Stated response and multiple discrete-continuous choice models

Analyses of residuals

Working Paper**Author(s):**

Jäggi, Boris; Axhausen, Kay W. Weis, Claude

Publication date:

2011-04

Permanent link:

<https://doi.org/10.3929/ethz-a-006686110>

Rights / license:

[In Copyright - Non-Commercial Use Permitted](#)

Originally published in:

Arbeitsberichte Verkehrs- und Raumplanung 661

Stated Response and Multiple Discrete-Continuous Choice Models: Analyses of Residuals

Boris Jäggi

Claude Weis

Kay W. Axhausen

Working paper

Transport Planning and Systems

661

April 2011

Working paper

661

Stated Response and Multiple Discrete-Continuous Choice Models: Analyses of Residuals

Boris Jäggi

IVT

ETH Zurich

CH-8093 Zurich

phone: +41-44-633 67 37

fax: +41-44-633 10 57

boris.jaeggi@ivt.baug.ethz.ch

Claude Weis

IVT

ETH Zurich

CH-8093 Zurich

phone: +41-44-633 39 52

fax: +41-44-633 10 57

weis@ivt.baug.ethz.ch

Kay W. Axhausen

IVT

ETH Zurich

CH-8093 Zurich

phone: +41-44-633 39 43

fax: +41-44-633 10 57

axhausen@ivt.baug.ethz.ch

April 2011

Abstract

In sophisticated transport models choice modeling is used to capture a wide range of behaviors, such as mode choice, fleet choice or route choice. A newly developed approach to improve realism is the multiple discrete-continuous extreme value (MDCEV) model, which allows to model the allocation of continuous amounts of a consumption good. Before using this models in overall frameworks, knowledge about the accuracy of the forecasting procedure is important. In this paper a MDCEV model of fleet choice based on data collected in a Stated Adaptation survey is presented. A forecast of the model predicting annual mileage of households to 17 different car types was made and the results were compared to the actual data calculating the residuals. The residual analysis shows that the model performs significantly better than a totally random model, but the share of wrongly allocated mileage, 70% of total, remains high. However an assessment of the result is difficult with only one model. The differences between two sub-models, one without public transport, another including it, regarding the distribution of the residuals indicate that the model specification has a big influence on its performance. Therefore, following work forecasting additional MDCEV models will be necessary to have a base for comparison. We compare two further MDCEV models to obtain a fuller understanding of their performance.

Keywords

Stated Preference, Stated Adaption, Fleet Choice, MDCEV, Forecast, Simulation, Residual, Validation

Preferred citation style

Jäggi, B., Weis, C. and Axhausen, K.W. (2011) Stated Response and Multiple Discrete-Continuous Choice Models: Analyses of Residuals, *Working paper*, **661**, Institute for Transport Planning and Systems (IVT), ETH Zurich, Zurich.

1 Introduction

In sophisticated transport models like the SACSIM model of the Sacramento Area, California (Bradley *et al.*, 2010), the ILUTE model in Toronto (Salvini and Miller, 2005) or the Albatross model from the Netherlands (Beckx *et al.*, 2009), choice modeling is used to capture a wide range of behaviors, such as mode choice, fleet choice or route choice. Discrete choice models in their standard formulations cannot integrate multivariate choices and associated continuous attributes of these choices. Still, there are a number of questions where this capability would allow the modeler to improve the realism of the description. One prime example is the composition of the fleet of mobility tools (Simm *et al.*, 2002; Simm and Axhausen, 2003, 2001b,a) and their associated mileage. The recent development of the MCDEV framework by Bhat (2005) offers a new approach to address this gap.

The overall transport model currently developed at IVT in collaboration with TU-Berlin is MATSim (Balmer, 2007; Meister *et al.*, 2009; Balmer *et al.*, 2008), an agent based micro-simulation tool for travel demand and traffic flow modeling. The present paper is part of the ongoing work to implement multiple discrete-continuous extreme value (MDCEV) models into the model frameworks of different fields. As our and the literature's experience in the use of this approach is, given its recent development, small, an evaluation of actual forecasting results and their residuals is necessary.

In MATSim travel demand is activity based and generated using activity chains from the Swiss national travel diary survey, the Mikrozensus (Swiss Federal Statistical Office, 2006). The Mikrozensus is conducted every five years. In the current version of MATSim, the agents can conduct activities (e.g. home, shopping, work, leisure, etc.) inside facilities (buildings). In the iterative solution process the agents optimize their given activity chain. The agents are not part of a household and they have no specific car type allocated to them yet. They only have an attribute that describes their car availability for the mode choice processes.

As a future development of MATSim, the agents shall be pooled in households and specific car fleets will be allocated to the households. This will not only enable analysis of energy consumption on a microscopic level, but also allow the implementation of a behavioral model to forecast the development of the car fleet and, as a result, of the energy consumption. The first step for these enhancements is a fleet ownership model.

Such a model is estimated using the MDCEV approach by Bhat (2005) on Stated Adaptation data collected in a survey conducted by Erath and Axhausen (2010). To test the performance of such a model for an application like MATSim, we applied a MDCEV model repeatedly and analyzed the residuals comparing the results to the actual choices in the experiment. Since 2004, when the MDCEV model was originally developed to analyze time use (Bhat, 2005), various researchers have used it to estimate preferences. (Sen, 2006) presented a MDCEV Model in the context of

examining vehicle type, model and usage decisions of households in his dissertation Bhat and Sen (2006). The impact of demographics, built environment attributes, vehicle characteristics and gasoline prices on the same issue are analyzed in (Bhat *et al.*, 2009). Pinjari *et al.* (2009) analyzed residential self-selection effects in time-use models and Spissu *et al.* (2009) presented an analysis of weekly out-of-home activity participation. Copperman and Bhat (2007) analyzed the determinants of childrens week end activity participation. In Pinjari and Bhat (2010b), the authors introduce the nested version of the MDCEV, the multiple discrete-continuous nested extreme value (MDCNEV) model and present an application on non-worker time-use behavior. A detailed description of the MDCEV and the role of its parameters can be found in (Bhat, 2008).

Pinjari and Bhat (2010a) presented an efficient forecasting procedure for such models. We intend to test the performance of MDCEV forecasting results with our Stated Adaption data sets about fleet choice, energy savings in household and private transport and induced demand. Unfortunately we could not find any useful literature on the topic of disaggregate validation on (multiple) discrete-continuous models. In most MNL models, validation happens on an aggregate level by comparing actual with predicted market shares, as done for MNL models since their introduction, for example in (Train, 1978). In the case of multiple discrete-continuous models however, a disaggregate validation is more valuable because it gives more insight in the model's characteristics.

2 Data

2.1 Survey

The primary data set used here was collected for a project of the Swiss Federal Office of Energy and the Federal Office for the Environment about long term fuel price elasticity and the effects on mobility tool ownership and residential location choice (Erath and Axhausen, 2010). In the survey, 409 households were questioned about their long term reactions to rising fuel costs. The survey was divided in a part on socioeconomic and mobility tool related questions and a three stage stated response survey. In the first part, the respondents are presented six scenarios of fuel prices ranging from CHF 1.5/l to CHF 5.5/l for gasoline. The survey was conducted in face-to-face interviews, in which the interviewer was equipped with a computer-software that simultaneously calculated the personalized mobility costs (fixed cost separate from variable cost) based on personal information collected previously. The respondents could choose their car fleet and annual mileage at a high level of detail including car type, engine size, drive-train, and if they would buy a new or a used car, while being supported by the real time calculations of the computer. They could also choose and/or change the mileage traveled by public transport.

In the second stage of the respondents were confronted with six different residential locations as well as varying fuel prices and were again asked to choose the preferred mobility tool (and mileage) for each situation. For the third stage of stated preference experiment another six choice situations were created. The choice sets in this consisted of two alternatives, one from both previous stages each. The data used in this particular paper comes from the first stage only.

2.2 Data Overview

The representativeness of the data in term of mobility tools, car ownership and socioeconomic variables are summarized in table 1. The column quota describes the targeted share according to the Swiss national transport survey (Swiss Federal Statistical Office, 2006), the column effect the actual share in the survey. In our data, male participants are slightly over-represented, as well as the age group of 36-50 years old. We have also a significantly higher share of single person household and persons without a public transport season ticket. In terms existing fleet composition, the most frequent car types, such as upper middle class, middle class, minivan and compact, are under-represented while the more special ones like sports car and micro are over-represented. The income distribution is matched reasonably well, although there are more high income households than expected.

3 Methodology

3.1 MDCEV

The methodology used for the fleet choice model used here and its analysis of residuals is the multiple discrete-continuous extreme value approach developed by Bhat (2005). In this paper he gives a detailed description and derivation of the model. The following section is a short summary of the third chapter of the paper, to give readers who are not familiar with the method an idea of the approach. The MDCEV was originally developed to estimate the influence of attributes on the decisions of allocating time (continuous) to activities (discrete) within a 24-hour budget. Because the various activities are equivalent and simultaneously chosen for each day the model considers multiple chosen alternatives. In the presented model, the discrete choices are car types, the continuous amount is annual mileage and it is a multiple discrete model because households can own more than one car.

Table 1: Data Overview

Variable	Characteristic	Quota [%]	Effect. [%]	Variable	Characteristic	Quota [%]	Effect. [%]
Sex	Male	48.3	55.0	Car availability	always	83.5	84.0
	Female	51.7	45.0		occasional	16.5	12.8
Age (years)	18 - 35	27.2	25	PT season ticket	none	63.2	73.8
	36-50	32.0	38.0		Half-Fare	29.9	16.1
	51-65	24.5	25.5		GA	6.9	4.1
	> 65	16.4	11.5				
Highest education level	Compulsory ed.	18.4	20.6	Cars in household	1	62.4	72.4
	Prof. School	55.6	57.6		2	31.0	16.1
	Tertiary ed.	26.0	21.8		> 2	6.6	4.1
Household income (CHF/m.)	< 2,000	1.8	2.4	Car type	Sports car	2.6	8.1
	2,000 - 4,000	14.6	12.5		Luxury /SUV	6.3	6.3
	4,000 - 6,000	28.6	25.7		Upper middle class	22.3	17.9
	6,000 - 8,000	23.6	17.6		Middle class	22.3	17.9
	8,000 - 10,000	14.3	15.4		Minivan/Van	14.1	13.3
	10,000-12,000	7.8	10.3		Compact	23.1	20.3
	>12,000	9.3	9.0		Subcompact	19.0	18.1
	n.a.	-	6.8		Micro	3.7	8.1
Persons per household	1	20.5	34.5				
	2	38.9	39.1				
	3	14.7	11.0				
	4	18.0	12.0				
	> 4	7.9	3.4				

Kim *et al.* (2002) defines the utility an individual obtains for his decisions as a sum over all j alternatives (in our case: car types):

$$U = \sum_{j=1}^K \psi(x_j, \varepsilon_j)(t_j + \gamma_j)^{\alpha_j} \quad (1)$$

In this utility structure, t_j is the continuous amount of annual mileage driven with car type j ($j = 1, 2, \dots, K$), γ_j and α_j are satiation parameter to estimated within the model. The function $\psi(x_j, \varepsilon_j)$ gives the baseline utility function for the mileage driven with car type j . In section 3.1

of his paper, Bhat (2005) presents a random utility function for the baseline utility:

$$\psi(x_j, \varepsilon_j) = \exp(\beta' x_j + \varepsilon_j) \quad (2)$$

In which β' is a vector of parameters that define the influence of the observed characteristics of the alternative x_j . ε_j captures the unobserved random utility. By combining the formulas (1) and (2) the overall random utility function for the MDCEV model can be defined as:

$$\bar{U} = \sum_{j=1}^K \exp(\beta' x_j + \varepsilon_j) \cdot (t_j + \gamma_j)^{\alpha_j} \quad (3)$$

By forming the Lagrangian and applying the Kuhn-Tucker conditions and assuming that the optimal allocation of annual mileage satisfies the budget constraint $\sum_{j=1}^K t_j^* = T$, the probability function can be derived. Bhat specifies a standard extreme value distribution for ε_j and assumes that it is independent from x_j as well as independently distributed across alternatives. The final result for the probability function is:

$$P(t_2, t_3, \dots, t_M, 0, 0, \dots, 0) = \left[\prod_{i=1}^M c_i \right] \left[\prod_{i=1}^M \frac{1}{c_j} \right] \left[\frac{\prod_{i=1}^M e^{V_i}}{(\sum_{j=1}^K e^{V_i})^M} \right] (M-1)! \quad (4)$$

whereas:

$$c_i = \left(\frac{1 - \alpha_i}{t_i + \gamma_i} \right) \quad (5)$$

M is the number of alternatives chosen by the individual. If only one alternative is chosen, the model collapses to the form of a standard Multinomial Logit model. Therefore this model is an extension of the standard MNL model, allowing multiple choices of continuous amounts.

3.2 Fleet Choice Model

The model presented in this paper is a fleet choice model that includes public transport as a 'car type'. The data used are from a stated adaption process, in which the respondents of the

survey described above, re-allocated their annual mileage between different car types and public transport as a reaction to a given gasoline (or diesel) price. The possibilities to change the fleet composition include, in addition to the car types shown in table 1, engine size, drive-train and whether a new or a used car is bought. If the respondents decided to keep their existing car it was labeled as used car. The combination of these features gives a huge number of discrete choices (=alternatives). Thus, the 2436 observations are categorized into 17 classes. The class 'no car' has an annual mileage of 0, but the alternative 'public transport' is added to the choice set, it consists of 17 alternatives, as shown in table 2. Observations mean the number of household that chose this alternative as their primary means of movement, i.e. with the highest mileage.

The model used for this paper has no outside good, meaning that there was no alternative that was chosen in every observation. It is obvious that there is no car type which has to be chosen by all households, as for example 'in home time' in time use models or 'housing costs' in financial allocation models.

For the estimation process, the Gauss code provided at Bhat's Web-page (Bhat, 2011) is used. The programm 'No Outside Good' is used and both configurations 'config' = 1 and 'config' = 4 were tested. 'config' = 1 means, that the estimated satiation parameters of the model are α parameters and the γ values are constraint to be equal to one for all goods. In this case, the specific utility function is:

$$U(t) = \sum_{j=1}^K \frac{1}{\alpha_j} \exp(\beta' x_j + \varepsilon_j) \cdot \{(t_j + 1)^{\alpha_j} - 1\} \quad (6)$$

The other configuration tested for the presented work is 'config' = 4, meaning that γ parameters are estimated while α values are fixed to be equal 0. In that case, the specific utility function is:

$$U(t) = \sum_{j=1}^K \gamma_j \cdot \exp(\beta' x_j + \varepsilon_j) \cdot \ln \left(\frac{x_j}{\gamma_j} + 1 \right) \quad (7)$$

The models estimated assumed satiation parameters that differ across individuals. The γ parameters are estimated as a function of household income, fuel price and a constant. We did not reach convergence for models using α parameters differing across individuals in the given time frame. Because of this, and because of the slightly better model fit, only results using γ satiation parameters are presented in this paper.

Table 2: Alternatives of the Fleet Choice Model

Alternative	New/Used	Drive-train	Car Types	Engine Size	Observations
A1	New	Diesel	Micro, Subcompact, Compact		160
A2	New	Diesel	Minivan, Middle-class Lower middle-class		89
A3	New	Diesel	SUV, Sports-car, Upper middle-class		76
A4	New	Gas, Hybrid Electric			119
A5	New	Gasoline	Micro, Subcompact, Compact	<1500 cm ³	279
A6	New	Gasoline	Micro, Subcompact, Compact	>1500 cm ³	110
A7	New	Gasoline	Minivan, Middle-class Lower middle-class		155
A8	New	Gasoline	SUV, Sports-car, Upper middle-class		123
Not included			No Car Car Sharing		94
A10	Used	Diesel	Micro, Subcompact, Compact		170
A11	Used	Diesel	Minivan, Middle-class Lower middle-class		141
A12	Used	Diesel	SUV, Sports-car, Upper middle-class		49
A13	Used	Gas, Hybrid Electric			136
A14	Used	Gasoline	Micro, Subcompact, Compact	<1500 cm ³	272
A15	Used	Gasoline	Micro, Subcompact, Compact	>1500 cm ³	135
A16	Used	Gasoline	Minivan, Middle-class Lower middle-class		198
A17	Used	Gasoline	SUV, Sports-car, Upper middle-class		130
A99			Public Transport		

3.3 Forecasting procedure

The forecasting procedure as well as the Gauss code used for this paper were developed and introduced by Pinjari and Bhat (2010a). In chapter four of his paper, two algorithms are presented, one for estimations with γ_j as satiation parameters and a fixed α , and one for estimations with α_j as satiation parameters. The algorithms presented in the paper are both for models with outside goods. On the Web-page of Bhat (2011), the Gauss code of the forecasting procedure can be downloaded.

In the first step of the algorithm, the baseline utilities $\exp(\beta'x_j + \varepsilon_j)$ of all alternatives are calculated out of the model parameters β , the input data and the Gumbel distributed unobserved utility ε . In the code developed by Pinjari and Bhat (2010a), the random numbers used are from a Halton-Sequence, after Halton (1960). In the current work, random numbers were drawn on the fly with the random function of the Gauss programm.

The 17 alternatives are then sorted in decreasing baseline utility and it is assumed that the first alternative is chosen. To check if a second alternative (with the second highest baseline utility) is chosen, an estimate of the overall budget is calculated. If the estimated budget is smaller than the actual budget, the next alternative in line is considered, until the budget is exhausted. If the estimated budget is higher than the actual, an iterative procedure, described in detail in (Pinjari and Bhat, 2010a), is used to find an adjusted baseline utility for the last alternative to meet the actual budget within a tolerance band.

After the exact number of chosen alternatives is determined, the mileage allocated to every of the chosen alternatives is calculated using the adjusted baseline utility of the previous steps and the satiation parameters γ .

The above algorithm calculates the consumption data (mileage) for one observation. It can be repeated multiple times to achieve more robust results. In the current work simulations with 10 repetitions and 50 repetitions are calculated, presented and compared.

4 Results

4.1 Estimation Results

Table 3 presents the estimated β -parameters for the alternatives in the standard model without the public transport (PT) alternative. Table 4 shows the estimated β -parameters for the model including public transport. Please note that **boldly** written numbers are significant at a 95% level and numbers written in *italic* are still significant at a 90% level. The Alternatives are labeled

according to table 2. For an easier and quicker reading of the table, an indication is given in the column "Alternative", where the size of the car is indicated. Alternatives A1 to A8 mean new cars, A10 to A17 mean used cars. A car type is considered "used", either if the respondent chooses to keep his current car, or if he buys a used car.

The satiation parameters presented in tables 3 and 4 are θ -parameters. To model γ parameters that differ across individuals, they are parameterized in the following form, dependent of monthly household income and given fuel price:

$$\gamma_j = \exp(\theta_C + \theta_I \cdot Income + \theta_F \cdot Fuelprice) \quad (8)$$

To evaluate the models in terms of model fit, the mean log likelihood values are compared with mean log likelihood value of the reference model, which estimates only the alternative specific constants. The mean log likelihood value for the reference model is -5.22, that for the model including PT is -3.41 and that for the model with only private transport options is - 2.94. The significantly higher model fit of the model without PT proves the existence of a big difference between intra-modal and inter-modal choices and that a mixture of these two kind of choices in one model gives worse results. However, looking at the results of the simulation, the residuals are not that different in terms of accuracy as the model fit would suggest.

The variable Const is the alternative specific constant. One can see that new gasoline cars, which are the majority of cars available on the market, have no significant negative value, in contrast to other car types. Luxurious cars are the only ones with significant negative values, as they are the also the most expensive ones.

Income is gross household income in 1'000 CHF per month. The significant effect of this variable is, that high income households are more likely to buy luxurious and middle class cars and lower income households favor small gasoline cars. Interestingly, cars with alternative fuels (A4, A13) are negatively influenced by income, although they are more expensive. This indicates that early adopters are not necessarily high income people.

Fuel is the fuel price, varying in the experiment from 1.5 CHF/l to 5.5 CHF/l, Fuel² is the square of the fuel price. The higher the fuel price, the less mileage is allocated to all car types, except for A4 and A13 which are alternative technology cars as well as A2, A3 and A11 which are middle to high class diesel cars, of which the fuel price influence is quadratic. In the case of rising fuel prices people with inefficient gasoline cars would switch to diesel until a certain fuel price is reached before adjusting the annual mileage or buying a hybrid or electric car.

Dist. is the respondent's distance between home and workplace in 100 km. The longer the commuting distance, the more mileage the respondents have to allocate. The high and significant

Table 3: Estimated β Parameters of Fleet Choice Model, mean log likelihood: -2.94

	Const	Income	Fuel	Fuel2	DistW	Age2	Age3	Age4	Male	Urban	Inertia	Ac.1	Ac.2	θ_C	θ_I	θ_F
A1 (small)														2.75	-0.13	0.19
A2 (middle)	-12.40	0.02	1.11	-0.19	0.59	10.28	10.34	10.98	-0.21	-0.08	1.11			16.70	-0.09	-2.92
A3 (lux)	-12.99	0.01	0.80	-0.14	-0.14	10.25	10.80	9.81	0.51	0.26	2.45			0.79	0.09	-0.01
A4 (hybrid)	-1.24	-0.03	0.14		0.40	0.67	1.14	2.64	-1.46	0.30	-10.27	0.06	0.12	0.84	3.54	2.39
A5 (small)	0.90	-0.06	-0.16		0.56	-0.09	-0.33	0.95	-0.43	0.05	1.51			1.53	-0.14	0.13
A6 (small)	0.35	0.02	-0.39		0.69	-0.68	-0.75	1.04	0.03	0.16	1.61			1.90	-0.09	-0.06
A7 (middle)	0.03	0.02	-0.39		0.37	-0.17	-0.38	1.04	-0.03	0.23	2.19			3.15	-0.13	-0.20
A8 (lux)	-0.57	-0.01	-0.59		0.69	-0.43	-1.71	-12.99	-0.29	-0.20	5.26			1.73	-0.06	-0.26
A10 (small)	-0.34	-0.03	-0.23		0.71	0.92	1.03	1.66	0.42	-0.25	1.63			5.47	-0.37	0.10
A11 (middle)	-12.47	0.14	0.33	-0.09	1.36	9.82	10.32	11.50	0.21	0.21	2.45			1.96	-0.07	0.05
A12 (lux)	-2.92	0.17	-0.27		1.67	-1.50	-1.90	0.92	1.33	0.82	3.12			1.16	0.01	0.08
A13 (hybrid)	0.30	-0.04	0.12		1.04	-1.11	-0.50	0.02	-0.75	0.39	4.84	0.06	0.12	0.49	2.55	1.78
A14 (small)	-0.12	-0.02	-0.32		1.44	-0.72	-0.58	0.59	-0.50	0.26	3.20			2.39	-0.14	-0.08
A15 (small)	-0.52	0.02	-0.32		1.62	-1.88	-0.67	0.57	0.32	0.12	2.36			7.13	-0.40	-0.12
A16 (middle)	0.11	0.10	-0.60		1.35	-1.93	-1.60	0.02	0.06	0.69	3.36			4.91	-0.33	0.12
A17 (lux)	-2.21	0.09	-0.69		1.10	-2.24	-2.77	0.30	0.88	-0.48	6.66			2.56	-0.21	0.22

Table 4: Estimated β Parameters of Fleet Choice Model incl. PT, mean log likelihood: -3.41

	Const	Income	Fuel	Fuel2	DistW	Male	Urban	Inertia	Ac.1	Ac.2	GA	HAT	SC	θ_C	θ_I	θ_F
A1 (small)														2.60	-0.14	0.18
A2 (middle)	-7.81	0.02	1.03	-0.18	0.52	-0.25	-0.07	1.15						18.19	-0.08	-3.29
A3 (lux)	-7.43	0.02	0.75	-0.14	0.24	0.52	0.34	2.76						1.02	0.09	-0.11
A4 (hybrid)	-1.02	-0.03	0.13		0.42	-1.41	0.23	-6.06	0.01	-0.01				0.64	2.61	1.69
A5 (small)	0.93	-0.06	-0.17		0.44	-0.40	0.04	1.43						1.42	-0.13	0.04
A6 (small)	0.31	0.02	-0.40		0.57	0.04	0.18	1.57						1.44	-0.08	-0.02
A7 (middle)	0.34	0.01	-0.38		0.29	-0.15	0.29	2.20						2.74	-0.15	-0.08
A8 (lux)	-0.34	0.00	-0.53		0.63	-0.43	-0.31	5.20						1.46	-0.04	-0.24
A10 (small)	-0.33	-0.05	-0.23		0.67	0.49	-0.22	2.10						5.28	-0.35	0.04
A11 (middle)	-7.32	0.14	0.18	-0.08	1.28	0.10	0.31	2.58						2.00	-0.08	-0.01
A12 (lux)	-2.67	0.17	-0.29		1.58	1.17	0.70	3.30						1.07	-0.01	0.16
A13 (hybrid)	0.62	-0.03	0.10		1.01	-0.77	0.28	3.51	0.01	-0.01				0.38	1.95	1.43
A14 (small)	0.00	-0.02	-0.34		1.31	-0.48	0.24	3.13						1.69	-0.11	-0.05
A15 (small)	-0.56	0.02	-0.32		1.57	0.30	0.10	2.30						5.06	-0.28	-0.19
A16 (middle)	0.55	0.07	-0.57		1.23	-0.13	0.72	3.27						0.89	2.42	1.76
A17 (lux)	-1.96	0.09	-0.63		1.04	0.65	-0.65	6.78						2.19	-0.19	0.19
A99 (PT)	4.40	-0.06	-0.29		1.02	-0.33	0.36		0.01	-0.01	2.24	1.06	1.42	-4.14	0.09	0.06

values are all for used cars, meaning that people with a longer commuting distance tend to be less willing to buy a new car. We can also see no effect, that people with longer commuting distances, and therefore higher consumption, would favor more efficient cars.

In influence of the age of the respondent is modeled using four age groups, with the youngest group as the reference group. The age groups are Age1 for persons younger than 26 years, Age2 for persons between 26 and 45 years, Age3 for persons between 46 and 65 years and Age4 for persons older than 65 years, which is also the legal retirement age in Switzerland. It is difficult to identify a pattern, but younger people reject used cars more than older. It is also interesting to see that people in the retirement age buy significantly more new alternative technology cars.

Male is a dummy for the gender of the respondent. Alternative drive trains and gasoline are preferred by women, while diesel is preferred by men. This is a quite interesting finding, because it rejects the general assumption, that men are more interested in, and therefore more open to new technologies.

Urban is a dummy which is one for people living in inner city or urban areas and zero for people in suburbs and rural areas. It is only significant for used cars. Resident location has not a big influence on fleet choice, but alternative technologies (A13) and luxurious diesel cars are favored in urban areas.

Inertia is a dummy to capture inertia effects. It is one if the chosen car type is the actual car type the respondent owns. This has the expected significant and substantial influence. Only the choice of a new hybrid, gas or electric car (A13) is not influenced by ownership of the same car type. This can be explained with the fact that the technology is so new that the advantage of a new car is small and owners of this car type have no need to buy a new one in the near future.

Acc.1 and Acc.2 are two variables for accessibility, coming from a factor analysis of private transport accessibility and public transport accessibility, based on a national aggregate transport model (Fröhlich *et al.*, 2006). Acc.1 stands for general accessibility of the respondent's home municipality and Acc.2 for any differences in public transport accessibility. The parameter is fixed for alternative fuel alternatives and public transport. These non-conventional alternatives are more often chosen in areas with higher accessibility in comparison to gasoline and diesel cars, but only in the standard model. In the model including public transport, there is no statistical significant effect. Differences between conventional alternatives regarding accessibility impacts are neither expected nor found.

The parameters for GA, HT and SC describe the influence of existing mobility tools for the public transport use in the model which includes public transport. GA (Generalabonnement) is a dummy for a season card for the whole of Switzerland, HT (Halbtax) one for a half-fare card for the whole of Switzerland and SC for a regional season card. The presence of such a mobility tool has the expected strong positive effect on annual public transport mileage.

The satiation parameters $\gamma = f(\text{Income, fuel price})$ describe the decreasing marginal utility with an increasing amount of traveled kilometer. The θ constant is highly significant. The more luxurious the car type, the lower that constant. That means that people are more likely to allocate their annual mileage in the more luxurious of two or more cars. For example: the main car, with a higher mileage, is the bigger, more comfortable car and the second car is for the case the first is not available. Cars with alternative fuel are not likely to be affected by reduction, meaning that if one has for example a hybrid car, the person is not likely to have a second car with which it drives even more. Income has the expected influence on the satiation such that higher income gives less satiation throughout all car types except alternative fuels. Fuel price has almost no significant impact on satiation which is surprising.

4.2 Forecasting Results

4.2.1 Stated Preference Models for Comparison

In this section two further models that also use Stated Adaption data sets for MDCEV modeling are briefly presented. The models differ in terms of context, survey methodology, sample size and, most importantly, number of available alternatives. This allows us to draw a broader picture of different possible outcomes of forecasting MDCEV models and their assessment. It is important to note that the assessment presented in the following sections is an assessment of the specific models and to some extent also of MDCEV models based on stated response data rather than an assessment of the MDCEV methodology per se.

The first model used for comparison is the Priority Evaluator (PE) model. The data set was collected in a survey of homeowners about investment in energy efficiency in housing and mobility, described in (Jäggi and Axhausen, 2010). The participants of this survey were presented with their annual energy consumption in an internet tool using the Priority Evaluator method by Hoinville (1977). The detailed energy consumption of every household was previously recorded in a paper and pen survey. The participants were asked to reduce the energy consumption of the household by selecting among different measures such as insulating the facade, replacing the windows, installing a heat pump, buy a more efficient car or fly less. Although the potential energy savings of every measures were pre-specified for each participant, it is assumed that the participants could determine the pattern of energy reduction well, thanks to the interactive nature of the internet tool. The continuous amount is the energy saved and the budget the total reduction. In this special form of MDCEV model, persons do not maximize utility by buying goods, but minimize damage by allocating bads. The number of alternatives in the model is 12, the number of observations is 197.

The second model used for comparison is the Induced Demand (ID) model. The data set was

collected in a survey about the reactions of travelers to changing travel times. The survey was described in Weis *et al.* (2010). The participants in this survey were confronted with significant changes in travel time (between -30 and + 90 minutes) to a previously reported, typical schedule. The survey was conducted using face-to-face interviews supported by computer software, so that the participants could adjust their activity schedules interactively. The adjustments are categorized in three alternatives: change of departure time, change of activity time and change of travel time. The continuous amount is time and the budget the total compensation of travel time stated by the respondents. The number of alternatives in the model is 3, the number of observations is 612.

4.2.2 Disaggregate Simulation Results

In this section, the results of the forecasting procedure for the above mentioned models are presented, compared and analyzed. The simulation results of the fleet choice model are eventually to be used to allocate car types to households and agents in the MATSim environment and thus are part of a wider transport modeling framework. In this context it is of importance to know about the accuracy of the predictions and to have indicators to evaluate and compare the models. The tools in this paper are hit ratio and the absolute and relative residuals, calculated as the differences between the forecast and the data used for the estimation. The predictions are also tested for their stability by comparing the results of forecasts with 10 or 50 repetitions.

The hit ratio gives the percentage of chosen alternatives in observed data, that are matched in the forecast (discrete choice only). The hit ratio is calculated for every repetition separately, and then the mean is taken. For example, if in the observed data mileage is allocated to alternatives A1, A2 and A3, and in the forecast to alternatives A3 and A4, then the hit ratio is 33.33% because one third of the observed choices are matched.

To calculate the residuals, the mean of the predicted mileage for every alternative is taken over all repetitions. The residuals are calculated with the formula

$$R_{abs} = \sum_{j=1}^K \frac{|t_j - \hat{t}_j|}{2} \quad (9)$$

while $\hat{t}_j$ is the predicted amount of mileage and t_j the observed value. The differences are divided by 2 because otherwise the mileage would show up twice in the residual, one time in the alternative it is falsely allocated and one time in the alternative it should have been allocated but was not.

Table 5: Model Simulation Overview

Model	Alt.	Sat.	Rep	Hit Ratio [%]			R_{rel} [%]		
				Mean	< 5%	> 50%	1. Quartile	Mean	3. Quart
Without PT	16	α	10	39.6	40.0	35.8	45.1	70.7	100.0
Without PT	16	α	50	39.6	39.0	37.8	43.4	70.5	100.0
Without PT	16	γ	50	36.8	41.1	34.2	43.1	69.3	100.0
With PT	17	α	10	25.3	11.1	1.2	57.6	71.9	90.0
With PT	17	α	50	25.3	8.3	1.4	57.6	71.3	87.6
With PT	17	γ	50	55.3	12.9	47.3	42.6	65.4	90.9
Ref. Model			50	8.7	14.3	0.0	85.4	88.9	92.9
PE Model	12	γ	50	81.8	0.0	96.9	41.4	54.8	66.0
PE Ref. Model			50	51.7	0.0	69.0	59.2	67.1	75.4
ID Model	3	γ	50	59.5	4.2	68.0	0.0	56.7	100.0
ID Ref. Model			50	58.7	10.5	54.7	31.9	65.9	100.0

Because the mileage budget differs among observations, in contrast to e.g. time-use models, only residuals relative to the total amount of annual mileage of the observation can be used for comparison, as shown in equation (10).

$$R_{rel} = \frac{R_{abs}}{\sum_{j=1}^K t_j} \quad (10)$$

This also allows a standardized comparison between different models with different budgets, as done the next section. Table 5 shows the hit ratio and the residuals for the forecasts of the two fleet choice models with different repetitions per observation as well as for different satiation parameters. The reference model of the last row of table 5 is derived from a forecast with all parameters set to zero. The same values are also calculated for the two comparison models, PE and ID. It is important to consider that these values should not be compared across models, but instead the improvement between reference and actual model.

What we can see is that despite the huge difference in model fit, the residuals of the initial model are not substantially smaller than the ones from the model including PT. The mean of R_{rel} between 70.5% and 72% means that only about 30% of the total mileage is correctly allocated

by the model forecast. Compared to a totally random model, which distributes only about 11% correctly this is still a fair improvement. One has to consider that chances of predicting the wrong alternative are relatively high with 16 alternatives in total. It is also clear that the number of repetition has not a big influence on the quality of the forecast.

In table 5 one can also see that the models using γ parameters have slightly smaller Residual than α models in the case of the standard model and a substantially smaller Residual as well as a substantially higher hit ratio in the model including PT. This is another reason why the paper focuses on γ models.

The hit ratios of the two comparison models are significantly higher, because less alternatives are involved. In the ID model with only three alternatives, even the reference model has a high hit ratio. This is also true for the PE model, where usually about three to 7 out of 12 alternatives are chosen. A overview over all models about the number of alternatives chosen and the percentage of corner solution is given in table 6. We can also see that for the standard fleet choice model, about 40% of all cases have a hit ratio of $< 5\%$ and can be considered as a total failure. On the other side, the relative residual is not determined by the number of alternatives and thus a better measures for comparisons. The best fleet choice model could achieve a reduction of 23.5% of mean relative residual compared to reference model. The PE model achieved a reduction of 12.3% and the ID model one of 9.2%. Whether this numbers can be considered as especially low or acceptable cannot be determined yet. But they give a first set for an assessment of further models.

Looking at the distribution of residuals supports the findings of the hit ratio. Figure 1 shows the distributions for reference model, the initial model and the PT model, all of them calculated using 50 repetitions.

The shapes of the distributions differ among the three models. The bars in the figure shows in each case how many predictions are made within a 2% range, e.g., in the reference model, in about 13% of the cases, the R_{rel} value is between 0.9 and 0.92. In this, totally random model, all of the forecasts are between 70% and 100% wrong, with the expected shape of a gumbel distribution. The estimated models have a totally different shape: Except the almost 40% completely wrong cases in the standard model (see hit ratio numbers) and the 16% of the model including PT, the relative residuals are fairly equally distributed. The big improvement of the PT model is the much smaller portion of completely false predictions Nevertheless, a share of 16% wrongly predicted observations is still high.

The shapes of the residual distributions of the comparison models differ significantly from the ones of the fleet choice models. While the ID model also has a large part of completely false forecasts (30%), the PE model has none. We assume that the portion of wrong forecasts depends heavily on the percentage of non-chosen alternatives (corner solutions). Table 6 gives

Figure 1: Distribution of Residuals: Fleet Choice Model

Figure 2: Distribution of Residuals: Priority Evaluator Model

an overview of this relationship for the presented models.

The more alternatives with zero consumption have to be predicted and thus are predicted, the higher is the portion of completely false predictions, for three out of the four models. That indicates that this relationship exists and that it has a substantial influence, but there are other factors contributing to the distribution of residuals. This makes sense, given that correctly predicting the sole chosen alternative out of 17 is particularly difficult, especially if the differences in baseline utility are small.

Another interesting point is whether there are substantial differences between the alternatives in

Figure 3: Distribution of Residuals: Induced Demand Model

Table 6: Relationship Corner Solutions - Forecast Accuracy

Model	Alt.	Alt. Chosen		Corner Solutions [%]		Completely False [%]
		Mean	Range	Data	Simulation	Simulation
FC Without PT	16	1.2	1 to 4	92.3	90.8	40.0
FC With PT	17	2.0	1 to 4	88.5	86.6	16.0
Induced Demand	3	1.6	1 to 3	48.4	56.0	30.0
Priority Evaluator	12	4.4	1 to 10	63.0	38.0	0.0

Figure 4: Relative Residual of Alternatives in the Fleet Choice Model incl. PT

terms of forecast quality. Figure 4 shows a box-plot of the relative residual of the alternatives.

For this figure only the residuals of observed chosen alternatives are used, because most of the residuals are zero as most alternatives have zero observed and zero predicted mileage (maximum 4 of 17 alternatives are chosen). The relative residual used in this figure is not exactly the same as in previous figures, but it is a valuable metric to make a comparison among the alternatives. Alternative fuel cars (A4, A13) as well as new small to middle sized diesel cars are more difficult to predict than other alternatives. These are exactly the car types that are expected to be switched to in the case of rising fuel prices. On the other hand, luxurious gasoline cars (A8, A17) are more often correctly predicted. The two most common alternatives, PT (A99) and used small gasoline car (A14) have also relatively low residuals. In the case of public transport, it is the qualitatively best predicted alternative, although its presence in the model causes a major change in the outcome.

In figure 5, the same analysis for the model without PT is shown.

Figure 5: Relative Residual of Alternatives in the Fleet Choice Model

One can see that the high percentage of totally wrong predicted cases can be explained by certain alternatives, that are much more difficult to forecast. Again, alternative technology (A4, A13) and small diesel cars (A1, A2, A10) are especially hard to predict. All other alternatives have similar residuals. The weakness of the model without public transport comes from its inability to predict these alternatives five alternatives. One reason could be that people choosing these types of cars are also frequent public transport users, and in the standard model a large part of their behavior is excluded.

Figure 6: Aggregate Shares of the Fleet Choice Model

4.3 Aggregate Simulation Results

This section contains a brief consideration of aggregate "market shares" of the car types from the fleet choice model. Figure 6 shows the market shares for the observed data, for the simulation and also for only the completely wrongly predicted cases. It is obvious, that alternatives with medium sized market shares are fairly well predicted, while alternatives with a high market share are over and alternatives with a low market share are under predicted. The high amount of completely false predictions stem to a good part from the highly over predicted alternatives.

5 Conclusion and Outlook

The results shown in this paper are very interesting yet fairly hard to judge because of the lack of comparable work in the literature. This is the first time that discrete continuous models of sophisticated stated adaption data are assessed in a disaggregate way and compared. It shows how models, established to predict changes in behavior in the case of high fuel prices, drastic energy laws or dramatic travel time changes, perform when implemented. The models and the forecasting procedure used for this paper are relatively new developments. However, the methodology is relatively easy to understand and very well documented and the needed software is freely available and we consider it as a very useful and promising technique in choice modeling. Because of the fact that the results of such model estimations are not particularly easy to interpret and an evaluation of the model fit as well as the suitability for implementation

are not obvious, an indicator set to analyze the residuals and compare different models is very useful. As this paper shows, the outcome of a forecast is on no account trivial or useless. The extreme differences in the distribution of the residuals and the unsatisfying overall accuracy show that such an analysis is needed not only to assess the specific models, but also to improve them in an iterative process. The analysis of residuals give much more insight to the usefulness of the models than the model fit parameters from the estimation. The differences in figure 5 also provide more detailed information on how the model may be improved, than is possibly be derived from the estimation results only. Therefore the presented simulation assessments are related to the presented models and can not be assigned to the methodology per se.

To give a concise assessment of the forecasting quality of stated preference models or even the MDCEV methodology, this analysis can be considered a first step. As a next step a different categorization for the fleet choice model will be tested and compared in terms of forecast accuracy. It would be also helpful to have models based on revealed preference data to asses in the same, disaggregate way and have more experience in judging the prediction quality. From an absolute view we think that the presented models are significantly better than randomness, but not as accurate as one could expect looking only at the estimation parameters and their significance. The models presented in this paper are all based on very complex, difficult and relatively new survey methods that require the participants to imagine very unfamiliar choice situations and thus have a much lower consistency than revealed preference data. Nevertheless, with a accurate assessment of the estimated models, they can give useful insights.

References

- Balmer, M. (2007) Travel demand modeling for multi-agent traffic simulations: Algorithms and systems, Ph.D. Thesis, ETH Zurich, Zurich, May 2007.
- Balmer, M., K. Meister, M. Rieser, K. Nagel and K. W. Axhausen (2008) Agent-based simulation of travel demand: Structure and computational performance of MATSim-T, paper presented at the *Innovations in Travel Modeling (ITM'08)*, Portland, June 2008.
- Beckx, C., T. A. Arentze, L. Int Panis, D. Janssens, J. Vankerkorn and G. Wets (2009) An integrated activity-based modelling framework to assess vehicle emissions: Approach and application, *Environment and Planning B*, **36** (6) 1086–1102.
- Bhat, C. R. (2005) A multiple discrete-continuous extreme value model: Formulation and application to discretionary time-use decisions, *Transportation Research Part B: Methodological*, **39** (8) 679–707.
- Bhat, C. R. (2008) The multiple discrete-continuous extreme value (mdcev) model: Role of utility function parameters, identification considerations, and model extensions, *Transportation Research Part B: Methodological*, **42** (3) 274–303.
- Bhat, C. R. (2011) Chandra bhat's web-page, webpage, February 2011, http://www.ce.utexas.edu/prof/bhat/FULL_PAPERS.htm.
- Bhat, C. R. and S. Sen (2006) Household vehicle type holdings and usage: An application of multiple discrete-continuous extreme value model, *Transportation Research Part B: Methodological*, **40** (1) 35–53.
- Bhat, C. R., S. Sen and N. Eluru (2009) The impact of demographics, built environment attributes, vehicle characteristics, and gasoline prices on household vehicle holdings and use, *Transportation Research Part B: Methodological*, **43** (1) 1–18.
- Bradley, M. A., J. L. Bowman and B. Griesenbeck (2010) SACSIM: An applied activity-based model system with fine-level spatial and temporal resolution, *Journal of Choice Modelling*, **3** (1) 5–31.
- Copperman, R. B. and C. R. Bhat (2007) An analysis of the determinants of children's weekend physical activity participation, *Transportation*, **34** (1) 67–87.
- Erath, A. and K. W. Axhausen (2010) Long term fuel price elasticity: Effects on mobility tool ownership and residential location choice, *Technical Report*, Swiss Federal Office of Energy (SFOE), Federal Office for the Environment (FOEN), IVT, ETH Zurich, Berne.
- Fröhlich, P., M. Tschopp and K. W. Axhausen (2006) Development of the accessibility of swiss municipalities: 1950 to 2000, *Raumforschung und Raumordnung*, **63** (6) 385–399.

- Halton, J. (1960) On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals, *Numerische Mathematik*, **2** (1) 84–90.
- Hoinville, G. (1977) The priority evaluator method, *Working Paper*, **3**, Social & Community Planning Research, London, September 1977.
- Jäggi, B. and K. W. Axhausen (2010) Surveying energy efficiency in housing and transport using a priority evaluator, *Working Paper*, **636**, IVT, ETH Zurich, Zurich.
- Kim, J., G. M. Allenby and P. E. Rossi (2002) Modeling consumer demand for variety, *Marketing Science*, **21** (3) 229–250.
- Meister, K., M. Rieser, F. Ciari, A. Horni, M. Balmer and K. W. Axhausen (2009) Anwendung eines agentenbasierten Modells der Verkehrsnachfrage auf die Schweiz, *Straßenverkehrstechnik*, **53** (5) 269–280.
- Pinjari, A. R. and C. R. Bhat (2010a) An efficient forecasting procedure for kuhn-tucker consumer demand model systems: Application to residential energy consumption analysis, *Technical Report*, University of South Florida, July 2010.
- Pinjari, A. R. and C. R. Bhat (2010b) A multiple discrete-continuous nested extreme value (mdcnev) model: Formulation and application to non-worker activity time-se and timing behavior on weekdays, *Transportation Research Part B: Methodological*, **44** (4) 562–583.
- Pinjari, A. R., C. R. Bhat and D. A. Hensher (2009) Residential self-selection effects in an activity time-use behavior model, *Transportation Research Part B: Methodological*, **43** (7) 729–748.
- Salvini, P. A. and E. J. Miller (2005) ILUTE: An operational prototype of a comprehensive microsimulation model of urban systems, *Networks and Spatial Economics*, **5** (2) 217–234.
- Sen, S. (2006) A joint multiple discrete continuous extreme value (mdcev) model and multinomial logit model (mnl) for examining vehicle type/vintage, make/model and usage decisions of the household, Ph.D. Thesis, University of Texas, Austin, August 2006.
- Simma, A. and K. W. Axhausen (2001a) Commitments and modal usage: An analysis of german and dutch panels, *Working Paper*, **98**, IVT, ETH Zurich, Zurich.
- Simma, A. and K. W. Axhausen (2001b) Successive days, related travel behaviour?, *Working Paper*, **62**, IVT, ETH Zurich, Zurich.
- Simma, A. and K. W. Axhausen (2003) Commitments and modal usage: Analysis of german and dutch panels, *Transportation Research Record*, **1854**, 22–31.
- Simma, A., R. Schlich and K. W. Axhausen (2002) Destination choice modelling for different leisure activities, *Working Paper*, **99**, IVT, ETH Zurich, Zurich.

- Spissu, E., A. R. Pinjari, C. R. Bhat, R. M. Pendyala and K. W. Axhausen (2009) An analysis of weekly out-of-home discretionary activity participation and time use behaviour, *Transportation*, **36** (5) 483–510.
- Swiss Federal Statistical Office (2006) *Ergebnisse des Mikrozensus 2005 zum Verkehrsverhalten*, Swiss Federal Statistical Office, Neuchatel.
- Train, K. E. (1978) A validation test of a disaggregate mode choice model, *Transportation Research*, **12** (3) 167–174.
- Weis, C., C. Dobler and K. W. Axhausen (2010) An interactive stated adaptation survey of activity scheduling decisions, *Working Paper*, **637**, IVT, ETH Zurich, Zurich.