LINKING CAR OWNERS´ AFFECTIVE AND INSTRUMENTAL MOTIVES TO THEIR CAR USE – AN APPLICATION OF AN INTEGRATED CHOICE AND LATENT VARIABLE MODEL

Sascha von Behren
Institute for Transport Studies, Karlsruhe Institute of Technology (KIT)
Kaiserstrasse 12, 76131 Karlsruhe, Germany
Email: sascha.vonbehren@kit.edu

Lisa Bönisch
Institute for Transport Studies, Karlsruhe Institute of Technology (KIT)
Kaiserstrasse 12, 76131 Karlsruhe, Germany
Email: lisa.boenisch@kit.edu

Ulrich Niklas
Brand Portfolio Strategy, Market Research, Competition
BMW AG
Petuelring 124-130, 80809 München, Germany
Email: ulrich.niklas@bmw.de

Bastian Chlond
Institute for Transport Studies, Karlsruhe Institute of Technology (KIT)
Kaiserstrasse 12, 76131 Karlsruhe, Germany
Email: bastian.chlond@kit.edu

Paper prepared for presentation and publication at the
99th Annual Meeting of the Transportation Research Board
6,493 words
4 tables (4 * 250 words = 1,000 words)
Total: 7,493 words (including abstract, text, references, tables)
Submission date: August 1, 2019

ABSTRACT

Car mobility in cities differs from that in suburban or rural areas, as there usually exist alternative modes such as public transit. Nevertheless, people own and use their cars. In addition to hard factors such as sociodemographics, motives play a particularly important role in urban areas. In the context of this study, we would like to examine the influence of soft factors on car use. This leads to our research question: what influence do the affective and instrumental motives have on the use frequency of cars? The data set used in this study was collected in Berlin and San Francisco. The respondents have provided insights on their everyday travel (e.g., use frequency of cars) and motives for car use. To investigate influences of motives, we applied motives into an integrated choice and latent variable model (ICLV) with an ordered probit kernel. The results indicate a high added value by integrating these soft factors. We can explain more of the overall heterogeneity and can give further insights to the decision-making process. Results show a high influence of the affective motive, whereby the instrumental motive does not matter. Considering the direct and indirect effects, we see that the influence of gender comes almost exclusively by the affective motive. Results suggest people are more likely to use cars for affective motives despite the city's adversities. For these people it is difficult to achieve a shift to alternative means of transport. The only way to intervene here is through regulatory intervention.

Keywords: Car motives, travel skeleton, integrated choice and latent variable model

INTRODUCTION

The private car has always played a decisive role in transportation research and planning. In cities in particular, the present focus is on reducing emissions. Apart from that the urban space consumption of cars has to be considered. The conventional car needs both disproportionate urban space for moving and additionally a parking lot for most of the time. At the same time, however, the car is a flexible and attractive means of transport. For research purposes, the motives for use are of particular interest, as there are often attractive mode alternatives available in cities, like cycling or using public transit.

In contrast to car sharing, the private car provides permanent availability. It gives their owners a comfortable and flexible opportunity to reach their target location without using diverse modes of transport, also in urban areas. For this reason, the literature contains a number of studies on motives for car use. As an obvious reason functionality and convenience are to be mentioned. People value car usage especially because of its flexibility, independence, availability, speed, reliability, safety, carrying capacity and comfort (*1*). Different functional reasons for car usage were identified in a qualitative study in Great Britain: transport of heavy goods, driving services (e.g., for family members), being short of time, convenience and pursuing trips (*2*). These subjective motives address the apparent instrumental benefits of using cars for individual travel. Next to this, people in general have a positive attitude towards using the car. Consequently, even psychological aspects assume a significant role in car travel behavior. In this context, cars can be attributed to symbolic and affective functions that affect car usage besides the functional aspects mentioned (*3*). The affective function represents an experiential value of vehicles for individuals. On the one hand, driving a car can be linked to positive emotions like driving pleasure. On the other hand, it can result in negative emotions because of stress while driving, e.g. due to traffic congestion (*4*). The symbolic value of a car addresses its social impact in terms of a status symbol and the influence on social identity. These two additional aspects can have an considerable impact on individual's emotions for car usage (*4*). In this research field the multiple investigations of Steg are to be emphasized. At this point, not all of her papers will be cited, but her elementary work (*3*) has to be highlighted. As a result she discovered attitudinal aspects to be a better predictor for car usage behavior than instrumental functions. She finally provides an international accepted and well-tested set of psychological items that refer to psychological motives for car use (instrumental, affective, symbolic).

This set was used as basis for related work of other authors who conducted research on car use behavior. Lois and Lopez-Saez (*5*) examined the relation between these three motives using structural equation model (SEM). They found that the affective element had direct influence on car usage, whereas symbolic and instrumental elements were only effecting the affective motivation and thus only had indirect influence on car usage. Also applying SEM in their study Šefara et al. (*6*) showed that personal motives for using cars have a notable impact on the preferences for car types and brands. For men, the influence of symbolic-affective motives for car use have greater impact on car preference than for women. Son and Yun (*7*) used the items of Steg (*3*) and Ellaway et al. (*8*) to verify the existence of car-dependent commuters with the application of support vector machine. Bergstadt et al. (*9*) investigated the influence of socio-demographic variables on daily car use, using OLS multiple linear regression analyses and the items of Steg. They found that affective-symbolic motives can partially describe the relationship between the number of weekly car trips and gender. In this context, they could evidence that psychological motives partially explain differences in car use regarding different sociodemographic groups. Nevertheless, Steg et al. (*1*) could confirm the impact of instrumental motives on the attractiveness of car use with different methods. Besides findings based on the items of Steg (*3*), Van and Fujii (*10*) investigated and compared the contribution of psychological factors in explaining mode choice in six Asian countries with another set of items similar to existent literature. Their findings confirm the significance of attitudinal factors on the behavioral intention to commute by car. Besides the familiar factors in literature, instrumental and symbolic affective components, they further complemented social orderliness as relevant factor to explain mode choice. Considering international cultural differences Belgiawan et al. (*11*) detected the symbolic-affective factor to be significant among attitudinal constructs in various countries when explaining the intention to purchase a car. Apart from that, an international comparison regarding car use frequency under consideration of attitudinal aspects is still lacking.

In our study, we want to close this gap by answering the following research questions: What influence do the affective and instrumental motives have on the use frequency of private cars? For this purpose, we analyzed data from a survey conducted in the urban area of San Francisco (U.S.) and Berlin (Germany) that examined the use of the private car and their motives. As the literature review has shown, the application of hybrid choice models (HCM) in this research field is still in its infancy. The HCM provides a framework to integrate unobservable, latent attitudes into the decision-making process. An example for a provisional implementation in the case of mode choice is provided by Habib et al. (*12*). They were able to prove that the incorporation of the full information affecting decision-making into a HCM allows to identify the real importance of relevant variables in the choice process. In our study, we investigate the influence of motives for car use applying an integrated choice and latent variable model (ICLV) with an ordered probit kernel in the choice component. The paper is structured as follows: First, we show a brief literature review on motives of car use. Second, we describe the used data and the methodology of our analysis. Third, we present results of the ICLV and interpret resulting factors that influence car usage in San Francisco and Berlin. Finally, we draw a conclusion, discuss the limits of our approach and refer to further work.

SURVEY DESIGN AND DATA COLLECTION

In the following, we first introduce the applied survey design with the different components to capture travel behavior and psychological factors. We will then describe how the survey was conducted.

Survey design of a travel skeleton

The data used in this research is based on the concept of a so-called travel skeleton. The essential of this concept is to capture typical elements of everyday travel. The travel skeleton provides a reasonable compromise between the level of detail needed and the effort required to capture travel behavior. The idea behind the development of this approach was to create a cost-effective survey alternative to longitudinal trip diaries. Besides everyday travel in different areas of life (commuting, chauffeuring, leisure, shopping, errands), this approach is also able to capture long-distance traveling, tech savviness and psychological factors of respondents. The concept of a travel skeleton was originally designed and tested for a study in Hamburg and Berlin (Germany) (*13*). For a more detailed description of the approach, we refer to von Behren et al. (*14*).

In our survey, we asked respondents, in detail about their typical private car use. In addition to travel behavior, we asked the respondents questions about their attitudes towards certain modes. In our study, we focus on 8 psychological items (indicators) that record the attitudes towards public transit and the car. With these items, we want to analyze how the attitudes towards the two main modes for commuting affect ODM use on the way to work. The items used in this study are shown in Figure 1. They are taken from two standardized psychological item sets by Hunecke et al. (*15*) and Steg (*3*), and are rated on a Likert scale from 1-5 ("does not apply" to "apply").

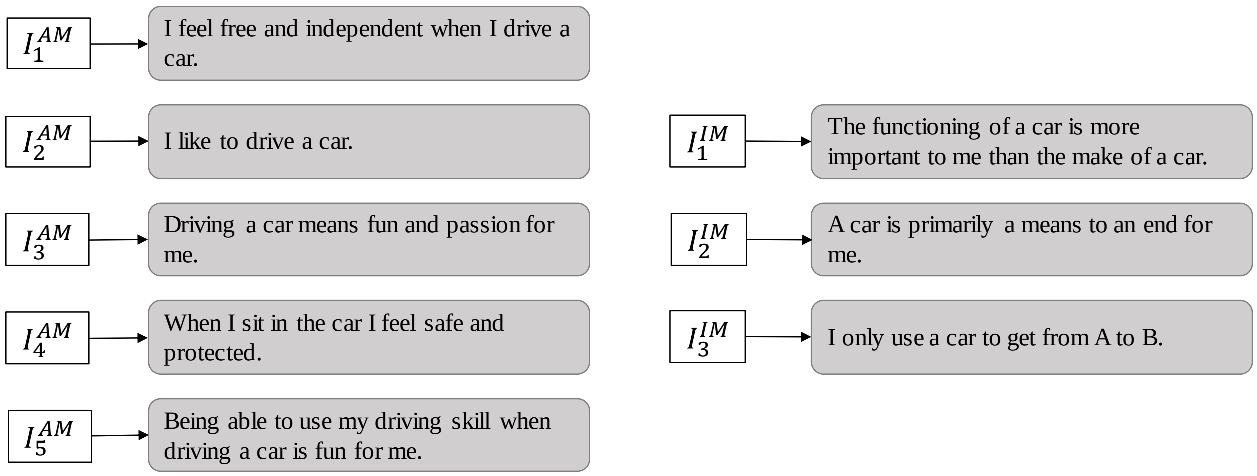

Figure 1. Psychological items (indicators) with their related questions

Data collection in Berlin and San Francisco

The research presented in this paper is based on data collected through two similar surveys, conducted in the urban area of Berlin (Germany) and San Francisco (U.S.) between October-2016 and January-2017. The two surveyed cities are well-developed and provide good public transport systems. In addition, each city has specific innovative transport services such as ODM (e.g., Uber, Didi or DriveNow). Berlin and San Francisco are regarded as “hybrid cities”, which exhibit dense public-transit-oriented urban cores, surrounded by low-density car-oriented suburban areas (*16*).

To generate comparable datasets from each city we used a standardized survey approach based on a computer assisted personal interview (CAPI). The face-to-face interviews had a duration of approximately 40 minutes. We also included questions concerning psychological factors, including attitudes towards means of transport, and social and individual norms. The complete sample size was 1,200 individuals with 600 respondents from each city. We conducted quota sampling regarding age, gender, household size, and net income to develop a representative survey group for each captured city. The survey was carried out by a professional market research firm (Spiegel Institut), using a slightly different approach in each city taking into consideration specific cultural particularities. In all cities, an access-panel with telephone screening and on-street recruitment was used. The aim of the surveys was to capture behavior and psychological factors for individuals above the age of 17 and, as far as possible, for the whole household. For our analysis of the motives of car use, we used only the data of people with a car in the household. As a result, we used 836 people from San Francisco and Berlin.

DESCRIPTIVE RESULTS

In our analysis, we examined people with a car in their household. However, the frequency of car use is not limited to a household's private car. The use of car-sharing such as DriveNow, Car2go or Zipcar was also considered. The requirement, however, was that the person must drive the car himself. In our study we did not find a person who never drives a car during the year. For our model, respondent's choices were divided into three categories, according to the frequency of car use:

- *occasional use*, corresponding to a usage of less than once a week (48.56%, 406 respondents),
- *regular use*, corresponding to a usage of at least once up to a maximum of several times per week (24.52%, 205 respondents),
- *daily use*, corresponding to day-to-day use of the car (26.91%, 225 respondents).

Table 1 shows the characteristics of the persons surveyed and the variables subsequently used for the model. Three classes were created to describe access to the car in the household. The variable "car disposal – sometimes" describes whether the person can only access the car by arrangement in the household. "Car disposal – always" describes a permanent access to a car in the household. For 46% of the respondents, no arrangement are necessary (see Table 1). In addition, we look on a household level, if the household possess a premium car from brands such as BMW, Mercedes, Audi or Tesla. 20% of the households in our sample have at least one premium car. We also show in Table 1 the amount of daytrips or vacation trips with overnight stays. High amount of daytrips means more than 12 trips per year and a high amount of vacation trips describe more than 8 trips per year. We also analyze, if person use in general their car for commuting (73.09%). In our sample 23.56% of the respondents use their car for long-distance leisure trips over 100 km. 29.19% of our respondents have a monomodal behavior, i.e., they use the same mode of transport in everyday travel. For the evaluation of car use it is important to take the spatial structure in consideration. Therefore, we look at two spatial information on zip code level of the residential location of the respondents. On the one hand, we use open street map (OSM) data to calculate built-up living space and data about the population from the data provider Nexiga to calculate the population density per built-up living. High population density includes all zip codes with more than 10,000 inhabitants per built-up space per square meter. On the other hand, we calculate through OSM data the public transit options and quality in the zip codes. When calculating the accessibility, we considered that rail-bound public transit is more valuable than buses. A high value indicates people have plenty of public transit options besides the car in this zip code. This applies to 25.84% of the respondents.

Table 1. Descriptive analysis of socio-demographic, mobility and spatial characteristics

Person characteristics				
	Yes	**No**		
Age over 30 years	78.23%	21.77%		
Fulltime job	57.30%	42.70%		
Male	52.27%	47.73%		
Own bicycle	48.92%	51.08%		
	Always	**Sometimes**	**Never**	
Share of car disposal	46.66%	23.20%	30.14%	
Household characteristics				
	Yes	**No**		
Premium car in household	20.45%	79.55%		
Household from Berlin	47.73%	52.27%		
	<2,500$	**2,500$-5,000$**	**5,001$-8,000$**	**>8,000$**
Share of income classes	16.75%	32.54%	31.34%	24.76%
	Type 1	**Type 2**	**Type 3**	**Type 4**
Share of household type	31.22%	18.30%	25.72%	24.76%
Mobility and car use characteristics				
	Yes	**No**		
High amount of daytrips	4.07%	95.93%		
High amount of vacation trips	8.73%	91.27%		
Commuting by car	73.09%	26.91%		
Long-distance trips by car	23.56%	76.44%		
Monomodal behavior	29.19%	70.81%		
Spatial characteristics				
	Yes	**No**		
High population density	39.71%	60.29%		
High public transit accesibility	25.84%	74.16%		

N=836

The answers to the attitudinal questions are summarized in Figure 2. In the case of the indicators describing the emotional aspects of driving (I_{1-5}^{AM}), almost 50% of the participants agree on all the questions. Over 60% agree with the statement: I like to drive a car (I_2^{AM}), where the question is very general. But for more than 30% driving a car does not mean fun or passion (I_3^{AM}). As can be seen, roughly more than 50% of the respondents rate the instrumental indicators (I_{1-3}^{IM}) positively. Especially with regard to I_1^{IM}, for nearly 70% is the function more important than the make.

As a preliminary analysis to our ICLV and to identify latent variables, we conducted an explorative factor analysis based on the presented attitudinal indicators. Table 2 shows the result of the principal factor analysis (PFA). Based on the scree plot (elbow criterion) and the Kaiser´s criterion, two factors can be extracted: one describing the affective motive (factor 1) and one describing the instrumental motive (factor 2) of car use. Factor 1 describes whether people like to drive a car and whether they feel free through the use. The instrumental motive describes whether people only use the car as a tool to satisfy their mobility needs.

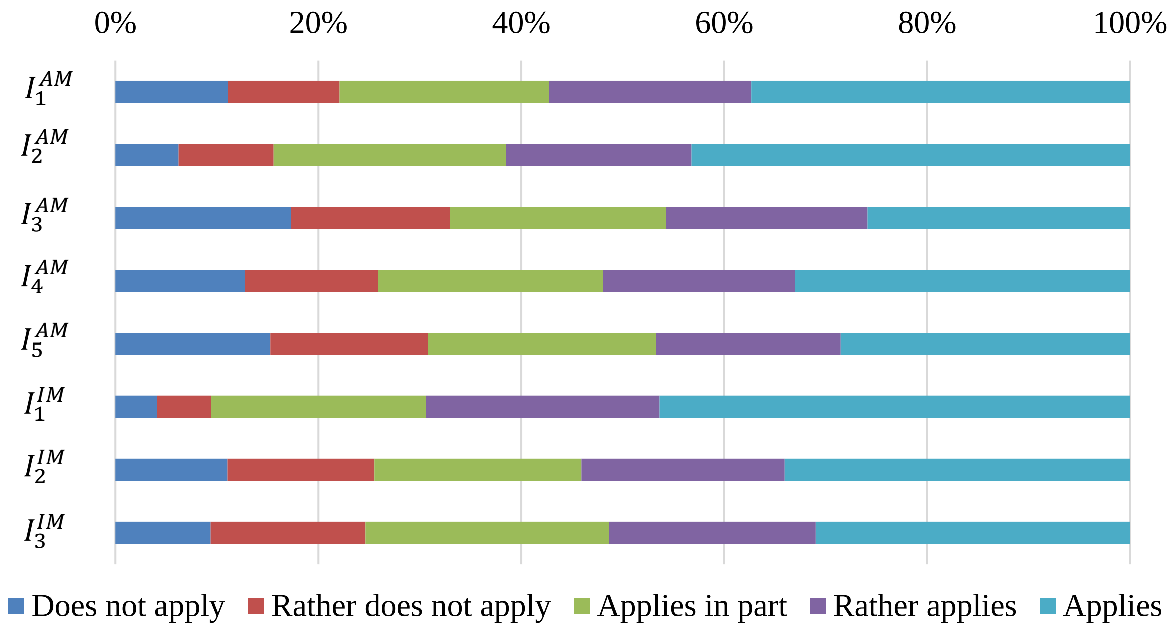

Figure 2. Motives of car use. 1-5 Likert scaled questions as described in Figure 1

Table 2. Principal Factor Analysis (PFA) - Varimax Rotated Factor Pattern

	Factors	
	Affective motive	*Instrumental motive*
Cronbach's Alpha	α = *0.91*	α = *0.71*
Indicators in PFA		
I_2^{AM}	0.8372	
I_1^{AM}	0.8278	
I_4^{AM}	0.8132	
I_5^{AM}	0.7868	
I_3^{AM}	0.7813	
I_3^{IM}		0.7502
I_2^{IM}		0.7108
I_1^{IM}		0.5550
		Printed is the maximum loading of each item
Criteria of extraction and quality for PFA		
Criteria of extraction	*# Factors*	
Kaiser`s criterion	2	
Scree Test	2	
Criteria of quality		*Value*
Kaiser´s Measure of Sampling Adequacy (MSA)		0.845>0.8 (meritorious)

N=836

MODELING METHODOLGY

In this section, we give a brief overview of the applied models. In our study, people's main choice is to reveal their regularity of using their car on an ordered scale (occasional/regular/daily use). We model this core choice with an ordered probit model, which will be topic of the first part of this section. Afterwards, we will present how to incorporate latent variables into the ordered probit model by means of an integrated choice and latent variable model (ICLV). By choosing an ordered probit model over an ordered logit model, we will be able to construct an identifiable reduced form model, which will be described in the final part of this section.

Ordered probit model

In an ordered regression model, individuals face a choice among J ordered alternatives, labelled 1 to J. For individual n ($n \in \{1, \dots, N\}$), we describe this choice by a random variable Y_n with sample space $\Omega = \{1, \dots, J\}$. To derive a distribution for these random variables, we use the setting of a random utility model. For each individual n, a latent variable, called utility, is defined that describes its propensity towards the considered quantity – in our case the benefit of frequent car usage – on a continuous scale. This utility is composed of a systematic part as well as a random part:

$$U_n = V(X_n, \beta) + \epsilon_n, \tag{1}$$

where U_n is the utility of individual n, $V(X_n, \beta)$ is its systematic part and ϵ_n are random disturbances. The systematic part considers the effects of a set of explanatory variables for each individual. It is determined by a vector of observable, explanatory variables X_n of the individual n and a vector of parameters β that are assumed to be identical for every individual. The random part however contains all other effects that cannot be explained by those variables. It is assumed to be independent and identically distributed over individuals. We consider a linear specification of the systematic part of utility, such that the parameters β describe the importance of the different explanatory variables respectively:

$$V(X_n, \beta) = \beta X_n. \tag{2}$$

Choices are now made based on the value of the utility. High values of utility imply decisions for higher levels in the sample space Ω. Using threshold values $\tau^{(0)}, \dots, \tau^{(J)}$, J ascending intervals are defined and the value of the variable Y_n is obtained by allocating utility to those intervals:

$$\begin{aligned} Y_n = 1 \quad &\Leftrightarrow \quad \tau^{(0)} < U_n \leq \tau^{(1)}, \\ Y_n = 2 \quad &\Leftrightarrow \quad \tau^{(1)} < U_n \leq \tau^{(2)}, \\ &\vdots \\ Y_n = J \quad &\Leftrightarrow \quad \tau^{(J-1)} < U_n \leq \tau^{(J)}, \end{aligned} \tag{3}$$

where $\tau^{(0)} = -\infty, \tau^{(J)} = \infty$. We call τ the vector of thresholds $(\tau^{(0)}, \dots, \tau^{(J)})$. Selecting a distribution assumption for the random components ϵ_n determines the specific ordered regression model. We assume a normal distribution, $\epsilon_n \sim N(0, \sigma^2)$, with variance σ^2, yielding an ordered probit model. The probability of individual n to choose alternative j, given the observable variables X_n and the parameters β and τ, is then given by

$$\begin{aligned} P(Y_n = j | X_n, \beta, \tau) &= P\left(\tau^{(j-1)} < U_n \leq \tau^{(j)} | X_n, \beta, \tau\right) \\ &= P\left(\tau^{(j-1)} - V_n < \epsilon_n \leq \tau^{(j)} - V_n | X_n, \beta, \tau\right) \\ &= P\left(\frac{\tau^{(j-1)} - V_n}{\sigma} < \frac{\epsilon_n}{\sigma} \leq \frac{\tau^{(j)} - V_n}{\sigma} | X_n, \beta, \tau\right) \\ &= \Phi\left(\frac{\tau^{(j)} - V_n}{\sigma}\right) - \Phi\left(\frac{\tau^{(j-1)} - V_n}{\sigma}\right), \end{aligned} \tag{4}$$

where Φ is the cumulative distribution function of the standard normal distribution.

Integrated choice and latent variable model (ICLV)

Thus far, we have only considered observable variables to describe people's choices. The ICLV gives a framework to incorporate unobservable, latent variables into our model of the decision-making process. We will only give the setting of the ICLV that is later being used in our study. As those variables cannot be observed directly, responses to a set of indicators on an ordered scale are asked in the survey. Those indicators are used to gain information on the latent variables describing people's motive of car use. They are included into the model by adding a latent variable model to the choice component described in the previous section. The resulting ICLV is illustrated in Figure

3. The model consists of a group of structural equations as well as a group of measurement equations, which are described in detail in the following.

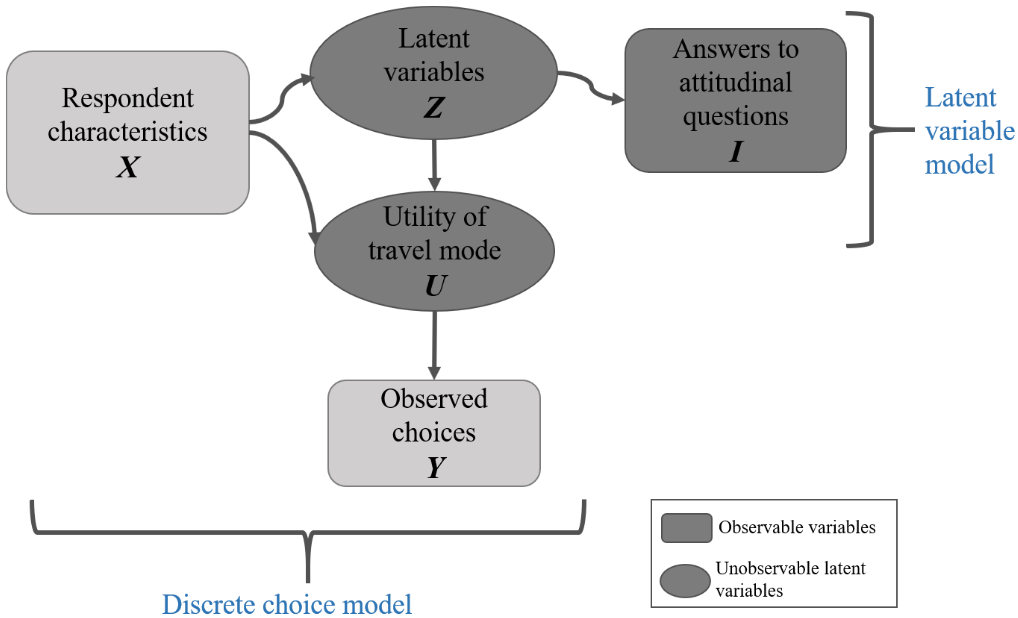

Figure 3. Integrated choice and latent variable model (ICLV) adapted to Ben-Akiva et al. (*17*)

By Z_n, we describe the vector of the L different latent variables of individual n. Each latent variable Z_{nl} is described by a linear combination of the individual's observable variables X_n, where the weights are given by the vectors α_l, as well as a random component ξ_{nl}, which is assumed to be normally distributed with mean zero and variance σ_l^2. Thus we obtain the following structural equation for the l-th latent variable:

$$Z_{nl} = \alpha_l X_n + \xi_{nl}, \qquad \xi_{nl} \sim N(0, \sigma_l^2). \tag{5}$$

In our model, we will assume independency of the latent variables given the observable variables and the parameters α_l. In short: ξ_{nl} is assumed to be independent for all $l \in \{1, \dots, L\}$ and $n \in \{1, \dots, N\}$. As scale and spread of the latent variables are arbitrary, no constant is needed in the specification of the latent variables and the variance can be set to one. The indicators, whose responses are given on an ordered scale represented by the values 1 to S, are manifestations of the latent variables. We denote the set of indicators of individual n by the vector I_n, containing the K single indicators I_{nk}, $k \in \{1, \dots, K\}$. We model each indicator by a continuous representation $\tilde{I}_{nk}$, which is composed of a linear combination of the latent variables with weights given by the vector ζ_k and an error component ψ_{nk} with mean zero:

$$\tilde{I}_{nk} = \zeta_k Z_n + \psi_{nk}. \tag{6}$$

The distribution of the Indicators I_{nk} is then defined by allocating the variables $\tilde{I}_{nk}$ to intervals given by the thresholds $\rho_k^{(0)}, \dots, \rho_k^{(S)}$:

$$\begin{aligned} I_{nk} = 1 \quad &\Leftrightarrow \quad \rho_k^{(0)} < \tilde{I}_{nk} \leq \rho_k^{(1)}, \\ I_{nk} = 2 \quad &\Leftrightarrow \quad \rho_k^{(1)} < \tilde{I}_{nk} \leq \rho_k^{(2)}, \\ &\vdots \\ I_{nk} = S \quad &\Leftrightarrow \quad \rho_k^{(S-1)} < \tilde{I}_{nk} \leq \rho_k^{(S)}, \end{aligned} \tag{7}$$

where $\rho_k^{(0)} = -\infty$ and $\rho_k^{(S)} = \infty$ for all $k \in \{1, \dots, K\}$. Assuming the error components ψ_{nk} are i.i.d. and follow a Gumbel distribution (also called extreme value type I distribution) for all n and k, we obtain an ordered logit model. Again, as scale and spread of the continuous representations $\tilde{I}_{nk}$ are arbitrary, no constant is necessary and we can set the scale parameter of the Gumbel distribution to one. The probability for a certain response s to the k-th indicator, given the latent variables and the parameters ζ_k, is thus given by

$$\begin{aligned} P(I_{nk} = s | Z_n, \zeta_k) &= P\left(\rho_k^{(s-1)} < \tilde{I}_{nk} \leq \rho_k^{(s)} | Z_n, \zeta_k\right) \\ &= P\left(\rho_k^{(s-1)} - \zeta_k Z_n < \psi_{nk} \leq \rho_k^{(s)} - \zeta_k Z_n | Z_n, \zeta_k\right) \\ &= \frac{\exp\left(\rho_k^{(s)} - \zeta_k Z_n\right)}{1 + \exp\left(\rho_k^{(s)} - \zeta_k Z_n\right)} - \frac{\exp\left(\rho_k^{(s-1)} - \zeta_k Z_n\right)}{1 + \exp\left(\rho_k^{(s-1)} - \zeta_k Z_n\right)}, \end{aligned} \tag{8}$$

where we used that the cumulative distribution function of a Gumbel distributed random variable is

$$F(x) = \frac{\exp(x)}{1 + \exp(\mathrm{x})}. \tag{9}$$

Finally, in the choice component of the model, we include the latent variables equivalently to the observable variables. The definition of utility is thus extended to

$$U_n = \beta X_n + \gamma Z_n + \epsilon_n \tag{10}$$

where γ is the vector of weights of the latent variables on utility. The choice probabilities, given all variables X_n and Z_n, is then calculated as before:

$$\begin{aligned} &P(Y_n = j | X_n, Z_n, \beta, \gamma, \tau) \\ &\qquad = \Phi\left(\frac{\tau^{(j)} - (\beta X_n + \gamma Z_n)}{\sigma}\right) - \Phi\left(\frac{\tau^{(j-1)} - (\beta X_n + \gamma Z_n)}{\sigma}\right). \end{aligned} \tag{11}$$

The random components ϵ_n include all unobserved heterogeneity among individuals, which appears random to us but does have its reasons in reality. By including the latent variables as in Eq. (10) and considering that they are random variables themselves, we hope to reduce the role of the purely random terms ϵ_n and to explain instead some of the heterogeneity through the latent variables. The full set of parameters is denoted by θ. It contains the parameters β, γ, τ and σ of the choice component, the parameters ζ and ρ of the measurement equations, as well as the parameters α of the structural equations of the latent variable component. To compute the probability of observing the full choice of individual n, including the alternative j as well as the responses $s_1, \dots, s_K$ to the indicators, given a set of parameters θ, we can use the law of total probability to condition on the latent variables and use the conditional independency of the choice variable Y_n and the indicators I_n:

$$\begin{aligned} P(Y_n = j, I_n = (s_1, \dots, s_k)|X_n, \theta) \\ &= \int_{\mathbb{R}^L} P(Y_n = j, I_n = (s_1, \dots, s_k)|Z_n = t, X_n, \theta) f_{Z_n}(t|X_n, \theta)\, dt \\ &= \int_{\mathbb{R}^L} P(Y_n = j|Z_n = t, X_n, \theta) \\ &\quad \cdot \prod_{k=1}^{K} P(I_{nk} = s_k|Z_n = t, \theta)\, f_{Z_n}(t|X_n, \theta)\, dt, \end{aligned} \quad \textbf{(12)}$$

where f_{Z_n} is the joint probability density function of the L latent variables Z_n of individual n. This probability can – theoretically – be evaluated using the distribution of the latent variables given by (5) and Eqs. (11) and (8). Further, we can calculate the probability of observing the choice for an alternative alone in the same manner:

$$P(Y_n = j|X_n, \theta) = \int_{\mathbb{R}^L} P(Y_n = j|Z_n = t, X_n, \theta) f_{Z_n}(t|X_n, \theta)\, dt. \quad \textbf{(13)}$$

Estimation

We want to gain estimations for the parameters using maximum likelihood estimation. Calling $y^{(n)}$ the observed choice of individual n and $s^{(n)}$ the vector of given responses to the Indicators of individual n, we can construct the likelihood-function, stating the probability of all observed choices in dependency of a set of parameters θ:

$$L(\theta|X_n) = \prod_{n=1}^{N} P\left(Y_n = y^{(n)}, I_n = s^{(n)} \middle| X_n, \theta\right). \quad \textbf{(14)}$$

Analogously, we can construct a likelihood function for the observed decisions for the alternatives alone in almost the same manner:

$$L_{Choice}(\theta|X_n) = \prod_{n=1}^{N} P(Y_n = y^{(n)}|X_n, \theta). \quad \textbf{(15)}$$

Again, in theory, those functions can be evaluated and thus maximised using Eq. (12) respectively Eq. (13), but it requires the evaluation of L dimensional integrals. Therefore an approximation method for the integrals in (12) is necessary. We approximate these integrals using draws according to the distribution of the latent variables. In doing so, constructing draws via Halton sequences has been shown to be far superior to purely random draws (see *18*). Calling $t^{(1)}, \dots, t^{(R)}$ the R L-dimensional draws, we approximate the probability of choosing an alternative j and giving the responses s to the indicators by

$$P(Y_n = j, I_n = (s_1, \dots, s_k) | X_n, \theta) \approx \frac{1}{R} \sum_{r=1}^{R} \left[P\left(Y_n = j | Z_n = t^{(r)}, X_n, \theta\right) \prod_{k=1}^{K} P\left(I_{nk} = s_k \middle| Z_n = t^{(r)}, X_n, \theta\right) \right]. \quad \textbf{(16)}$$

The probabilities of the choice component alone can be approximated identically. Using approximations of that kind in maximizing the likelihood function is known as maximum simulated likelihood estimation.

Reduced form model

Following the procedure of Vij and Walker (*19*), we can construct a reduced form model of the HCM. Replacing the latent variables Z_n in the definition (10) of utility by their structural equations (5), we obtain

$$\begin{aligned} U_n &= \beta X_n + \gamma Z_n + \epsilon_n = \beta X_n + \gamma_1 Z_{n1} + \dots + \gamma_L Z_{nL} + \epsilon_n \\ &= \beta X_n + \gamma_1(\alpha_1 X_n + \xi_{n1}) + \dots + \gamma_L(\alpha_L X_n + \xi_{nL}) + \epsilon_n \\ &= (\beta + \gamma_1\alpha_1 + \dots + \gamma_L\alpha_L)X_n + \gamma_1\xi_{n1} + \dots + \gamma_L\xi_{nL} + \epsilon_n \\ &= \beta_{RFM} X_n + \epsilon_{RFM} \end{aligned} \quad \textbf{(17)}$$

with $\beta_{RFM} = \beta + \gamma_1\alpha_1 + \dots + \gamma_L\alpha_L$ and $\epsilon_{RFM} = \gamma_1\xi_{n1} + \dots + \gamma_L\xi_{nL} + \epsilon_n$. The error components $\xi_{nl}, l \in \{1, \dots, L\}$ of the latent variables are all independently standard normally distributed and thus $\gamma_l\xi_{nl}$ are independently normally distributed with variance γ_l^2. Since we used an ordered probit model in the choice component, i.e. the error component ϵ_n of utility is also independently normally distributed with variance σ^2, ϵ_{RFM} is a sum of $L+1$ independent normally distributed random variables, and thus again a normally distributed error component with variance $\sigma_{RFM}^2 \coloneqq \gamma_1^2 + \dots + \gamma_L^2 + \sigma^2$. Altogether, the HCM is again an ordered probit model with the special weights β_{RFM} for the observable variables and with variance σ_{RFM}^2. Consequently, we can only expect the choice component of the HCM to be as good as an ordered probit model in terms of the overall likelihood value. In general, the overall likelihood of a simple ordered probit model will be at least as good as the likelihood of the choice component of the HCM using the same observable variables. That is because in the former, parameters are only estimated to describe the choices for the alternatives as good as possible, whereas in the latter, parameters are estimated to describe choices for the alternatives simultaneously to the responses to the indicators. Ideally, the HCM can describe choices in almost the same manner but gives us a more detailed explanation of how decisions are made by decreasing the importance of the purely random component ϵ_{RFM}, and instead interpreting some of the overall variance through the latent variables.

MODEL SPECIFICATION

The integrated choice and latent variable model (ICLV) was defined according to the equations in the methodology section. Figure 4 shows the variables used in our model. After extensive specification testing in the "base" ordered probit model, we have used the explanatory variables shown in Figure 4. The following variables were used as dummy variables (as described in the section descriptive results): age, fulltime job, male, premium car in household, from Berlin, daytrips, vacation trips, commuting by car, long-distance trips by car, monomodal, population density and public transit accessibility. The dummy variable low income describes people with a monthly household net income under 2,500$. We also include the car disposal (split into three categories, taking never as the base) and the household type (split into three categories, taking household type 1 and 2 as one- or two-person households as the base). Consequently, 18 β-parameters and 2 γ-parameters for the influence of the variables on utility, the parameter σ^2 describing the variance of the error component in utility, 12 α-parameters for the structural equations of the latent variables affective and instrumental motive respectively, as well as the parameters ζ and ρ for the indicators had to be estimated.

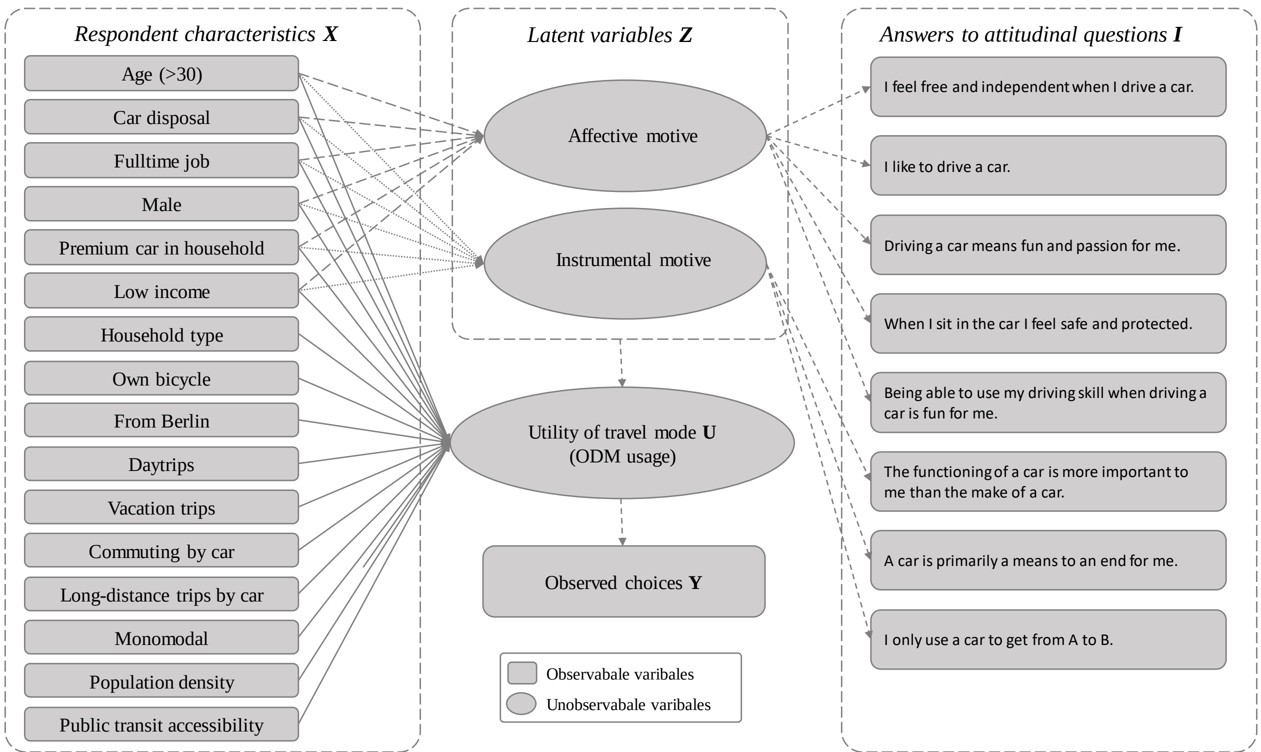

Figure 4. Specification of the integrated choice and latent variable (ICLV) model

To set the scale of utility, one of the parameters included in the specification of utility has to be fixed. Usually, this is done by setting the variance σ^2 to one, but we chose to fix one of the β parameters instead, such that we can compare the variance of the purely random component in the ICLV and in the reduced form model in the end. Ultimately, the parameter for the variable describing if people use their cars for commuting has been set to one. For estimation we used the adapted CMC choice modelling code for R (*20*) with 2,500 Halton draws for each latent variable.

RESULTS

The integrated choice and latent variable model (ICLV) helps to illustrate how various sociodemographics as well as motives influence car use frequency. First, we look at the direct impact of the respondents' characteristics in the utility function (see Table 3 (A)). Not surprising is the high influence of age. People over the age of 30 are more likely to use their car more often. The highest influence comes from car availability. People who have permanent access to a car in their household are more likely to use it. We cannot see this high influence if the person can only use their car in arrangement with the household. The results show no influence over the place of residence. Car owners from Berlin do not use their cars more often than owners from San Francisco. In San Francisco and Berlin, people with lower incomes use their cars less frequently than people from higher income groups. An interesting result can be observed in monomodal behavior. People who use rather few different means of transport are also less likely to use their cars more often. This is surprising, as frequent car users in particular are monomodal persons. In the case of spatial structures, there is only an influence from the public transit offer. People with alternatives e.g. by public transit use their cars less frequently. Table 3 (B) presents the findings for the structural model for the two latent variables (LV). For younger car owners a higher value for affective motives and a lower for instrumental motives can be observed. The higher influence of the instrumental motives of older people is in contrast to findings of Van and Fujii (*10*). This implies if young adults have a car, then they have stronger affective motives than older ones. Person from household with premium cars have higher affective motives and lower instrumental motives than person without a premium car. This finding can be confirmed by the work of Sefara et al. (*6*) who determined the impact of personal motives on preferences regarding car type. Next, we look at the impact of the LVs in the utility function (see Table 3 (A)). We see a significant positive influence of the affective motive. Respondents with a high value regarding the affective motive have a greater utility for car use. In contrast, the second LV (instrumental motive) has no influence. This result is in line with existing research about motives (*5*).

The ICLV provides the advantage of allowing us to examine the split into direct effects and effects through the LVs. This is given by the vectors β, $\gamma_{emotional}\alpha_{emotional}$ and $\gamma_{instrumental}\alpha_{instrumental}$ respectively. Therefore, we followed the procedure of Vij and Walker (*19*) and constructed a reduced form model of the ICLV. The direct influences and the influences through the LVs on the choice are summarized in Table 4. The effects through the LVs also show no effect of the instrumental motive. By looking at the affective motive, it becomes clear that the influence of gender only arises through this LV. So people with a higher affective motive are more likely to use cars. In addition, the influences of car availability and the presence of a premium vehicle in the household are reinforced by the affective motive. The effect of the premium vehicle and the availability through arrangements in the household is even doubled. As we have already seen in Table 3 (A), a low income has a negative effect on car use. However, this effect is reduced if the person has an affective motive. This suggests that people dispense with other things to be able to use their car, even if it is more expensive than public transit or cycling.

Table 3. Main parameter estimates. (A) Parameters of the structural equation of the choice model. (B) Parameters of the structural equations of the latent variables.

Log-Likelihood	-8,778.87
Log-Likelihood of choice component	-557.13
Log-Likelihood of null model (choice component)	-918.43
McFadden Pseudo-R^2	0.39
N	836

Thresholds of the choice component

$\tau^{(1)}$	1.780
$\tau^{(2)}$	3.033

(A)		(B)		
Parameter	***Value***	***Parameter α from variable***	***on latent variable***	
			affective motive	***instrumental motive***
$\beta_{age>30}$	0.463***	Age > 30 years	-0.171**	0.232***
$\beta_{fulltime}$	0.153*	Fulltime job	0.189**	-0.004
β_{male}	0.008	Male	0.393***	-0.293***
$\beta_{bicycle}$	-0.097			
$\beta_{car-sometimes}$	0.185	Car disposal - sometimes	0.637***	0.310***
$\beta_{car-always}$	1.198***	Car disposal - always	1.118***	0.224***
$\beta_{premiumcar}$	0.250**	Premium car in household	0.592***	-0.124*
β_{berlin}	0.080			
$\beta_{lowincome}$	-0.362**	Low income	0.403***	0.700***
$\beta_{hhtype3}$	0.320***			
$\beta_{hhtype4}$	0.202*			
$\beta_{highdaytrips}$	0.587**			
$\beta_{highvacation}$	-0.104			
$\beta_{commuting\_car}$	1.000			
$\beta_{long-distance\_car}$	0.308***			
$\beta_{monomodal}$	-0.207**			
$\beta_{density}$	-0.044			
$\beta_{pt\_accessibility}$	-0.107			
$\gamma_{affective}$	0.399***			
$\gamma_{instrumental}$	0.030			
σ	0.963***			

Parameters marked with ***, **, and * are significant at the 1%, 5%, and 10% level, respectively

Table 4. Direct influences vs. influences through the LVs on the choice

Variable	*Direct effect*	*Effect via LV affective motive*	*Effect via LV instrumental motive*	*Effect via LVs combined*	*Overall effect*	*Overall effect in an ordered probit*
Age > 30 years	0.463	-0.068	0.007	-0.061	0.402	0.347
Fulltime job	0.153	0.075	0.000	0.075	0.228	0.206
Male	0.008	0.157	-0.009	0.148	0.156	0.139
Own bicycle	-0.097				-0.097	-0.146
Car disposal - sometimes	0.185	0.254	0.009	0.264	0.449	0.355
Car disposal - always	1.198	0.446	0.007	0.453	1.651	1.414
Premium car in household	0.250	0.236	-0.004	0.233	0.483	0.434
From Berlin	0.080				0.080	0.121
Low income	-0.362	0.161	0.021	0.182	-0.180	-0.157
Household type 3	0.320				0.320	0.233
Household type 4	0.202				0.202	0.228
High daytrips	0.587				0.587	0.559
High vacation trips	-0.104				-0.104	-0.135
Commuting by car	1.000				1.000	1.000
Long-distance trips by car	0.308				0.308	0.305
Monomodal behavior	-0.207				-0.207	-0.180
High population density	-0.044				-0.044	0.022
High public transit accessibility	-0.107				-0.107	-0.159

Looking at the two last columns in Table 4 we notice that the overall effects of the observable variables on utility in the HCM are, as expected, almost identical to the effects in a separately estimated “base” ordered probit model using the same observable variables. The same holds for the threshold values. Also, the overall variance in the ICLV, as discussed earlier, is almost identical to the variance in the pure ordered probit model. This gives strong confidence into the parameter estimates. Small deviations are to be expected due to simulation noise using the maximum simulated likelihood method.

An important aspect of the ICLV is to gain insight into what share of heterogeneity in the model can actually be linked to the LVs. The share of variance is thus

$$\frac{\gamma_{affective}^2 + \gamma_{instrumental}^2}{\sigma^2 + \gamma_{affective}^2 + \gamma_{instrumental}^2} = 14.72\%.$$

Hence, we were able to explain 14.72% of the overall heterogeneity in a pure ordered probit model by including LVs.

CONCLUSIONS

In urban areas, car use is often hindered by parking problems and congestion. In addition, good alternatives such as public transit and cycling are often existing. Nevertheless, many people own a car and use it regularly. With this study, we have not only considered hard factors such as age, gender and income, but also soft factors such as the motives for using a car to investigate their impacts on car use frequency.

To uncover the different effects, we used an ordered probit as well as an integrated choice and latent variable model (ICLV) with an ordered probit kernel. With the ICLV we could consider hard as well as soft factors in the model. Regarding the hard factors, the results show that people over the age of 30, who permanently have a car, are more likely to use it frequently. Above all, the fact whether a premium vehicle is in the household increases the probability of its use. An interesting fact is that there is no difference between people from Berlin and San Francisco, certainly not among car owners. With regard to spatial effects, it can be seen, that a good public transit service reduces the usage. However, a very high population density at the place of residence has no influence. The indirect effects through the latent variables (LV) demonstrate that only the affective motives have an influence and increase the probability of car use. This finding is only partially in line with previous research mentioned in the introduction section. In comparison to existing literature, the added insight of this research is the outcome that affective motives are the unique influencing factor when considering car use frequency. This also leads to people on lower incomes driving their car more frequently if they have an emotional connection to it. Results have thus demonstrated that emotional aspects play a decisive role in the frequency of car use. Besides the further insights regarding the indirect effects in the ICLV, the model offers the possibility to explain a part of the heterogeneity through the LVs. The results show that the added value of the LVs as 14.72% of the unexplained variance can be explained by the deterministic variance of the LVs. The comparison of the overall effects between the "base" ordered probit and the ICLV shows the similarity of the parameter values despite the maximum simulated likelihood method. This guarantees the necessary certainty in the interpretation of the results.

The results have clearly highlighted the added value of the ICLV when considering car use. Further research could be conducted under inclusion of autonomy into the models. Autonomy could describe whether people, from their point of view, have another possibility to use a car. This addresses the consumed perception of car drivers of the available alternatives.

ACKNOWLEDGMENTS

This paper presents results of a project funded by BMW AG.

AUTHOR CONTRIBUTIONS

The authors confirm contribution to the paper as follows: study conception and design: von Behren; literature review: Bönisch; data preparation: von Behren; data analysis: von Behren; interpretation of results: von Behren, Niklas, Chlond; draft manuscript preparation: von Behren. All authors reviewed the results and approved the final version of the manuscript.

REFERENCES

1. Steg, L., C. Vlek, and G. Slotegraaf. Instrumental- Reasoned and Symbolic-Affective Motives for Using a Motor Car. *Transportation Research Part F4*, 2001.
2. Mackett, R. L. Why Do People Use Their Cars for Short Trips?, 2003.
3. Steg, L. Car Use: Lust and Must. Instrumental, Symbolic and Affective Motives for Car Use. *Transportation Research Part A: Policy and Practice*, Vol. 39, 2-3, 2005, pp. 147–162.
4. Hunecke. Mobilitätsverhalten Verstehen Und Verändern. Psychologische Beiträge Zur Interdiszipliniaren Mobilitätsforschung, 2015.
5. Lois, D., and M. López-Sáez. The Relationship Between Instrumental, Symbolic and Affective Factors as Predictors of Car Use: A Structural Equation Modeling Approach. *Transportation Research Part A: Policy and Practice*, Vol. 43, 9-10, 2009, pp. 790–799.
6. Sefara, D., M. Franek, and V. ZUBR. Socio-Psychological Factors That Influence Car Prefernce in Undergraduate Students: The Case of the Czech Republic. *Technological and Economic Development of Economy*, Vol. 21, No. 4, 2015, pp. 643–659.
7. Sohn, K., and J. Yun. Separation of Car-Dependent Commuters from Normal-Choice Riders in Mode-Choice Analysis. *Transportation*, Vol. 36, No. 4, 2009, pp. 423–436.
8. Ellaway, A., S. Macintyre, R. Hiscock, and A. Kearns. In the Driving Seat: Psychosocial Benefits from Private Motor Vehicle Transport Compared to Public Transport. *Transportation Research Part F: Traffic Psychology and Behaviour*, Vol. 6, No. 3, 2003, pp. 217–231.
9. Bergstad, C. J., A. Gamble, O. Hagman, M. Polk, T. Gärling, and L. E. Olsson. Affective–symbolic and Instrumental–independence Psychological Motives Mediating Effects of Socio-Demographic Variables on Daily Car Use. *Journal of Transport Geography*, Vol. 19, No. 1, 2011, pp. 33–38.
10. Van, H. T., and S. Fujii. A Cross Asian Country Analysis in Attitudes Toward Car and Public Transport. *J. East. Asia Soc. Transp. Stud.*, Vol. 9, 2011, pp. 411–421.
11. Belgiawan, P. F., J.-D. Schmöcker, M. Abou-Zeid, J. Walker, T.-C. Lee, D. F. Ettema, and S. Fujii. Car Ownership Motivations Among Undergraduate Students in China, Indonesia, Japan, Lebanon, Netherlands, Taiwan, and USA. *Transportation*, Vol. 41, No. 6, 2014, pp. 1227–1244.
12. Nurul Habib, K., A. Tudela, J. Carrasco, and A. Idris. *Incorporating the Explicit Role of Psychological Factors on Mode Choice: A Hybrid Mode Choice Model by Using Data from an Innovative Psychometric Survey*.
13. von Behren, S., C. Minster, M. Magdolen, B. Chlond, M. Hunecke, and P. Vortisch. Bringing Travel Behavior and Attitudes Together: An Integrated Survey Approach for Clustering Urban Mobility Types. In *TRB 97th Annual Meeting Compendium of Papers*, Washington, D.C., 2018.

14. von Behren, S., C. Minster, J. Esch, M. Hunecke, P. Vortisch, and B. Chlond. Assessing Car Dependence: Development of a Comprehensive Survey Approach Based on the Concept of a Travel Skeleton. *Transportation Research Procedia*, Vol. 32, 2018, pp. 607–616.
15. Hunecke, M., S. Haustein, S. Böhler, and S. Grischkat. Attitude-Based Target Groups to Reduce the Ecological Impact of Daily Mobility Behavior. *Environment and Behavior*, Vol. 42, No. 1, 2010, pp. 3–43.
16. Institute for Mobility Research (ifmo). *Megacity Mobility Culture*. How Cities Move on in a Diverse World. Springer, Berlin, Heidelberg, 2013.
17. Ben-Akiva, M., D. Mcfadden, T. Gärling, D. Gopinath, J. Walker, D. Bolduc, A. Börsch-Supan, P. Delquié, O. Larichev, T. Morikawa, A. Polydoropoulou, and V. Rao. Extended Framework for Modeling Choice Behavior. *Marketing Letters*, Vol. 10, No. 3, 1999, pp. 187–203.
18. Bhat, C. R. Quasi-Random Maximum Simulated Likelihood Estimation of the Mixed Multinomial Logit Model. *Transportation Research Part B: Methodological*, Vol. 35, No. 7, 2001, pp. 677–693.
19. Vij, A., and J. L. Walker. How, When and Why Integrated Choice and Latent Variable Models Are Latently Useful. *Transportation Research Part B: Methodological*, Vol. 90, 2016, pp. 192–217.
20. CMC. *CMC Choice Modelling Code for R.* Choice Modelling Centre, University of Leeds, 2017.