



HAL
open science

Integrated choice and latent variable models: A literature review on mode choice

Hélène Bouscasse

► **To cite this version:**

Hélène Bouscasse. Integrated choice and latent variable models: A literature review on mode choice. 2018. hal-01795630

HAL Id: hal-01795630

<https://hal.science/hal-01795630v1>

Preprint submitted on 18 May 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**Integrated choice and latent variable
models: a literature review on mode
choice**

Hélène Bouscasse

May 2018

JEL: C25, D9, R41

<https://gael.univ-grenoble-alpes.fr/accueil-gael>

contact : agnes.vertier@inra.fr



Integrated Choice and Latent Variable Models: a Literature Review on Mode Choice

Hélène Bouscasse ^a

^aUniv. Grenoble Alpes, CNRS, INRA, Grenoble INP, GAEL, 38000 Grenoble, France.
helene.bouscasse@univ-grenoble-alpes.fr

ABSTRACT

Mode choice depends on observable characteristics of the transport modes and of the decision maker, but also on unobservable characteristics, known as latent variables. By means of an integrated choice and latent variable (ICLV) model, which is a combination of structural equation model and discrete choice model, it is theoretically possible to integrate both types of variables in a psychologically and economically sound mode choice model. To achieve such a goal requires clear positioning on the four dimensions covered by ICLV models: survey methods, econometrics, psychology and economics. This article presents a comprehensive survey of the ICLV literature applied to mode choice modelling. I review how latent variables are measured and incorporated in the ICLV models, how they contribute to explaining mode choice and how they are used to derive economic outputs. The main results are: 1) the latent variables used to explain mode choice are linked to individual mental states, perceptions of transport modes, or an actual performed behaviour; 2) the richness of structural equation models still needs to be explored to fully embody the psychological theories explaining mode choice; 3) the integration of latent variables helps to improve our understanding of mode choice and to adapt public policies.

KEYWORDS

Mode choice; Survey; Integrated choice and latent variable model; Structural equation modelling; Behavioural theories; Economic outputs

1. Introduction

Mode choice is a multiparametric decision process determined by the characteristics of the transport alternatives and a combination of individual socio-economic, psychological and spatial factors. Psychological factors are unobservable, latent, variables and remain among the less studied factors in discrete choice models (DCMs), even though, when integrated, they often prove to be significant (De Witte, Hollevoet, Dobruszkes, Hubert, & Macharis, 2013). At the same time, psychological theories propose a behaviourally rich representation of mode choice by modelling relationships between latent variables, such as norms, values or attitudes. Structural equation models (SEMs) are a powerful tool to model these relationships. However, since they are not based on economic theory, SEMs cannot be used in transport models or to derive economic outputs, such as value of time or elasticities, necessary for project appraisal. One solution is to incorporate behavioural theories in discrete choice models by means of an Integrated Choice and Latent Variable (ICLV) model, which theoretically allows the analyst to benefit from the economic and behavioural foundations of both approaches.

ICLV models involve, at the same time, survey theory, econometric theory, psychological theory and economic theory, and the promise of ICLV models is to embrace and integrate these multiple dimensions in an integrative model. Previous work has provided insight into some specific dimensions, for instance, Vij and Walker (2014)

focus on identification issues, [Bhat and Dubey \(2014\)](#) on estimation methods, and [Vij and Walker \(2016\)](#) on the operational use of ICLV models. However, there has been no attempt to address the multidimensional nature of ICLV models and to combine theory and practice for each dimension with a synthesis and analysis of the results produced by ICLV models. This paper intends to fill that gap.

The aim of this paper is to provide, for the survey, econometric, psychological and economic dimensions of ICLV models, a comprehensive review of the literature on how mode choice and the determinants of mode choice can be better modelled. All dimensions involved in ICLV models are addressed to emphasize and structure the progress that has been made in the literature in recent years, as well as to identify future research needs. Specifically, the contribution of this paper to the current knowledge on mode choice is three-fold. Firstly, this paper is the first to review how latent variables that explain mode choice are measured and to propose the use of psychological and management science methods to improve this measurement. Secondly, this paper is the first to review the different structures of SEMs that can be incorporated in ICLV models to improve the behavioural representation of mode choice. Thirdly, this paper is the first to propose an analysis grid to structure the latent variables used in ICLV models and to review how these latent variables explain mode choice.

The rest of the paper is organized in five sections. Section 2 presents the four dimensions of ICLV models and the search strategy. Section 3 is dedicated to the tool itself with the econometric dimension. Section 4 discusses the psychological dimension of ICLV models with a specific focus on latent variables. Section 5 reports on the use made of ICLV models and, in particular, the economic dimension. Section 6 is the conclusion.

2. Methodology

2.1. Overview of the dimensions addressed in ICLV models

To develop an ICLV model requires clear positioning on four dimensions. Firstly, the data collection process has to be adapted to collect the indicators measuring the latent variables. Special attention also needs to be paid to the sample size, since estimated parameters are more numerous and may result in identification issues.

Secondly, ICLV models require an estimation of a SEM and a DCM. Depending on the articles, different terminologies, such as latent variable model, MIMIC model¹ or LISREL models,² are used to refer to the SEM. In this article, the terminology “SEM” is intentionally used to reflect the richness of the models that can be included in ICLV model and to build the bridge with psychological and management science models. Structural equation modelling is a methodology for representing, estimating, and testing a network of relationships between variables (measured or latent variables, exogenous or endogenous variables). They are able to represent a wide variety of relationships between variables, in particular mediation and moderation processes ([Hayes, 2013](#)).

Thirdly, the latent variables included in ICLV models are drawn from psychological theories, which describe the cognitive mechanisms resulting in travel mode choice. These latent variables and theories are various and their integration in the determin-

¹Multiple Indicators Multiple Causes (MIMIC) models are specific SEMs.

²Linear Structural RELations (LISREL) is a statistical software package used for structural equation modelling.

istic part of the utility is supposed to improve the behavioural realism of mode choice models.

Fourthly, ICLV models are modelled using economic random utility theory. The decision rule, that is choosing the option that maximizes the utility, remains the same as in the other DCMs. Because of their economically sound formulation, ICLV models can be used to derive economic policies or economic outputs, such as elasticities, willingness-to-pay indicators and market shares.

2.2. *Search strategy*

The search for articles was conducted with the aim of being as exhaustive as possible as regards ICLV models on mode choice. The sources used to search for literature are web-based search tools (Web of Science, Researchgate, Google Scholar and the database of the French National Center for Scientific Research). After this first step, cited references were tracked to create a domino effect. The search was conducted in August and September 2016 using the search terms (“mode choice” or “modal choice”) and (“integrated choice and latent variable(s)” or “hybrid choice model(s)” or ‘hybrid modelling approach’ or “hybrid choice modelling”). The resulting output was filtered with a focus on articles containing an application of ICLV models (with real or simulated data) concerned with individual inland mode choice.³ Freight, boat and air applications have been excluded.

Eventually, 35 articles were retained for further analysis and processed for this review. They were published or presented in conferences between 1998 and 2016 but 50% of them were published after 2013. All articles contain an application but three use only simulated data (Bhat & Dubey, 2014; Daziano & Bolduc, 2013; Raveau, Yáñez, & Ortúzar, 2012). These articles have been retained because they contain valuable insight on the method.

3. Survey dimension

This section provides an overview of the surveys used in ICLV models before focusing on the collection of data to measure the latent variables.

3.1. *Overview of the surveys*

Two thirds of the surveys used for ICLV models are revealed preferences surveys, the remaining third being stated preferences surveys. Some datasets are common to several studies. For instance, Glerum, Atasoy, and Bierlaire (2014), Atasoy, Glerum, and Bierlaire (2013) and Fernández-Antolín, Guevara-Cue, de Lapparent, and Bierlaire (2016) all use the same Swiss survey. Yanez, Raveau, and Ortúzar (2010) and Raveau, Álvarez-Daziano, Yáñez, Bolduc, and Ortúzar (2010) use the same Chilean survey. Datasets are heterogeneous in terms of the number of persons interviewed (see Figure 1). The median is around 800 with a mean around 2,400 due to two large RP surveys (La Paix Puello & Geurs, 2015; Roberts, Popli, & Harris, 2014), which include, respectively, 12,000 and 13,141 observations.

³In one article, mode choice is made conjointly with purchase intentions of alternative fuel vehicles (Daziano, 2015). Another article models mode shift defined as the choice of departing from the car mode (Idris, Habib, & Shalaby, 2015).

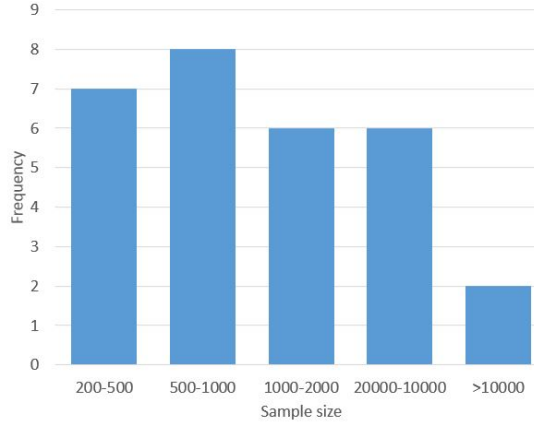


Figure 1. Distribution of the size of samples (persons interviewed) used for ICLV models on mode choice (excluding simulated and duplicated datasets)

All datasets were collected in western countries (Canada, Chile, Europe), with the exception of one collected in China (J. Walker, Li, Srinivasan, & Bolduc, 2013). Some studies collected data only for one city (Habib & Zaman, 2012) or a specific inter-urban corridor (Daziano & Rizzi, 2015; Johansson, Heldt, & Johansson, 2006). Some other studies concern all cities in a country (Daziano, 2015) or the whole country (Roberts et al., 2014). Some surveys specifically restrict their field of study to employees of a university (e.g., Córdoba & Jaramillo, 2012; Tudela, Habib, Carrasco, & Osman, 2011), which limits the generalization of their results to the whole population.

With regard to journey purposes, four situations emerge from the review: 1) all purposes are considered (e.g., Glerum et al., 2014; Yanez et al., 2010); 2) the focus is on work trips (e.g., Daziano, 2015; Idris et al., 2015; Johansson et al., 2006); 3) only trips to school or university are considered (e.g., Kamargianni, Dubey, Polydoropoulou, & Bhat, 2015), and 4) only trips to go out with friends are considered (Scagnolarie, Walker, & Maggi, 2015). Depending on the purposes, respondents to the surveys are also naturally heterogeneous in terms of age.

The selection of respondents can also be made according to their (potential) use of specific modes. For instance, Glerum et al. (2014) only interviewed people living in towns or villages connected to the postal buses. Sottile, Meloni, and Cherchi (2013) restricted participation in their survey to the (potential) users of the park and ride facility they study. The car plays a key role in choice sets, whether or not a distinction is made between car as driver and car as passenger. Apart from Daziano and Rizzi (2015), all choice sets include the car alternative, either as a specific mode or as part of an aggregate. The car is often chosen as the reference alternative. Other modes considered are:

- private motorized modes, such as carpool (Daziano, 2015; Idris et al., 2015) or powered two wheelers / motorcycles (Kamargianni et al., 2015);
- public transport modes, such as park and ride (Sottile et al., 2013), bus or trains (Anwar et al., 2014);
- active modes: walking and cycling (Daziano, 2015; Idris et al., 2015).

The articles reviewed on mode choice therefore account for a wide variety of contexts and represent what has been done so far, in terms of survey method, geographic area and transport modes.

3.2. *Psychometrics: theory and application*

Psychometrics is the field of study concerned with the theory and technique of psychological measurement. Measurements are made by constructing a scale that assigns numbers to objects or events according to rules (Tinsley & Brown, 2000). The typical scale consists of a set of stimuli (items) designed to elicit particular kinds of responses (the indicators used in the SEM), a set of response alternatives and a set of directions on how to respond to items. Over time, different types of scales have been described in the literature (Zikmund, Babin, Carr, & Griffin, 2013), among them multiple items scales.

An overwhelming majority of studies use four- to ten- points Likert scales (Likert, 1932) to measure latent variables. Anchors of the scale can denote agreement (e.g. Kamargianni, Ben-Akiva, & Polydoropoulou, 2014), frequency (Sottile, Cherchi, & Meloni, 2014), similarity (Temme, Paulssen, Dannewald, et al., 2007), importance (Daziano, 2015), or good/bad representation (Morikawa & Sasaki, 1998). If the rating scale is anchored at each end by antonym adjectives, then they are named semantic differential scales (Osgood & Luria, 1954). To measure attitudes towards alcohol consumption, Scagnolarie et al. (2015) uses a seven-point semantic differential scale with adjectives, such as pleasant/unpleasant, funny/boring.

Two studies (Idris et al., 2015; Tudela et al., 2011) use the Verplanken scale to measure habits (Verplanken, Aarts, van Knippenberg, & van Knippenberg, 1994). This scale measures habitual behaviour intensity by counting the number of times a given behaviour would be repeated under specific circumstances without too much reasoning. For instance, for non-work activities, Idris et al. (2015) ask the respondent to provide the mode she would use and then count how many times she had mentioned each mode to perform different activities.

Besides structured scales with closed questions, spontaneous answers to open-questions can be used to construct scales. A common question in French travel surveys is to ask for three spontaneous adjectives best describing a transport mode. Numeric data, and therefore scales, can be generated from these words (e.g. Kaufmann, Tabaka, Jean-Marie, & Louvet, 2010). Although they do not mention it explicitly in their study, Glerum et al. (2014) use the method proposed by Thurstone (1927) to construct a scale based on adjectives. Using a Swiss survey, they asked 25 evaluators to rate the adjectives related to comfort on public transport. A positive number is associated with a positive representation of the latent variable and a negative number implies a negative representation. The scale is a five-point scale from -2 to 2.

The technique to define an appropriate scale to measure latent variables is a broad area of research in psychology and management research (Dunlap, Liere, Mertig, & Jones, 2000; John & Benet-Martínez, 2000). If there is no adequate or appropriate existing scale to measure a latent variable, then a systematic seven-step process has to be followed (Hinkin, Tracey, & Enz, 1997): item generation; test for conceptual consistency; administration; purification of the measure with dimension reduction techniques, such as explanatory factor analysis, and confirmatory factor analysis; internal consistency assessment with reliability indicators (Cronbach α , Joreskog ρ); construct validity (discriminatory and convergent), and replication with a new data set. Failure to carefully develop and validate the measurement scale can result in invalid and unintegratable data (Hinkin et al., 1997).

With the exception of the Verplanken scales measuring habits, to our knowledge, the scales used in ICLV studies on mode choice are either constructed by the authors or the authors use a part, a compilation or an adaptation of existing scales. For instance,

Paulssen, Temme, Vij, and Walker (2014) measured respondents' value orientations based on the portraits value questionnaire from Schwartz et al. (2001). This questionnaire originally identifies ten motivationally distinct value constructs but the authors restricted their attention to three: power, hedonism and security. To measure variety-seeking, Rieser-Schüssler and Axhausen (2012) construct a scale that includes some of the items in Niininen, Szivas, and Riley (2004) and Mokhtarian and Salomon (2001). Sottile et al. (2014) worked with a team of environmental psychologists to measure all items included in the Theory of Planned Behaviour.

Among the tools designed to construct and validate scales, three are used in ICLV models on mode choice. Some studies report the use of principal component analysis (Sottile et al., 2014) or exploratory factor analysis (e.g. Johansson et al., 2006; La Paix Puello & Geurs, 2015; Scagnolarie et al., 2015) for dimension reduction. Confirmatory factor analysis, that is, the measurement model of the SEM, are reported by Johansson et al. (2006), Tudela et al. (2011) and Roberts et al. (2014). Concerning reliability, Johansson et al. (2006), Scagnolarie et al. (2015) and Politis, Papaioannou, and Basbas (2012) use Chronbach's α to check reliability.

4. Econometric dimension

This section describes the ICLV tool and reviews the general formulation of ICLV models, before describing some components of the model. It also discusses the identification issue and the methods used to estimate ICLV models.

4.1. General formulation ICLV models

I describe the general specification of ICLV models, before clarifying some hypotheses that can be conducted on error terms, the measurement model of the SEM and the interactions terms in the structural equations.

4.1.1. General specification

Mathematically, ICLV models consist of two components: a DCM (Equations 1 and 4) and a SEM (Equations 2 and 3):

$$u_n = Bx_n + \Gamma x_n^* + \epsilon_n, \quad (1)$$

$$x_n^* = Ax_n + v_n, \quad (2)$$

$$i_n^* = Dx_n^* + \eta_n, \quad (3)$$

$$y_{nj} = \begin{cases} 1 & \text{if } u_{nj} \geq u_{nj'} \text{ for } j' \in \{1, \dots, J\} \\ 0 & \text{otherwise.} \end{cases} \quad (4)$$

Equation 1 is the structural equation of the DCM. u_n is the $(J \times 1)$ vector containing the J utilities of the alternatives faced by each individual n ($n \in \{1, \dots, N\}$). Alternatives are described by K observable explanatory variables in the $(K \times 1)$ vector x_n . The M latent explanatory variables are in the $(M \times 1)$ vector x_n^* . B is the $(J \times K)$ matrix with the coefficients of the observable variables. Some terms may be equal to zero if the set of observable variables x_n is not the same in Equation 1 and in Equation 2. Γ is the $(J \times M)$ matrix with coefficients of the latent variables. ϵ_n stands for the $(J \times 1)$ vector composed of error terms.

Equation 2 is the structural equation of the SEM. A is the $(M \times K)$ matrix with the coefficients denoting the effects of observable variables on latent variables. Some terms may be equal to zero if the set of observable variables x_n are different in Equation 1 and 2. v_n is the $(M \times 1)$ vector composed of M independantly and identically distributed (i.i.d) normal distributed error terms with ϕ the covariance matrix.

Equation 3 is the measurement equation of the SEM. i_n^* is the $(R \times 1)$ indicators' vector used to measure the latent variables. They are assumed to be continuous and normally distributed. D is a $(R \times M)$ matrix with the coefficients associated with the latent variables. η_n is the $(R \times 1)$ vector composed of R i.i.d. normal distributed error terms with Ψ the covariance matrix. The error terms ϵ_n , v_n and η_n are considered as mutually independent.

Equation 4 is the measurement equation of the discrete choice part of the model. y_{nj} is the choice indicator, which is equal to one if alternative j is chosen and 0 otherwise.

The model is sometimes written in its reduced form, by integrating Equation 2 into Equation 1 :

$$\begin{aligned} u_n &= Bx_n + \Gamma(Ax_n + v_n) + \epsilon_n \\ &= (B + \Gamma A)x_n + \Gamma v_n + \epsilon_n. \end{aligned} \quad (5)$$

The joint density of y_n and i_n is

$$f_{y,i} = \int_{x^*} f_y(y_n|x_n, x_n^*; B, \Gamma) f_{i^*}(i_n^*|x_n^*; D, \Psi) f_{x^*}(x_n^*|x_n; A, \Psi) dx^*. \quad (6)$$

The expressions of the density functions f_y , f_i and f_{x^*} depend on the hypotheses made on the associated error terms ϵ_n , v_n and η_n .

4.1.2. Distribution of the error terms

Although not always explicitly mentioned in the articles reviewed, a common choice is to define the elements of ϵ_n as i.i.d. Gumbel distributed across alternatives and individuals (Atasoy et al., 2013; Raveau et al., 2012). This hypothesis leads to the well-known multinomial logit kernel. Alternative distributions lead to mixed logit (Anwar et al., 2014; Idris et al., 2015; Maldonado-Hinarejos, Sivakumar, & Polak, 2014; Scagnolarie et al., 2015; Sottile et al., 2014; Tudela et al., 2011; Yanez et al., 2010), cross-nested logit (Habib, Tian, & Zaman, 2011; Habib & Zaman, 2012; Paulssen et al., 2014), and probit models (Bhat & Dubey, 2014; Daziano, 2015; Johansson et al., 2006; Kamarigianni et al., 2015; Politis et al., 2012; Roberts et al., 2014).

The mixed logit specification makes it possible to associate random coefficients with both observable and latent variables, whose effect on mode choice may be particularly heterogeneous in the population. For instance, Yanez et al. (2010) estimate three models, a mixed logit without latent variables, a mixed ICLV model with a random parameter associated with cost and a mixed ICLV model with random parameters associated with cost and the two latent variables (accessibility and comfort/safety). The last in the list proved to be the best in terms of estimation and forecasting (market shares reproduction). Tudela et al. (2011) associate random parameters with habits, attitudes and affective latent variables and estimate two ICLV models including, or not, the affective variable. They note that when all three latent variables are integrated, the standard deviation of the random parameters is smaller, which means that the mean value is more certain and randomness is captured by the affective factor.

Regarding probit specifications, two reasons may explain this choice. Firstly, some studies (e.g. [Politis et al., 2012](#); [Roberts et al., 2014](#)) use software originally developed for SEM estimations (M+, AMOS), which usually support only normal error terms. Secondly, recent articles ([Bhat & Dubey, 2014](#); [Daziano, 2015](#); [Kamargianni et al., 2015](#)) use normally distributed error terms in order to 1) reduce the estimation time with alternative estimation methods (bayesian estimator or maximum approximate composite marginal likelihood), and 2) to enable more flexible substitution patterns across alternatives by specifying a general covariance matrix ([Kamargianni et al., 2015](#)).

4.1.3. Ordered probit in the measurement model

The indicators measuring the latent variables (Equation 3) are almost exclusively measured with scales leading to ordinary variables. Nevertheless, the ordinal nature of the indicators is usually overlooked and they are considered as continuously normally distributed (e.g. [Glerum et al., 2014](#); [Habib & Zaman, 2012](#); [Idris et al., 2015](#); [Paulssen et al., 2014](#); [Scagnolarie et al., 2015](#)). The usual framework in the SEM literature ([K. Bollen & Hoyle, 2012](#)) is to add a threshold model equation to the measurement model of the SEM. In Equation 7, i_{nr}^* is a latent variable reflecting the “true” value accorded to the evaluated statement and i_{rn} is its discrete translation

$$i_{nr} = \begin{cases} 1 & \text{if } i_{nr}^* \leq \tau_1 \\ 2 & \text{if } \tau_1 < i_{nr}^* \leq \tau_2 \\ \dots & \\ M & \text{if } \tau_{M-1} < i_{nr}^*. \end{cases} \quad (7)$$

M is the total number of categories of i_{nr} and the τ 's parameters are thresholds or cutoff points for i_{nr}^* that determine the probabilities of observing each category of i_{rn} , with $\tau_1 \leq \tau_2 \leq \dots \tau_{M-1}$. The probability that the indicator takes the value m is therefore equal to

$$\begin{aligned} P(i_{nr} = m) &= P(\tau_{m-1} < i_{nr}^* \leq \tau_m) \\ &= F_\eta(\tau_m) - F_\eta(\tau_{m-1}). \end{aligned}$$

Assuming that the latent response variables are normally distributed, the corresponding model is an ordered probit model. Only a few studies applying ICLV models on mode choice use this framework ([Abou-Zeid & Ben-Akiva, 2011](#); [Kamargianni et al., 2015](#); [Sottile et al., 2013](#)).

4.1.4. Interaction terms in structural equations

Interaction terms between latent variables or between latent and observed variables can be included in the two structural sub-models. In Equation 1, the utility may include cross variables to study how latent variables impact the effect of an observed variable. For instance, [Fernández-Antolín et al. \(2016\)](#) define a cross-variable time \times car loving attitude and study how value of time varies according to attitude.

In Equation 2, latent variables do not interact with each other and depend only on observed variables, as is generally the case in the reviewed ICLV models. However, most psychological and management science theories do imply interactions between

latent variables, which requires Equation 2 to be changed to Equation 8:

$$x_n^* = Ax_n + A'x_n^* + v_n. \quad (8)$$

A' denotes the $(M \times M)$ matrix of coefficients translating the effect of some latent variables on other latent variables. The diagonal coefficients must be fixed to 0. All structures of SEMs may be estimated with this specification, in particular the full mediation, partial mediation and moderation models (see Figure 2 and corresponding matrix), which are extensively used in behavioural theories. The reader is referred to Little, Card, Bovaird, Preacher, and Crandall (2007) and Hayes (2013) for more details on these models. In these three models, there is no feedback between latent variables, A' is thus a strictly lower triangular matrix. If feedbacks were allowed, then only off-diagonal elements would be non-zero. This could be the case, for instance, in a study of social interactions within networks.

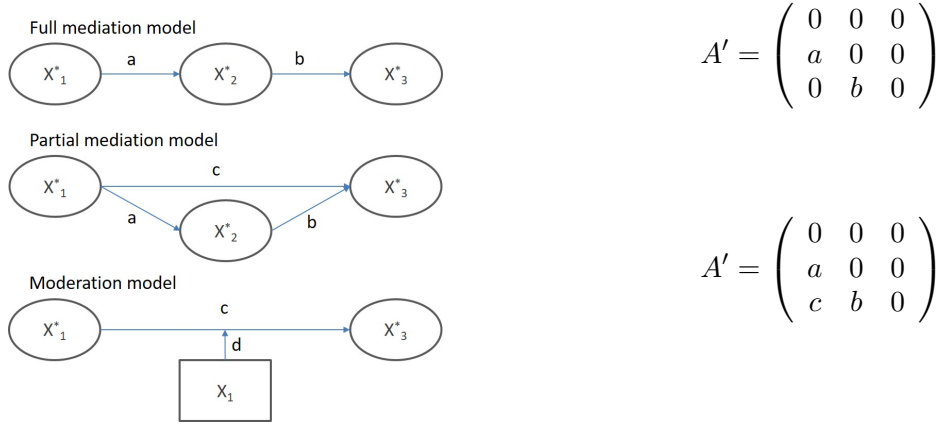


Figure 2. Mediation and moderation structures in SEMs

4.2. Identification issues

Given the degree of complexity of ICLV models, particular attention needs to be paid to their theoretical and practical identification. A model is theoretically identifiable if there are no two distinct sets of parameters that generate the same probability distribution of observable data. Generally, constraints have to be imposed to obtain a unique vector of consistent parameter estimates. Theoretical identifiability has been extensively studied for both DCMs (Train, 2009) and SEMs (K. A. Bollen, 2014; O'Brien, 1994). Vij and Walker (2014) and Daziano and Bolduc (2013) bridge the gap and derive sufficient conditions to ensure theoretical identifiability for ICLV models. However, necessary identification conditions for ICLV models still have to be developed. Even if a model is theoretically identified, a lack of data variability may lead to an empirically unidentified model. Using simulated data, Raveau et al. (2012) show that the discrete choice part of the model is more robust than the SEM model. When variability is low, the parameters obtained in the structural component of the SEM are

either underestimated or overestimated. Wrong sign problems may even occur, leading to counterintuitive relationships between variables. The sample size and sampling method are therefore critical to ensure empirical identification.

4.3. *Estimation methods*

Best practice is to estimate ICLV models with a simultaneous approach (J. L. Walker, 2001), that is, jointly estimating Equations 1 through 4. The simulated maximum likelihood estimation method is used by a majority of studies (e.g., Paulssen et al., 2014; Scagnolarie et al., 2015; J. Walker et al., 2013), since, for a large number of latent variables, the computational complexity increases exponentially and numerical integration is not feasible. However, the sequential approach, which first estimates the latent variables and then includes them in the DCM, provides inconsistent estimates with measurement errors, since it treats the fitted latent variables as non-stochastic. The sequential approach is still sometimes preferred (e.g., Anwar et al., 2014; Johansson et al., 2006; Maldonado-Hinarejos et al., 2014) due to the high computational burden of the simultaneous approach.

To address this issue, Bhat and Dubey (2014) develop the maximum approximate composite marginal likelihood inference approach with a probit specification. The dimensionality of integration in the composite marginal likelihood function is independent of the number of latent variables or the number of indicators in the measurement equations and only depends on the number of alternatives. The models estimated on simulated mode choice data converge in minutes against hours with the standard frequentist estimator. The authors note, however, that larger sample sizes are required to best recover the effects of the latent variables on choice. This approach was further applied by Scagnolarie et al. (2015), with no convergence problems. In Daziano and Bolduc (2013) and Daziano (2015), Bayesian estimation methods are developed. Daziano proposes a Bayesian estimator that accommodates a panel structure and latent variables that are endogenous and manifested through effect indicators that are discrete, continuous, or both. This estimator “avoids the curse of dimensionality, and addresses other issues, such as having exact (small-sample) properties” (Daziano, 2015, p.2). The statistical properties of this estimator are not affected by the number of latent attributes, with estimation time in the order of minutes.

Roberts et al. (2014) and Politis et al. (2012) rely exclusively on traditional SEM software (e.g., M+, AMOS) and estimate the whole model (including the DCM component) as a SEM. An advantage of this approach is that it allows for alternative estimation methods that are dedicated to ordinal data, such as the weighted least squares method with a mean and variance correction. However, this approach lacks flexibility since, for instance, it requires that all error terms are normally distributed.

5. Psychological dimension

After reviewing the use of behavioural theories to explain mode choice, this section proposes a grid analysis to categorize the latent variables used in ICLV models and to discuss how latent variables are explained through the structural model and how latent variables influence mode choice.

5.1. *Linking ICLV models to behavioural theories*

ICLV models were primarily developed to better understand how behaviours are formed and to uncover the “black box” of decision processes. Following an analysis of case studies, [Ben-Akiva et al. \(2002, p.35\)](#) cautioned users of ICLV models to “first think clearly about the behavioural hypotheses behind the choices, then develop the framework, and then design a survey to support the model”. A key step in developing ICLV models is therefore to learn from the theories developed by behavioural scientists to study mode choice.

Behavioural scientists mainly use theories developed to explain pro-environmental behaviour ([Gifford, Steg, & Reser, 2011](#)). Three types of theories may be distinguished (Table 1): theories that explain behaviour with own personal interest and are based on a rational decision process; theories based on values and norms, and theories based on habits. A further distinction depends on whether the theory only relies on internal mental states, as in the Norm-Activation Model, or also takes into account contextual factors, as in the Attitude-Behaviour-Context theory. The contextual factors may act negatively (for instance, budget constraint) or positively (for instance, monetary incentives) on behaviour.

Five of the previously mentioned theories are used, or at least mentioned, in ICLV models on mode choice. However, even when clear reference is made to a specific theory, the SEM component generally simplifies it and incorporates only some variables directly in the utility of the DCM. The mediation and/or moderation structure of the psychological theories are generally omitted. An exception is [Paulssen et al. \(2014\)](#) who estimate a whole value-attitude-behaviour model by means of a SEM with a full mediation structure, with no direct action of values on behaviour.⁴ Two other theoretical examples of mediation or moderation structures can be found in the ICLV models, but the corresponding models are not estimated. [Abou-Zeid and Ben-Akiva \(2011\)](#) develop a theoretical model in which observed actions of others impact some latent variables, such as attitudes, perceptions or personality. These latent variables, in turn, impact utility for each transport mode, either directly or with the mediation of comparative happiness. In the theory of interpersonal behaviour on which [Tudela et al. \(2011\)](#) build their simplified model, behaviour is determined by intention. The relationship between both variables is moderated by the habit latent variable. That means that depending on habits, intention has a greater or smaller impact on the effective observed behaviour.

[Paulssen et al. \(2014, p.875\)](#) draw the same conclusions as us, noting that previous studies “tended to simplify significantly the cognitive theories motivating the use of these models, and much of the behavioural richness captured originally in these theories through the complex interplay between different latent psychological constructs has often been lost as a consequence of these simplifications.” [Roberts et al. \(2014, p.3\)](#) also note that “ICLV model studies of mode choice generally devote little or no attention to the theoretical model of decision making that underlies the empirical work”.

The underuse of behavioural theories may have two explanations. Firstly, incorporating behavioural theories is not always possible, since current estimation methods require long estimation times and raise convergence issues depending on the number of latent variables and indicators. However, the developments mentioned in the previous section seem to overcome these problems and may allow for a better inclusion of behaviourally sound SEMs in ICLV models. Secondly, some surveys may not have been developed with the specific purpose of estimating ICLV models. Econometric improve-

⁴Another exception, regarding departure time choice, is [Thorhauge, Haustein, and Cherchi \(2016\)](#).

Table 1. Behavioural theories used in ICIV models.

Theory	Summary	Used to explain mode choice in the psychological literature	Mentioned or (partly) used (*) in the ICIV literature.
Rational theories			
Theory of Planned Behaviour (TPB, Ajzen, 1991)	Rests on the assumption that the best predictor of a behaviour is intention, which is determined by attitudes, perceived social norms and perceived behavioural control with respect to the behaviour.	Bamberg, Ajzen, and Schmidt (2003); Bamberg, Hunecke, and Blöbaum (2007); Bamberg and Schmidt (2003); De Groot and Steg (2007); Heath and Gifford (2002); Wall, Devine-Wright, and Mill (2007)	Roberts et al. (2014), Paulssen et al. (2014), Sotfile et al. (2014*)
Attitude-Behaviour-Context (ABC, Guagnano, Stern, & Dietz, 1995) Theory	Behaviour is an interactive product of attitudinal variables and contextual factors. The attitude-behaviour association is strongest when contextual factors are neutral and it approaches zero when contextual forces are strongly positive or negative, effectively compelling or prohibiting the behaviour in question.		Roberts et al. (2014*)
Norms and values based theories			
Norm-Activation Model (NAM, Schwartz, 1977)	Originally developed in the context of altruistic behaviour. Behaviour is predicted by personal norms which are actively experienced “as feelings of moral obligation not as intentions”. Personal norms are determined by two factors: awareness that performing the particular behaviour has certain consequences, and feelings of responsibility for performing this specific behaviour	Bamberg et al. (2007); Bamberg and Schmidt (2003); Wall et al. (2007)	Roberts et al. (2014)
Value-Belief-Norm Theory (VBN, Stern, Dietz, Abel, Guagnano, & Kalof, 1999)	Is an adaptation of the NAM applied to pro-environmental behaviours. Makes a direct link between values (socio-altruist, egoistic and biospheric) and environmental beliefs which, in turn, determine personal norms and finally act on behaviour.	Collins and Chambers (2005); Nordlund and Garvill (2003)	
Value-attitude-behaviour Theory (Homer & Kahle, 1988)	Links values to attitudes and attitudes to behaviour in a full mediation model		Paulssen et al. (2014*)
Habits based theory			
Theory of Interpersonal Behaviour (TIB, Triandis, 1977)	As in TPB, intentions are immediate antecedents of behaviour. In the TIB, habits determine behaviour as well as contextual factors. Intention is shaped by three factors: attitude, social factors and affect	Bamberg and Schmidt (2003)	Paulssen et al. (2014), Roberts et al. (2014), Tudela et al. (2011*)

ments will therefore be fruitful only in cross-disciplinary work, involving psychologists, economists, and transport specialists, from the development of the questionnaire to the interpretation of the results.

5.2. *Categories of latent variables*

Latent variables included in ICLV models on mode choice are related to a wide variety of themes, reflecting the wealth and extent of the mode choice decision process. Almost 60 variables have been identified (Figure 3). They refer to personal values (Paulssen et al., 2014), comfort (Glerum et al., 2014), safety and security (Raveau et al., 2010), environmental concern (Kamargianni et al., 2015), convenience and accessibility (Yanez et al., 2010), flexibility (Anwar et al., 2014), reliability (Raveau et al., 2010), habits (Idris et al., 2015), social-comparison (Abou-Zeid & Ben-Akiva, 2011), information on stress (Sottile et al., 2014), affective attitude and addiction to the car (Polydoropoulou, Kamargianni, & Tsirimpa, 2014). These latent variables are very heterogeneous, in terms of meaning and measurement, even when a common term is used. As a consequence, comparisons between studies and results are difficult to make since they refer to concepts that are either too dissimilar or that have different measurements. An interpretative framework, an analysis grid, is therefore necessary to get a deeper understanding of how latent variables explain mode choice.

Following an in-depth qualitative examination of the literature on ICLV models applied to mode choice, a three-entries grid seems well adapted. The first entry relies on the topic to which the latent variables refer. Eight main topics are identified: environmental concern, comfort, safety, reliability, flexibility, convenience, habits and personality trait (Figure 3), the first three being incorporated in a third of the ICLV models on mode choice.

The second entry draws upon psychological concepts and distinguishes three categories of latent variables (Table 2).

- (1) The first category of latent variables describes the internal mental state of the individual (norms, values, needs), and the variables are similar to socio-economic variables. These variables are not related to a specific transport mode and they are measured with indicators, such as “How important is it for you to use a convenient and comfortable mode? A stress-free and relaxed mode? A mode on which you don’t have to worry about anything while using it?” (Johansson et al., 2006; Temme et al., 2007).
- (2) The second category of latent variables describes the individual attitudes⁵ towards or perceptions of the alternatives; here, the transport modes. In addition to be individual-specific, these variables are also alternative-specific and are similar to the objective attributes describing the alternatives. They are measured with indicators, such as “How do you rate the comfort during your trip by train?” or “In your opinion, how convenient is the coach schedule?” (Daziano & Rizzi, 2015; Raveau et al., 2010, e.g.).
- (3) The third category of latent variables describes an actual behaviour, either related to transport or to other areas, such as waste recycling. These variables are measured with items, such as “Do you wear a bicycle helmet when cycling?” or “Do you compost kitchen refuse?” (Johansson et al., 2006).

⁵The concept of attitude can be defined as the disposition to respond favourably or unfavourably to a specific object or class of objects (Oskamp, 1977).

The third entry divides the latent variables according to whether the individual is considered in isolation or taking into account her social context. Although the psychological literature underlines the importance of the social context in the shaping of behaviours by means of variables, such as subjective norms in the theory of planned behaviour or social factors in the theory of interpersonal behaviour, almost all reviewed variables fall into the first individual category. The only application of ICLV models on mode choice to take into account social context variables is [Kamargianni et al. \(2014\)](#) who investigate the impact of parents’ walking habits on teenagers’ mode choice. [Abou-Zeid and Ben-Akiva \(2011\)](#) study the effect of social comparisons on travel happiness and behaviour but do not estimate the whole ICLV model.

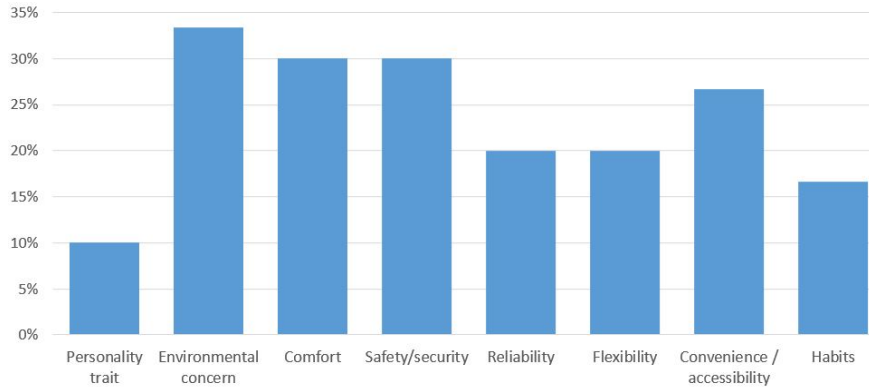


Figure 3. Topics related to latent variables included in ICLV mode choice models

5.3. *Explaining the latent variables*

The effect of latent variables on mode choice depends on the variables chosen in the structural model (Equation 5). However, “in practice, ICLV models often suffer from weak structural equations where the observable explanatory variables are poor predictors of the latent variables, and the latent variables are, in a sense, truly latent” ([Vij & Walker, 2016](#), p.197). The literature review confirms this finding. Firstly, as already mentioned, there are generally no relationships between the latent variables, such that socio-economic variables are the only determinants of latent variables. Secondly, the structural models offer contradictory results. For instance, need for comfort is explained by age and gender in three studies ([Anwar et al., 2014](#); [Johansson et al., 2006](#); [Paulssen et al., 2014](#)). Depending on the study, both variables are found to be associated with a negative, positive or non-significant parameter. Another example is income, which has either a positive ([Anwar et al., 2014](#)) or a non-significant ([Paulssen et al., 2014](#)) effect on need for comfort. These results suggest caution in the interpretation of structural models.

5.4. *Effects of latent variables on mode choice*

In most of the articles reviewed, the latent variables studied are significant determinants of mode choice with intuitive effects. The effects of latent variables on mode choice is discussed for the eight topics previously identified.

Table 2. Latent variables grid with examples

Topic	Internal mental states	Perceived alternative-specific variables	Actual behaviour
Personality trait	Temme et al. (2007) Power: concern for social status and prestige, and control or dominance over people and resource: e.g. appreciation of the statement "She/he always wants to be the one who makes the decisions. She/He likes to be the leader."	-	-
Environmental concern	Sottile et al. (2014) Personal norm regarding the environment: e.g. appreciation of the statement "Regardless of what other people do, I feel a moral duty to travel in an environmentally more sustainable way."	-	Johansson et al. (2006) Pro-environmental behaviour: e.g. the respondent's habit of composting kitchen refuse
Comfort	Johansson et al. (2006) Preferences for comfort: e.g. appreciation of the importance of travelling in a calm, non-noisy environment	Daziano and Rizzi (2015) Bus/Train comfort: e.g. appreciation of overall comfort (quality of seats, roominess, etc.)	-
Safety / security	Daziano (2015) Pro-safety consumers: appreciation of safety as a relevant aspect of travel decisions	Kamargianni et al. (2015) Safety Consciousness: e.g. appreciation of the statement "I feel safe when I use the bus"	Johansson et al. (2006) Safety behaviour: e.g. the respondent's habit of wearing a helmet when cycling
Reliability	Anwar et al. (2014) Reliability: e.g. appreciation of the importance of punctuality	Raveau et al. (2010) Reliability: e.g. appreciation of the possibility of calculating travel time before the trip for different modes (car, bus...)	-
Flexibility	Temme et al. (2007) Attitude towards flexibility: e.g. appreciation of the importance of using a means of transport that can be used spontaneously and without planning	-	-
Convenience / accessibility	Johansson et al. (2006) Preferences for convenience: e.g. appreciation of the importance of avoiding queues and congestion while travelling to/from work	La Paix Puello and Geurs (2015) Perception of connectivity: e.g. appreciation of connection to other modes public transport by train	-
Habits	Rieser-Schüssler and Axhausen (2012) Interest in varying ones daily routine: e.g. appreciation of the statement "I like to experience novelty and change in my daily life."	Idris et al. (2015) Habits: for non-work activities (e.g. to visit friends/family, to go to the movies, etc.) ask to provide the mode they would use. Counting how many times the respondent had mentioned each mode to perform different activities.	Kamargianni et al. (2014) Parent's walking habits: e.g. appreciation of the statement "My mother walks for her short-distance trips."

5.4.1. *Personality traits*

Two articles are specifically concerned with personality traits and personal values. [Paulssen et al. \(2014\)](#) find that personal values (power, hedonism and security) have an impact on attitudes, which in turn explain mode choice but these values have no direct impact on mode choice. [Córdoba and Jaramillo \(2012\)](#) use the 16 pf psychometric test to measure personality traits, such as emotional stability or spontaneity. With the hypothesis that they took the car as a reference alternative, it seems that people who choose an alternative mode are more pragmatic, shy, insecure and stressed and have a higher level of emotional instability than those who choose the car.

5.4.2. *Environmental concern*

As expected, the overall finding is that people with a higher level of environmental concern or environmental awareness are more likely to choose public transport, park and ride facilities or active transport modes ([Atasoy et al., 2013](#); [Polydoropoulou et al., 2014](#); [Rieser-Schüssler & Axhausen, 2012](#); [Sottile et al., 2014](#)). However, [Sottile et al. \(2014\)](#) notice that individuals are less sensitive to long-term effects, such as environmental pollution, than to effects that have immediate repercussions, such as stress. When environmental concern is measured through actual behaviour, the conclusion is the same. A green lifestyle favours the choice of bus and reduces the probability of choosing motorized modes, such as the car and powered two wheelers ([Kamargianni et al., 2015](#)), and increases the likelihood of choosing train over bus ([Johansson et al., 2006](#)).

[Politis et al. \(2012\)](#) and [Sottile et al. \(2013\)](#) are the two exceptions to these general results. In particular, [Sottile et al. \(2013\)](#) find that environmental concern, measured as propensity to adopt environmentally friendly behaviour (waste, energy, food, technologies but not transport), favours the car alternative over the park and ride alternative. The authors justify this contradictory result with cognitive dissonance, which is the inconsistency between attitudes and behaviour. Cognitive dissonance is common in a transportation context. One explanation is that pro-environmental motivation is not not enough to make people adopt behaviour that replaces or reduces private car use in favour of sustainable travel modes, since mode shift is not sufficiently easy. Indeed, individuals behave pro-environmentally only when an action is easy to perform. Another explanation is that environmental concern is measured with behavioural indicators related to non-transport fields. However, in the transportation field, environmental behaviour is generally moderately or poorly correlated with other kinds of environmentally friendly behaviour. Compensation effects may strengthen this explanation. Indeed, people with high environmental concern may ease their sense of guilt about using the car with increased environmental behaviours in other areas of life ([Johansson et al., 2006](#)).

In a RP survey, the same authors ([Sottile et al., 2014](#)) consider environmental personal norms as part of the theory of planned behaviour. After giving some information about reduction of CO2 emissions and stress, they find that high environmental moral rules and obligations lead individuals to behave in a pro-environmental way, producing an increase in the probability that they will choose park and ride instead of the car. This survey tends to prove an absence of cognitive dissonance. Three differences between the two surveys may explain these contradictory results. The first difference concerns the type of survey (RP *versus* SP). The second difference concerns the type of information provided (economic and health information for the RP survey or stress and CO2 information for the SP survey). The third difference concerns the measurement of

the latent variables, and that emphasizes the importance of clearly defining in which of the previously defined categories (internal mental states, perceived alternative-specific variables or actual behaviour) the latent variable falls.

5.4.3. *Comfort*

Despite heterogeneity in measures, the underlying idea behind the comfort latent variable is mainly to assess the extent to which public transport offers opportunities for alternative activities, such as relaxation and work. However, [Atasoy et al. \(2013\)](#) measure comfort with items translating difficulties of taking public transport (travelling with heavy luggage or children, having transfers, etc.). This can be approached by the notions of convenience or accessibility which are often paired with the comfort measure (e.g., [Morikawa & Sasaki, 1998](#); [Raveau et al., 2010](#)).

All studies involving comfort, either as a need or a perception, converge in a common finding. Comfort is an important factor in explaining mode choice. More precisely, having high comfort (and convenience) needs increase the probability of choosing a public transport mode ([Temme et al., 2007](#)). This can be explained by fatigue due to driving, lack of parking space, and the possibility of performing activities during travel by public transport.

Among public transport modes, trains are perceived as more comfortable than buses, which favours the choice of train over bus ([Daziano & Rizzi, 2015](#)). Indeed, in trains, it is generally easier to work or rest, since the seats are more comfortable, there is more space and there are fewer shocks and loads. A positive perception of public transport modes increases their utility ([Glerum et al., 2014](#)) and sensitivity to travel time decreases with a better perception of comfort on public transport. There also seems to be considerable heterogeneity with regard to the effect of comfort on the choice process since it varies importantly among individuals ([Yanez et al., 2010](#)). Moreover, the impact of latent variables on mode choice, and in particular comfort, may increase over time ([Anwar et al., 2014](#), with a comparison between years 2008/2009 and 2010/2011).

5.4.4. *Safety and security*

[Paulssen et al. \(2014\)](#) use the importance of owning the transport mode as a proxy for the safety value. The ownership variable tends to disadvantage public transit modes in comparison to the car. However, with regards to the choice between only ever using the car and driving to access public transport, the effect of ownership is statistically insignificant, since both alternatives require access to a car. For [Kamargianni et al. \(2015\)](#), safety consciousness increases the probability of choosing the car and decreases the probabilities of bus and walking.

[Yanez et al. \(2010\)](#), [Raveau et al. \(2010\)](#), [Anwar et al. \(2014\)](#) and [Daziano and Rizzi \(2015\)](#) evaluate the level of security for each mode integrated in their model. As expected, they find that individuals tend to choose those modes perceived as safer. For instance, the train is preferred to the bus ([Daziano & Rizzi, 2015](#)).

In [Johansson et al. \(2006\)](#), the safety variable, measured in terms of behaviour, is found to be insignificant in both the choice between car and bus and the choice between train and bus. One explanation is that safety and security constitute a combination of two variables: preferences for personal security and traffic safety. These two components work in opposite directions, since the car is usually perceived as being safer in terms of theft and personal safety, while public transport is more reliable in terms of traffic safety issues. Future research will therefore need to distinguish between these

two effects.

5.4.5. *Reliability*

Raveau et al. (2010) and Yanez et al. (2010) define reliability as the possibility of calculating travel and waiting time before the trip, and as the availability of suitable information. They find the intuitive result that the higher a mode is valued in terms of reliability, the greater the probability that individuals will choose that mode. La Paix Puello and Geurs (2015) show that a positive perception of connectivity, including punctuality and frequency of trains, is positively associated with bicycle use to go to railway stations. Johansson et al. (2006) include both concepts (punctuality and prior knowledge of travel time) in a global variable but they do not find any significant effect on mode choice. On the contrary, in Anwar et al. (2014), reliability is found to have the highest impact on mode choice, along with safety.

5.4.6. *Flexibility*

Latent variables specifically concerned with flexibility measure requirements for a mode that is available at any time, that can be used regardless of weather conditions, and that offers the possibility of shopping or transporting children during a commute to work. Need for flexibility tend to favour the car and metro over carpooling and other public transport modes (Johansson et al., 2006; Politis et al., 2012).

5.4.7. *Convenience and accessibility*

Convenience and accessibility are measured with indicators such as ease of access, frequency of public transport, the presence of queues, wait time and whether the mode is going to the required destination. It is always associated with indicators related to comfort, reliability or flexibility. The general conclusion is that modes that are perceived as convenient are more likely to be chosen (Anwar et al., 2014; Raveau et al., 2010; Yanez et al., 2010).

5.4.8. *Habits*

Car habits, measured with the Verplanken scale, result either in a non-significant effect on mode choice (Tudela et al., 2011) or in a positive effect that favours the car at the expense of the other modes (Idris et al., 2015). The latter result means that the choice of car as a transport mode is mainly determined by habits rather than traditional individual variables and modal attributes. The habits of relatives also determine mode choice. In particular, if parents tend to walk, teenagers are more prone to walk to school (Kamargianni et al., 2014).

Rieser-Schüssler and Axhausen (2012) approach the concept of habits by measuring the extent to which individuals are interested in varying their daily routine. They find that a greater desire for variation in the daily routine is associated with a higher probability of choosing public transport modes over the car.

In conclusion, the more individuals or their relatives are accustomed to a mode, the more likely they are to use that mode. This is especially true for car users, since they seem to be more attached to their routine than public transport users.

6. Economic dimension

ICLV models may be used with three objectives: 1) to explain behaviour and derive public policies; 2) to analyze economic outputs, and 3) to forecast market shares.

6.1. *Explaining behaviour and deriving public policies*

ICLV models generally allow for a better description of behaviours than the models to which they are compared (multinomial logit or probit models). Firstly, they outperform the model without latent variables in terms of fit to the data. Several articles also report a few shifts in the value and significance of the alternative-specific parameters and state that ICLV models make it possible to detect the real role of these variables (Raveau et al., 2010; Tudela et al., 2011).

Even if the statistical effects may be due to the inclusion of additional variables in the structural model of the SEM component (Vij & Walker, 2016), ICLV models still depict mode choice behaviour with a broader set of determinants. Understanding how these determinants influence mode choice is useful for deriving accurate public policies in two ways. Firstly, the individual variables related to internal mental states may help to understand how individuals will react to a policy and can be used to support the design of an efficient policy. For instance, Temme et al. (2007) report that the German rail operator changed its pricing policy in 2002. After some adjustments, it offered a discount card with a slightly higher discount rate (55% instead of 50%) but made it a requirement to book a seat seven days in advance. The operator did not expect the drop in market share that followed. The impact of the lack of flexibility due to the booking obligation was indeed underestimated if not actually overlooked. With a model integrating a latent measure of the need for flexibility, the outcome could have been anticipated and the pricing policy adapted to promote train use.

Secondly, the perception of the alternative-specific variables may be as important, or even more important than the objective levels of such variables (Bhattacharya, Brown, Jaroszynski, & Batuhan, 2014). Indeed, individual perception of the quality of public transport modes differs from the objective measures of service quality (for instance, travel time and reliability) used by planners to make and evaluate decisions. Kaufmann (2002), for instance, shows that the alleged slowness of public transport is not only a question of the objective time it takes to make a journey, but also of the perception of that time. So, the time it takes to travel by car is often underestimated, whereas the time it takes by public transport is overestimated. Measuring and taking into account both objective and subjective measures is therefore necessary to an understanding of mode choice. Moreover, some alternative-specific variables, such as safety or comfort, are difficult to measure objectively. It is therefore all the more important to be able to measure and integrate individual perceptions of the travel mode characteristics. ICLV models could help to understand how the negative perceptions of public transport modes affect their modal shares and set the priorities for improvement.

6.2. *Deriving economic outputs*

Since they are based on random utility theory, ICLV models may be used to derive economic outputs: market shares, elasticities and willingness-to-pay indicators, such as Value of Time (VoT). Atasoy et al. (2013) compare market shares across three models (ICLV, latent class and multinomial logit models) and observe that they do not vary

much. They draw the same conclusion regarding VOT. Conversely, [La Paix Puello and Geurs \(2015\)](#) find that the multinomial logit model overestimates the market share of bicycles in comparison to the ICLV model. They conclude that “missing important unobserved effects may lead to overestimation of demand solution in travel behaviour”. [Johansson et al. \(2006\)](#) also find notable differences between a multinomial logit model and a ICLV model in terms of VOT. [Yanez et al. \(2010\)](#) find inconsistent VOTs with a mixed logit without latent variables, reaffirming the importance of including such variables in the model formulation. [Glerum et al. \(2014\)](#) highlight the impact of ratings from different evaluations on the demand by exploring the three indicators, market shares, values of time and elasticities. [Polydoropoulou et al. \(2014\)](#) calculate ratios for public transport and private motorized vehicles regarding in-vehicle travel time and out-of-vehicle travel time. These examples show the range of economic outputs that are elicited with ICLV models and highlight the differences from multinomial logit models that might occur.

6.3. Forecasting market shares

ICLV models allow for an improvement in the forecast of modal shares, but only on two conditions. The first condition is that forecasts for measurement indicators are available but that condition is often not met. The second condition is that, in the structural model of the SEM, the observable explanatory variables are poor predictors of the latent variables ([Vij & Walker, 2016](#)), which is often the case. If forecasts for measurement items are expected to be unavailable, then ICLV models do not provide further insight for forecasting, compared to mixed logit models ([Vij & Walker, 2016](#)). Nevertheless, even if forecasting by means of an ICLV model has no added value from an econometric point of view, it may be interesting from a behavioural point of view. For instance, [Politis et al. \(2012\)](#) simulate the effect of a change of behavioural stage, that is, peoples’ awareness/readiness for changing travel habits, on utility. [Daziano \(2015\)](#) determines that women have higher pro-environment attitudes than men. In the forecasting exercise, they predict that public transport market shares would increase by 16.9% if men are represented as having the same attitudes as women. [Paulssen et al. \(2014\)](#) compare the impact on market shares of two scenarios: a 10% increase of car travel time or a 10% increase of attitudes regarding comfort and convenience. Their model predicts that the first scenario would reduce the market share for the car by 8.4%. In comparison, the second scenario is expected to reduce the market share for the car by only 0.9%. Another approach is to artificially change the values of the observable variables in the structural equations of the SEM. [Glerum et al. \(2014\)](#) decrease the number of cars in each household by one. This leads to a better perception on comfort on public transport and slightly increases (1.8%) the market share of public transport. [Anwar et al. \(2014\)](#) and [Yanez et al. \(2010\)](#) produce a similar argument and they vary the level of income. In conclusion, the simulation of scenarios could help to provide a measure of the scope for improvement of the modal share of public transport modes, or a measure of the respective roles of latent and observable variables in the achievement of a goal, such as the reduction of car use.

7. Conclusion

ICLV models are rich and complex models involving multiple dimensions and which therefore represent a fruitful tool for interdisciplinary programs. This review of liter-

ature details and explores how the survey, econometric, psychological and economic dimensions of ICLV models may serve to better explain mode choice. The econometric dimension is well addressed by the literature, and recent developments will certainly help to better integrate behavioural theories in discrete choice models. At the current time, although some of these theories are beginning to be explored, there is still room for further integration. There remains considerable scope in research for integrating more complex SEMs based on models previously estimated in travel behaviour research. Moreover, latent variables used in ICLV models on mode are heterogeneous in terms of definition and measure. To structure and facilitate comparisons, an analysis grid is proposed. I also suggest adapting the data collection process, with special attention to that part of the questionnaire dedicated to the measurement of the latent variables. It would also be interesting for future research to broaden the latent variables included in ICLV models on mode choice and include more widely social context variables. The usefulness of ICLV models for deriving public policies has been challenged in previous research ([Chorus & Kroesen, 2014](#); [Vij & Walker, 2016](#)). I provide examples to show that, although forecasting is indeed difficult with ICLV models, they are still a useful tool to explain mode choice, to identify which latent variables exert an influence on mode choice and to suggest how public policies may be designed to promote public transport use.

References

- Abou-Zeid, M., & Ben-Akiva, M. (2011). The effect of social comparisons on commute well-being. *Transportation Research Part A: Policy and Practice*, *45*(4), 345–361.
- Ajzen, I. (1991). The theory of planned behavior. *Organizational behavior and human decision processes*, *50*(2), 179–211.
- Anwar, A. M., Tieu, A. K., Gibson, P., Berryman, M. J., Win, K. T., McCusker, A., & Perez, P. (2014). Temporal and parametric study of traveller preference heterogeneity using random parameter logit model. *Journal of Travel Behaviour and Transportation*, *4*(4), 437–455.
- Atasoy, B., Glerum, A., & Bierlaire, M. (2013). Attitudes towards mode choice in Switzerland. *disP - The Planning Review*, *49*(2), 101–117.
- Bamberg, S., Ajzen, I., & Schmidt, P. (2003). Choice of travel mode in the theory of planned behavior: The roles of past behavior, habit, and reasoned action. *Basic and applied social psychology*, *25*(3), 175–187.
- Bamberg, S., Hunecke, M., & Blöbaum, A. (2007). Social context, personal norms and the use of public transportation: Two field studies. *Journal of Environmental Psychology*, *27*(3), 190–203.
- Bamberg, S., & Schmidt, P. (2003). Incentives, morality, or habit? predicting students car use for university routes with the models of ajzen, schwartz, and triandis. *Environment and behavior*, *35*(2), 264–285.
- Ben-Akiva, M., Walker, J., Bernardino, A. T., Gopinath, D. A., Morikawa, T., & Polydoropoulou, A. (2002). Integration of choice and latent variable models. *Perpetual motion: Travel behaviour research opportunities and application challenges*, 431–470.
- Bhat, C. R., & Dubey, S. K. (2014). A new estimation approach to integrate latent psychological constructs in choice modeling. *Transportation Research Part B: Methodological*, *67*, 68–85.
- Bhattacharya, T., Brown, J., Jaroszynski, M., & Batuhan, T. (2014). The effects of perception vs. reality on travel behavior after a major transit service change: The case of Tallahassee, Florida. *Journal of Public Transportation*, *17*(2), 1.
- Bollen, K., & Hoyle, R. (2012). Latent variables in structural equation modeling. In *Handbook of structural equation modeling*. Guilford Press.
- Bollen, K. A. (2014). *Structural equations with latent variables*. John Wiley & Sons.
- Chorus, C. G., & Kroesen, M. (2014). On the (im-) possibility of deriving transport policy implications from hybrid choice models. *Transport Policy*, *36*, 217–222.
- Collins, C. M., & Chambers, S. M. (2005). Psychological and situational influences on commuter-transport-mode choice. *Environment and behavior*, *37*(5), 640–661.
- Córdoba, J., & Jaramillo, G. (2012). Inclusion of the latent personality variable in multinomial logit models using the 16pf psychometric test. *Procedia-Social and Behavioral Sciences*, *54*, 169–178.
- Daziano, R. A. (2015). Inference on mode preferences, vehicle purchases, and the energy paradox using a bayesian structural choice model. *Transportation Research Part B: Methodological*, *76*, 1–26.
- Daziano, R. A., & Bolduc, D. (2013). Covariance, identification, and finite-sample performance of the msl and bayes estimators of a logit model with latent attributes. *Transportation*, *40*(3), 647–670.
- Daziano, R. A., & Rizzi, L. I. (2015). Analyzing the impact of a fatality index on a discrete, interurban mode choice model with latent safety, security, and comfort. *Safety science*, *78*, 11–19.
- De Groot, J., & Steg, L. (2007). General beliefs and the theory of planned behavior: The role of environmental concerns in the TPB. *Journal of Applied Social Psychology*, *37*(8), 1817–1836.
- De Witte, A., Hollevoet, J., Dobruszkes, F., Hubert, M., & Macharis, C. (2013). Linking modal choice to motility: A comprehensive review. *Transportation Research Part A: Policy and Practice*, *49*, 329–341.

- Dunlap, R., Liere, K., Mertig, A., & Jones, R. E. (2000). Measuring endorsement of the new ecological paradigm: A revised nep scale. *Journal of Social Issues, 56*(3), 425–442.
- Fernández-Antolín, A., Guevara-Cue, A., de Lapparent, M., & Bierlaire, M. (2016). Correcting for endogeneity due to omitted attitudes: Empirical assessment of a modified mis method using rp mode choice data. *Journal of Choice Modelling, 20*, 1–15.
- Gifford, R., Steg, L., & Reser, J. P. (2011). Environmental psychology. In *Iaap handbook of applied psychology* (pp. 440–470). Wiley Online Library.
- Glerum, A., Atasoy, B., & Bierlaire, M. (2014). Using semi-open questions to integrate perceptions in choice models. *Journal of Choice Modelling, 10*, 11–33.
- Guagnano, G. A., Stern, P. C., & Dietz, T. (1995). Influences on attitude-behavior relationships a natural experiment with curbside recycling. *Environment and behavior, 27*(5), 699–718.
- Habib, K. M. N., Tian, Y., & Zaman, H. (2011). Modelling commuting mode choice with explicit consideration of carpool in the choice set formation. *Transportation, 38*(4), 587–604.
- Habib, K. M. N., & Zaman, M. H. (2012). Effects of incorporating latent and attitudinal information in mode choice models. *Transportation Planning and Technology, 35*(5), 561–576.
- Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford Press.
- Heath, Y., & Gifford, R. (2002). Extending the theory of planned behavior: predicting the use of public transportation1. *Journal of Applied Social Psychology, 32*(10), 2154–2189.
- Hinkin, T. R., Tracey, J. B., & Enz, C. A. (1997). Scale construction: Developing reliable and valid measurement instruments. *Journal of Hospitality & Tourism Research, 21*(1), 100–120.
- Homer, P. M., & Kahle, L. R. (1988). A structural equation test of the value-attitude-behavior hierarchy. *Journal of Personality and social Psychology, 54*(4), 638.
- Idris, A. O., Habib, K. M. N., & Shalaby, A. (2015). An investigation on the performances of mode shift models in transit ridership forecasting. *Transportation Research Part A: Policy and Practice, 78*, 551–565.
- Johansson, M. V., Heldt, T., & Johansson, P. (2006). The effects of attitudes and personality traits on mode choice. *Transportation Research Part A: Policy and Practice, 40*(6), 507–525.
- John, O. P., & Benet-Martínez, V. (2000). *Measurement: Reliability, construct validation, and scale construction*. Cambridge University Press.
- Kamargianni, M., Ben-Akiva, M., & Polydoropoulou, A. (2014). Incorporating social interaction into hybrid choice models. *Transportation, 41*(6), 1263–1285.
- Kamargianni, M., Dubey, S., Polydoropoulou, A., & Bhat, C. (2015). Investigating the subjective and objective factors influencing teenagers school travel mode choice—an integrated choice and latent variable model. *Transportation Research Part A: Policy and Practice, 78*, 473–488.
- Kaufmann, V. (2002). Temps et pratiques modales. le plus court est-il le mieux? *Recherche-Transports-Sécurité, 75*, 131–143.
- Kaufmann, V., Tabaka, K., Jean-Marie, G., & Louvet, N. (2010). *Et si les français n’ avaient plus seulement une voiture dans la tête?* (Tech. Rep.). Certu.
- La Paix Puello, L., & Geurs, K. (2015). Modelling observed and unobserved factors in cycling to railway stations: application to transit-oriented-developments in the netherlands. *European journal of transport and infrastructure research, 15*(1), 27–50.
- Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology, 22*(140), 1–55.
- Little, T. D., Card, N. A., Bovaird, J. A., Preacher, K. J., & Crandall, C. S. (2007). Structural equation modeling of mediation and moderation with contextual factors. In *Modeling contextual effects in longitudinal studies* (pp. 207–230). Guilford Press.
- Maldonado-Hinarejos, R., Sivakumar, A., & Polak, J. W. (2014). Exploring the role of in-

- dividual attitudes and perceptions in predicting the demand for cycling: a hybrid choice modelling approach. *Transportation*, 41(6), 1287–1304.
- Mokhtarian, P. L., & Salomon, I. (2001). How derived is the demand for travel? some conceptual and measurement considerations. *Transportation Research Part A: Policy and Practice*, 35(8), 695–719.
- Morikawa, T., & Sasaki, K. (1998). *Discrete choice models with latent variables using subjective data*.
- Niininen, O., Szivas, E., & Riley, M. (2004). Destination loyalty and repeat behaviour: an application of optimum stimulation measurement. *International Journal of Tourism Research*, 6(6), 439–447.
- Nordlund, A. M., & Garvill, J. (2003). Effects of values, problem awareness, and personal norm on willingness to reduce personal car use. *Journal of environmental psychology*, 23(4), 339–347.
- O'Brien, R. M. (1994). Identification of simple measurement models with multiple latent variables and correlated errors. *Sociological methodology*, 24, 137–170.
- Osgood, C. E., & Luria, Z. (1954). A blind analysis of a case of multiple personality using the semantic differential. *The Journal of Abnormal and Social Psychology*, 49(4p1), 579.
- Oskamp, S. (1977). *Attitudes and opinions*. Englewood Cliffs N J: Prentice-Hall.
- Paulssen, M., Temme, D., Vij, A., & Walker, J. L. (2014). Values, attitudes and travel behavior: a hierarchical latent variable mixed logit model of travel mode choice. *Transportation*, 41(4), 873–888.
- Politis, I., Papaioannou, P., & Basbas, S. (2012). Integrated choice and latent variable models for evaluating flexible transport mode choice. *Research in Transportation Business & Management*, 3, 24–38.
- Polydoropoulou, A., Kamargianni, M., & Tsirimpa, A. (2014). *Car use addiction vs. ecological consciousness: Which one prevails on mode choice behavior*. Lulu Publishers, USA.
- Raveau, S., Álvarez-Daziano, R., Yáñez, M., Bolduc, D., & Ortúzar, J. d. D. (2010). Sequential and simultaneous estimation of hybrid discrete choice models: some new findings. *Transportation Research Record: Journal of the Transportation Research Board*(2156), 131–139.
- Raveau, S., Yáñez, M. F., & Ortúzar, J. d. D. (2012). Practical and empirical identifiability of hybrid discrete choice models. *Transportation Research Part B: Methodological*, 46(10), 1374–1383.
- Rieser-Schüssler, N., & Axhausen, K. (2012). Investigating the influence of environmentalism and variety seeking on mode choice. *Transportation Research Record: Journal of the Transportation Research Board*(2322), 31–41.
- Roberts, J., Popli, G., & Harris, R. J. (2014). Do environmental concerns affect commuting choices? hybrid choice modelling with household survey data. *The Sheffield Economic Research Paper Series (SERPS)*.
- Scagnolarie, S., Walker, J., & Maggi, R. (2015). Young drivers night-time mobility preferences and attitude toward alcohol consumption: A hybrid choice model. *Accident Analysis & Prevention*, 83, 74–89.
- Schwartz, S. H. (1977). Normative influences on altruism. *Advances in experimental social psychology*, 10, 221–279.
- Schwartz, S. H., Melech, G., Lehmann, A., Burgess, S., Harris, M., & Owens, V. (2001). Extending the cross-cultural validity of the theory of basic human values with a different method of measurement. *Journal of cross-cultural psychology*, 32(5), 519–542.
- Sottile, E., Cherchi, E., & Meloni, I. (2014). Measuring soft measures within a stated preference survey: the effect of pollution and traffic stress on mode choice. In (Vol. 11, pp. 434–451). Elsevier.
- Sottile, E., Meloni, I., & Cherchi, E. (2013). A hybrid discrete choice model to assess the effect of awareness and attitude towards environmentally friendly travel modes. In (Vol. 5, pp. 44–55). Elsevier.
- Stern, P. C., Dietz, T., Abel, T. D., Guagnano, G. A., & Kalof, L. (1999). A value-belief-norm theory of support for social movements: The case of environmentalism. *Human Ecology*

- Review*, 6(2), 81.
- Temme, D., Paulssen, M., Dannewald, T., et al. (2007). Integrating latent variables in discrete choice models—how higher-order values and attitudes determine consumer choice. *Berlin, Germany: Humboldt University*.
- Thorhauge, M., Haustein, S., & Cherchi, E. (2016). Accounting for the theory of planned behaviour in departure time choice. *Transportation Research Part F: Traffic Psychology and Behaviour*, 38, 94–105.
- Thurstone, L. L. (1927). A law of comparative judgment. *Psychological review*, 34(4), 273.
- Tinsley, H. E., & Brown, S. D. (2000). *Handbook of applied multivariate statistics and mathematical modeling*. Academic Press.
- Train, K. E. (2009). *Discrete choice methods with simulation*. Cambridge university press.
- Triandis, H. C. (1977). *Interpersonal behavior*. Brooks/Cole Pub. Co.
- Tudela, A., Habib, K., Carrasco, J., & Osman, A. (2011). Incorporating the explicit role of psychological factors on mode choice: A hybrid mode choice model by using data from an innovative psychometric survey. In *2nd international choice modelling conference, leeds*.
- Verplanken, B., Aarts, H., van Knippenberg, A., & van Knippenberg, C. (1994). Attitude versus general habit: Antecedents of travel mode choice. *Journal of Applied Social Psychology*, 24(4), 285 - 300.
- Vij, A., & Walker, J. (2014). *Hybrid choice models: The identification problem*. Edward Elgar Publishing Limited.
- Vij, A., & Walker, J. L. (2016). How, when and why integrated choice and latent variable models are latently useful. *Transportation Research Part B: Methodological*, 90, 192–217.
- Walker, J., Li, J., Srinivasan, S., & Bolduc, D. (2013). Travel demand models in the developing world: Correcting for measurement errors. *Transportation Letters*.
- Walker, J. L. (2001). *Extended discrete choice models: integrated framework, flexible error structures, and latent variables* (Unpublished doctoral dissertation). Massachusetts Institute of Technology.
- Wall, R., Devine-Wright, P., & Mill, G. A. (2007). Comparing and combining theories to explain proenvironmental intentions: The case of commuting-mode choice. *Environment and behavior*, 39(6), 731–753.
- Yanez, M. F., Raveau, S., & Ortúzar, J. d. D. (2010). Inclusion of latent variables in mixed logit models: modelling and forecasting. *Transportation Research Part A: Policy and Practice*, 44(9), 744–753.
- Zikmund, W. G., Babin, B. J., Carr, J. C., & Griffin, M. (2013). *Business research methods*. Cengage Learning.