	(-3.38)	<b>-0.0164</b>	(-3.38)
PS * Male	Utility PHEV	<b>-0.0155</b>	(-3.16)	<b>-0.0162</b>	(-3.26)
PS * Male	Utility EV	<b>-0.00302</b>	(-0.55)	<b>0.000491</b>	(0.09)
MidAge	Utility HEV	<b>-0.266</b>	(-2.5)	<b>-0.287</b>	(-2.65)
MidAge	Utility PHEV	<b>-0.383</b>	(-3.46)	<b>-0.374</b>	(-3.48)
MidAge	Utility EV	<b>-0.784</b>	(-4.55)	<b>-0.946</b>	(-4.91)
Old	Utility HEV	<b>-1.18</b>	(-5.82)	<b>-1.19</b>	(-5.63)
Old	Utility PHEV	<b>-1.44</b>	(-6.24)	<b>-1.35</b>	(-6.07)
Old	Utility EV	<b>-2.55</b>	(-9.02)	<b>-2.88</b>	(-8.77)
DLV Green	Utility HEV	<b>3.48</b>	(6.86)	<b>3.19</b>	(6.56)
DLV Green	Utility PHEV	<b>3.25</b>	(6.08)	<b>2.71</b>	(4.84)
DLV Green	Utility EV	<b>5.77</b>	(8.5)	<b>5.89</b>	(9.83)
RA	Utility EV	<b>0.00444</b>	(6.25)	<b>0.00459</b>	(6.81)
LSMid	Utility EV	<b>0.0121</b>	(0.07)	<b>0.0426</b>	(0.25)
LSHigh	Utility EV	<b>0.661</b>	(4.22)	<b>0.71</b>	(4.43)
IM3	Utility EV	<b>0.306</b>	(2.19)	<b>0.305</b>	(2.14)
Log-likelihood					
Overall Model		<b>-16,177.759</b>		<b>-16,169.571</b>	
Log-likelihood					
DC Component		<b>-6,612.876</b>		<b>-6,605.847</b>	

considers an additional non-zero error component, which is unnecessary for model estimation. Along this line, the improvement in goodness-of-fit is mostly explained by the discrete choice component, and this is based on the fact that the categorization (and its additional error term) affects only the utility functions.

### 3.2. Case 2 – modal choice in Germany

In this SP experiment respondents were asked to choose between different interurban public transport alternatives in Germany (regional and intercity trains, and interurban

**Table 3.** Definition of the variables considered in the model. Case 2.

Variable	Definition
<i>Inertia</i>	Dummy variable taking the value of one for the <i>status quo</i> alternative.
<i>FVZ</i>	Dummy variable taking the value of one if the alternative considers the use of intercity trains.
<i>RE</i>	Dummy variable taking the value of one if the alternative considers the use of regional trains.
<i>LB</i>	Dummy variable taking the value of one if the alternative considers the use of coaches.
<i>Old</i>	Dummy variable indicating individuals older than 50 years.
<i>Bahncard</i>	Dummy variable indicating ownership of a Deutsche Bahn yearly discount card.
<i>Woman</i>	Dummy variable indicating feminine gender.
<i>VeryLowIncome</i>	Dummy variable indicating a net income under 700€ p.m.
<i>LowIncome</i>	Dummy variable indicating a net income between 700€ and 1,500€ p.m.
<i>MiddleIncome</i>	Dummy variable indicating a net income between 1,500€ and 2,500€ p.m.
<i>HighIncome</i>	Dummy variable indicating a net income over 2,500€ p.m.
<i>Price</i>	Travel fare in €.
<i>TravelTime</i>	Travel time in minutes.
<i>Transfers</i>	Number of transfers.
<i>SafetyLevel</i>	Number of severely injured passengers and the number of fatalities in the overall network over a year.
<i>LV TrainFan</i>	Latent variable TrainFan.
<i>DLV TrainFan</i>	Discretized latent variable accounting for individuals with high affinity for trains (two categories only).
<i>DLV TrainFan2</i>	Discretized latent variable accounting for individuals with medium affinity for trains (3 categories only).
<i>DLV TrainFan3</i>	Discretized latent variable accounting for individuals with high affinity for trains (3 categories only).
<i>HSTrains</i>	Attitudinal Indicator for 'Investing on the development of high-speed trains should be encouraged'.
<i>RailLines</i>	Attitudinal Indicator for 'New high-speed rail lines should be built'.

coaches). The experiment was carried out in three waves, contacting students and employees of two universities in Berlin as well as employees of member institutions of the Leibniz-Gemeinschaft (Bahamonde-Birke et al. 2014). Respondents were required to choose between a first pivotal alternative, representing a trip previously described, and a new one. Alternatives were described in terms of their travel time, fare, number of transfers, safety level and mode of transport: regional trains (RE), intercity trains (FVZ) and coaches (LB).

The original study considered several indicators, associated with different and complex latent variables. For the purposes of this work we only considered one LV (TrainFan), associated with the following two indicators: 'Investing on the development of high-speed trains should be encouraged' (HSTrains) and 'New high-speed rail lines should be built' (RailLines). Originally, the indicators were stated in a 10-point Likert scale. For computational simplicity, in this application we reduced it to only five, aggregating consecutive levels. Table 3 presents an overview of the variables considered.

As in the previous case, the *LV TrainFan* was used to construct the latent classes with the dichotomous variable *DLV TrainFan* taking a positive value if *LV TrainFan* was larger than a threshold. The only difference in the specification of the utility functions of both classes is that the valuation of the alternative FVZ is class-specific (as it interacts with *DLV TrainFan*).

Models following the LVLC and DCLV approaches were again estimated using Python-Biogeme and considering numerical integration for the computation of the likelihood function. Indicators were considered discrete outcomes (OLM). Table 4 presents the results for both models (again only two statistically different classes were identified). The structure of the table is the same as in the previous case (results for the measurement equations are presented, once more, in the appendix).

While in Case 1 all estimators (apart from the threshold) were not statistically different in both models, in Case 2 the parameters associated with the discretized LV are significantly different. As the discretized LV is considered in conjunction with the modal parameter of the intercity trains, it also affects the remaining modal parameters. This difference may be

**Table 4.** Parameter estimates. Case 2.

Variable	Equation	LVLC		DCLV	
<i>Old</i>	<i>S.E. LV TrainFan</i>	<b>0.207</b>	(0.36)	<b>-0.093</b>	(-0.41)
<i>Bahncard</i>	<i>S.E. LV TrainFan</i>	<b>0.684</b>	(5.41)	<b>0.673</b>	(5.46)
<i>Woman</i>	<i>S.E. LV TrainFan</i>	<b>-0.686</b>	(-5.64)	<b>-0.663</b>	(-5.75)
<i>MiddleIncome</i>	<i>S.E. LV TrainFan</i>	<b>0.368</b>	(3.19)	<b>0.426</b>	(3.24)
<i>HighIncome</i>	<i>S.E. LV TrainFan</i>	<b>0.515</b>	(2.53)	<b>0.606</b>	(2.82)
<i>Threshold</i>	C.F	<b>0.622</b>	(1.44)	<b>-0.346</b>	(-0.59)
<i>Scale Parameter</i>	C.F	<b>10</b>	(fixed)	<b>1</b>	(fixed)
<i>Inertia</i>	<i>Utility Alternative 1</i>	<b>0.336</b>	(13.48)	<b>0.345</b>	(13.48)
<i>FVZ</i>	<i>Utility Alternative 1 and 2</i>	<b>0</b>	(fixed)	<b>0</b>	(fixed)
<i>LB</i>	<i>Utility Alternative 1 and 2</i>	<b>-1.17</b>	(-12.65)	<b>-0.559</b>	(-12.62)
<i>RE</i>	<i>Utility Alternative 1 and 2</i>	<b>-0.167</b>	(-9.59)	<b>0.448</b>	(9.52)
<i>Travel Time</i>	<i>Utility Alternative 1 and 2</i>	<b>-1.49</b>	(-26.57)	<b>-1.56</b>	(-26.57)
<i>Ln(Price) * Very Low Income</i>	<i>Utility Alternative 1 and 2</i>	<b>-5.29</b>	(-31.91)	<b>-5.41</b>	(-31.9)
<i>Ln(Price) * Low Income</i>	<i>Utility Alternative 1 and 2</i>	<b>-4.7</b>	(-26.28)	<b>-4.85</b>	(-26.26)
<i>Ln(Price) * Middle Income</i>	<i>Utility Alternative 1 and 2</i>	<b>-3.63</b>	(-14.98)	<b>-3.8</b>	(-15)
<i>Ln(Price) * High Income</i>	<i>Utility Alternative 1 and 2</i>	<b>-2.54</b>	(-7.81)	<b>-2.67</b>	(-7.84)
<i>Safety Level</i>	<i>Utility Alternative 1 and 2</i>	<b>-0.00374</b>	(-4.59)	<b>-0.00408</b>	(-4.59)
<i>Transfers</i>	<i>Utility Alternative 1 and 2</i>	<b>-0.446</b>	(-18.25)	<b>-0.46</b>	(-18.25)
<i>FVZ * DLV TrainFan</i>	<i>Utility Alternative 1 and 2</i>	<b>0.716</b>	(4.93)	<b>1.42</b>	(5.27)
<i>Log-likelihood Overall Model</i>		<b>-10,543.642</b>		<b>-10,535.42</b>	
<i>Log-likelihood DC Component</i>		<b>-7,461.029</b>		<b>-7,453.871</b>	

attributed to the fact that changing the distribution (the variability) of the categorizing function may allow for identifying different groups of people. In fact, taking a look at the threshold parameters, it seems (accounting for the wider variability of the LVLC) that the thresholds have been set at a different level.

Notoriously in this case, when following the LVLC approach, the scale parameter  $\lambda$  of the categorizing function diverged (at least in computational terms). For that reason, it was necessary to fix it as 10 (larger values led to computational problems when computing the Fisher information matrix, but neither the estimates nor the final value of the log-likelihood were significantly affected).

The DCLV approach offered, again, a better adjustment than the LVLC model to the data; this may be mainly related to a better explanation of the discrete choices. Interestingly in this case, the scale parameter of the LVLC model suggests that the categorizing function should be as close to a Dirac delta function (Dirac 1958) as possible. In this limit, the LVLC model would collapse to the DCLV. Nevertheless, it is not possible to achieve this limit (or get close to it) due to computational limitations (variable overflow); therefore, it is impossible to get rid of the error induced through the LVLC approach.

In both cases studies we were only able to identify two statistically different classes; notwithstanding, and for illustrative purposes, we also present here a model considering three classes. In this case, the *LV TrainFan* was used to construct the dichotomous variables *DLV TrainFan1* – identifying a latent class of individuals moderately supportive of trains, and *DLV TrainFan2* – identifying a latent class of individuals highly supportive of trains. The third latent class corresponded to individuals for which none of the DLV are active, and identifies individuals with a low support for trains. The difference between the three latent classes is



**Table 5.** Parameter estimates. Case 2. Three categories.

Variable	Equation		DCLV
<i>Old</i>	<i>S.E. LV TrainFan</i>	<b>-0.107</b>	(-0371)
<i>Bahncard</i>	<i>S.E. LV TrainFan</i>	<b>0.648</b>	(6.28)
<i>Woman</i>	<i>S.E. LV TrainFan</i>	<b>-0.664</b>	(-6.27)
<i>MiddleIncome</i>	<i>S.E. LV TrainFan</i>	<b>0.426</b>	(3.97)
<i>HighIncome</i>	<i>S.E. LV TrainFan</i>	<b>0.603</b>	(3.40)
<i>Threshold1</i>	C.F	<b>-0.761</b>	(fixed)
<i>Threshold2</i>	C.F	<b>0.442</b>	(0.42)
<i>Scale Parameter</i>	C.F	<b>1</b>	(fixed)
<i>Inertia</i>	<i>Utility Alternative 1</i>	<b>0.344</b>	(13.17)
<i>FVZ</i>	<i>Utility Alternative 1 and 2</i>	<b>0</b>	(fixed)
<i>LB</i>	<i>Utility Alternative 1 and 2</i>	<b>-0.447</b>	(-1.77)
<i>RE</i>	<i>Utility Alternative 1 and 2</i>	<b>0.559</b>	(2.42)
<i>Travel Time</i>	<i>Utility Alternative 1 and 2</i>	<b>-1.55</b>	(-25.01)
<i>Ln(Price) * Very Low Income</i>	<i>Utility Alternative 1 and 2</i>	<b>-5.40</b>	(-30.42)
<i>Ln(Price) * Low Income</i>	<i>Utility Alternative 1 and 2</i>	<b>-4.85</b>	(-25.17)
<i>Ln(Price) * Middle Income</i>	<i>Utility Alternative 1 and 2</i>	<b>-3.81</b>	(-14.86)
<i>Ln(Price) * High Income</i>	<i>Utility Alternative 1 and 2</i>	<b>-2.68</b>	(-7.92)
<i>Safety Level</i>	<i>Utility Alternative 1 and 2</i>	<b>-0.00407</b>	(-4.77)
<i>Transfers</i>	<i>Utility Alternative 1 and 2</i>	<b>-0.459</b>	(-17.90)
<i>FVZ * DLV TrainFan2</i>	<i>Utility Alternative 1 and 2</i>	<b>1.03</b>	(1.76)
<i>FVZ * DLV TrainFan3</i>	<i>Utility Alternative 1 and 2</i>	<b>1.63</b>	(5.56)
<i>Log-likelihood Overall Model</i>		<b>-10,534.62</b>	
<i>Log-likelihood DC Component</i>		<b>-7,453.107</b>	

only the valuation of intercity trains (FVZ). In this case, it was necessary to fix the first threshold following the heuristic proposed in section 2 (for that purpose, a model considering four classes was previously calibrated). The results are presented in Table 5 (we only include the DCLV model, as fixing the first threshold does not allow for a fair comparison between both approaches<sup>10</sup>).

As can be observed, besides the latent class specific parameters, the remaining estimates do not divert significantly from the previous model. Further, including a third class does not represent a statistical improvement. In fact, performing a likelihood-ratio test between both DCLV models yields:

$$D = -2 \cdot (\Delta L) = 1.6 < \chi_{1,90\%}^2 = 2.706 \quad (19)$$

As a consequence, it is not possible to reject the null hypothesis of equality between both models even at a significance level of 10%. Hence, only two categories should be considered.

#### 4. Conclusions

Classifying individuals, based on their unobserved characteristics, offers significant opportunities from a modelling perspective, as well as a chance to gain significant insights into their decision-making process. Nowadays, the dominant approach to probabilistically categorize individuals based on unobserved characteristics is the latent class approach. However, this method has been subject to criticism due to the intrinsic difficulties associated with linking latent classes (LC), as identified by the model, with behavioural elements.

This criticism may be overruled when LC are used for identifying objective properties (such as missing information, lexicographic respondents, etc.), and the analyst can specify different utility functions for different LC, aiding (and guiding) their identification. However, when attempting to model attitudes, perceptions, and other latent characteristics of individuals, the analyst cannot control or guide the identification process of the latent classes by using different (class-specific) specifications. Additionally, the classical latent class approach does not take advantage of additional information, such as the existence of psychometric indicators.

In this work, we explored three ways to improve the categorization of individuals. First, we considered the LCPI approach, which improves the categorization by considering exogenous information (psychometric indicators), but it is guilty of the same problem associated with the LC models described previously. Note that, by definition, this approach should offer a superior goodness-of-fit than alternative approaches, as the class-membership function is built to maximize fit and not to satisfy a given specification. This fact is a double-edged sword, as, on the one hand, it may offer more insights into the actual underlying decision-making process, but it also poses significant difficulties for the analyst in terms of testing a specific theory or identifying a given behavioural pattern.

In the second place, we examined the LVLC approach, which combines the HDC and LC frameworks, by considering latent variables into the class-membership function. Unfortunately, this approach is guilty of an incorrect specification of the error terms, leading to an artificial decrease in the model's adjustment. Finally, we proposed an alternative way to attempt a direct categorization of individuals based on latent variables (DCLV). This approach overcomes the error issues of the LVLC.

We further illustrated the differences between the LVLC and DCLV approaches with the help of two case studies (we did not consider the LCPI approach, as given the different nature of the models no fair comparison was possible). In the first example, we could not establish the existence of significant differences in the estimated parameters (aside from the threshold parameter). However, in the second example, the differences in variability of the class-membership function led to a different categorization. In line with our expectations, the DCLV approach offered a superior goodness-of-fit than the LVLC (even though the latter had one extra degree-of-freedom); this is due to the incorrect specification of the error terms. The improvements in goodness-of-fit were mainly explained by a better adjustment of the discrete choice component.

Based on our theoretical analysis and empirical results, we recommend the DCLV approach over the LVLC method. It offers a consistent treatment of the error term, per underlying theory, and also offers a better performance in terms of goodness-of-fit. No definite recommendations can be made regarding the use of the DCLV or LCPI approaches, as both respond to different mechanisms; however, if the modeller aims at testing a specific attitudinal or perceptual characteristic, the DCLV offers significant advantages, as it facilitates specifying the required structure to be tested, and therefore, the identification of the latent construct.

Future research should aim at establishing for which kind of LV and latent phenomena a categorical non-parametric treatment offers significant advantages over continuous parametric treatments. We know that a non-linear effect can be assumed (as done in this work) based on theoretical considerations. It would be advantageous to establish which kind of latent phenomena are unlikely to be accurately captured by parametric representations (at

least regarding their impact on utility functions) and which phenomena can eventually be approximated using parametric representations, on an empirical basis. Furthermore, future research should be conducted to evaluate the use of non-simulated estimation techniques to allow for more flexible error term structures in the context of the DCLV.

## Notes

1. Modellers mostly use psychometric data to target unobserved/unobservable characteristics of the individuals, such as attitudes or perceptions, the aim of this work. Notwithstanding, they can also use the HDC framework to address other kinds of unobserved/unobservable information of the individuals or the alternatives. In this case, the indicator does not necessarily have to be psychometric data.
2. Even though Walker and Ben-Akiva (2002) postulate a potential link between both components, they did not address it in their work.
3. We use the same matrix of variables  $X_q$  in all equations of the paper for notational purposes. It does not imply that all columns of  $X_q$  would necessarily impact the outcome. The explicit inclusion of  $\eta_q^*$  in (2) simply acknowledges the fact that the structural equations may be subject to endogeneity. When considering a set of simultaneous structural equations, identifiability must be established on a case-by-case basis.
4. Notorious exceptions of the former would be given by situations when LC are used to test the availability of a given alternative (classical use of the LC framework), or when one class is expected to ignore a given attribute (lexicographic preferences), where the modeller indeed specifies the utility functions differently. However, when dealing with attitudes or perceptions that would result in systematic taste variations, the modeller cannot specify the structure differently in most cases, as the difference between classes is merely given by different values for the estimators.
5. Note that Hess, Shires, and Jopson (2013) just considered two classes, not specifying whether they were dealing with a MNL or an Ordinal model (as both specifications are equivalent for two classes). As we aim at latent classes as an expression of an underlying latent variable, we extended their framework as an Ordinal model. This extension does not affect the assessment of the model.
6. In their original specification, Hess, Shires, and Jopson (2013) multiplied  $\eta_c$  by a parameter to be estimated and fixed ( $\lambda$ ); however, it is easy to see that both specifications are equivalent.
7. When considering the model sequentially by integrating over the domain of the LV perfect convergence is achieved, but the threshold remains unidentified and must be fixed *a priori* (Bahamonde-Birke et al., 2017).
8. Just in the case of the LV to be discretized. The errors of any other LV may follow any desired distribution.
9. Bahamonde-Birke and Hanappi (2016) considered further variables and estimated more involved models than those considered in this study. For the purposes of this work, this specification was considered appropriate, as additional complexity would only add noise and computational complexity to our analysis.
10. Given their different structure, the heuristic value used in both models would be necessarily different. While the difference is indeed small, it does not allow for a fair comparison.

## Acknowledgments

This paper is partially based on scientific work done during the DEFINE (Development of an Evaluation Framework for the Introduction of Electromobility) project. We gratefully acknowledge the funding for DEFINE (<https://www.ihs.ac.at/projects/define/>) as part of the ERA-NET Plus Electromobility+ call by the EU-Commission and national funding institutions: the Ministry for Transport, Innovation and Technology (Austria), the Federal Ministry of Transport and Digital Infrastructure, formerly Federal Ministry for Transport, Building and Urban Development (Germany), and the National Centre for Research and Development (Poland). We are also grateful to the Institute in Complex Engineering

Systems (CONICYT: FB0816), the 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). The authors would also like to thank Prof. Michel Bierlaire and five anonymous referees for their useful comments and insights on an earlier version of the paper. All errors are the authors' sole responsibility.

## Disclosure statement

No potential conflict of interest was reported by the author(s).

## ORCID

Juan de Dios Ortúzar  <http://orcid.org/0000-0003-3452-3574>

## References

- Ashok, K., W. Dillon, and S. Yuan. 2002. "Extending Discrete Choice Models to Incorporate Attitudinal and Other Latent Variables." *Journal of Marketing Research* 39: 31–46.
- Bahamonde-Birke, F. J., and T. Hanappi. 2016. "The Potential of Electromobility in Austria: Evidence from Hybrid Choice Models Under the Presence of Unreported Information." *Transportation Research Part A: Policy and Practice* 83: 30–41.
- Bahamonde-Birke, F. J., U. Kunert, H. Link, and J. de D. Ortúzar. 2014. "Bewertung der Angebotsmerkmale des Personenfernverkehrs vor dem Hintergrund der Liberalisierung des Fernbusmarktes." *Zeitschrift für Verkehrswissenschaft* 85 (2): 107–123.
- Bahamonde-Birke, F. J., U. Kunert, H. Link, and J. de D. Ortúzar. 2017. "About Attitudes and Perceptions – Finding the Proper Way to Consider Latent Variables in Discrete Choice Models." *Transportation* 44: 475–493.
- Bahamonde-Birke, F. J., and J. de D. Ortúzar. 2014a. "On the Variability of Hybrid Discrete Choice Models." *Transportmetrica A: Transport Science* 10 (1): 74–88.
- Bahamonde-Birke, F. J., and J. de D. Ortúzar. 2014b. "Is Sequential Estimation a Suitable Second Best for Estimation of Hybrid Choice Models?" *Transportation Research Record: Journal of the Transportation Research Board* 2429: 51–58.
- Ben-Akiva, M., and B. Boccara. 1995. "Discrete Choice Models with Latent Choice Sets." *International Journal of Research in Marketing* 12: 9–24.
- Ben-Akiva, M., D. McFadden, K. Train, J. Walker, C. Bhat, M. Bierlaire, D. Bolduc, et al. 2002. "Hybrid Choice Models: Progress and Challenges." *Marketing Letters* 13: 163–175.
- Bhat, C. R. 2011. "The Maximum Approximate Composite Marginal Likelihood (MACML) Estimation of Multinomial Probit-Based Unordered Response Choice Models." *Transportation Research Part B: Methodological* 45 (7): 923–939.
- Bierlaire, M. 2003. "BIOGEME: A Free Package for the Estimation of Discrete Choice Models." *3rd Swiss Transportation Research Conference*, Ascona, Switzerland.
- Bujosa, A., A. Riera, and R. L. Hicks. 2010. "Combining Discrete and Continuous Representations of Preference Heterogeneity: A Latent Class Approach." *Environmental and Resource Economics* 47 (4): 477–493.
- Cherchi, E., and J. de D. Ortúzar. 2008. "Empirical Identification in the Mixed Logit Model: Analysing the Effect of Data Richness." *Networks and Spatial Economics* 8: 109–124.
- Chia-Jung, C., and C. Pei-Chun. 2014. "Preferences and Willingness to Pay for Green Hotel Attributes in Tourist Choice Behaviour: The Case of Taiwan." *Journal of Travel & Tourism Marketing* 31: 937–957.
- DeSarbo, W. S., V. Ramaswamy, and S. H. Cohen. 1995. "Market Segmentation with Choice-Based Conjoint Analysis." *Marketing Letters* 6: 137–147.
- Dirac, P. 1958. *The Principles of Quantum Mechanics*. Oxford: Clarendon Press.
- Domencich, T., and D. McFadden. 1975. *Urban Travel Demand – A Behavioural Analysis*. Oxford: North Holland Publishing.

- El Zarwi, F., A. Vij, and J. L. Walker. 2018. "A Discrete Choice Framework for Modelling and Forecasting the Adoption and Diffusion of New Transportation Services." *Transportation Research Part C: Emerging Technologies* 79: 207–223.
- Hess, S., and N. Beharry-Borg. 2012. "Accounting for Latent Attitudes in Willingness-to-Pay Studies: The Case of Coastal Water Quality Improvements in Tobago." *Environmental and Resource Economics* 52: 109–131.
- Hess, S., J. Shires, and A. Jopson. 2013. "Accommodating Underlying Pro-Environmental Attitudes in a Rail Travel Context: Application of a Latent Variable Latent Class Specification." *Transportation Research Part D: Transport and Environment* 25: 42–48.
- Humlum, M. K., K. J. Kleinjans, and H. S. Nielsen. 2012. "An Economic Analysis of Identity and Career Choice." *Economic Inquiry* 50: 39–61.
- Hurtubia, R., and M. Bierlaire. 2014. "Estimation of Bid Functions for Location Choice and Price Modelling with a Latent Variable Approach." *Networks and Spatial Economics* 14: 47–65.
- Hurtubia, R., M. H. Nguyen, A. Glerum, and M. Bierlaire. 2014. "Integrating Psychometric Indicators in Latent Class Choice Models." *Transportation Research Part A: Policy and Practice* 64: 135–146.
- Kamakura, W. A., and G. Russell. 1989. "A Probabilistic Choice Model for Market Segmentation and Elasticity Structure." *Journal of Marketing Research* 26: 379–390.
- Kim, J., S. Rasouli, and H. Timmermans. 2016. "A Hybrid Choice Model with a Nonlinear Utility Function and Bounded Distribution for Latent Variables: Application to Purchase Intention Decisions of Electric Cars." *Transportmetrica A: Transport Science* 12 (10): 909–932.
- Lam, S. K., M. Ahearne, Y. Hu, and N. Schillewaert. 2010. "Resistance to Brand Switching When a Radically New Brand is Introduced: A Social Identity Theory Perspective." *Journal of Marketing* 74: 128–146.
- Likert, R. 1932. "A Technique for the Measurement of Attitudes." *Archives of Psychology* 140: 1–55.
- Link, H. 2015. "Is car Drivers' Response to Congestion Charging Schemes Based on the Correct Perception of Price Signals?" *Transportation Research Part A: Policy and Practice* 71: 96–109.
- Luo, L., P. K. Kannan, and B. T. Ratchford. 2008. "Incorporating Subjective Characteristics in Product Design and Evaluations." *Journal of Marketing Research* 45: 182–194.
- McCullagh, P. 1980. "Regression Models for Ordinal Data." *Journal of the Royal Statistical Society. Series B (Methodological)* 42: 109–127.
- McFadden, D. 1974. "Conditional Logit Analysis of Qualitative Choice Behaviour." In *Frontiers in Econometrics*, edited by P. Zarembka, 105–142. New York: Academic Press.
- McFadden, D. 1986. "The Choice Theory Approach to Market Research." *Marketing Science* 5: 275–297.
- Motoaki, Y., and R. A. Daziano. 2015. "Assessing Goodness of Fit of Hybrid Choice Models: An Open Research Question." *Transportation Research Record: Journal of the Transportation Research Board* 2495: 131–141.
- Nylund, K. L., T. Asparouhov, and B. O. Muthén. 2007. "Deciding on the Number of Classes in Latent Class Analysis and Growth Mixture Modeling: A Monte Carlo Simulation Study." *Structural Equation Modeling: A Multidisciplinary Journal* 14 (4): 535–569.
- Ortúzar, J. de D., and L. G. Willumsen. 2011. *Modelling Transport*. 4th ed. Chichester: John Wiley and Sons.
- Train, K. E., D. L. McFadden, and A. A. Goett. 1987. "Consumer Attitudes and Voluntary Rate Schedules for Public Utilities." *The Review of Economics and Statistics* 69: 383–391.
- van Acker, V., P. L. Mokhtarian, and F. Witlox. 2011. "Going Soft: On How Subjective Variables Explain Modal Choices for Leisure Travel." *European Journal of Transport and Infrastructure Research* 11: 115–147.
- Vij, A., and J. Walker. 2014. "Hybrid Choice Models: The Identification Problem." In *Handbook of Choice Modelling*, edited by S. Hess and A. Daly, 519–564. Cheltenham: Edward Elgar Publishing.
- Vij, A., and J. L. Walker. 2016. "How, When and Why Integrated Choice and Latent Variable Models are Latently Useful." *Transportation Research Part B: Methodological* 90: 192–217.
- Walker, J., and M. Ben-Akiva. 2002. "Generalized Random Utility Model." *Mathematical Social Sciences* 43: 303–343.
- Zellner, A. 1970. "Estimation of Regression Relationships Containing Unobservable Variables." *International Economic Review* 11: 441–454.

**Appendix 1. Parameter estimates for the measurement equations. Case 1.**

Variable	Equation	LVLC		DCLV	
<i>Scale Parameter</i>	<i>M.E. EcAwareness</i>	<b>0.959</b>	(16.14)	<b>0.954</b>	(16.11)
<i>Threshold 1</i>	<i>M.E. EcAwareness</i>	<b>-3.25</b>	(-15.18)	<b>-3.18</b>	(-15.29)
<i>Threshold 2</i>	<i>M.E. EcAwareness</i>	<b>0.0152</b>	(0.09)	<b>0.0719</b>	(0.42)
<i>Threshold 3</i>	<i>M.E. EcAwareness</i>	<b>3.05</b>	(12.85)	<b>3.1</b>	(13.22)
<i>Threshold 4</i>	<i>M.E. EcAwareness</i>	<b>5.06</b>	(14.74)	<b>5.1</b>	(14.9)
<i>Threshold 5</i>	<i>M.E. EcAwareness</i>	<b>7.01</b>	(12.11)	<b>7.05</b>	(12.16)
<i>Scale Parameter</i>	<i>M.E. LocalFood</i>	<b>1.18</b>	(14.89)	<b>1.22</b>	(14.89)
<i>Threshold 1</i>	<i>M.E. LocalFood</i>	<b>-1.76</b>	(-9.58)	<b>-1.68</b>	(-9.51)
<i>Threshold 2</i>	<i>M.E. LocalFood</i>	<b>0.655</b>	(3.74)	<b>0.692</b>	(4.11)
<i>Threshold 3</i>	<i>M.E. LocalFood</i>	<b>2.59</b>	(12.3)	<b>2.59</b>	(12.79)
<i>Threshold 4</i>	<i>M.E. LocalFood</i>	<b>4.23</b>	(15.2)	<b>4.2</b>	(15.63)
<i>Threshold 5</i>	<i>M.E. LocalFood</i>	<b>7.17</b>	(10.56)	<b>7.09</b>	(10.7)
<i>Scale Parameter</i>	<i>M.E. EcoFriendly</i>	<b>0.988</b>	(15.94)	<b>1.02</b>	(16.06)
<i>Threshold 1</i>	<i>M.E. EcoFriendly</i>	<b>-2.68</b>	(-12.98)	<b>-2.58</b>	(-13.01)
<i>Threshold 2</i>	<i>M.E. EcoFriendly</i>	<b>-0.727</b>	(-4.12)	<b>-0.664</b>	(-3.92)
<i>Threshold 3</i>	<i>M.E. EcoFriendly</i>	<b>1.28</b>	(6.9)	<b>1.3</b>	(7.32)
<i>Threshold 4</i>	<i>M.E. EcoFriendly</i>	<b>2.88</b>	(12.85)	<b>2.88</b>	(13.32)
<i>Threshold 5</i>	<i>M.E. EcoFriendly</i>	<b>6.77</b>	(12.65)	<b>6.69</b>	(12.87)
<i>Scale Parameter</i>	<i>M.E. Protection</i>	<b>0.442</b>	(13.06)	<b>0.446</b>	(13.05)
<i>Threshold 1</i>	<i>M.E. Protection</i>	<b>-7.7</b>	(-13.42)	<b>-7.59</b>	(-13.39)
<i>Threshold 2</i>	<i>M.E. Protection</i>	<b>-3.33</b>	(-11.57)	<b>-3.25</b>	(-11.54)
<i>Threshold 3</i>	<i>M.E. Protection</i>	<b>1.3</b>	(5.53)	<b>1.34</b>	(5.83)
<i>Threshold 4</i>	<i>M.E. Protection</i>	<b>4.12</b>	(10.97)	<b>4.14</b>	(11.15)
<i>Threshold 5</i>	<i>M.E. Protection</i>	<b>8.04</b>	(11.84)	<b>8.03</b>	(11.91)
<i>Scale Parameter</i>	<i>M.E. CO2Footprint</i>	<b>0.951</b>	(16.84)	<b>0.95</b>	(16.77)
<i>Threshold 1</i>	<i>M.E. CO2Footprint</i>	<b>-4.39</b>	(-17.52)	<b>-4.32</b>	(-17.61)
<i>Threshold 2</i>	<i>M.E. CO2Footprint</i>	<b>-1.92</b>	(-10.27)	<b>-1.86</b>	(-10.29)
<i>Threshold 3</i>	<i>M.E. CO2Footprint</i>	<b>0.276</b>	(1.56)	<b>0.329</b>	(1.92)
<i>Threshold 4</i>	<i>M.E. CO2Footprint</i>	<b>1.89</b>	(9.4)	<b>1.93</b>	(9.85)
<i>Threshold 5</i>	<i>M.E. CO2Footprint</i>	<b>5.31</b>	(14.73)	<b>5.35</b>	(14.89)

**Appendix 2. Parameter estimates for the measurement equations. Case study 2.**

Variable	Equation	LVLC		DCLV	
<i>Scale Parameter</i>	<i>M.E. HSTrains</i>	<b>1.13</b>	(8.48)	<b>1.09</b>	(7.6)
<i>Threshold 1</i>	<i>M.E. HSTrains</i>	<b>-3.74</b>	(-13.3)	<b>-3.77</b>	(-12.42)
<i>Threshold 2</i>	<i>M.E. HSTrains</i>	<b>-2.21</b>	(-11.8)	<b>-2.21</b>	(-11.3)
<i>Threshold 3</i>	<i>M.E. HSTrains</i>	<b>-0.179</b>	(-1.47)	<b>-0.16</b>	(-1.37)
<i>Threshold 4</i>	<i>M.E. HSTrains</i>	<b>1.43</b>	(10.12)	<b>1.47</b>	(10.56)
<i>Scale Parameter</i>	<i>M.E. RailLines</i>	<b>2.03</b>	(3.9)	<b>2.23</b>	(2.84)
<i>Threshold 1</i>	<i>M.E. RailLines</i>	<b>-2.83</b>	(-12.73)	<b>-2.74</b>	(-11.66)
<i>Threshold 2</i>	<i>M.E. RailLines</i>	<b>-1.5</b>	(-10.01)	<b>-1.44</b>	(-9.44)
<i>Threshold 3</i>	<i>M.E. RailLines</i>	<b>0.0287</b>	(0.26)	<b>0.0539</b>	(0.53)
<i>Threshold 4</i>	<i>M.E. RailLines</i>	<b>1.4</b>	(9.91)	<b>1.4</b>	(9.81)

### Appendix 3. Parameter estimates for the measurement equations. Case study 2. Three categories.

Variable	Equation		DCLV
<i>Scale Parameter</i>	<i>M.E. HSTrains</i>	<b>1.09</b>	(7.44)
<i>Threshold 1</i>	<i>M.E. HSTrains</i>	<b>-3.77</b>	(-12.25)
<i>Threshold 2</i>	<i>M.E. HSTrains</i>	<b>-2.21</b>	(-11.18)
<i>Threshold 3</i>	<i>M.E. HSTrains</i>	<b>-0.157</b>	(-1.36)
<i>Threshold 4</i>	<i>M.E. HSTrains</i>	<b>1.47</b>	(10.58)
<i>Scale Parameter</i>	<i>M.E. RailLines</i>	<b>2.26</b>	(2.72)
<i>Threshold 1</i>	<i>M.E. RailLines</i>	<b>-2.73</b>	(-11.49)
<i>Threshold 2</i>	<i>M.E. RailLines</i>	<b>-1.43</b>	(-9.33)
<i>Threshold 3</i>	<i>M.E. RailLines</i>	<b>0.057</b>	(0.57)
<i>Threshold 4</i>	<i>M.E. RailLines</i>	<b>1.4</b>	(9.74)