Incorporating Latent Variables into Discrete Choice Models –
A Simultaneous Estimation Approach Using SEM Software

Dirk Temme, Marcel Paulssen, Till Dannewald

Accounting | Finance | Management | **Marketing** | Operations and Information Systems

Incorporating Latent Variables into Discrete Choice Models – A Simultaneous Estimation Approach Using SEM Software

Dirk Temme, Institute of Marketing, Humboldt University of Berlin, Germany, E-Mail: temme@wiwi.hu-berlin.de
Marcel Paulssen, HEC Hautes Etudes Commerciales, Université de Genève, Switzerland, E-Mail: Marcel.Paulssen@unige.ch
Till Dannewald, Infas TTR Frankfurt, Germany, E-Mail: till.dannewald@infas-ttr.de

Abstract

Integrated choice and latent variable (ICLV) models represent a promising new class of models which merge classic choice models with the structural equation approach (SEM) for latent variables. Despite their conceptual appeal, applications of ICLV models in marketing remain rare. We extend previous ICLV applications by first estimating a multinomial choice model and, second, by estimating hierarchical relations between latent variables. An empirical study on travel mode choice clearly demonstrates the value of ICLV models to enhance the understanding of choice processes. In addition to the usually studied directly observable variables such as travel time, we show how abstract motivations such as power and hedonism as well as attitudes such as a desire for flexibility impact on travel mode choice. Furthermore, we show that it is possible to estimate such a complex ICLV model with the widely available structural equation modeling package Mplus. This finding is likely to encourage more widespread application of this appealing model class in the marketing field.

Keywords: Hybrid choice models, Mode choice, Mplus, Value-attitude hierarchy

1. Introduction

Discrete choice models are extensively used in various academic fields to analyze a huge range of choices between mutually exclusive alternatives (e.g., brands, service providers, travel modes, financial investments, residences, political parties, or strategies). Traditionally, these models have directly mapped observed features of alternatives and observed characteristics of decision makers to overt choice behavior. For instance, in order to explain travel mode choice for daily work trips, modal attributes (e.g., travel time and cost) as well as commuter socio-demographics (e.g., household income and number of drivers in a household) have been considered (e.g., Train 1978). The decision maker's internal processes during preference formation and notably the role of factors that are not directly ob-

servable, such as attitudes or lifestyle preferences, remain unexplained in a so-called "black box" in traditional discrete choice analysis.

Meanwhile, researchers have increasingly recognized that decision makers differ significantly in psychological constructs such as attitudes, perceptions, values, or lifestyle preferences and that these factors affect an alternative's utility in a systematic way (Ben-Akiva, McFadden, Train, Walker, Bhat, Bierlaire, Bolduc, Boersch-Supan, Brownstone, Bunch, Daly, De Palma, Gopinath, Karlstrom, and Munizaga 2002; Walker and Ben-Akiva 2002). Mode choice decisions, for example, might not only depend on objective criteria (e.g., time, income) but also on commuters' preferences for convenience, safety, or flexibility (e.g., Vredin Johansson, Heldt, and Johansson 2006). Two otherwise identical commuters differing in their desire for flexibility

might thus choose different travel modes. Extending choice models with latent variables representing attitudes or values can therefore lead to a deeper understanding of the choice processes taking place in the consumer's "black box" and at the same time should provide greater explanatory power. Therefore, integrated choice and latent variable (ICLV) models which merge classic choice analysis with the structural equation approach (SEM) for latent variables represent a promising new class of models. Recently, some encouraging applications of ICLV models have appeared in the literature: Explaining prototype choice in conjoint analysis by incorporating subjective product characteristics (i.e., perceptions) (Luo, Kannan, and Ratchford 2008); analyzing private asset investments taking into account factors such as individual risk attitude and impatience (Eyman, Börsch-Supan, and Euwals 2002); and modeling the impact of lifestyle preferences on residence choice (Walker and Li 2007).

Despite their conceptual appeal, there are still relatively few applications of ICLV models in marketing and related fields. The major reason for their lack of popularity is most likely the fact that full information estimation of these models is rather involved and hitherto it was been required that researchers develop their own programs (Ben-Akiva, McFadden, Train, Walker, Bhat, Bierlaire, Bolduc, Börsch-Supan, Brownstone, Bunch, Daly, De Palma, Gopinath, Karlstrom, and Munizaga 2002).[1] Most of the rare current applications are restricted to binary choice and, with the noticeable exception of the paper by Dellaert and Stremersch (2005), only consider direct effects of latent variables on choice (e.g., Ben-Akiva, Walker, Bernardino, Gopinath, Morikawa, and Polydoropoulou 2002; Ashok, Dillon, and Yuan 2002). Thus, causal relationships between latent variables commonly investigated in structural equation modeling are neglected. In contrast, we test a behavioral theory, the value-attitude hierarchy, which proposes hierarchical relationships between latent variables in a discrete choice analysis. Furthermore, by applying the program Mplus (Muthén and Muthén 1998–2007), one of the most comprehensive software packages for SEM, we present a powerful and very flexible option for estimating ICLV models which has not been considered so far.

To sum up, our paper primarily provides a methodical contribution. We extend previous ICLV applications by first estimating a multinomial choice model and, second, by estimating hierarchical relations between latent variables. Thus, unlike previous applications of the ICLV model, we do not only include latent variables as an additional set of predictors (e.g., Ben-Akiva, Walker, Bernardino, Gopinath, Morikawa, and Polydoropoulou 2002; Ashok, Dillon, and Yuan 2002). Furthermore, our paper extends the transportation choice literature by testing a value-attitude hierarchy with the impact of commuters' personal values on "soft" choice criteria and on subsequent mode choice.

The remaining part of the paper is structured as follows: First, we introduce the general structure of ICLV models and discuss their estimation with the Mplus software. Then, we illustrate the applicability of Mplus in an empirical study on travel mode choice. A hierarchical behavioral model is tested in which we include values and attitudes as well as traditional alternative-specific and socio-demographic variables such as explanatory variables. Next, we discuss the implications of the comprehensive ICLV framework in general as well as in the context of our empirical study. We conclude by addressing some limitations and by providing avenues for further research.

2. The integrated choice and latent variable model

In the general formulation of the ICLV model, two components can be distinguished: a multinomial discrete choice model and a latent variable model (see Figure 1). Each of these submodels consists of a structural and a measurement part. In the discrete choice part, the alternatives' utilities may depend on both observed and latent characteristics of the options and decision makers. Utility as a theoretical construct is operationalized by a single nominal indicator representing the observed choice for each individual. The latent variable part is rather flexible in that it allows for both simultaneous relationships between the latent variables and MIMIC-type models (Jöreskog and Goldberger 1975) where observed exogenous variables influence the latent variables. Such a specification enables the researcher to disentangle the direct and indirect effects of observed as
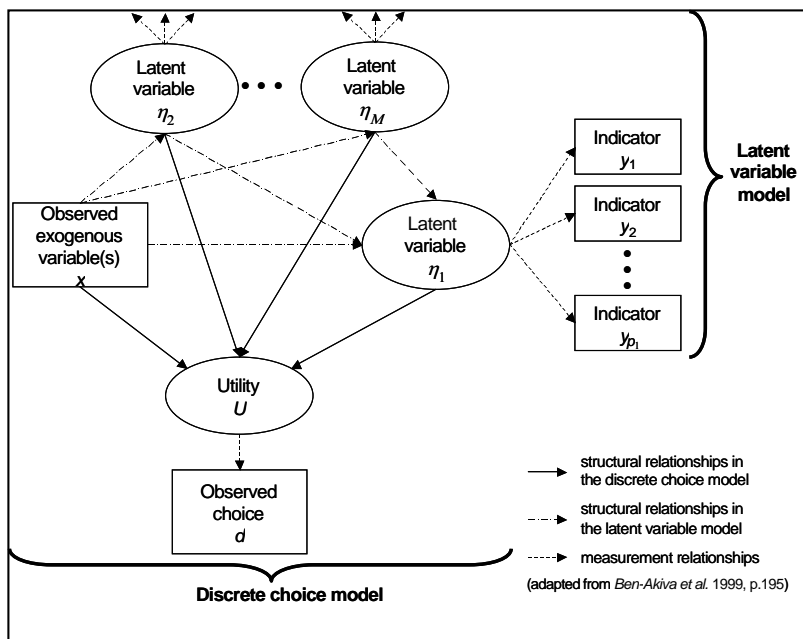
---

[1] Vredin Johansson, Heldt, and Johansson (2006), for instance, avoided the simultaneous estimation approach because it was considered rather cumbersome.

well as latent variables on the alternatives' utilities. The latent variables themselves are assumed to be measured by multiple indicators representing, for example, the respondents' answers to survey items.

2.1 Model specification

The structure and full information estimation of the ICLV model will now be discussed in more detail. For alternative treatments refer, for example, to Ashok, Dillon, and Yuan (2002), Walker and Ben-Akiva (2002), and Bolduc, Ben-Akiva, Walker and Michaud (2005).

Figure 1: Framework for integrated choice and latent variable models

(adapted from *Ben-Akiva et al.* 1999, p.195)

Discrete Choice Model: The random utility component is based on the assumption that a decision maker n ($n = 1,..., N$), faced with a finite set C_n of mutually exclusive alternatives i ($i = 1,..., I_n$), chooses the option i which provides the greatest utility U_{ni}. Each alternative's utility is described as a function of explanatory variables forming the systematic part of the utility, $V(\cdot)$, and disturbances, v_{ni}, representing the stochastic utility component:

$$(1) \qquad U_{ni} = V\left(x_{ni}, \eta_{ni}; \beta\right) + v_{ni},$$

where x_{ni} is a $(K \times 1)$ vector of observed variables and η_{ni} is a $(M \times 1)$ vector of latent variables. These

variables represent either (latent) characteristics of the decision maker (x_{sn}, η_{sn}) or (latent) attributes of the alternatives (x_{zni}, η_{zni}). The importance of the explanatory variables on the utility of the options is reflected in the $((K+M) \times 1)$ vector β. By assuming, for example, that each v_{ni} is independently, identically distributed (i.i.d.) extreme value, the widely used multinomial logit (MNL) model results (e.g., Ben-Akiva and Lerman 1985):

$$(2) \qquad P(u_{ni} = 1 \mid x_{ni}, \eta_{ni}; \beta) = \frac{e^{V(x_{ni}, \eta_{ni}; \beta)}}{\sum_{j \in C_n} e^{V(x_{nj}, \eta_{nj}; \beta)}}$$

As is common practice in choice modeling, the representative utility $V(\cdot)$ is specified to be linear in parameters:

$$(3) \qquad V_{ni} = \beta_x' x_{ni} + \beta_\eta' \eta_{ni},$$

where β_x and β_η is a $(K \times 1)$ and a $(M \times 1)$ vector, respectively.

Latent Variable Model: Model identification typically requires that the unobserved ηs are operationalized by multiple manifest variables, y.[2] In the simplest case, a linear factor model is appropriate for describing the mapping of the indicators onto the latent variables, leading to the following measurement equation:

$$(4) \qquad y_{ni} = \Lambda \eta_{ni} + \varepsilon_{ni},$$

where y_{ni} is a $(P \times 1)$ vector, Λ is a $(P \times M)$ matrix of factor loadings, and ε_{ni} is a $(P \times 1)$ vector of measurement errors which are i.i.d. multivariate normal.[3] Our structural model for the latent variables integrates alternative formulations by Ashok, Dillon, and Yuan (2002) and Walker and Ben-Akiva (2002) by allowing for interrelationships among the latent variables as well as for the influence of observed

---

[2] This has been established in further analyses.

[3] Alternatively, for ordinal indicators, a factor model with continuous latent response variables might be specified (Muthén 1983, 1984).

explanatory variables $\boldsymbol{x}_{ni}$ on the latent variables:[4]

(5) $\qquad \boldsymbol{\eta}_{ni} = \mathbf{B}\boldsymbol{\eta}_{ni} + \boldsymbol{\Gamma}\boldsymbol{x}_{ni} + \boldsymbol{\zeta}_{ni},$

where the $(M \times M)$ matrix $\mathbf{B}$ and the $(M \times L)$ matrix $\boldsymbol{\Gamma}$ contain unknown regression coefficients. The $(M \times 1)$ vector $\boldsymbol{\zeta}_{ni}$ represents random disturbances assumed to be i.i.d. multivariate normal.

<u>Likelihood Function</u>: Since all information about the latent variables is contained in the multiple observed indicators, the joint probability of the choice and latent variable indicators conditioned on the exogenous variables is considered. Assuming that the random errors ν, ε, and ζ are independent, integrating over the joint distribution of the latent variables leads to the following multidimensional integral:[5]

(6) $\qquad P(u_i = 1 \mid x, \theta) = \int_{R_\eta} P_u(u_i = 1 \mid x, \eta, \beta, \Sigma_\nu)$

$\qquad \times f_y(y \mid \eta, \Lambda, \Sigma_\varepsilon) f_\eta(\eta \mid x; B, \Gamma, \Sigma_\zeta) d\eta$

where θ represents the model parameters, P_u denotes the probability function of observing the choice of a specific alternative (2), the density function f_y for the latent variable indicators relates to the measurement model (4), and the density function f_η of the latent variables corresponds to the structural model (5). Σ_ν, Σ_ε, and Σ_ζ indicate covariance matrices for the residuals. R_η denotes that integration is over the range space of the vector of latent variables that have a direct impact on the choice decision. If maximum likelihood techniques are applied to estimate the parameter vector θ in (6), for any particular individual we obtain the following likelihood function:

(7) $\qquad L = \prod_{i \in C_n} P(u_i = 1, y \mid x, \theta)^{u_i}$

$\qquad = \int_{R_\eta} \prod_{i \in C_n} P(u_i = 1 \mid x, \eta, \beta, \Sigma_\nu)^{u_i} f_y(y \mid \eta, \Lambda, \Sigma_\varepsilon)$

$\qquad \times f_\eta(\eta \mid x; B, \Gamma, \Sigma_\zeta) d\eta$

where $u_i = 1$ if the decision maker chooses i and zero otherwise.

2.2 Estimation

The simplest way to include psychological constructs in a discrete choice analysis is to perform a sequential estimation (e.g., Ashok, Dillon, and Yuan 2002). First, the latent variable part of the ICLV model is estimated (Eqs. (4) and (5)) and factor scores are computed. Second, factor scores substitute the latent variables in the discrete choice model (Eq. (3)) as additional error-free exogenous variables $\boldsymbol{x}_{ni}$. This two-step limited information estimation is straightforward, using standard software for both discrete choice (e.g., SAS, NLOGIT, or STATA) and SEM (e.g., LISREL, AMOS, or EQS) analysis. However, the approach is deficient in the sense that (1) it leads to inconsistent and biased estimates for the random utility part (e.g., Walker and Ben-Akiva 2002)[6] and (2) does not allow to test behavioral theories including more complex relationships between latent predictors and revealed choice (e.g., direct and indirect effects for multiple layers of latent variables).

Full information estimation, on the other hand, is rather involved due to the multidimensional integral in Eq. (6). For a restricted number of latent variables (typically three or fewer variables) entering the utility function, numerical integration methods like Gaussian quadrature are feasible (e.g., Ashok, Dillon, and Yuan 2002). With an increasing number of latent variables, the computational complexity rises exponentially. Hence, in the case of more than three latent variables, other techniques like Monte Carlo integration are found to be more appropriate (for a discussion see Judd (1998)). So far, researchers conducting full information estimation of an ICLV model have developed their own routines in flexible statistic software such as GAUSS (e.g., Ashok, Dillon, and Yuan 2002). A more convenient way proposed here is to use the SEM software package Mplus (Muthén and Muthén 1998-2007) the capabilities of which make it suitable for a broad range of applications of the ICLV approach.[7] Besides offering

---

[4] For simplicity, it is assumed that the same set of observed exogenous variables $\boldsymbol{x}$ as in the random utility equation (1) enters the structural model for the latent variables. Of course, different variable sets can be defined for both equations.

[5] The index denoting the individual decision maker has been dropped for notational simplicity.

[6] For a sequential estimation approach, which leads to consistent but not fully efficient estimators, see Morikawa, Ben-Akiva, and McFadden (2002).

[7] Alternatively, ICLV models might be estimated using the software programs GLAMM (Skrondal and Rabe-Hesketh 2004) or Latent GOLD Choice (Vermunt and Magidson 2005).

the full flexibility of an SEM program to specify complex structures of latent variables, both numerical and Monte Carlo integration are available for simultaneously estimating a multinomial logit model with latent predictors. Mplus allows one to perform both conventional as well as robust maximum likelihood estimation. In the latter case, the asymptotic standard errors are corrected for non-severe violations of the distributional assumptions associated with the stochastic disturbances. We will now briefly outline the peculiarities of the Mplus approach for the analysis of ICLV models.

Whereas SEM in its narrowest formulation assumes continuous indicators, Mplus allows for a proper treatment of various other types of observed data (e.g., ordered and unordered categorical). In the present case, we make use of Mplus' ability to handle nominal indicators representing the observed choice for each decision maker. Even though specifying an unconditional MNL model in Mplus is straightforward, estimating a conditional MNL model including alternative-specific variables is slightly more difficult.[8] The reason for this is that for the last alternative no explicit utility function can be specified. Since only differences in utility matter (e.g., Train 2003), this issue can be solved by expressing the utilities of the other choice options as differences with respect to the last alternative's utility. For that purpose, the model constraint option implemented in Mplus is used. To be more specific, an alternative-specific variable describing the last choice option enters all other utility functions with the opposite sign by specifying the required linear constraint for the parameter of that variable. Because the effect of individual-specific variables is not identified for all alternatives, for each such variable one parameter has to be normalized to zero;[9] the same also applies to the alternative-specific constants which capture the average influence of unobserved effects on an alternative's utility. In Mplus, these constraints are imposed by default, as all parameters in the utility function of the last choice option are fixed to zero.

Since Mplus' model specification underlying the analysis of nominal variables is not well documented, we assessed the validity of our approach in two ways: First, estimation results for a conventional conditional MNL model were compared to those of standard software (NLOGIT and STATA); no differences to the Mplus parameter estimates were revealed. Second, a Monte Carlo simulation study including four observed and two latent explanatory variables was conducted using both numerical and Monte Carlo integration. The performance of Mplus has been assessed in terms of the mean absolute relative bias (MARB) of both the parameter estimates and the standard errors, and the 95% coverage rate. The simulation results provide the confidence to conclude that Mplus performs reasonably well at sample sizes of 500 (under this condition the MARB for the coefficient estimates only slightly exceeds the threshold value of 0.025 suggested by Boomsma and Hoogland 2001). Excellent results emerge for a sample size of $n = 1,000$.[10] In the following chapter, we apply Mplus to estimate an ICLV model with hierarchical relations between latent variables to explain travel mode choice.

3. Empirical analysis of an integrated choice and latent variable model of (travel mode) choice

3.1 Model development

In traditional travel choice models, individual mode choice is analyzed both as a function of individual characteristics of the decider such as income, employment status, gender, number of children, etc., and of attributes of the travel mode alternatives such as travel time, travel cost, availability, etc. However, in the last 10 years, many researchers have criticized this approach and called for the inclusion of unobservable or latent variables such as preferences for convenience, flexibility, or safety into models of mode choice (e.g., McFadden 1986; Ashok, Dillon, and Yuan 2002; Morikawa, Ben-Akiva, and McFadden 2002). The overall idea is that the inclusion of latent variables, mirroring an individual's preference or attitudes, is a more adequate representation of behavior and helps to gain valuable insight into the decision-making process of the individual (Vredin Johansson, Heldt, and Johansson 2006). In the following, we develop an extended

---

[8] For a discussion of both the unconditional and the conditional logit model see, for example, Green (2008).

[9] For alternative identification constraints refer to Train (2003).

[10] More detailed results of the Monte Carlo study can be requested from the first author.

choice model of travel mode choice based on an individual's values, attitudes, and demographics as well as typical characteristics of the traffic mode alternatives such as time and availability. Recent research indicates that more abstract constructs such as values, lifestyle orientation, and personality traits might also impact travel mode choice (Choo and Mokhtarian 2004; Nordlund and Garwill 2003; Collins and Chambers 2005). Especially values[11] have received some attention, and several researchers have investigated their role in travel mode choice (Bamberg 1996; Bamberg and Kühnel 1998; Choo and Mokhtarian 2004; Lanzendorf 2002). None of those studies have, however, developed and tested a model on how values impact actual transport mode choice. Since results for direct value-behavior relationships are often disappointing (e.g., Kassarjian and Sheffet 1991; Kristiansen and Hotte 1996), researchers have proposed that values impact specific behaviors through intervening constructs. Empirical validations of a mediated impact on behavior (intentions, preferences) through the so-called value-attitude hierarchy have been conducted by, among others, McCarthy and Shrum (1994) and Thøgersen and Grunert-Beckmann (1997). Building on this research, we propose that values affect the classic attitudes towards mode choice, such as preferences for comfort/convenience, flexibility, and safety:

P[1]: Respondents' value orientations affect their attitudes towards mode choice.

As stated above, recent research has shown that the inclusion of attitudes in models of transport mode choice leads to substantial improvements in terms of model fit as well as explanation and further provides a more satisfying representation of behavior (Choo and Moktarian 2004; Ben-Akiva, Walker, Bernardino, Gopinath, Morikawa, and Polydoropoulou 2002; Vredin Johansson, Heldt, and Johansson 2006). We build on recent research by Vredin Johansson, Heldt, and Johansson (2006) and include attitudes towards mode choice such as convenience and flexibility in our model. Following the results of the cited studies, these attitudes are proposed to influence mode choice:

P[2]: Attitudes towards mode choice affect mode choice.

What are the sources of value orientation? People's demographic characteristics such as age, gender, and income largely determine people's life circumstances in terms of their socialization, their social roles, their life stage, and their expectations. Differences in life circumstances in turn affect the salience of values (Schwartz 2003). Examples for age and gender might illustrate this point. With increasing age, the ability to cope with change is waning and security values become more relevant. Socialization leads boys and girls to adopt different social roles with different life goals and orientations. Women, being more relational and communal than men, tend to attribute more importance to benevolence values and less importance to power values (Prince-Gibson and Schwartz 1998; Schwartz and Rubel 2005). We accordingly propose:

P[3]: Socio-demographic characteristics (age, income, gender) affect values.

In their behavioral framework for choice models with latent variables, Ben-Akiva, Walker, Bernardino, Gopinath, Morikawa, and Polydoropoulou (2002) proposed that socio-demographic characteristics of an individual affect his/her attitudes (e.g., the relevance of flexibility of a transport mode depends on having children or not). Vredin Johansson, Heldt, and Johansson (2006) tested that proposition and demonstrated that demographic variables impacted attitudes of flexibility and comfort. We accept these results and propose:

P[4]: Socio-demographic characteristics affect attitudes towards mode choice.

Most empirical models on travel mode choice use modal attributes such as travel time and cost, as well as individual socio-demographic characteristics such as age and education as explanatory variables for mode choice behavior (e.g., Vredin Johansson, Heldt, and Johansson 2006). We expect similar effects in our study and propose:

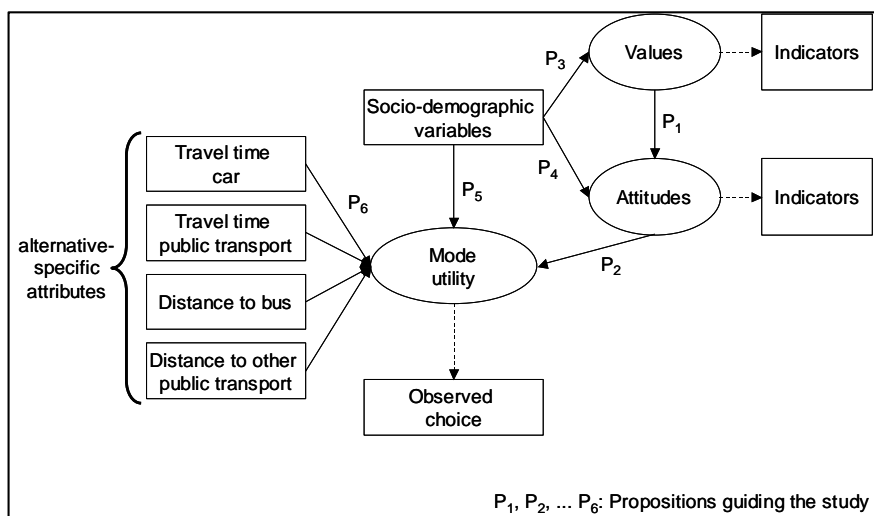P[5]: Socio-demographic characteristics (e.g., age, income, gender) affect mode choice.

P[6]: Traffic mode attributes (e.g., travel time) affect mode choice.

In summary, our application specifically investigates the influence of psychological factors (individ-

---

[11] See, e.g., Schwartz and Bilsky 1990; Bardi and Schwartz 2003 for a definition of the value construct and a discussion regarding its impact on behavior.

ual values, attitudes) in concert with known observed variables (access, time, age, gender) on commuter mode choice (see Figure 2). Six propositions were derived from the literature review and will now be tested in an empirical analysis of commuter mode choice.

Figure 2: Structure of the integrated choice and hierarchical latent variable model on mode choice.

nearest stations of various public transport alternatives (if available) and the time needed for daily trips to work with public transport as well as by car had to be estimated. In the third section, we asked about attitudes towards transport modes for daily trips to work. This section was modeled after a recent paper by Vredin Johansson, Heldt, and Johansson (2006) and a previously published working paper (Vredin Johansson, Heldt, and Johansson 2005), where respondents had to rate attitudinal questions relating to modal comfort, convenience, and flexibility on five-point scales ranging from not important at all to very important. In the fourth section, we asked respondents about their mode choice for daily trips to work or education. Respondents had to indicate whether they predominantly used car, public transport, or a combination of both. The survey closed with a section where we measured respondents' value orientations with the Portraits Value Questionnaire (PVQ) from Schwartz, Melech, Lehmann, Burgess, Harris, and Owens (2001). Respondents had to indicate their similarity to person descriptions (portraits); gender-matched with the respondent on six-point rating scales ranging from very unlike to very much alike (Schwartz, Melech, Lehmann, Burgess, Harris, and Owens 2001).

3.2 Data and methods

Data for our analysis of travel mode choice came from a sample of German consumers between 14 and 75 years of age. Following a survey pre-test with 20 subjects, 907 respondents were drawn from a consumer panel of a major international market research company. The survey was administered in a computer-aided telephone interview. Panelists were recruited following a demographic quota sampling approach based on age, profession as a proxy for status, gender, household size, and size of city/place of residence. The sample distribution on demographic variables did not significantly deviate from the population distribution. The overall survey response rate of the invited panellists was 45%. The questionnaire consisted of five major parts: In the first section, respondents were asked demographic questions needed for quota sampling. The second section included questions about personal mobility. Respondents were asked about the possession of a driver's license, of seasonal tickets for public transport alternatives (bus, streetcar, integrated public transport system, railroad), and about possession of cars. Furthermore, the distance to the

Table 1: Descriptive statistics stratified by travel mode choice

	Public transport $n = 50$		Car + Public transport, $n = 57$		Car only, $n = 412$	
	M	*SD*	*M*	*SD*	*M*	*SD*
Gender (females = 1)	0.52	0.51	0.40	0.50	0.50	0.50
Age (years)	39.50	13.81	40.32	14.71	40.12	13.04
Income (in Euro)	2,242	979	2,395	1,205	2,640	1,352

Notes: n = number of observations for each travel mode, M = mean, SD = standard deviation

For 43% of the respondents in our sample, daily trips to work/education were not applicable (e.g., they were housewives/househusbands, unemployed or retired) or alternative travel modes did not exist (e.g., they did not possess a driver's license or had no car in the household). After deletion of these cases (see Vredin Johansson, Heldt, and Johansson (2006) for a similar approach), our analytic sample thus consisted of $N = 519$ respondents (for descriptive statistics see Table 1).

3.3 Results

3.3.1 Confirmatory factor analysis

We followed the two-step approach in structural equation modeling (e.g., Anderson and Gerbing 1988) and first tested the reliability and validity of the measurement models used in the study by means of confirmatory factor analysis. Item formulations for both the attitude and value constructs are reported in Appendix A1. As stated in the previous section, building on Vredin Johansson, Heldt, and Johansson (2006) we developed items for the three dimensions flexibility, convenience/comfort, and safety. Unfortunately, the measurement model for safety did not work as expected. The reason for this result may be that the three original items were a mixture of personal safety and traffic safety attitudes in mode choice. Since also Vredin Johansson, Heldt, and Johansson (2006) reported that the differences of different modes with respect to traffic safety are negligible, we decided to keep the possession item (see Appendix A1) as an admittedly suboptimal measure of attitude towards personal safety in mode choice. In prior qualitative interviews, some of the respondents mentioned that in public transport they felt threatened or uneasy due to the presence of unwanted others. In contrast, for cars, possession allowed them to be on their own or to select the persons they travel with. Reliability for this single item has been fixed to a value of .80. Although Schwartz's Portraits Questionnaire provides well-established and validated scales, results of separate confirmatory factor analyses for the focal value con-

structs prompted us to eliminate two further items: one for power and one for security (see Appendix A1).

Our final confirmatory factor model for attitudes towards mode choice and values has been estimated with the robust WLSMV estimator implemented in the Mplus software (Muthén and Muthén 1998–2007, Version 5.1). Goodness-of-fit statistics for this model indicate an acceptable overall fit to the data (χ^2=131.28, df=57, NNFI=.91, CFI=.90, RMSEA=.05, weighted RMSR=.90).[12] Convergent validity is established by statistically significant factor loadings with t-statistics ranging from 5.07 to 10.14. Completely standardized factor loadings range from .47 to .70 for the attitude and from .50 to .79 for the value measures. Except for the factor convenience/comfort, all construct reliabilities (see Table 2) are above a recommended threshold of .60 (Bagozzi and Yi 1988).

Since the squared correlation between the two attitude constructs flexibility and convenience/comfort is larger than the average variance extracted for both factors (thus indicating a possible violation of discriminant validity (Fornell and Larcker 1981); see Table 2), we estimated a modified factor model where (1) the correlation between both factors has been fixed to unity and (2) the correlations of both constructs with like factors have been constrained to be equal (van der Sluis, Dolan, and Stoel 2005). The highly significant chi-square difference ($\Delta\chi^2$=54.31, df=5, p=.000) provides support for the discriminant validity of both constructs.

Table 2: Construct reliability and validity measures

Construct	1	2	3	4	5	6
1. Flexibility	.67, .41					
2. Conv./Comf.	.45	.53, .27				
3. Safety	.13	.09	.80, .80[a]			
4. Power	.06	.07	.01	.78, .58		
5. Hedonism	.02	.13	.02	.03	.70, .44	
6. Security	.05	.31	.08	.03	.03	.62, .42

Notes: Entries on the diagonal represent (1) Bagozzi's (1980) construct reliability ρ_c and (2) Fornell and Larcker's (1981) average variance extracted ρ_{ave}.
The off-diagonal elements are squared correlations among the constructs.
All intercorrelations are significant at $p<0.05$.
[a] Single-indicator construct with fixed measurement error variance.

---

[12] Degrees of freedom for WLSMV are estimated according to a formula given in the Technical Appendices of Mplus (Muthén 1998–2007, p. 20). Using the alternative WLSM estimator leads to a chi-square statistic of 207.28 where ordinary de-

3.3.2 Classic MNL and ICLV analysis

In order to test our ICLV model of travel mode choice and to assess to what extent the latent value-attitude hierarchy provides additional explanatory power, we first estimated a classic MNL model. This model only contains directly observed variables describing the choice alternatives (e.g., travel time) and the decision makers (e.g., age). The utilities of the three mode choice alternatives (public transport only (*PTO*), car + public transport (*CPT*), and car only (*CO*)) are given by:[13]

$$(8) \quad U_{PTO} = V(x_{PTO}, \beta_{x_{PTO}}) + v_{PTO}$$
$$= \beta_{0,PTO} + \beta_1 TT_{PT,PTO} + \beta_3 DB + \beta_5 DOPT$$
$$+ \beta_7 AGE + \beta_8 GEND + \beta_9 INC + v_{PTO}$$

$$U_{CPT} = V(x_{CPT}, \beta_{x_{CPT}}) + v_{CPT}$$
$$= \beta_{0,CPT} + \beta_1 TT_{PT,CPT} + \beta_2 TT_{C,CPT} + \beta_4 DB$$
$$+ \beta_6 DOPT + \beta_{10} AGE + \beta_{11} GEND$$
$$+ \beta_{12} INC + v_{CPT}$$

$$U_{CO} = V(x_{CO}, \beta_{x_{CO}}) + v_{CO} = \beta_2 TT_{C,CO} + v_{CO}$$

with TT_{PT} = travel time public transport, TT_C = travel time car, DB = distance to next bus stop, $DOPT$ = distance to other public transport, AGE = age, $GEND$ = gender, and INC = net monthly income. For identification purposes, the alternative-specific constant (i.e., $\beta_{0,CO}$) as well as the coefficients of the socio-demographic variables are set to zero for mode car. McFadden's pseudo R^2 for this model is 7% (see Table 3). Given the fact that in contrast to many other studies our analysis does not focus on commutes in a specific area (e.g., Train 1978) or between specific cities (e.g., Vredin Johansson, Heldt, and Johansson 2006), this result can be considered as reasonable. Our sample was drawn from across Germany, thus commuters' mode choices occurred under very different circumstances (e.g., concerning the quality and safety of public transport systems). This substantial heterogeneity is likely to considerably reduce the explanatory power of our model. Results for the traditional MNL model are in line with published research on travel mode

choice (e.g., Vredin Johansson, Heldt, and Johansson 2006; Keijer and Rietveld 2000; Loutzenheiser 1997; O'Sullivan and Morrall 1996; McFadden 1974; Train 1980; Kitamura 1989). Except for distance to the next bus station, all parameter estimates for attributes characterizing the choice options are significantly different from zero at $p<.05$ and also show the expected negative signs (see Table 3).

Next, we present the empirical findings of our proposed ICLV model (see Figure 1). The model consists of an MNL part, where following proposition P$_2$ attitudes towards mode choice have been included as additional explanatory latent variables, and a latent variable model that captures the effects of values on attitudes as well as the effects of socio-demographic variables on both types of latent variables. Expanding the conventional MNL part in (8) by individual-specific attitudes towards mode choice leads to the following utility functions:

$$(9) \quad U_{PTO} = V(x_{PTO}, \beta_{x_{PTO}}) + \beta_{13} FLEX + \beta_{14} CC$$
$$+ \beta_{15} SAFE + v_{PTO},$$

$$U_{CPT} = V(x_{CPT}, \beta_{x_{CPT}}) + \beta_{16} FLEX + \beta_{17} CC$$
$$+ \beta_{18} SAFE + v_{CPT},$$

$$U_{CO} = V(x_{CO}, \beta_{x_{CO}}) + v_{CO,}$$

with *FLEX* = flexibility, *CC* = convenience/comfort, and *SAFE* = safety. Model identification is established by fixing the latent variable coefficients for mode car to zero.

Both the discrete choice and the hierarchical latent variable model have been estimated simultaneously using Mplus (see Appendix A2 for the corresponding input file). Next, we first briefly discuss the results for the MNL part (see Table 3) and subsequently elaborate on the results of the latent variable model (see Tables 4 and 5).

Comparing the traditional and the ICLV model shows that the latter indeed provides greater explanatory power regarding mode choice. McFadden's pseudo R^2 has improved by more than 60%.[14]

---

grees of freedom are 90. Yu (2002) suggests that a weighted RMSR equal to or below .90 indicates a good model fit.

[13] For simplicity, the person index has been dropped.

[14] In order to calculate McFadden's pseudo R^2, a null model with constants only for the discrete choice part including the measurement and structural models for the latent variables has been estimated. The difference in the log-likelihood values of this null model and the ICLV model was then compared to the log-likelihood value of the null model without the latent variable part.

Table 3: Robust ML parameter estimates for the traditional and ICLV model – Discrete choice part

Explanatory variable/ parameter	Traditional MNL model		ICLV MNL model	
	Estimate	*t*-statistic	Estimate	*t*-statistic
Flexibility$_1$	—	—	−0.44***	−2.81
Flexibility$_2$	—	—	−0.39**	−2.09
Conv./Comf.$_1$	—	—	0.61**	2.06
Conv./Comf.$_2$	—	—	0.54*	1.85
Safety$_1$	—	—	−0.57***	−2.98
Safety$_2$	—	—	−0.19	−1.02
Travel time PT$_{1,2}$	−1.16***	−2.64	−1.24***	−2.64
Travel time car$_{2,3}$	−1.62***	−2.64	−1.51**	−2.31
Distance bus$_1$	−0.09	−0.72	−0.10	−0.74
Distance bus$_2$	−0.03	−0.71	−0.03	−0.58
Distance other PT$_1$	−0.12**	−2.46	−0.11**	−2.34
Distance other PT$_2$	−0.13**	−2.13	−0.13**	−2.11
Age$_1$	−0.00	−0.13	−0.00	−0.24
Age$_2$	0.00	0.06	−0.00	−0.34
Gender$_1$	0.06	0.19	0.12	0.36
Gender$_2$	−0.41	−1.35	−0.40	−1.26
Income$_1$	−0.11**	−1.99	−0.03	−0.44
Income$_2$	−0.08	−1.53	−0.02	−0.33
Mode constant$_1$	−0.72	−1.00	−0.89	−1.10
Mode constant$_2$	−0.62	−1.00	−0.64	−0.96
LL	−11,766		−11,752	
McFadden's R^2	7%		11.5%	
AIC	23,741		23,724	
BIC^{adj}	23,853		23,843	

Notes: Variable subscripts denote travel modes: 1 = public transport only, 2 = car + public transport (PT), 3 = car only. Travel time by public transport differs for modes 1 and 2; travel time by car differs for modes 2 and 3.

Coding of the gender variable: 0=male, 1=female.

A constants-only model was used to determine McFadden's pseudo R^2.

*** $p<0.01$, ** $p<0.05$, * $p<0.10$

Likewise, the information criteria as well as a significant chi-square difference ($\Delta\chi^2$=28.61, *df*=6, *p*=.00007) support the assumption that including attitudes toward mode choice in the MNL part of the model leads to a better explanation. In addition, all attitudes significantly impact mode choice thus corroborating our proposition P$_2$. Note also that consistent with our model, the inclusion of values in the MNL part neither increased model fit nor was any effect on mode choice significant.[15] Concerning the effect of attitudes on mode choice, we find that the desire for flexibility significantly increases the propensity to avoid any means of public transport and to exclusively use the car for daily work trips. In turn, importance of a convenient and comfortable commute increases the probability of choosing public transport. If a commuter finds it important to own the transport mean – our proxy variable for personal safety – this decreases the probability of using public transport either alone or in combination with the car (the latter effect, however, is not significant). Our results concerning flexibility and convenience/comfort attitudes are in line with those of Vredin Johansson, Heldt, and Johansson (2006). Further results concerning the effects of the observed variables describing the choice alternatives and the decision makers are very close to the results of the traditional MNL model and, as stated before, largely consistent with the published literature (see Table 3). Thus, our propositions P$_5$ and P$_6$ were also supported in our extended model. A remarkable exception is the vanishing effect of income on avoiding public transport in the extended model. The impact of income is absorbed by the attitudinal variables, especially the desire for flexibility. In turn, our results show that income has a strong effect on flexibility. Thus, income determines transport mode through its strong positive effect on the desire for flexibility.

The results of our latent variable model clearly confirm that personal values indeed impact attitudes towards mode choice (see Table 4) and thereby provide strong empirical support for our proposition P$_1$.

---

[15] This has been established in further analyses.

Table 4: Robust ML parameter estimates for the effects of personal values and socio-demographic variables on attitudes toward mode choice – Latent variable part of the ICLV model

Explanatory variable	Dependent variable	Estimate	*t*-statistic
Power	Flexibility (R^2=.23)	0.25***	2.82
Hedonism		0.15**	1.97
Security		0.36***	2.90
Age		–0.16*	–1.78
Gender		0.05	0.82
Income		0.23***	3.57
Power	Conv./Comf. (R^2=.44)	0.21**	2.10
Hedonism		0.28**	2.46
Security		0.62***	3.59
Age		–0.08	–0.82
Gender		0.04	0.64
Income		0.01	0.16
Power	Safety (R^2=.12)	0.10	1.62
Hedonism		0.16**	2.36
Security		0.25***	2.87
Age		0.10	1.42
Gender		0.10*	1.86
Income		0.05	1.03

Notes: Standardized parameter estimates; *** p<0.01, ** p<0.05, * p<0.10

Hedonism has its strongest positive impact on convenience/comfort but to a lesser extent also drives our measure for personal safety. The central motivational goal of hedonism is pleasure and self-gratification (Schwartz and Bilsky 1990). Respondents for whom hedonism is a salient motivation put a higher emphasis on convenience/comfort in mode choice – a result that has high face validity. The explanation of the effect on personal safety is somewhat less straightforward. Respondents who put a high relevance on owning the transport mode might also associate other, more pleasure-related aspects and activities with it (e.g., enjoy driving the vehicle they own, being undisturbed by unwanted others, etc.). The main motivational concern expressed through security is safety, stability, and harmony of the self, of society, and of relationships. Security orientation positively impacts all three attitudes towards transport mode choice at p<.05. This result makes sense since all three attitudes safety, convenience/comfort, and flexibility prevent the individual from making unexpected, potentially undesirable experiences in transport mode choice. Respondents for whom power is a particularly salient value put a higher relevance on both flexibility and convenience/comfort. Again, this result has face validity since power values express a desire for social status, prestige as well as control or dominance over people and resources (Schwartz and Bilsky 1990). Thus, the salience of the power value should be positively related to flexibility, since flexibility increases control over resources (time, cost). Interestingly, the effect on security is not significant. On the other hand, safety concerns are less relevant for those with a strong inclination to control and dominate. Except for the one-indicator construct safety, the explained variance in attitudes is substantial in an absolute sense, with values of 23% and 44%. To summarize, our results concerning the value-attitude relationships possess face validity and clearly support proposition P1.

In our model, the relation between attitudes and socio-demographic variables is rather weak (see Table 4). We only find one significant effect from income on flexibility. Furthermore, even though only significant at p<.10, gender impacts the relevance of safety. For women, personal safety of a transport mode is of higher importance than for men (see, for example, Vredin Johansson, Heldt, and Johansson (2006) for a similar result). The fact that socio-demographic variables are sources of value priorities and thereby impact attitudes towards mode choice via values might explain our somewhat weaker results for the direct effects put forth in proposition P4.

As proposed in proposition P3, socio-demographic variables possess some interesting effects on personal values (see Table 5). Power is clearly more salient for men than for women, a result that is consistent with research in psychology (Schwartz and Rubel 2005). Both age and income are negatively related to hedonism as a guiding personal value. Again, this result is consistent with published research (Schwartz and Rubel 2005). Furthermore, the strong positive effect of age on security supports the contention that age is positively related to conservation values. This hypothesis derives from the

fact that older people are more likely to be embedded in social networks, to have developed habitual behaviors that they adhere to, and are less likely to seek exciting changes and challenges (Schwartz 2003). Our results concerning sources of value priorities support proposition P_3 and are consistent with previously confirmed or hypothesized effects in the psychological literature.

Table 5: Robust ML parameter estimates for the effects of socio-demographic variables on personal values – Latent variable part of the ICLV model

Explanatory variable	Dependent variable	Estimate	*t*-statistic
Age	Power (R^2=.04)	0.08	1.44
Gender		−0.15***	−2.96
Income		0.06	1.03
Age	Hedonism (R^2=.10)	−0.25***	−4.45
Gender		−0.09*	−1.72
Income		−0.14***	−2.89
Age	Security (R^2=.22)	0.47***	5.82
Gender		0.04	0.66
Income		−0.05	−0.92

Notes: Standardized parameter estimates; *** p<0.01, ** p<0.05, * p<0.10

4. Discussion and implications

The goal of this research project was to make a methodical and a theoretical contribution. Concerning our methodical contribution, we have extended existing research in two major ways: To the best of our knowledge, this is the first application of an ICLV model to multinomial choice and hierarchical latent variable relationships both in transportation research and marketing. Previous studies in marketing have only analyzed binary choice situations where respondents were asked to indicate their behavioral intentions after certain experimental manipulations (Ashok, Dillon, and Yuan 2002; Dellaert and Stremersch 2006). Our model, in contrast, examines real-world behavioral data on decisions to either use a car, some kind of public transport, or a combination of both for daily trips to work or education. The extended choice model clearly outperforms a traditional MNL model on several accounts

and provides valuable insights into the motivational processes that determine mode choice.

In addition, our model extends previous accounts of ICLV models by providing a general framework that allows for any interrelationship between latent variables to be specified. Furthermore, latent variables can also be predicted by observed explanatory variables. Since latent variables such as attitudes or values cannot be easily forecasted, the relation of these constructs to observed socio-demographic variables may aid in forecasting such variables (Vredin Johansson, Heldt, and Johansson 2006). Both selected latent and observed variables can enter the multinomial logit model as direct determinants of choice.

A further contribution of this paper lies in the fact that it suggests and demonstrates a convenient alternative for estimating ICLV models with an SEM software package. From a substantial point of view, ICLV models can be considered as one of the most interesting advances in discrete choice modeling in the last decade. Still, applications in marketing and related fields are scarce. The major reason for this lack of popularity is most likely the fact that researchers consider the full information estimation of ICLV models too complicated. We have shown and validated in a separate Monte Carlo study that ICLV models can be estimated with the Mplus program (Muthén and Muthén 2007).

With respect to the theoretical contribution, we set out to develop a more comprehensive model of choice that also maps the impact of such abstract motivational constructs as values on consumers' real choices. The general structure of our ICLV model consists of a discrete choice part where latent variables – in our example attitudes – enter a multinomial logit model in addition to the observed attributes of the different choice options as well as attributes of the decision maker. The latent variable part of the model allows for relations between the latent variables and observed variables, as well as causal relationships between the latent variables. Additionally, socio-demographics are included as explanatory variables both in the discrete choice and the latent variable model in order to control for observed heterogeneity and to aid in forecasting the latent variables. In our empirical example, a hierarchical model where personal values determine attitudes which in turn impact on actual behavior was proposed and validated. Note here that the notion of hierarchical goal structures and their impact on

consumer behavior is a current topic in the marketing field (e.g., Paulssen and Bagozzi 2006; Yang, Allenby, and Fennell 2002). However, to date, research in marketing has not investigated the impact of the value-attitude hierarchy on actual choice but only on intentions (McCarthy and Shrum 1994; Thøgersen and Grunert-Beckmann 1997). Only the recent methodical advances make the inclusion of these latent variables in a choice model possible. We hope that the current paper will contribute to more applications of the integrated choice and latent variable model for substantive research questions in the marketing field and beyond.

5. Limitations and further research

As shown in our paper, Mplus offers a convenient way to simultaneously estimate both the discrete choice and the latent variable part of the ICLV model. Due to its origin in structural equation modeling, the program's particular strengths reside in the latent variable part where it offers great flexibility in terms of models and estimators. Compared to that, modeling discrete choice in Mplus currently seems somewhat restricted. As laid out in Section 2, our basic specification of the random utility function relies on the MNL model, and thus is subject to the same limitations as the MNL approach in general (e.g., Train 2003). First, the logit model cannot properly handle variations in attribute preferences that are either purely random or associated with unobserved factors. Though including latent variables like attitudes and values can mitigate this issue to a certain extent, other important variables may still remain undetected. Second, the IIA ("independence of irrelevant alternatives") assumption associated with the logit specification does not allow for disproportionate changes in the choice probabilities of alternatives (e.g., in the case of a new option).

In order to overcome the aforementioned restrictions, various alternative models (e.g., probit model, nested logit model, mixed logit model) have been proposed in the literature (for an overview see, for example, Train (2003)). At present, Mplus offers the opportunity to specify a mixed logit model where the mixing distribution for the random utility parameters is discrete. Hence, it is assumed that unobserved variations in attribute preferences can be adequately captured by estimating specific sets of parameters for a discrete and finite number of latent

classes or market segments. [16] Finite mixture ICLV models as proposed by Ashok, Dillon, and Yuan (2002) as well as Walker and Ben-Akiva (2002) can therefore be estimated with the Mplus program. In principle, we could have used the latent class option in Mplus to extend our analysis of the empirical ICLV model on travel mode choice. According to the results of our Monte Carlo study, however, the number of observations seems to be just large enough for a single-sample analysis. Additionally, mode choices exhibit a strong asymmetry in favor of cars which further amplifies the sample size issue. Although we tried to control for heterogeneity by including socio-demographic variables, there is still some risk that unobserved heterogeneity has biased our results. Thus, we see the finite mixture ICLV model as an opportunity for future research on travel mode choice.

Acknowledgements

---

[16] This has been established in further analyses.

Appendix:

A.1 Measures used in the study

Attitudes (new scales based on Vredin Johansson, Heldt and Johansson 2006)
Flexibility (3 measures)
- That a means of transport is available right away is... (N=519, M=4.3, SD=0.77, ρ_{ii}=0.49)
- That a means of transport can be used spontaneously and without planning is... (N=519, M=4.3, SD=0.82, ρ_{ii}=0.42)
- That a means of transport reaches its final destination without a detour or change is... (N=519, M=4.2, SD=0.86, ρ_{ii}=0.32)
Convenience/Comfort (3 measures)
- That a means of transport is exceedingly convenient and comfortable is... (N=519, M=3.6, SD=0.97, ρ_{ii}=0.38)
- That using a means of transport is stress-free and relaxed is... (N=519, M=3.9, SD=0.86, ρ_{ii}=0.23)
- That you do not have to worry about anything while using the means of transport is... (N=519, M=3.5, SD=1.07, ρ_{ii}=0.22)
Safety (1 measure)
- That you own the means of transport is...(N=519, M=3.6, SD=1.22, ρ_{ii}=0.80[fixed])
- That a means of transport is as secure as possible is... (eliminated)
- That a means of transport can be used alone or with friends is... (eliminated)
Personal values (based on Schwartz, Melech, Lehmann, Burgess, Harris and Owens 2001)
Power (2 measures)
- She/he always wants to be the one who makes the decisions. She/He likes to be the leader. (N=516, M=3.6, SD=1.33, ρ_{ii}=0.63)
- It is important to her/him to be in charge and tell others what to do. She/He wants people to do what she/he says. (N=516, M=3.3, SD=1.32, ρ_{ii}=0.53)
- It is important to her/him to be rich. She/He wants to have a lot of money and expensive things. (eliminated)
Hedonism (3 measures)
- She/He seeks every chance she/he can to have fun. It is important to her/him to do things that give her/him pleasure. (N=515, M=4.6, SD=1.09, ρ_{ii}=0.52)
- She/He really wants to enjoy life. Having a good time is very important to her/him. (N=514, M=4.6, SD=1.13, ρ_{ii} = 0.46)
- Enjoying life's pleasures is important to her/him. She/He likes to 'spoil' herself/himself. (N=516, M=4.5, SD=1.25, ρ_{ii}=0.37)
Security (4 measures)
- It is very important to her/him that her/his country be safe. She/He thinks the state must be on watch against threats. (N=515, M=4.5, SD=1.24, ρ_{ii}=0.39)
- It is important to her/him to live in secure surroundings. She/He avoids anything that might endanger her/his safety. (N=515, M=4.3, SD=1.22, ρ_{ii}=0.28)
- Having a stable government is important to her/him. She/He is concerned that the social order be protected. (N=515, M=4.5, SD=1.15, ρ_{ii}=0.27)
- It is important to her/him that things be organized and clean. She/He really does not like things to be a mess. (N=515, M=4.5, SD=1.35, ρ_{ii}=0.25)
- She/He tries hard to avoid getting sick. Staying healthy is very important to her/him. (eliminated)
Notes: Six-point scale, very dissimilar to very similar
N=number of observations, M=sample mean, SD=standard deviation, ρ_{ii}=indicator reliability

A.2 Mplus input file for the ICLV model on travel mode choice

```
TITLE:        ICLV model on travel mode choice
DATA: FILE = mode_choice_data.dat;
VARIABLE: NAMES =
        Choice                              !revealed mode choice
                                            !1 = public transport only (PTO)
                                            !2 = car + public transport (CPT)
                                            !3 = car only (CO)
        tt_c_co tt_c_cpt                    !travel time by car for CO; travel time by car for CPT
        tt_pt_pto tt_pt_cpt                 !travel time by public transport for PTO;
                                            !travel time by public transport for CPT
        dist_bus dist_opt                   !distance to next bus stop;
                                            !distance to other public transport
        age gender income                   !socio-demographic variables
        y1 y2 y3 y4 y5 y6 y7                !indicators for attitudes
        y8 y9 y10 y11 y12 y13 y14;          !indicators for personal values
NOMINAL = choice;
CATEGORICAL = y1 y2 y3 y4 y5 y6;
MISSING = ALL (999);
ANALYSIS:
        ESTIMATOR = MLR;                    !robust ML estimation
        INTEGRATION = GAUSSHERMITE;         !numerical integration
MODEL:
        !Latent variable models for attitudes
        flex BY y1@1 y2 y3;
        cc BY y4@1 y5 y6;
        safe BY y7@1;
        y7@.2962;
        !Latent variable models for values
        power BY y8@1 y9;
        hedon BY y10@1 y11 y12;
        secur BY y13@1 y14 y15 y16;
        !Regressions of attitudes on values and socio-demographics
        flex ON power hedon secur age gender income;
        cc ON power hedon secur age gender income;
        safe ON power hedon secur age gender income;
        !Regressions of values on socio-demographics
        power ON age gender income;
        hedon ON age gender income;
        secur ON age gender income;
        !Multinomial logit model
        choice#1 ON tt_pt_pto (1);          !equality constraint
        choice#1 ON tt_c_co (p2);           !linear constraint
        choice#1 ON dist_bus dist_opt;
        choice#1 ON age gender income;
        choice#1 ON flex cc safe;

        choice#2 ON tt_pt_cpt (1);          !equality constraint
        choice#2 ON tt_c_co (p2);           !linear constraint
        choice#2 ON tt_c_cpt (p1);          !linear constraint
        choice#2 ON dist_bus dist_opt;
        choice#2 ON age gender income;
        choice#2 ON flex cc safe;
        [choice#1 choice#2];                !alternative-specific constants
MODEL CONSTRAINT:
p2 = -p1;                                   !linear constraint
OUTPUT: STANDARDIZED RESIDUAL;
```

References

Anderson, James C. and David W. Gerbing (1988): Structural equation modeling in practice: A review and recommended two-step approach, *Psychological Bulletin*, 103 (3): 411–423.

Andrews, Rick L., Andrew Ainslie, and Imran S. Currim (2002): An empirical comparison of logit choice models with discrete versus continuous representations of heterogeneity, *Journal of Marketing Research*, 39 (4): 479–487.

Ashok, Kalidas, William R. Dillon, and Sophie Yuan (2002): Extending discrete choice models to incorporate attitudinal and other latent variables, *Journal of Marketing Research*, 39 (1): 31–46.

Bagozzi, Richard P. (1980): *Causal models in marketing*, John Wiley & Sons : New York.

Bagozzi Richard P. and Youjae Yi (1988): On the evaluation of structural equation models, *Journal of the Academy of Marketing Science*, 16 (1): 74–94.

Bamberg, Sebastian (1996): Allgemeine oder spezifische Einstellungen bei der Erklärung umweltschonenden Verhaltens? [General or specific attitudes as predictors of environmentally friendly behavior?], *Zeitschrift für Sozialpsychologie*, 27 (1): 47–60. (in German)

Bamberg, Sebastian and Peter Schmidt (1994): Auto oder Fahrrad? Empirischer Test einer Handlungstheorie zur Erklärung der Verkehrsmittelwahl [Car or bicycle? An empirical test of a behavioral theory on travel mode choice], *Kölner Zeitschrift für Soziologie und Sozialforschung*, 46 (1): 80–102. (in German)

Bamberg, Sebastian and Steffen Kühnel (1998): Überzeugungssysteme in einem zweistufigen Modell rationaler Handlungen [Systems of conviction in a two-stage model of rational behavior], *Zeitschrift für Soziologie*, 27 (4): 256–270. (in German)

Bardi, Anat and Shalom H. Schwartz (2003): Values and behaviour: Strength and structure of relations, *Personality and Social Psychology Bulletin*, 29 (10): 1207–1220.

Ben-Akiva, Moshe and Steven R. Lerman (1985): *Discrete choice analysis: Theory and application to travel demand*, Cambridge: MIT Press.

Ben-Akiva, Moshe, Mark Bradley, Taka Morikawa, Julian Benjamin, Thomas Novak, Harmen Oppewal, and Vithala Rao (1994): Combining revealed and stated preferences data, *Marketing Letters*, 5 (4): 335–349.

Ben-Akiva, Moshe, Daniel McFadden, Tommy Gärling, Dinesh Gopinath, Joan Walker, Denis Bolduc, Axel Börsch-Supan, Philippe Delquié, Oleg Larichev, Taka Morikawa, Amalia Polydoropoulou, and Vithala Rao (1999): Extended framework for modeling choice behavior, *Marketing Letters*, 10 (3): 187–203.

Ben-Akiva, Moshe, Daniel McFadden, Kenneth Train, Joan Walker, Chandra Bhat, Michel Bierlaire, Denis Bolduc, Axel Börsch-Supan, David Brownstone, David S. Bunch, Andrew Daly, Andre De Palma, Dinesh Gopinath, Anders Karlstrom, and Marcela A. Munizaga (2002): Hybrid choice models: Progress and challenges, *Marketing Letters*, 13 (3): 163–175.

Ben-Akiva, Moshe, Joan Walker, Adriana T. Bernardino, Dinesh A. Gopinath, Taka Morikawa, and Amalia Polydoropoulou (2002): Integration of choice and latent variable models, in: Hani S. Mahmassani (ed.): *In perpetual motion: Travel behaviour research opportunities and application challenges*, Elsevier, Amsterdam, 431–470.

Bolduc, Denis, Moshe Ben-Akiva, Joan Walker, and Alain Michaud (2005): Hybrid choice models with logit kernel: Applicability to large scale models, in: Martin Lee-Gosselin and Sean T. Doherty (eds.): *Integrated land-use and transportation models*, Elsevier, Amsterdam, 275–302.

Boomsma, Anne and Jeffrey J. Hoogland (2001): The robustness of LISREL modeling revisited, in: Robert Cudeck, Steven du Toit, and Dag Sörbom (eds.): *Structural equation models: Present and future. A Festschrift in honor of Karl Jöreskog*, Scientific Software International, Lincolnwood, IL, 139–168.

Choo, Sangho and Patricia L. Mokhtarian (2004): What type of vehicle do people drive? The role of attitude and lifestyle in influencing vehicle type choice, *Transportation Research Part A: Policy and Practice*, 38 (3): 201–222.

Collins, Christy M. and Susan M. Chambers (2005): Psychological and situational influences on commuter-transport-mode-choice, *Environment and Behavior*, 37 (5): 640–661.

Dellaert, Benedict G. C. and Stefan Stremersch (2005): Marketing mass-customized products: Striking a balance between utility and complexity, *Journal of Marketing Research*, 42 (2): 219–227.

Elrod, Terry (1991): Internal analysis of market structure: Recent developments and future prospects, *Marketing Letters*, 2 (3): 253–266.

Elrod, Terry and Michael P. Keane (1995): A factor-analytic probit model for representing the market structure in panel data, *Journal of Marketing Research*, 32 (1): 1–16.

Eymann, Angelika, Axel Börsch-Supan, and Rob Euwals (2002): Risk attitude, impatience, and asset choice, *SFB 504 Discussion Paper*, No. 02-28, University of Mannheim.

Fornell, Claes and David L. Larcker (1981): Structural equation models with unobservable variables and measurement error, *Journal of Marketing Research*, 18 (1): 39–50.

Green, William H. (2008): *Econometric analysis*, 6th ed., Prentice Hall: Englewood Cliffs.

Jöreskog, Karl G. and Arthur S. Goldberger (1975): Estimation of a model with multiple indicators and multiple causes of a single latent variable, *Journal of the American Statistical Association*, 70 (351): 631–639.

Judd, Kenneth L. (1998): *Numerical methods in economics*, Cambridge: MIT Press.

Kassarjian, Harold H. and Mary J. Sheffet (1991): Personality and consumer behavior: An update, in: Harold H. Kassarjian and Thomas S. Robertson (eds.): *Handbook of Consumer Behavior*, Prentice-Hall, New Jersey, 281–301.

Keijer, M. J. N. and Piet Rietveld (2000): How do people get to the railway station? The Dutch experience, *Transportation Planning and Technology*, 23 (3): 215–235.

Kitamura, Ryuichi (1989): A causal analysis of car ownership and transit use, *Transportation*, 16 (2): 155–73.

Kristiansen, Connie M. and Alan M. Hotte (1996): Morality and the self: Implications for when and how of value-attitude behavior relations, in: Clive Seligman, James M. Olson, and Mark P. Zanna (eds.): *The Ontario Symposium: Vol. 8 – The psychology of values*, Lawrence Erlbaum, Hillsdale, NJ, 77–106.

Lanzendorf, Martin (2002): Mobility styles and behaviour – Application of a lifestyle approach to leisure travel, *Transportation Research Record*, 1807: 163–173.

Loutzenheiser, David R. (1997): Pedestrian access to transit: Model of walk trips and their design and urban form determinants around bay area rapid transit stations, *Transportation Research Record*, 1604: 40–49.

Luo, Lan, P. K. Kannan, and Brian T. Ratchford (2008), Incorporating subjective characteristics in product design and evaluations, *Journal of Marketing Research*, 45 (2): 182–194.

McCarthy, John A. and L. J. Shrum (1994): The recycling of solid wastes: Personal value orientation, and attitudes about recycling as antecedents of recycling behavior, *Journal of Business Research*, 30 (1): 53–62.

McFadden, Daniel L. (1974): The measurement of urban travel demand, *Journal of Public Economics*, 3 (4): 303–328.

McFadden, Daniel L. (1986): The choice theory approach to marketing research, *Marketing Science*, 5 (4): 275–297.

Morikawa, Taka, Moshe Ben-Akiva, and Daniel L. McFadden (2002): Discrete choice models incorporating revealed preferences and psychometric data, in: Philip H. Franses and Alan L. Montgomery (eds.): *Econometric Models in Marketing*, Elsevier, Amsterdam, 29–55.

Muthén, Bengt O. (1983): Latent variable structural equation modeling with categorical data, *Journal of Econometrics*, 22 (1-2): 43–65.

Muthén, Bengt O. (1984): A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators, *Psychometrica*, 49 (1): 115–132.

Muthén, Linda K. and Bengt O. Muthén (1998–2007): *Mplus User's Guide. 5th ed.*, Muthén & Muthén: Los Angeles.

Nordlund, Annika M. and Jörgen Garwill (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'Sullivan, Sean and John F. Morrall (1996): Walking distances to and from light-rail transit stations, *Transportation Research Record*, 1538: 19–26.

Paulssen, Marcel and Richard P. Bagozzi (2006): Goal hierarchies as antecedents of market structure, *Psychology & Marketing*, 23 (8): 689–709.

Prince-Gibson, Eetta and Shalom H. Schwartz (1998): Value priorities and gender, *Social Psychology Quarterly*, 61 (1): 49–67.

Schmid, Klaus-Peter (2003, May 15): Begrenzt einsatzfähig [Limited fitness for use]. *Die Zeit*, 21. (in German)

Schwartz, Shalom H. (2003): A proposal for measuring value orientations across nations, in: European Social Survey: *The questionnaire development package of the European Social Survey*, Chapter 7, 259–319.

Schwartz, Shalom H. and Wolfgang Bilsky (1990): Toward a theory of the universal content and structure of values: Extensions and cross-cultural replications, *Journal of Personality and Social Psychology*, 58 (5): 878–891.

Schwartz, Shalom H., Gila Melech, Arielle Lehmann, Steven Burgess, Mari Harris, and Vicki Owens (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.

Schwartz, Shalom H. and Tammy Rubel (2005): Sex differences in value priorities: Cross-cultural and multi-method studies, *Journal of Personality and Social Psychology*, 89 (6): 1010–1028.

Skrondal, Anders and Sophia Rabe-Hesketh (2004): *Generalized latent variable modelling*, Chapman & Hall: Boca Raton.

Thøgersen, John (2006): Understanding repetitive travel mode choices in a stable context: A panel study approach, *Transportation Research Part A: Policy & Practice*, 40 (8): 621–638.

Thøgersen, John and Suzanne C. Grunert-Beckmann (1997): Values and attitude formation towards emerging attitude objects: From recycling to general waste minimizing behavior, *Advances in Consumer Research*, 24 (1): 82–189.

Train, Kenneth E. (1978): A validation test of a disaggregate mode choice model: *Transportation Research Part A: Policy and Practice*, 12 (2): 167–174.

Train, Kenneth E. (1980): A structured logit model of auto ownership and mode choice, *Review of Economic Studies*, 28 (2): 357–70.

Train, Kenneth E. (2003): *Discrete choice with simulations*, Cambridge University Press: Cambridge, UK.

van der Sluis, Sophie, Conor V. Dolan, and Reinoud D. Stoel (2005): A note on testing perfect correlations in SEM, *Structural Equation Modeling*, 12 (4): 551–577.

Vermunt, Jeroen K. and Jay Magidson (2005): *Latent Gold Choice 4.0 User's Manual*, Statistical Innovations: Belmont, MA.

Vredin Johansson, Maria, Tobias Heldt, and Per Johansson (2005): Latent variables in a travel mode choice model: Attitudinal and behavioural indicator variables, Working Paper 2005:5, Uppsala University.

Vredin Johansson, Maria, Tobias Heldt, and Per Johansson (2006): The effects of attitudes and personality traits on mode choice, *Transportation Research Part A: Policy and Practice*, 40 (6): 507–525.

Walker, Joan L. and Moshe Ben-Akiva (2002): Generalized random utility model, *Mathematical Social Sciences*, 43 (3): 303–343.

Walker, Joan L. and Jieping Li (2007): Latent lifestyle preferences and household location decisions, *Journal of Geographical Systems*, 9,(1): 77–101.

Wedel, Michel, Wagner Kamakura, Neeraj Arora, Albert Bemmaor, Jeongwen Chiang, Terry Elrod, Rich Johnson, Peter Lenk, Scott Neslin, and Carsten S. Poulsen (1999): Discrete and continuous representations of unobserved heterogeneity in choice modeling, *Marketing Letters*, 10 (3): 219–232.

Yang, Sha, Greg M. Allenby, and Geraldine Fennell (2002): Modeling variation in brand preference: The roles of objective environment and motivating conditions, *Marketing Science*, 21 (1): 14–31.

Yu, Ching-Yun (2002): *Evaluating cutoff criteria of model fit indices for latent variable models with binary and continuous outcomes*, Dissertation, University of California: Los Angeles.

Biographies

Dirk Temme is Associate Professor of Marketing at Humboldt University Berlin since 2007. He holds a master degree in Business Administration from the University of Bielefeld and a doctoral degree from Humboldt University Berlin. In his doctoral thesis, he studied the effect of unobservable variables ("invisible

assets") in structural models of business performance. In 2006, he received his Habilitation (lecturing qualification) with a postdoctoral thesis on the application of advanced structural equation models in quantitative marketing research. His research interests include: marketing strategy-performance link, invisible assets, brand management, and measurement theory. His research has appeared in various national and international journals.

Marcel Paulssen is Professor of Marketing at the HEC Geneva since 2008. His academic credentials include a Bachelor of Business Administration, a master degree in Management Science from the Stevens Institute of Technology, a master degree in Business Administration from the Technical University of Berlin, and a Ph.D. in Marketing from the Technical University of Berlin. He has been Assistant Professor at Humboldt University Berlin and Visiting Professor at the University of Michigan Business School. Prior to his academic career, Dr. Paulssen gained several years of management experience in the area of strategic marketing at Matshushita and Daimler. His research interests comprise the role of goals in decision making, the impact of attachment orientations on customer behavior in relationships to companies as well as antecedents and consequences of customer-company identification. His research on these topics has appeared in various national and international journals.

Till Dannewald is currently working as Research Consultant at Infas TTR in Frankfurt. He earned a master degree in economics from the University of Regensburg in 2001. Between 2002 and 2007 he was Research Assistant at the Institute of Marketing at Humboldt University Berlin. He finished his doctoral thesis on "The modeling of strategic interactions on international markets" in 2007. His research interests are the application of choice models in various fields of business administration and the analysis of competitive behavior in fast-moving consumer goods markets.