Considering latent attitudes in mode choice: The case of Switzerland

Ricardo Hurtubia, Transport and Mobility Laboratory, Ecole Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland, ricardo.hurtubia@epfl.ch

Bilge Atasoy, Transport and Mobility Laboratory, Ecole Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland, bilge.kucuk@epfl.ch

Aurelie Glerum, Transport and Mobility Laboratory, Ecole Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland, aurelie.glerum@epfl.ch

Anne Curchod, Transport and Mobility Laboratory, Ecole Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland, anne.curchod@epfl.ch

Michel Bierlaire, Transport and Mobility Laboratory, Ecole Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland, michel.bierlaire@epfl.ch

Abstract

The paper presents the preliminary results of a recent study on mode choice for Switzerland, where psychometric indicators about attitudes and perceptions were collected. The attitude against public transportation is modeled and included as an explanatory variable in the choice model.

keywords: discrete choice, latent variable, mode choice.

1 Introduction

Most travel demand models consider main modal attributes and individual socioeconomic characteristics as the principal variables that explain mode choice (Ben-Akiva and Lerman, 1985). However, there are more complex, unobserved factors that may have a relevant effect on travel behavior. Examples of these are the individuals lifestyle, personal attitudes or perceptions. The way an individual perceives a transport mode is not observable, but might explain her preferences beyond traditional attributes like cost, travel time or comfort. For example, an individual might have a bias against public transport that is explained by her own perception of the safety or flexibility of this mode. In a similar way, an individual might prefer to use the car because she perceives that it is an indicator of social status.

Although not observable, these perceptions can be revealed by psychometric indicators. These indicators usually take the form of complementary questions in a survey, where the respondent is asked about her level of agreement with a series of statements. The indicators can be used to relate the individual's characteristics with the unobserved factor (or latent variable), allowing to build structural relations that can be later used to include the latent variable in a choice model.

This paper presents a mode choice model that includes attitude as latent variables. The characteristic feature of this work is its structure, which combines qualitative and quantitative methods in order to come up with stronger models. The research was done in the context of a collaborative work between EPFL's Transportation Center (TraCe) and Car Postal, the public transport branch of the Swiss Postal Service. The Car Postal system is characterized for serving low-density and isolated areas where, in practice, most individuals choose the car as their transport mode, probably as consequence of unobserved phenomena as the previously mentioned issues. The goal of this work is to understand the role of the unobserved variables to, eventually, develop strategies oriented to increase the market share of public transport.

The paper is organized as follows: in section 2 we summarize the theoretical formulation for discrete choice models including latent variables. Section 3 describes the data collection campaign, which included a qualitative and a quantitative survey. Section 4 explain the process to identify the relevant latent variables that will be included in the mode choice model, described in section 5. Finally, section 6 concludes the paper and identifies possible further work.

2 Integrated choice and latent variable model

The model presented next is based on the extended framework for integrated discrete choice and latent variables models proposed by Ben-Akiva et al. (1999) and generalized by Walker and Ben-Akiva (2002). Figure 1 shows the general structure of the model, where boxes represent observable variables and ovals represent unobservable variables.

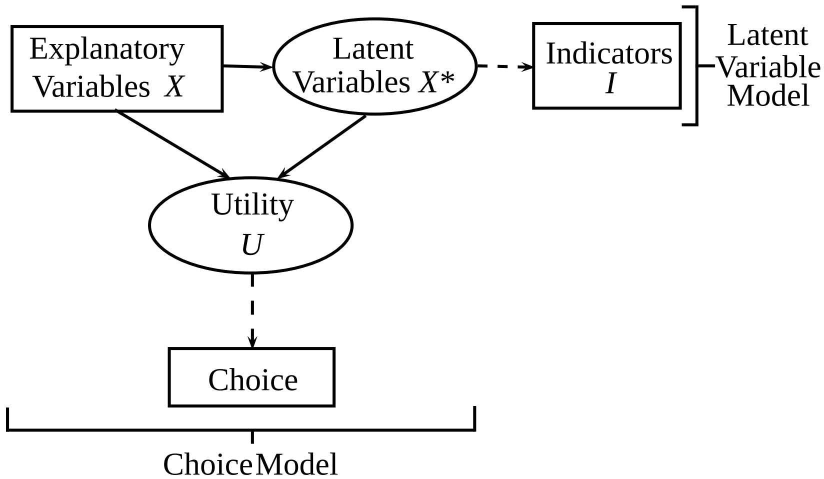

Figure 1: Integrated choice and latent variable model

Latent variables, denoted by X_n^* for individual n in our formulation, are unobserved constructs that capture subjective perceptions or attributes of individuals. They can't be directly measured, but the effect they have on measurable indicators (for example, survey questions regarding attitudes and perceptions) can be observed and therefore used to build a latent variable measurement equation:

$$I_n = I(X_n^*; \alpha) + \upsilon_n \tag{1}$$

where the indicator (I_n) is a function of the latent variables, a set of parameters (α) and an error term (υ_n).

Simultaneously, the latent variable can be related to observable explanatory variables (X_n); for example the characteristics of individual n. This relation is described by the following structural equation:

$$X_n^* = X^*(X_n; \lambda) + \omega_n \tag{2}$$

where λ is a set of parameters and ω_n is an error term.

The utility of choosing an alternative i which is also a latent quantity, can be defined as a function of the characteristics of the decision maker, the attributes of the alternatives and the latent variables:

$$U_{in} = V(X_n, X_i, X_n^*; \beta) + \varepsilon_{in} \tag{3}$$

where β is a set of parameters and ε_{in} is an error term. In a utility maximization framework, the measurement equation giving the probability of individual n choosing alternative i from a set of alternatives C_n is:

$$P(i|X_{n,}X_i, X_n^*; \beta, \theta_\varepsilon) = \mathsf{Prob}[U_{in} \geq U_{jn}, \forall j \in C_n] \tag{4}$$

where θ_ε is a vector of parameters of the error term in equation 3. Since X_n^* is not observable, it is necessary to integrate over the distribution of the latent variables and the indicators. The density function for the latent variables, $f(X^*|X_n; \lambda, \theta_\omega)$, is obtained from assumption on the distribution of ω_n in equation (2) while the density function of the indicators, $f(I_n|X^*; \alpha, \theta_\upsilon)$ is obtained from assumptions about the distribution of υ_n in equation (1). The parameters θ_ω and θ_υ are related to the distributional assumptions about the error terms of the structural equation and the measurement equation respectively. Incorporating the density functions, it is possible to write the joint probability of observing choice i and indicator I_n as:

$$P(i, I_n|X_{n,}X_i; \beta, \alpha, \lambda, \theta_\varepsilon, \theta_\upsilon, \theta_\omega) =$$

$$\int_{X^*} P(i|X_n, X_i, X^*; \beta, \theta_\varepsilon) f(I_n|X^*; \alpha, \theta_\upsilon) f(X^*|X_n; \lambda, \theta_\omega) dX^* \tag{5}$$

Estimating the parameters for this probability involves maximizing the likelihood function of observed choices and indicators. If we define y_{in} as a variable that is 1 if the observed choice of individual n is alternative i (and zero otherwise), the log likelihood function (L) can be written as:

$$L = \sum_n \sum_{i \in C_n} y_{in} \log P(i, I_n, |X_{n,}X_i; \beta, \alpha, \lambda, \theta_\varepsilon, \theta_\upsilon, \theta_\omega). \tag{6}$$

Once the model is estimated, the measurement equations are not used anymore for application, and the following specification is used:

$$P(i|X_n, X_i) = \int_{X^*} P(i|X_n, X_i, X^*) f(X^*|X_n) dX^*. \tag{7}$$

The previous probability accounts directly for the effect of the latent variable through the estimated parameters and can be directly applied to a data set containing only observable variables ($\mathsf{X_n}$,$\mathsf{X_i}$).

In the literature, several studies have applied the latent variable approach to choice models in the transport context. For example, for mode choice (Ben-Akiva and Boccara, 1995; Espino et al., 2006), car choice (Bolduc et al., 2008) or residential location choice (Walker and Li, 2007), just to name a few. All the studies found in the literature report an improvement in the quality of the estimates and the achievement of more realistic models when including unobserved factors through the latent variable approach.

3 Data Collection

The data collection campaign considered two surveys in the area of study (non-urban areas served by Car Postal). First, a qualitative survey (informal interviews) was performed in order to identify potential unobserved variables that affect travel behavior. This, in combination with examples found in the literature, was used to build a set of potential latent variables. The second survey registered revealed preferences (RP) regarding travel behavior and a set of psychometric indicators to measure the latent variables. This section describes the data gathering process, focusing in the construction of a set of latent variables and psychometric indicators.

In order to identify potential latent variables to measure, a qualitative survey was conducted. This consisted in interviews to 20 individuals in the Swiss canton of Vaud, focusing on residential choice, mobility biography, and mobility habits.

In addition to this, each of the 20 respondents were asked to carry a GPS with them for seven days, recording all their movements. The geocoded results were shown afterward to the respondents, were they identified the transport modes and purposes associated to each trip. During this part, additional (and informal) questions were made in order to complement the information already collected in the first part of the interviews.

The interviews were analyzed and a set of potential latent variables was identified (for a detailed description of the analysis see (Doyen, 2010)). This set of latent variables candidates, was contrasted and completed with latent variables used in similar studies found in the literature: Kitamura et al. (1997), Bagley and Mokhtarian (2002), Ory and Mokhtarian (2005) and Espino et al. (2006). This process allowed to identify the following set of latent variables to measure in the upcoming (quantitative) survey.

- **Spontaneity in travel behavior:** The interviews revealed that mode choice and travel preferences are not always "everyday decisions". Some individuals consider themselves to be captive to a specific transport mode while others are more flexible. This will depend on how spontaneous an individual is when it comes to mode choice or general travel behavior

- **Trip constraints:** Individuals show different levels of constraints when they travel. This is observed in specific needs that trigger the choice of specific transport modes. Examples of this are the requirement to travel at hours where the public transport supply is scarce or nonexistent, or the need to transport people (e.g. children) or objects.

- **Benefit from travel:** Some individuals consider their travel (especially their commuting) as a useful transition period from one activity to the other. Others use their travel time to perform activities (working, reading, socializing). The different levels of perception of benefit from travel are likely to affect the choice of the transport mode (e.g. using the train allows to read or work during the travel).

- **Predispositions toward specific transport modes:** Individuals have preconceived perceptions of the different transport modes. This generates (in the case of negative perceptions) the exclusion of some alternatives from the choice set. In the case of positive perceptions the choice of a transport mode might be influenced by these perceptions in a way that is not necessarily related with standard utility maximization.

- **Pro high density:** Regardless of their actual living location, many respondents report that they prefer to develop their activities in areas with high or low density. This (usually unobserved) preference has an indirect effect on mode choice since the convenience of using a specific transport mode may depend on the density of the travel destination (e.g. is preferable to go downtown by bus than by car)

- **Mutual aid:** Several respondents reported a strong use of their social networks to fulfill their transport needs. Examples of this are spontaneous car-pooling with neighbors or providing mutual help to transport objects or persons.

- **Personalized service:** Some individuals reported having a familiar relation with the drivers of the public transport system, especially in the case of isolated areas where the driver can sometimes deviate from his route to leave a passenger or wait for someone arriving late to the bus-stop. For these individuals, the possibility of having this form of "personalized service" was a strong factor in their preference for public transport.

- **Mobility habits (history):** The long term habits of a person may generate inertia in their decisions. Long-time individuals of a particular mode may have a biased preference for similar modes that is explained mostly for this habit.

- **Mobility skills:** The ability of understanding the transport system affects the mode choice. Individuals that feel stress for driving a car may prefer public transport, while individuals that are not familiar with the public transport system's schedule are less likely to choose it.

- **Environmental concern:** Some individuals base their transport mode choice on how environmental-friendly the different alternatives are perceived. People with high environmental concern might have a bias towards public or non-motorized transport

- **Status Seeking:** A status seeking person is likely to make choices considering how they are observed by the rest of society more than for the practical attributes of the alternatives themselves. This is likely to affect the activities a person performs, as well as his mode choice.

- **Lifestyle:** Several different types of lifestyles are reported in the literature. Examples of this are "Family oriented", "Workaholic", "Spontaneous" or "Conservative" people. The hypotheses on how these latent variables affect travel behavior are numerous. An example is the hypothesis that workaholics might have a high willingness to pay to reduce their travel time to the minimum or will prefer transport modes providing facilities to work during the trip. On the other hand, family oriented people might show more complicated travel patterns or needs, since they are likely to travel with their family. Spontaneous or conservative people is likely to extend this attribute to the way they choose their activities, the travel mode and even the route.

All the potential latent variables listed above where considered in the design of a list of psychometric indicators to measure unobserved factors that affect travel behavior.

The indicators were constructed based on examples found in the literature (Redmond, 2000; Vredin Johansson et al., 2006; Kitamura et al., 1997; Ory and Mokhtarian, 2005) and built as statements related with each of the potential latent variables. The indicators were selected in order to capture heterogeneity in the responses, trying to propose an equal number of negative and positive statements for each variable.

For example, for the predisposition or attitude towards public transport, some of the proposed statements were the following:

- *Its hard to take public transportation when I travel with bags or luggage.*
- *I know well which bus or train I must take, regardless of where I'm going.*

Each of the previous statements tries to capture a negative or positive perception of using public transport. The rest of the latent variables were related with statements in a similar way. A large list of statements, with several different types of response scales, was tested in a small group of voluntary respondents who were interviewed afterward in order to identify those that were confusing or unclear.

A final list of 52 statements with a 5 level Likert scale to indicate the level of agreement (from total disagreement to total agreement) was generated and included in a revealed preferences survey which also included:

- a trip diary where respondents registered all the trips performed during a day. This included origin destination, cost, travel time, chosen mode and activity at the destination,
- a set of questions on the socioeconomic of the respondent and his/her household.

At the moment of writing this paper, 1124 completed surveys were received and processed. For each respondent, cyclic sequences of trips (starting and ending at the same location) were detected and their main transport mode identified. Cycles with a main mode different from car or public transport were discarded since they were out of the scope of this work. In the remaining observations 32% of the cycles were performed by public transport while 68% were performed by car. Modal attributes like cost and travel time were calculated directly from the trip diary; for the non-chosen alternative, estimates of cost and travel time were inputted.

The data was used to generate the estimation database, with 1051 observations relating sequences of trips, psychometric indicators and socioeconomic attributes.

4 Latent variable selection

The collected answers for the level of agreement on the indicators were analyzed in order to confirm their relation with the hypothetical latent variables. An exploratory factor analysis was performed in order to identify the correlations between indicators, grouping them into factors that represent the latent variables. The relation between the indicators (I_k) and the unobserved factors (F_j) is given by the following equation:

$$I_k = \bar{I}_k + \sum_j \rho_{kj} F_j + \varphi_k \tag{8}$$

where $\bar{I}_k$ is the mean value of the answer for indicator k and φ_k is an error term following a normal distribution. The factor loadings (ρ_{kj}) quantify the correlation between the indicators and the factors.

The factor analysis generated 17 relevant factors with an eigenvalue greater than 1, that together explain 57% of the total variance among indicators.

Table 1 shows the resulting factor loadings for the 3 most relevant factors (those that explain most of the variability) and their related indicators. The table only shows the indicators with an absolute value greater than 0.3 and significant at the 95% level.

By analyzing the involved indicators and their signs, it is possible to infer the latent variable. For example, the two biggest positive factor loadings for Factor 1 are those of indicators 16 and 22, while the two most negative are those of indicators 1 and 25. Each of these indicators measured the level of agreement with the following statements:

- I_{16} : (+) *It's hard to take public transportation when I travel with my children.*
- I_{22} : (+) *I don't like to change transport modes when I travel.*
- I_1 : (-) *We should increase the price of gasoline to reduce congestion and air pollution.*
- I_{25} : (-) *I know well which bus or train I must take, regardless of where I'm going.*

A high level of agreement with the statements of indicators 16 and 22 shows dislike of public transport and multimodal trips. Indicator 1 is related with environmental concern and a willingness to reduce the use of car while indicator 25 shows knowledge or skills in using the public transport system. Extending this analysis to the rest of the indicators, and considering the sign

Table 1: Factor loadings

Indicator	Factor 1	Factor 2	Factor 3
1	-0.333	0.499	-
2	-	0.362	-
3	-	-0.533	-
4	-	-0.649	-
5	-	0.619	-
6	-	0.692	-
14	0.392	-	-
15	-	-	-
16	0.696	-	-
17	0.640	-	-
22	0.644	-	-
23	0.532	-	-
24	0.327	-	-0.338
25	-0.331	-	0.629
26	-	-	0.728
28	-	-	-0.458
29	-	-	0.314
30	-	-	0.552
32	-	-	0.338

of each factor loading, it is possible to conclude that Factor 1 represents an *attitude against public transport*.

A similar analysis performed with Factors 2 and 3 allowed to conclude that they represents latent variables related with *environmental concern* and *public transport awareness* respectively. The list of statements for all the indicators of Table 1 is included in the appendix.

5 Model specification and estimation results

A logit model including latent variables was specified and estimated following the framework described in section 2.

The choice is the main mode for a cycle of trips (starting and ending in the same location) and the alternatives are car or public transport. An indirect utility function was defined for each alternative as follows:

$$V_{CAR} = ASC_{CAR} + \beta_{cost} C_{CAR} + \beta_{TT_{CAR}} TT_{CAR} + \beta_{Att} Att\, TT_{CAR} + \sum_s \beta_s X_s \quad (9)$$

$$V_{PT} = ASC_{PT} + \beta_{cost} C_{PT} + \beta_{TT_{PT}} TT_{PT} + \beta_w \delta_{work} + \beta_{freq} F_{PT} \quad (10)$$

where C_{CAR}, C_{PT}, TT_{CAR} and TT_{PT} are the cost (in CHF) and travel time (in minutes) for car and public transport respectively. In the utility function for car, X_s is a vector of socioeconomic attributes, including number of cars in the household and dummy variables for university-level education, presence of children, and location of the household (French or German speaking region of Switzerland). Att is the latent variable for *attitude against public transport* described in section 4 and identified as the most relevant unobserved factor. High values for this variable imply a strong structural dislike or bias against public transport. In this specification the latent variable is interacted with the travel time for car.

In the utility function for public transport, δ_{work} is a dummy variable that is 1 if the sequence of trips involves only one destination and the performed activity is work (this attempts to identify simple and "compulsory" trips). F_{PT} is the average frequency (veh/h) of the different (public) transport modes involved in the cycle of trips.

In order to build the choice model, the utility functions of equations (9) and (10) need to be associated with an error term, as shown in equation (3). If we assume the error term to be Extreme Value distributed, a (binary) logit model is obtained, where the probabilities of choosing car or public transport are given by the following equations:

$$P_{CAR} = \frac{\exp(V_{CAR})}{\exp(V_{CAR}) + \exp(V_{PT})}; \;\; P_{PT} = \frac{\exp(V_{PT})}{\exp(V_{CAR}) + \exp(V_{PT})}. \quad (11)$$

The structural equation, relating the latent variable (Att) and the individual's socioeconomics is the following:

$$Att = \overline{Att} + \lambda_{cars} N_{cars} + \lambda_{educ} \delta_{educ} + \omega, \quad (12)$$

where $\overline{Att}$ is the mean value of the latent variable (to be estimated), N_{cars} is the number of cars in the households and δ_{educ} is a dummy variable that assumes the value of 1 if the individual has higher education degree. The error term ω is distributed Normal with mean 0 and standard deviation θ_ω

Finally, measurement equations were built for the three most relevant indicators of the considered latent variable (Indicators 16, 17 and 22 in Table 1 and Table 4). In order to capture the same effects already detected in the factor analysis, the measurement equations follow a structure similar to the structure of equation (8):

$$I_k = a_k + \alpha_k Att + \upsilon_k \qquad \forall k = 16, 17, 22, \quad (13)$$

where a_k and α_k are parameters to be estimated and Att is the latent variable defined by equation (12). The error term is normally distributed with mean 0 and standard deviation θ_{υ_k}.

Replacing equation (11) and the probability density functions obtained from equations (12) and (13) in equation (5) a likelihood function is obtained and the parameters are estimated using an extended version of the software package BIOGEME (Bierlaire, 2003). As only one latent variable is included in the model, numerical integration has been used. Estimation results, including reference results for a logit model, are shown in table 2

Table 2: Estimation results

	Affected utility		Latent var model		Logit	
Parameter	V_{CAR}	V_{PT}	**Value**	**t-test**	**Value**	**t-test**
ASC_{CAR}	x		-0.336	-0.75*	-0.229	-0.57*
ASC_{PT}		x	0**	-	-	-
β_{cost}	x	x	-0.118	-4.21	-0.058	-4.64
$\beta_{TT_{CAR}}$	x		-0.185	-3.77	-0.033	-4.4
$\beta_{TT_{PT}}$		x	-0.019	-3.64	-0.014	-3.21
β_{freq}		x	0.562	1.75*	0.488	1.81*
β_{w}		x	0.607	2.82	0.633	3.39
$\beta_{N_{cars}}$	x		0.691	3.29	0.702	3.48
$\beta_{children}$	x		0.444	1.96	0.328	1.63*
β_{French}	x		0.996	3.36	1.150	4.55
β_{educ}	x		0.672	2.68	0.390	1.92*
β_{Att}	x		0.473	3.4	-	-
$\overline{Att}$	x		2.850	38.07	-	-
λ_{cars}	x		0.121	2.9	-	-
λ_{educ}	x		-0.175	-2.84	-	-
a_{16}			0**	-	-	-
a_{17}			0.805	2.47	-	-
a_{22}			0.617	1.77*	-	-
α_{16}			1**	-	-	-
α_{17}			0.879	7.98	-	-
α_{22}			1.060	9.19	-	-
θ_{ω}			-0.519	-6.62	-	-
$\theta_{\upsilon_{16}}$			-0.166	-4.41	-	-
$\theta_{\upsilon_{17}}$			-0.012	-0.43*	-	-
$\theta_{\upsilon_{22}}$			-0.149	-3.74	-	-

(* Statistical significance < 95%)
(** Fixed parameter)

All the estimated parameters have the expected sign. Cost and travel time have a negative effect on the utility for both car and public transport. Variables like number of cars, presence of children, higher education or living in the French speaking region of Switzerland have a positive effect on the utility for car. On the other hand, the frequency of the public transport system has a positive effect on this mode's utility and trips with work as the only purpose are more likely to be performed by public transport, probably because trips with different or multiple purposes have more complex patterns and, therefore, are more easily performed by car.

The values of the parameters for the measurement equations (a_k, α_k) are all positive. This, and the fact that all considered indicators measured the agreement to negative statements for public transport, confirms the hypothesis of a latent variable (Att) measuring a bias *against* public transport.

As expected, the latent variable has a positive effect in the utility for car, which increases with longer trips. Since the travel time for car and public transport are positively correlated, this means that a person with a bias against public transport not only prefers the car, but this preference is even stronger when the trip is long.

An interesting result is the effect of the higher education degree. As mentioned before, people with a high education degree are more likely to choose car, compared with people with lower education ($\beta_{\mathsf{educ}} > 0$). However, the same variable has a negative parameter in the latent variable's structural equation (λ_{educ}). This means that there is a trade off regarding this variable: on one hand it linearly increases the utility for car while, simultaneously, it decreases the *attitude against public transport*, therefore reducing the utility for car in a non linear fashion (increasing with travel time). The net effect of having a higher education degree is positive for car when travel time is lower than 8.12 minutes; longer trips imply a negative net effect for this variable, therefore making people with high education less likely to choose car.

Something similar happens with the number of cars in the household: it has a linear positive effect in the utility of car and, simultaneously, has a (also positive) increasing effect as result of the interaction of the latent variable with the travel time. This type of result reveals a complex behavioral explanation of the mode choice process, which is only possible by using a latent variable model.

When comparing results with the logit model, it is possible to see that there are some significant differences in the values of the estimated parameters, most notably in the parameters for cost and travel time. This yields different estimated values of travel time savings (VoT) between both models. The logit model overestimates the value of time savings for car when comparing with previously estimated reference values for Switzerland (Axhausen et al., 2008), as seen in Table 3.

Table 3: Value of time

	VoT car (CHF/h)
Latent variable model	25.5
Logit model	34.32
Reference value*	20.98

(* overall value of travel time savings for all purposes)

The latent variable model is able to obtain more realistic values for the willingness to pay for travel time savings, thanks to the interaction of the latent variable with the travel time for car. This can be understood as the latent variable explaining a less relevant negative-effect of the travel time for car.

The final log likelihood for the latent variable model (L=-347.8) is bigger than the log likelihood of the MNL (L=-361.1) indicating a better fit for the latent variable model. Also, in general, the statistical significance of the parameters is worse in the logit model, having two parameters not significant at the 95% level ($\beta_{\mathsf{children}}$ and β_{educ}) that are significant in the latent variable model.

6 Conclusions

The inclusion of latent variables in choice models, specifically in mode choice, increases the quality of the estimates and provides a deeper understanding of the underlying dynamics of the choice process. In the particular case of a negative attitude towards public transport, the latent variable model was able to explain more complex phenomena than a logit model. This includes the double effect of the number of cars in the household, as an incentive for the use of car and a explanatory variable for the attitude against public transport, and the role of the education level, as a variable that has a positive effect in the utility for car for short trips and a negative one for long trips.

The latent variable model also produced better, more realistic, estimates. This is confirmed when analyzing the willingness to pay for travel time savings.

The paper presented a series of potential latent variables and identified those that are more relevant in terms of explanatory power. However, the estimated model included only one of these variables. Further model estimation should analyze the effect of other latent variables and their potential simultaneous inclusion in the utility functions of the choice model. This would require the use of simulation techniques for the estimation of the model.

The use of a more extensive set of indicators for each latent variable should also be explored, attempting to identify the optimal number of indicators to consider in this type of specification.

7 Acknowledgments

The authors would like to thank Car Postal for funding this study. Special thanks go to Gregor Ochsenbein for his comments and feedback during the whole study.

The identification of the latent variables and their indicators was a joint work with the Urban Sociology Laboratory (LASUR) and the Urban and Regional Planning Laboratory (CEAT) at EPFL; special thanks go to Vincent Kaufmann, Martin Schuler and Etienne Doyen.

The authors would also like to thank Sonia Lavadinho for her help with the survey design and Antonin Danalet for his help with data collection and preparation.

References

Axhausen, K. W., Hess, S., Koenig, A., Abay, G., Bates, J. J. and Bierlaire, M. (2008). Income and distance elasticities of values of travel time savings: New Swiss results, *Transport Policy* **15**(3): 173–185.

Bagley, M. N. and Mokhtarian, P. L. (2002). The impact of residential neighborhood type on travel behavior: A structural equations modeling approach, *The Annals of Regional Science* **36**(2): 279–297.

Ben-Akiva, M. E. and Boccara, B. (1995). Discrete choice models with latent choice sets, *International Journal of Research in Marketing* **12**: 9–24.

Ben-Akiva, M. and Lerman, S. R. (1985). *Discrete Choice Analysis: Theory and Application to Travel Demand*, MIT Press Series in Transportation Studies, The MIT Press, Cambridge, MA.

Ben-Akiva, M., McFadden, D., Garling, T., Gopinath, D., Walker, J., Bolduc, D., Boersch-Supan, A., Delquie, P., Larichev, O., Morikawa, T., Polydoropoulou, A. and Rao, V. (1999). Extended framework for modeling choice behavior, *Marketing Letters* **10**(3): 187–203.

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

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

Doyen, E. (2010). Résultats d'une enquete qualitative avec suivi GPS sur 20 personnes, *Technical report*, Urban Sociology Laboratory (LASUR), Ecole Polytechnique Fédérale de Lausanne, Switzerland.

Espino, R., Roman, C. and Ortuzar, J. d. D. (2006). Analyzing demand for suburban trips: A mixed RP/SP model with latent variables and interaction effects, *Transportation* **33**(3): 241–261.

Kitamura, R., Mokhtarian, P. L. and Laidet, L. (1997). A micro-analysis of land use and travel in five neighborhoods in the San Francisco Bay Area, *Transportation* **24**: 125–158.

Ory, D. and Mokhtarian, P. (2005). Dont work, work at home or commute? discrete choice model of the decision for san francisco bay area residents, *Technical report*, Institute of Transportation Studies, University of California, Davis.

Redmond, L. (2000). *Identifying and analyzing travel-related attitudinal, personality, and lifestyle clusters in the San Francisco Bay Area*, Master's thesis, University of California, Davis, USA.

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

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

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

Appendix:

Table 4: Psychometric Indicators

Ind	Statement	Factor 1	Factor 2	Factor 3
1	We should increase the price of gasoline to reduce congestion and air pollution	-0.333	0.499	-
2	We need more public transport, even if it means higher taxes	-	0.362	-
3	Environmentalism harms the small businesses	-	-0.533	-
4	People and employment are more important than the environment	-	-0.649	-
5	I am concerned about global warming	-	0.619	-
6	We must act and make decisions to reduce emissions of greenhouse gases	-	0.692	-
14	I am not comfortable when I travel with people I do not know near	0.392	-	-
15	Taking the bus helps to make the city more comfortable and welcoming	-	-	-
16	It's hard to take public transportation when I travel with my children	0.696	-	-
17	It's hard to take public transportation when I travel with bags or luggage	0.640	-	-
22	I don't like to change transport modes when I travel	0.644	-	-
23	If I use public transport instead of my car, I have to cancel some activities	0.532	-	-
24	The bus schedule is sometimes hard to understand	0.327	-	-0.338
25	I know well which bus or train I must take, regardless of where I'm going	-0.331	-	0.629
26	I know the bus schedule by heart	-	-	0.728
28	I feel very disoriented when I'm in a place I don't know	-	-	-0.458
29	I use the Internet to find out about the bus or train schedule	-	-	0.314
30	I have used public transport all my life	-	-	0.552
32	I know some of the drivers of the buses I take	-	-	0.338