Mode choice with attitudinal latent class: a Swiss case-study

Bilge Atasoy, Aurélie Glerum, Michel Bierlaire

Transport and mobility laboratory
EPFL

2^{nd} International Choice Modeling Conference, Leeds
July 6, 2011

Outline

Introduction

- The research is carried out with CarPostal: bus service in peri-urban and suburban areas of Switzerland.
- The aim of this study is to ...
  - understand the travel behavior in the area of interest,
  - come up with a latent segmentation of the population through attitudinal attributes,
  - design more efficient public transport policies specific to each segment,
  - improve the market share of public transport.

Data collection

This study is based on the data obtained by a revealed preferences (RP) survey:

- Travel diary for a predefined day
- Socio-economic characteristics
- Travel habits
- Psychometric indicators $\Rightarrow$ attitudes

Dataset consists of 2265 observations.

Representativity of the data - Federal Census 2000

Category		Sample	Population
Education	University	14.2%	6.2%
	Vocational university	16.2%	10.6%
	Vocational Training and Education	61.0%	50.9%
	Compulsory school	7.6%	27.6%
	No school diploma	1.0%	4.7%
Age	16-19 years	2.3%	8.2%
	20-39 years	21.2%	33.4%
	40-64 years	55.9%	41.6%
	65-79 years	18.7%	12.7%
	80 years and above	1.8%	4.1%
Gender	Male	53.0%	49.0%
	Female	47.0%	51.0%

- Individuals with high education, males and middle-aged respondents are oversampled.
- Observations are weighted for a better representativity of the data.

Factor analysis

Exploratory analysis on the potential segmentation of the population

	Factor 1	Factor 2
Choice PT		**0.250**
Socio-economics		
$N_{children}$	**0.517**	
Student/trainee	0.117	**0.770**
HighIncome	**0.252**	
Age $\geq$ 60	**-0.375**	
Couple without children	**-0.606**	
Couple with children	**0.927**	**-0.368**
Living with parents	0.159	**0.956**
Single	**-0.371**	
Psychometric indicators		
Ind1 - PT children		
Ind2 - Flexibility car		-0.130
Ind3 - Family oriented	0.135	

$\geq$ 0.1: presented, $\geq$ 0.25: highlighted

Segment 1: **Independent**

- Families with children
- Higher number of children
- High income
- Younger than 60 years old

Segment 2: **Dependent**

- Student/Trainee
- Living with parents
- Choice of public transport

Factor analysis

- Ind1: It's hard to take public transportation when I travel with my children.
- Ind2: With my car, I can go where I want when I want.
- Ind3: I would like to spend more time with my family and friends.

Psychometric Indicators	Factor 1	Factor 2
Ind1 - PT children		
Ind2 - Flexibility car		-0.130
Ind3 - Family oriented	0.135	

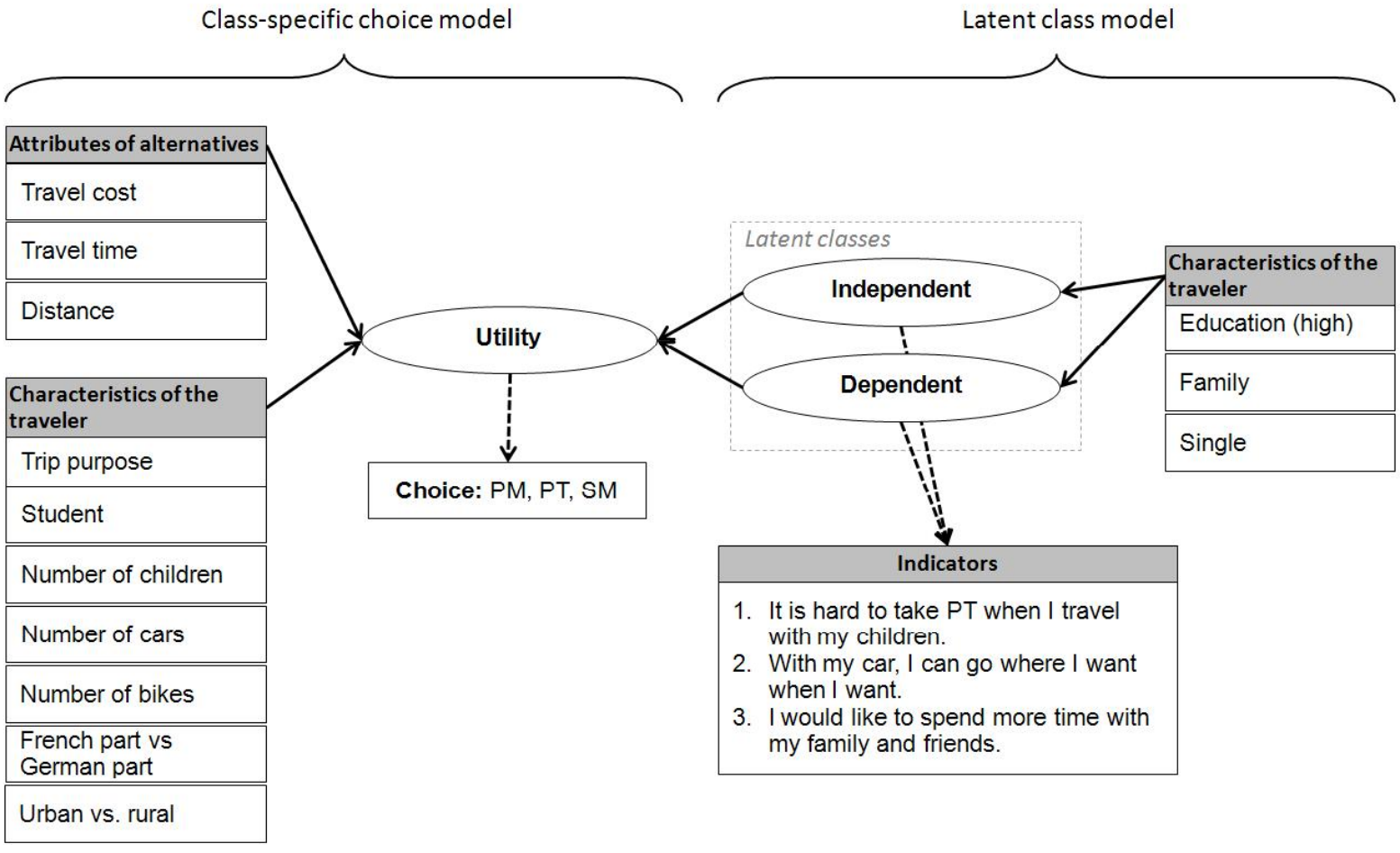
- Factor loadings are not high for psychometric indicators when estimated together with the socio-economic characteristics.
- However, since family status plays a key role, these indicators are found to be important in the model.

Latent class choice model

- Latent class model
  - Class membership model: Learnings from the factor analysis
  - Measurement model for psychometric indicators
- Class-specific choice models
  - Choice between the alternatives of:
    - Private mode (PM): car, taxi, car-sharing, motorbike
    - Public transport (PT): train, bus, metro etc.
    - Soft mode (SM): walking and bike
  - Explained by:
    - Attributes of the alternatives
    - Characteristics of the travelers

Soft mode alternative is not available for the latent class *dependent*.

Integrated model framework



Maximum likelihood estimation

- Maximum likelihood estimation is used for the simultaneous estimation of latent class model and the class-specific choice models. BIOGEME is used as a software.
- Likelihood is written over the joint probability of choice and responses to indicators:

$$L = \sum_n \sum_{i \in C_s} \log P(i_n, I_n | X_n, X_i; \beta, \gamma, \theta_\varepsilon)$$

where:

$$P(i_n, I_n | X_n, X_i; \beta, \gamma, \theta_\varepsilon) = \sum_{s \in S} P(i_n | X_n, X_i, s; \beta^s, \theta_\varepsilon^s) P(I_n | s) P(s | X_n; \gamma).$$

P(choice) P(indicator) P(class)

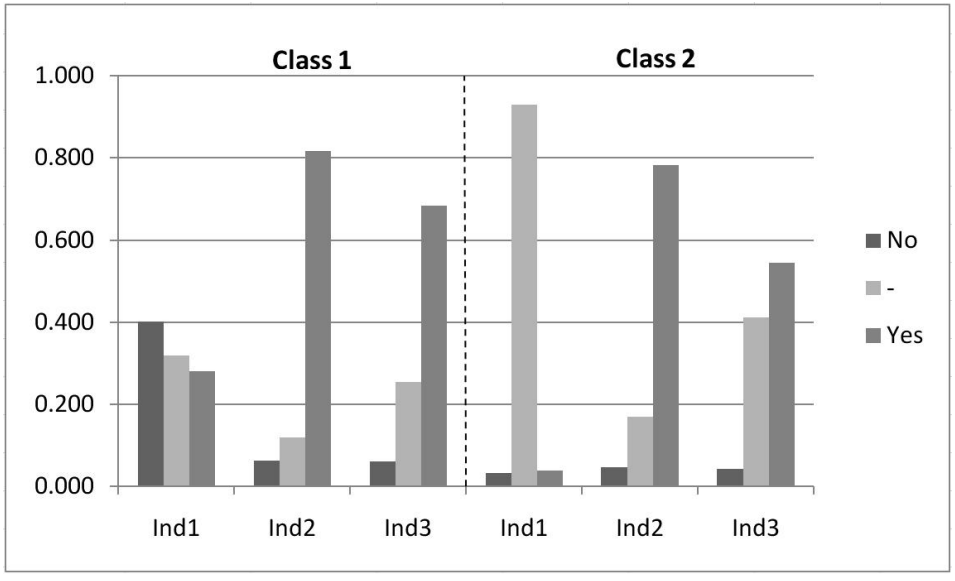
Estimation results - Class membership model

	Class		Estimated results	
Parameter	1- Independent	2- Dependent	Value	t-test
ASC_{in}	x		-0.642	-3.470
α_{family}	x		4.010	4.570
α_{income}	x		0.479	2.110
α_{single}		x	0.861	3.880

- All parameters are significant and have the expected signs.
- Families with children and individuals with high income have higher probability to be in Class 1.
- Individuals living single have higher probability to be in Class 2.

Estimation results - Psychometric indicators

- 5 levels from strong disagreement (1) to strong agreement (5).
- Probability of responding $(1 + 2) \Rightarrow$ *No*, $(4 + 5) \Rightarrow$ *Yes*



- Ind1: It's hard to take public transportation when I travel with my children.
- Ind2: With my car, I can go where I want when I want.
- Ind3: I would like to spend more time with my family and friends.

- Second class of people are neutral to the indicator related to children.
- First class has a higher probability to agree with the 2^{nd} and 3^{rd} statements due to their family situation.

Estimation results - Class-specific choice models

	Affected utility					Estimated results	
Parameter	V^1_{PM}	V^1_{PT}	V^1_{SM}	V^2_{PM}	V^2_{PT}	Value	t-test
ASC^1_{PM}	×					-1.100	-3.750
ASC^2_{PM}				×		0.274	1.210*
ASC^1_{SM}			×			0.529	1.320*
β^1_{cost}	×	×				-0.0229	-2.520
β^2_{cost}				×	×	-0.321	-4.720
$\beta^1_{TT_{PM}}$	×					-0.0159	-2.740
$\beta^2_{TT_{PM}}$				×		-0.108	-5.280
$\beta^1_{TT_{PT}}$		×				-0.00643	-2.190
$\beta^2_{TT_{PT}}$					×	-0.0433	-4.810
$\beta_{distance}$			×			-0.200	-3.680
$\beta_{N_{cars}}$	×					1.340	8.050
β_{urban}		×				0.366	2.200
β_{work}	×					-0.807	-4.970
$\beta_{language}$	×					1.320	6.500
$\beta_{N_{children}}$	×					0.395	4.800
$\beta_{N_{bikes}}$			×			0.206	3.460
$\beta_{student}$					×	3.820	5.100

(* Statistical significance $<$ 95%)

- All the parameters are significant and have the expected signs.
- Estimated parameters are used to make analysis on the market shares, value of time, elasticities of demand etc.

Value of time

Aggregate value of time (VOT) for alternative i is calculated as:

$$VOT_i = \sum_{s=1}^{S} P(s) VOT_i^s, \qquad \text{where } VOT_i^s = \frac{\beta_{time}^{i,s}}{\beta_{cost}^{i,s}}$$

and $P(s)$ is the aggregate class membership probability.

	VOT_{PM} (CHF/h)	VOT_{PT} (CHF/h)
Class 1: Independent	41.66	16.85
Class 2: Dependent	20.19	8.09
Aggregate	31.72	12.80

- First class of individuals have higher VOT compared to the second class.
- Individuals have higher VOT for private mode.
- At an aggregate level, travelers are ready to pay 32 CHF to decrease the travel time by one hour for private mode and 13 CHF for public transport.

Market shares

Estimated market share for alternative i is given by:

$$P_i = \frac{\sum_{n=1}^{N} \sum_{s=1}^{S} w_n P_n(s) \cdot P_n(i|s)}{\sum_{n=1}^{N} \sum_{s=1}^{S} w_n P_n(s)}$$

where

- w_n is the sample weight for individual n,
- $P_n(s)$ is the class membership probability for individual n,
- $P_n(i|s)$ is the choice probability for i given that n belongs to class s.

	PM	PT	SM
Class 1: Independent	61.23%	29.17%	9.60%
Class 2: Dependent	67.73%	32.27%	-
Aggregate	62.70%	32.35%	4.95%
Weighted sample	61.30%	33.95%	4.74%

Time and cost elasticities of demand

Aggregate elasticity of demand for alternative i is given by:

$$E_x^i = \frac{\sum_{n=1}^{N}\sum_{s=1}^{S} w_n P_n(i|s)\cdot P_n(s)\cdot E_{x_n}^{i,s}}{\sum_{n=1}^{N}\sum_{s=1}^{S} w_n P_n(i|s)\cdot P_n(s)}$$

where

- w_n is the sample weight for individual n,
- $P_n(s)$ is the class membership probability for individual n,
- $P_n(i|s)$ is the choice probability for i given that n belongs to class s,
- $E_{x_n}^{i,s}$ is the disaggregate elasticity of demand given by:

$$E_{x_n}^{i,s} = \frac{\partial P_n(i|s)}{\partial x_n}\frac{x_n}{P_n(i|s)}.$$

Time and cost elasticities of demand

% decrease in the market share when there is a 1% increase in time/cost.

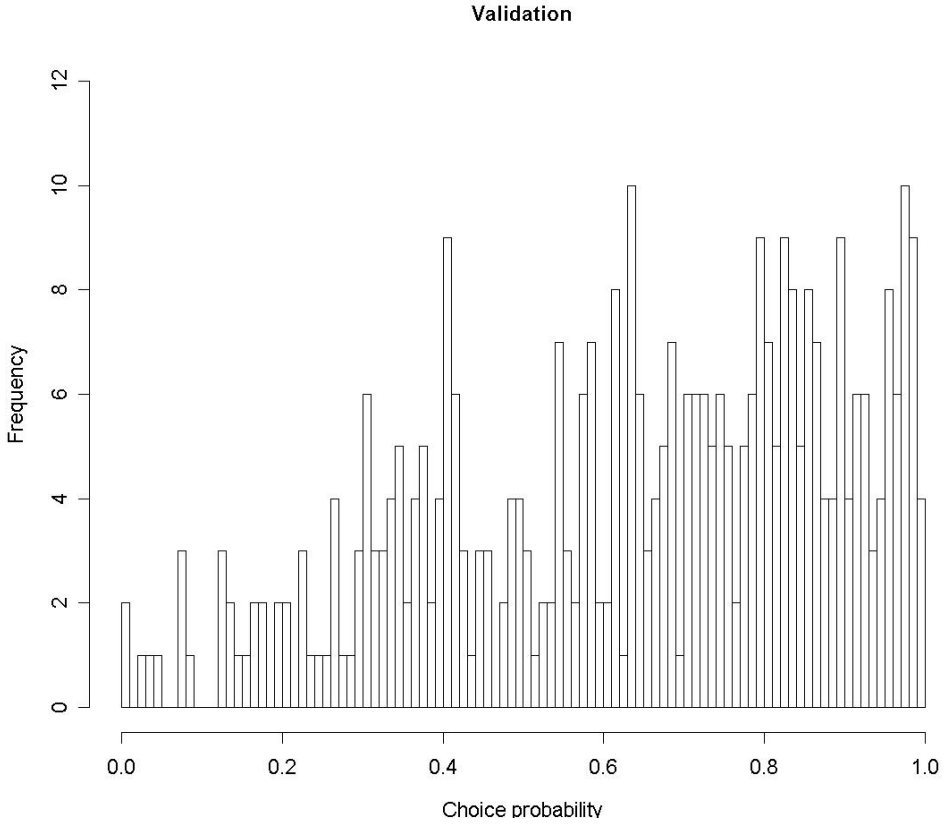
		Time elasticity	Cost elasticity
Class 1: Independent	PM	-0.17	-0.03
	PT	-0.34	-0.12
Class 2: Dependent	PM	-0.41	-0.15
	PT	-0.90	-0.48
Aggregate	PM	-0.28	-0.09
	PT	-0.59	-0.28

- Two segments of individuals have different elasticities:
  - *Independent* class has a lower elasticity
- Price elasticity is lower than the time elasticity.
- Elasticity is higher for public transport.

Validation

Validation is done by estimating the model on 80% of the data and predicting the remaining 20%.

- 70% of the estimated choice probabilities are above 0.5.
- 16% of the estimated choice probabilities are above 0.9.



Conclusions

- Attitudes enable us to understand the travel behavior more deeply.
- Identified latent classes have different choice sets and taste parameters.
  - Class 1: Independent $\Rightarrow$ more flexible alternatives should be offered due to the family situation.
  - Class 2: Dependent $\Rightarrow$ price reduction for students and old people as in the existing offers.

As future research...

- Analysis with more than two classes is expected to enable to design more specific offers to each segment of the population.
- Analysis by including more psychometric indicators is expected to help to better identify the segments.

Thank you for your attention !