

INTEGRATED CHOICE AND LATENT VARIABLE MODELS

APPLICATIONS TO VEHICLE AND MODE CHOICE MODELING

Aurélie Glerum, EPFL

IRE seminar, USI
10th October 2013

A HYBRID CHOICE MODEL TO FORECAST THE DEMAND FOR ELECTRIC VEHICLES

FROM SURVEY DESIGN TO MODEL APPLICATION

Aurélie Glerum
Lidija Stankovikj
Michaël Thémans
Michel Bierlaire

Introduction & motivation

Data collection

Methodology

Model

- SP model
- Choice model for forecasting

Forecasting analysis

Conclusion

Aim

- Develop a **comprehensive methodology** to forecast demand for a new technology: **electric vehicles**

Context

- **Current situation:**
 - Alternative fuel vehicles (LPG, CNG, etc.) on the car market
 - Electric vehicles (EV) being released
- **Collaborative project** EPFL-Renault Suisse:
 - Renault has launched Zero Emission (Z.E.) product line in 2011-2013
 - **Aim:** analyze demand for two EV models for **private use**



Zoé



Fluence Z.E.

Literature

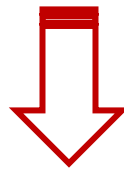
- **SP survey design:**
 - Personalized choice situations (Bunch et al., 1993, Achtnicht et al., 2008, etc.)
 - Fractional factorial designs (Brownstone et al., 1996, Ewing and Sarigöllü, 2000, Horne et al., 2005)
- **Choice models for demand for EVs or alternative-fuel vehicles:**
 - Widely applied (Brownstone and Train, 1999, Dagsvik et al., 2002, Mueller and de Haan, 2009, etc.)
 - Integrated choice and latent variable (ICLV) models for environmental concern (Alvarez-Daziano and Bolduc, 2009)
- **Model application:**
 - Models developed on SP data need adjustments before application (Brownstone et al., 1996)
 - Joint RP-SP estimations (e.g. Brownstone et al., 2000)
 - Lack of examples of applications of models designed to evaluate demand for new alternatives (Daly and Rohr, 1998)

Main features of the model

- Customized choice situations using iterative proportional fitting (IPF)
- Include attitudinal dimensions
- Specify model for the whole market, from a model based on SP data

Main features of the model

- Customized choice situations using iterative proportional fitting (IPF)
- Include attitudinal dimensions
- Specify model for the whole market, from a model based on SP data



**COMPREHENSIVE
FRAMEWORK**

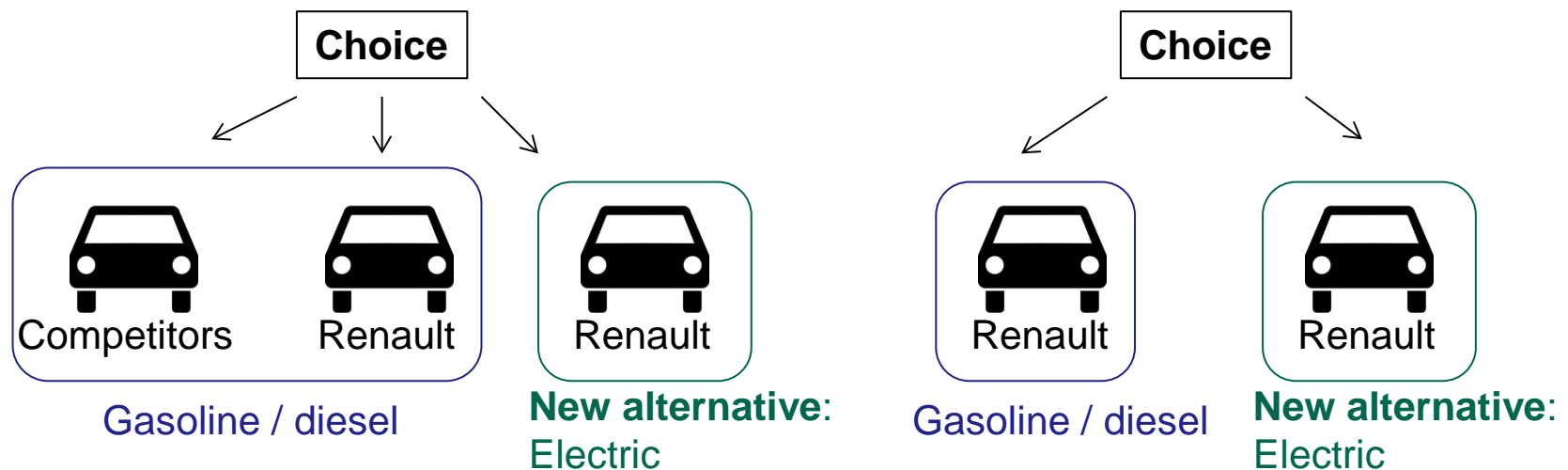
Type of survey: **stated preference (SP)** survey

Within same car segment: hypothetical choices between

Own car

Renault – gasoline (if own car is not Renault)

Renault – electric



STRUCTURE OF THE SURVEY

2 phases:

Phase I:

- Characteristics of respondent's car(s)
- Socio-economic information
- Mobility habits

Phase II:

- Choice situations
- Opinions on topics related to EV
- Perceptions of four categories of vehicles

STRUCTURE OF THE SURVEY

2 phases:

Phase I:

Characteristics of respondent's car(s)

Socio-economic information

Mobility habits



Phase II:

Choice situations

Opinions on topics related to EV

Perceptions of four categories of vehicles

...used to design...



STRUCTURE OF THE SURVEY

Opinions on themes related to electric vehicles

- **Environmental concern** (5 statements)
Example: An electric car is a 100% ecological solution.
- **Attitude towards new technologies** (5 statements)
Example: A control screen is essential in my use of a car.
- **Perception of the reliability of an electric vehicle** (5 statements)
Example: Electric cars are not as secure as gasoline cars.
- **Perception of leasing** (5 statements)
Example: Leasing is an optimal contract which allows me to change car frequently.
- **Attitude towards design** (5 statements)
Example: Design is a secondary element when purchasing a car, which is above all a practical transport mode.

Ratings

- Total disagreement (1)
- Disagreement (2)
- Neutral opinion (3)
- Agreement (4)
- Total agreement (5)
- I don't know (6)

5 types of respondents sampled in Switzerland:

- Recent buyers
- Prospective buyers
- Renault customers
- Pre-orders
- Newsletter

5 types of respondents sampled in Switzerland:

- Recent buyers
 - Prospective buyers
 - Renault customers
 - Pre-orders
 - Newsletter
- } **Sampling protocol**
- } All available

Sampling protocol → representativity from:

- 3 language regions of Switzerland (German, French, Italian)
- Gender
- Age category (18-35 years, 36-55 years, 56-74 years)

GfK

Ask GfK

0% 25% 50% 75% 100%

Situation de choix 4 de 5

Vous avez ici la description de votre véhicule actuel ainsi que celle de véhicules similaires, thermique et électrique, de la marque Renault. Compte tenu des caractéristiques de chacun de ceux-ci, laquelle des trois solutions choisiriez-vous, si vous deviez changer de voiture aujourd'hui ?

Les valeurs indicatives de leasing sont calculées sur la base d'un apport initial de 20%, d'un kilométrage annuel de 10'000 km et d'une durée de financement de 48 mois.

Caractéristiques	Votre véhicule	Véhicule thermique Renault	Véhicule électrique Renault
Marque	SEAT	RENAULT	RENAULT
Modèle	LEON	MEGANE	FLUENCE
Carburant	Diesel	Diesel	Electricité
Prix d'achat (en CHF)	37510	42739	34008
Prime du gouvernement (en CHF)	0	0	0
Prix total à l'achat (en CHF)	37510	42739	34008
OU : Prix mensuel du leasing (en CHF)	402	435	404
Coûts d'entretien (en CHF par 30'000 km)	850	850	425
Coût en carburant/électricité par 100 km (en CHF)	9.65	10.8	3.55
Leasing de la batterie (en CHF par mois)	0	0	105

précédent

suivant

An example of choice experiment

Reported by
respondent

Characteristics	Your vehicle	Renault vehicle with combustion engine	Renault electric vehicle
Make	Audi	Renault	Renault
Model	A4	Laguna	Fluence
Fuel	Petrol	Petrol	Electricity
Purchase price (in CHF)	42'400	37'200	56'880
Incentive (in CHF)	0	0	-1'000
Total purchase price (in CHF)	42'400	37'200	55'880
OR: Monthly leasing price (in CHF)	477	399	693
Maintenance costs (in CHF for 30'000 km)	850	850	425
Cost in fuel/electricity for 100 km (in CHF)	11.70	13.55	3.55
Battery lease (in CHF per month)	0	0	125
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

An example of choice experiment

Characteristics	Your vehicle	Renault vehicle with combustion engine	Renault electric vehicle
Make	Audi	Renault	Renault
Model	A4	Laguna	Fluence
Fuel	Petrol	Petrol	Electricity
Purchase price (in CHF)	42'400	37'200	56'880
Incentive (in CHF)	0	0	-1'000
Total purchase price (in CHF)	42'400	37'200	55'880
OR: Monthly leasing price (in CHF)	477	399	693
Maintenance costs (in CHF for 30'000 km)	850	850	425
Cost in fuel/electricity for 100 km (in CHF)	11.70	13.55	3.55
Battery lease (in CHF per month)	0	0	125

Deduced from segment of owned car

An example of choice experiment

Characteristics	Your vehicle	Renault vehicle with combustion engine	Renault electric vehicle
Make	Audi	Renault	Renault
Model	A4	Laguna	Fluence
Fuel	Petrol	Petrol	Electricity
Purchase price (in CHF)	42'400	37'200	56'880
Incentive (in CHF)	0	0	-1'000
Total purchase price (in CHF)	42'400	37'200	55'880
OR: Monthly leasing price (in CHF)	477	399	693
Maintenance costs (in CHF for 30'000 km)	850	850	425
Cost in fuel/electricity for 100 km (in CHF)	11.70	13.55	3.55
Battery lease (in CHF per month)	0	0	125

Obtained from data base of cars currently sold on market

An example of choice experiment

Characteristics	Your vehicle	Renault vehicle with combustion engine	Renault electric vehicle
Make	Audi	Renault	Renault
Model	A4	Laguna	Fluence
Fuel	Petrol	Petrol	Electricity
Purchase price (in CHF)	42'400	37'200	56'880
Incentive (in CHF)	0	0	-1'000
Total purchase price (in CHF)	42'400	37'200	55'880
OR: Monthly leasing price (in CHF)	477	399	693
Maintenance costs (in CHF for 30'000 km)	850	850	425
Cost in fuel/electricity for 100 km (in CHF)	11.70	13.55	3.55
Battery lease (in CHF per month)	0	0	125

Fixed attributes

An example of choice experiment

Characteristics	Your vehicle	Renault vehicle with combustion engine	Renault electric vehicle
Make	Audi	Renault	Renault
Model	A4	Laguna	Fluence
Fuel	Petrol	Petrol	Electricity
Purchase price (in CHF)	42'400	37'200	56'880
Incentive (in CHF)	0	0	-1'000
Total purchase price (in CHF)	42'400	37'200	55'880
OR: Monthly leasing price (in CHF)	477	399	693
Maintenance costs (in CHF for 30'000 km)	850	850	425
Cost in fuel/electricity for 100 km (in CHF)	11.70	13.55	3.55
Battery lease (in CHF per month)	0	0	125

Design variables

Design variables

EV variable	Level 1	Level 2	Level 3	Level 4
Purchase price	$(P_{\text{own}} + 5'000) * 0.8$	$(P_{\text{own}} + 5'000) * 1$	$(P_{\text{own}} + 5'000) * 1.2$	-
Governmental incentive	- 0 CHF	- 500 CHF	- 1'000 CHF	- 5'000 CHF
Cost of fuel/electricity for 100 km	1.70 CHF	3.55 CHF	5.40 CHF	-
Battery lease	85 CHF	105 CHF	125 CHF	-

EXPERIMENTAL DESIGN

Fractional factorial design with sampling weights

Fractional factorial design

- Orthogonal
- Size = 64 (full factorial design has size 108)

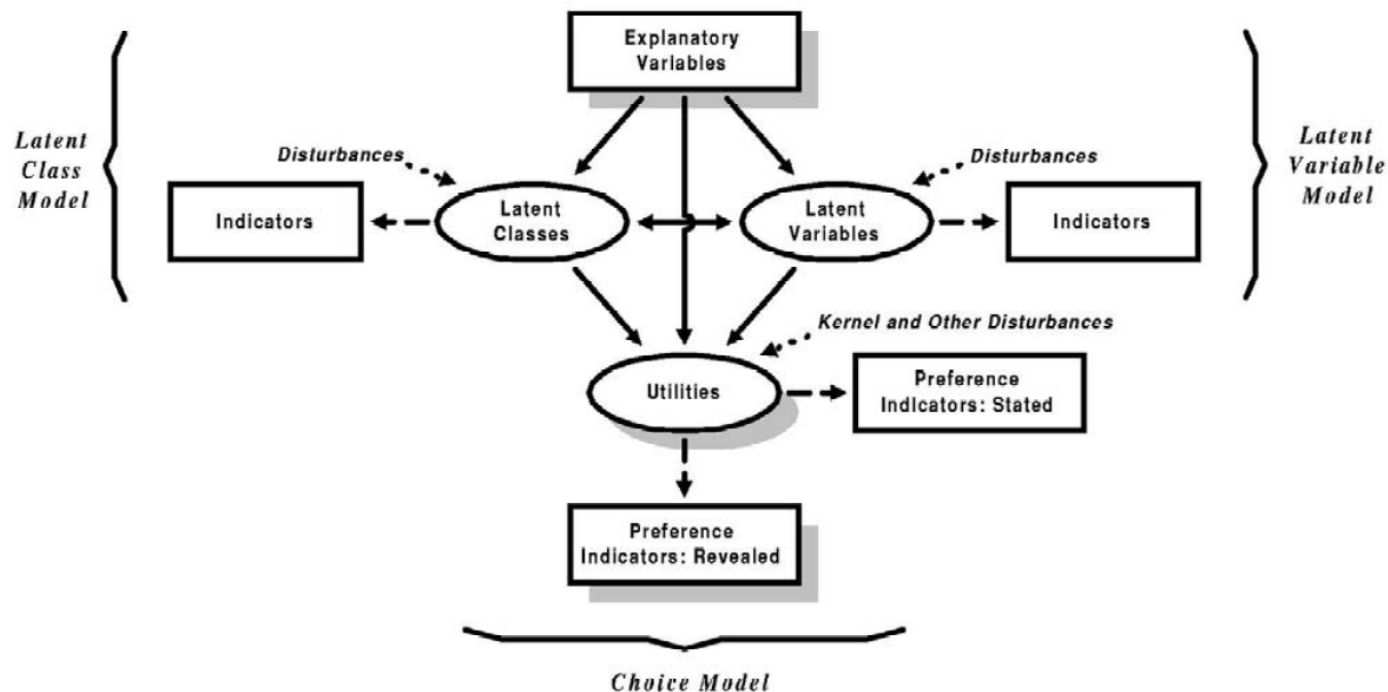
Sampling weights:

- Correct for **oversampling** of some levels
- Weights computed with **iterative proportional fitting (IPF)**

	Incentive	Price	Fuel cost of 100 km	Battery lease
1	0	0.80	1.70	85
2	0	1.00	3.55	125
3	0	1.00	5.40	105
4	0	1.20	3.55	105
5	-500	0.80	1.70	125
6	-500	1.00	3.55	85
7	-500	1.00	5.40	105
8	-500	1.20	3.55	105
9	-1000	0.80	3.55	105
10	-1000	1.00	5.40	105
11	-1000	1.00	3.55	85
12	-1000	1.20	1.70	125
13	-5000	0.80	3.55	105
14	-5000	1.00	5.40	105
15	-5000	1.00	3.55	125
16	-5000	1.20	1.70	85

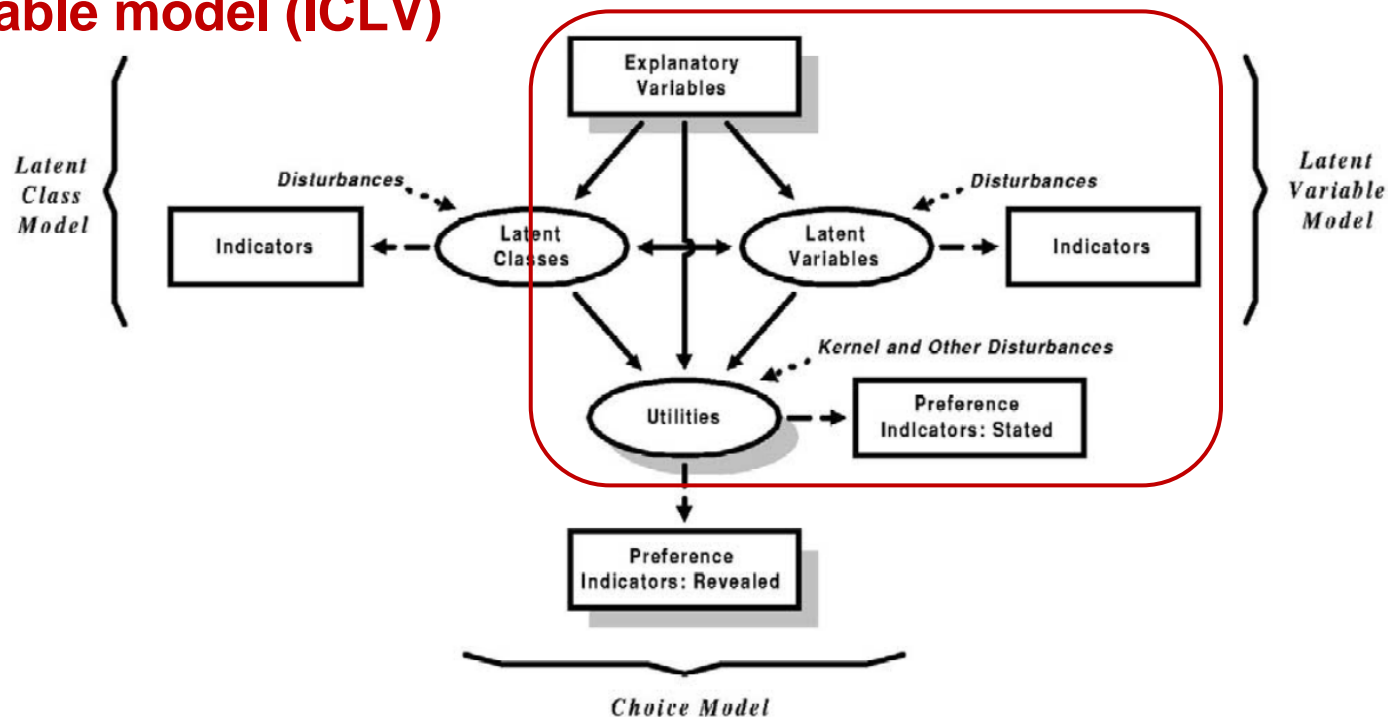
Hybrid choice model (HCM): DCM with latent constructs.

Allows to capture e.g. **attitudes et perceptions**



Hybrid choice model (HCM): DCM with latent constructs.

In this research: focus on the **integration of choice model and latent variable model (ICLV)**



Hybrid choice model specification

Structural equations:

Choice model:

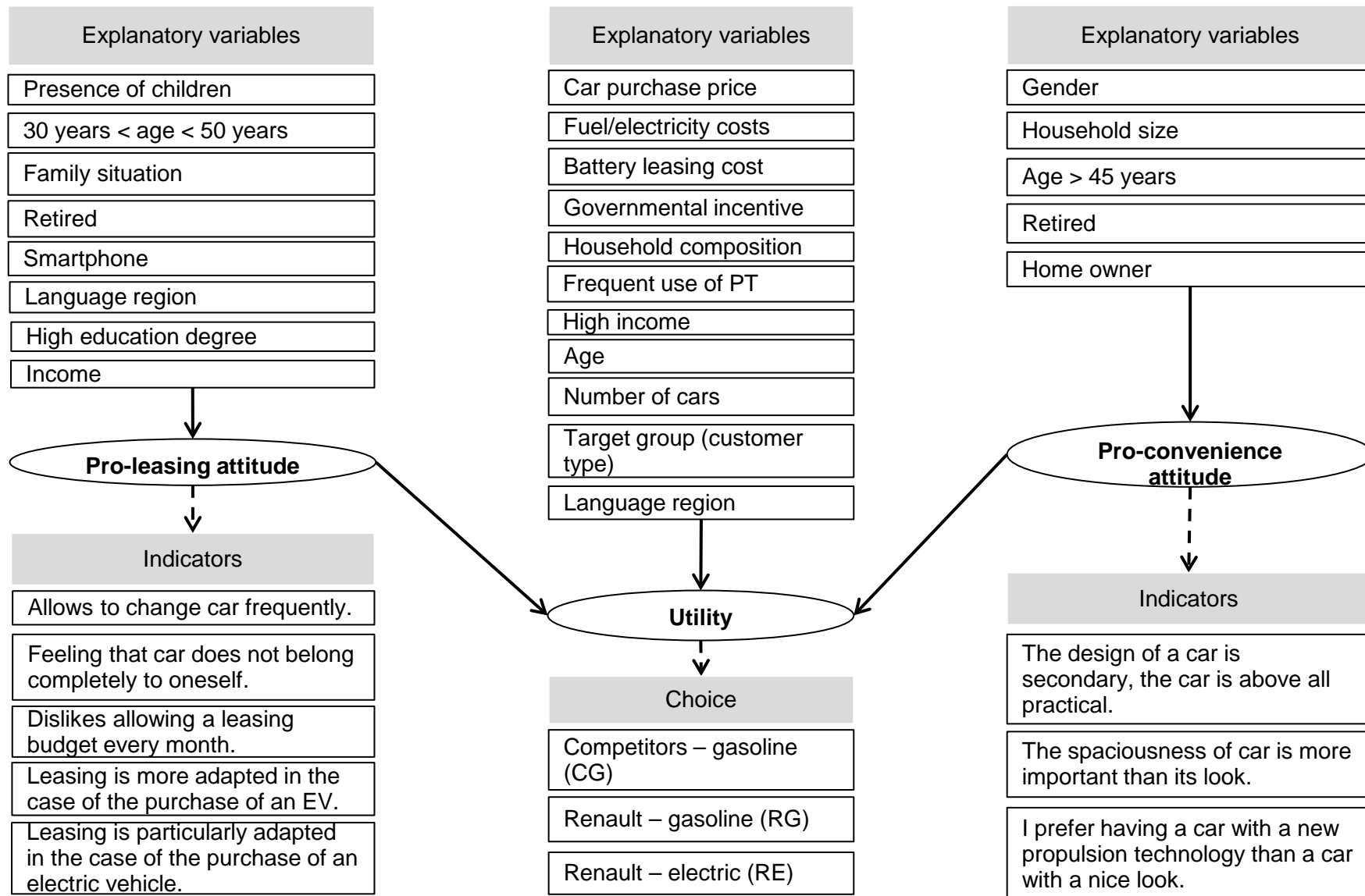
$$U_{in} = V(X_{in}, X_n^*; \beta) + \varepsilon_{in} \quad \text{with} \quad \varepsilon_{in} \sim EV(0,1)$$

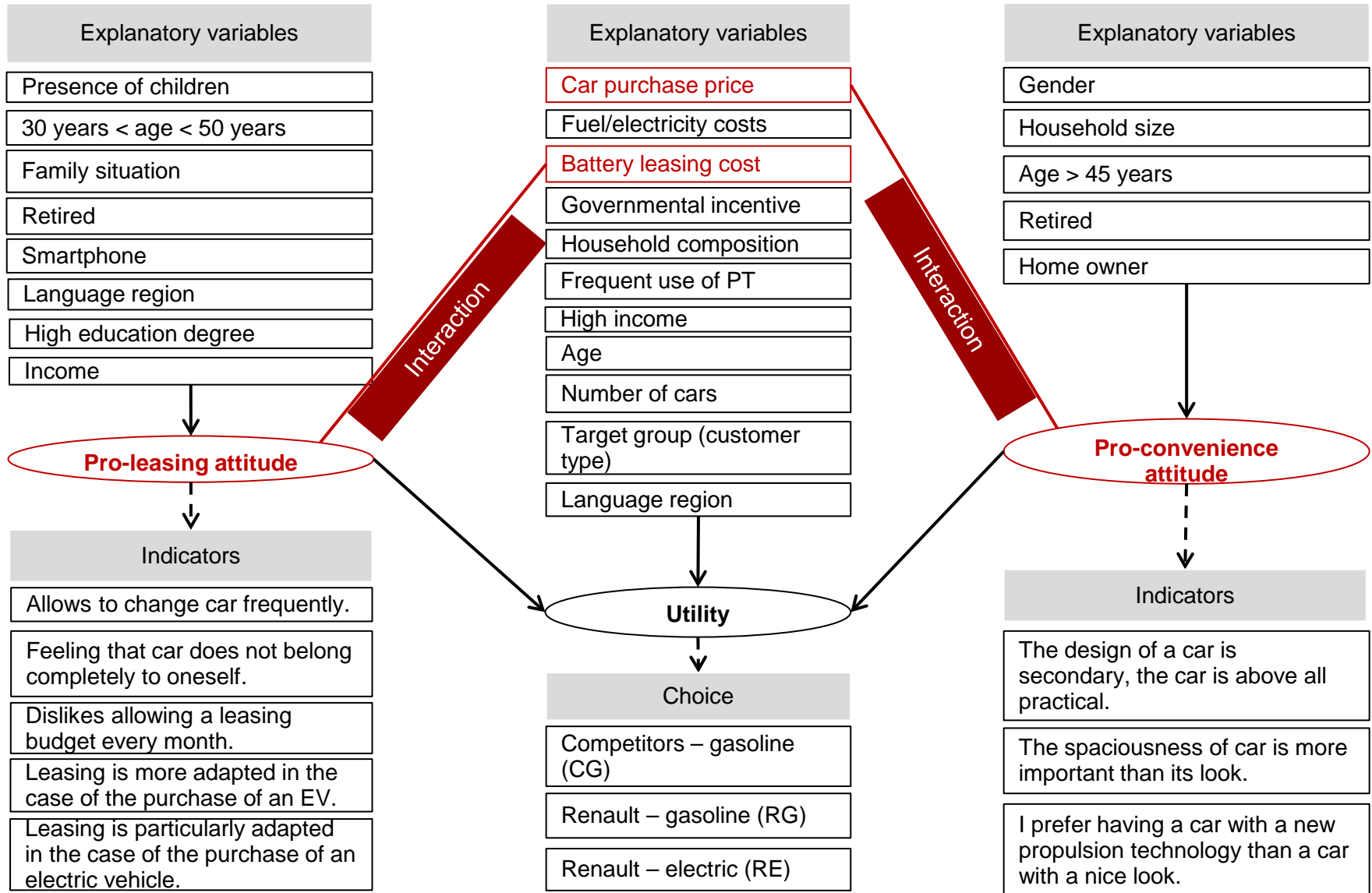
Latent variable model:

$$X_n^* = h(X_{in}; \lambda) + \omega_n \quad \text{with} \quad \omega_n \sim N(0, \sigma_\omega)$$

Measurement equations (continuous):

$$I_n^* = m(X_n^*; \alpha) + \nu_n \quad \text{with} \quad \nu_n \sim N(0, \sigma_\nu)$$





Structural equations:

Choice model:

$$\begin{aligned}
 U_{CG} &= -\exp(\beta_{price_{CG}} + \beta_{AttC} AttC) price_{CG} + \sum_k \beta_k X_k + \varepsilon_{CG,n} \\
 U_{RG} &= -\exp(\beta_{price_{RG,TG1245}} TG1245 + \beta_{price_{RG,TG3}} TG3 + \beta_{AttC} AttC) price_{RG} + \sum_l \beta_l X_l + \varepsilon_{RG,n} \\
 U_{RE} &= -\exp(\beta_{price_{RE,TG12}} TG12 + \beta_{price_{RG,TG3}} TG3 + \beta_{price_{RG,TG45}} TG45 + \beta_{AttC} AttC) price_{RE} \\
 &\quad - \exp(\beta_{Battery} + \beta_{AttL} AttL) Battery + \sum_m \beta_m X_m + \varepsilon_{RE,n} \quad \text{with } \varepsilon_{in} \sim EV(0,1)
 \end{aligned}$$

Latent variable model:

$$\begin{aligned}
 AttL &= \beta_{Mean1} + \sum_i \beta_{1,i} \cdot X_{1,i} + \exp(\nu_1) \cdot \Omega_1 \quad \text{with } \Omega_1 \sim N(0,1) \\
 AttC &= \beta_{Mean2} + \sum_i \beta_{2,i} \cdot X_{2,i} + \exp(\nu_2) \cdot \Omega_2 \quad \text{with } \Omega_2 \sim N(0,1)
 \end{aligned}$$

Measurement equations (continuous):

$$\begin{aligned}
 I_{1,k} &= \alpha_{1,k} + \lambda_{1,k} \cdot AttL + \exp(\sigma_{1,k}) \Omega_{1,k} \quad \text{with } \Omega_{1,k} \sim N(0,1), \text{ for } k = 1, \dots, 5 \\
 I_{2,k} &= \alpha_{2,k} + \lambda_{2,k} \cdot AttC + \exp(\sigma_{2,k}) \Omega_{2,k} \quad \text{with } \Omega_{2,k} \sim N(0,1), \text{ for } k = 1, 2, 3
 \end{aligned}$$

ESTIMATION RESULTS

Name	Value	t-test	Name	Value	t-test
<i>Parameters in linear terms</i>			<i>Parameters in linear terms (ctd)</i>		
ASC_{CG}	-2.71	-4.77	$\beta_{Income_{CG}}$	-0.223*	-1.92
ASC_{RG}	-2.17	-3.63	$\beta_{Income_{RG}}$	-0.259	-2.25
$\beta_{UseCostGasoline}$	-0.0469**	-1.41	$\beta_{French_{CG}}$	0.373	2.94
$\beta_{UseCostElecHighFluence}$	-0.264	-2.20	$\beta_{French_{RG}}$	0.0254**	0.19
$\beta_{UseCostElecHighZoé}$	-0.802	-4.82	$\beta_{Age_{CG}}$	0.0172	3.65
$\beta_{UseCostElecMedZoé}$	-0.514	-3.21	$\beta_{Age_{RG}}$	-0.00210**	-0.43
$\beta_{IncentiveHigh}$	0.799	6.21	$\beta_{TG12_{CG}}$	1.60	4.57
$\beta_{IncentiveMed}$	0.0538**	0.40	$\beta_{TG12_{RG}}$	0.664*	1.89
$\beta_{IncentiveLow}$	0.0164**	0.12	$\beta_{TG3_{CG}}$	0.104**	0.11
$\beta_{PT_{CG,TG1245}}$	-0.259	-1.96	$\beta_{TG3_{RG}}$	2.63	5.18
$\beta_{PT_{RG,TG1245}}$	-0.577	-3.67	<i>Parameters in non-linear terms</i>		
$\beta_{PT_{CG,TG3}}$	-2.64	-3.85	$\beta_{price_{CG}}$	-3.60	-4.77
$\beta_{PT_{RG,TG3}}$	-1.17	-4.40	$\beta_{price_{RG,TG1245}}$	-1.39	-4.33
$\beta_{FamSit_{CG}}$	-0.157**	-1.37	$\beta_{price_{RG,TG3}}$	-0.290**	-1.06
$\beta_{FamSit_{RG}}$	0.183**	1.56	$\beta_{price_{RE,TG12}}$	-0.365	-2.57
$\beta_{NbCars_{CG,TG1245}}$	-0.207	-2.75	$\beta_{price_{RE,TG3}}$	0.342	2.10
$\beta_{NbCars_{RG,TG1245}}$	-0.193	-2.32	$\beta_{price_{RE,TG45}}$	-0.152**	-1.33
$\beta_{NbCars_{CG,TG3}}$	-0.664*	-1.88	β_{AttC}	-0.142	-4.93
$\beta_{NbCars_{RG,TG3}}$	-0.945	-6.24	$\beta_{Battery}$	2.17	5.87
			β_{AttL}	-0.193*	-1.74

ESTIMATION RESULTS

Name	Value	t-test	Name	Value	t-test
<i>Parameters in linear terms</i>			<i>Parameters in linear terms (ctd)</i>		
ASC_{CG}	-2.71	-4.77	$\beta_{Income_{CG}}$	-0.223*	-1.92
ASC_{RG}	-2.17	-3.63	$\beta_{Income_{RG}}$	-0.259	-2.25
$\beta_{UseCostGasoline}$	-0.0469**	-1.41	$\beta_{French_{CG}}$	0.373	2.94
$\beta_{UseCostElecHighFluence}$	-0.264	-2.20	$\beta_{French_{RG}}$	0.0254**	0.19
$\beta_{UseCostElecHighZoé}$	-0.802	-4.82	$\beta_{Age_{CG}}$	0.0172	3.65
$\beta_{UseCostElecMedZoé}$	-0.514	-3.21	$\beta_{Age_{RG}}$	-0.00210**	-0.43
$\beta_{IncentiveHigh}$	0.799	6.21	$\beta_{TG12_{CG}}$	1.60	4.57
$\beta_{IncentiveMed}$	0.0538**	0.40	$\beta_{TG12_{RG}}$	0.664*	1.89
$\beta_{IncentiveLow}$	0.0164**	0.12	$\beta_{TG3_{CG}}$	0.104**	0.11
$\beta_{PT_{CG,TG1245}}$	-0.259	-1.96	$\beta_{TG3_{RG}}$	2.63	5.18
$\beta_{PT_{RG,TG1245}}$	-0.577	-3.67	<i>Parameters in non-linear terms</i>		
$\beta_{PT_{CG,TG3}}$	-2.64	-3.85	$\beta_{price_{CG}}$	-3.60	-4.77
$\beta_{PT_{RG,TG3}}$	-1.17	-4.40	$\beta_{price_{RG,TG1245}}$	-1.39	-4.33
$\beta_{FamSit_{CG}}$	-0.157**	-1.37	$\beta_{price_{RG,TG3}}$	-0.290**	-1.06
$\beta_{FamSit_{RG}}$	0.183**	1.56	$\beta_{price_{RE,TG12}}$	-0.365	-2.57
$\beta_{NbCars_{CG,TG1245}}$	-0.207	-2.75	$\beta_{price_{RE,TG3}}$	0.342	2.10
$\beta_{NbCars_{RG,TG1245}}$	-0.193	-2.32	$\beta_{price_{RE,TG45}}$	-0.152**	-1.33
$\beta_{NbCars_{CG,TG3}}$	-0.664*	-1.88	β_{AttC}	-0.142	-4.93
$\beta_{NbCars_{RG,TG3}}$	-0.945	-6.24	$\beta_{Battery}$	2.17	5.87
			β_{AttL}	-0.193*	-1.74

- $\beta_{AttC} < 0$ and significant:
pro-convenience individuals
less price-sensitive

Name	Value	t-test	Name	Value	t-test
<i>Parameters in linear terms</i>			<i>Parameters in linear terms (ctd)</i>		
ASC_{CG}	-2.71	-4.77	$\beta_{Income_{CG}}$	-0.223*	-1.92
ASC_{RG}	-2.17	-3.63	$\beta_{Income_{RG}}$	-0.259	-2.25
$\beta_{UseCostGasoline}$	-0.0469**	-1.41	$\beta_{French_{CG}}$	0.373	2.94
$\beta_{UseCostElecHighFluence}$	-0.264	-2.20	$\beta_{French_{RG}}$	0.0254**	0.19
$\beta_{UseCostElecHighZoé}$	-0.802	-4.82	$\beta_{Age_{CG}}$	0.0172	3.65
$\beta_{UseCostElecMedZoé}$	-0.514	-3.21	$\beta_{Age_{RG}}$	-0.00210**	-0.43
$\beta_{IncentiveHigh}$	0.799	6.21	$\beta_{TG12_{CG}}$	1.60	4.57
$\beta_{IncentiveMed}$	0.0538**	0.40	$\beta_{TG12_{RG}}$	0.664*	1.89
$\beta_{IncentiveLow}$	0.0164**	0.12	$\beta_{TG3_{CG}}$	0.104**	0.11
$\beta_{PT_{CG,TG1245}}$	-0.259	-1.96	$\beta_{TG3_{RG}}$	2.63	5.18
$\beta_{PT_{RG,TG1245}}$	-0.577	-3.67	<i>Parameters in non-linear terms</i>		
$\beta_{PT_{CG,TG3}}$	-2.64	-3.85	$\beta_{price_{CG}}$	-3.60	-4.77
$\beta_{PT_{RG,TG3}}$	-1.17	-4.40	$\beta_{price_{RG,TG1245}}$	-1.39	-4.33
$\beta_{FamSit_{CG}}$	-0.157**	-1.37	$\beta_{price_{RG,TG3}}$	-0.290**	-1.06
$\beta_{FamSit_{RG}}$	0.183**	1.56	$\beta_{price_{RE,TG12}}$	-0.365	-2.57
$\beta_{NbCars_{CG,TG1245}}$	-0.207	-2.75	$\beta_{price_{RE,TG3}}$	0.342	2.10
$\beta_{NbCars_{RG,TG1245}}$	-0.193	-2.32	$\beta_{price_{RE,TG45}}$	-0.152**	-1.33
$\beta_{NbCars_{CG,TG3}}$	-0.664*	-1.88	β_{AttC}	-0.142	-4.93
$\beta_{NbCars_{RG,TG3}}$	-0.945	-6.24	$\beta_{Battery}$	2.17	5.87
			β_{AttL}	-0.193*	-1.74

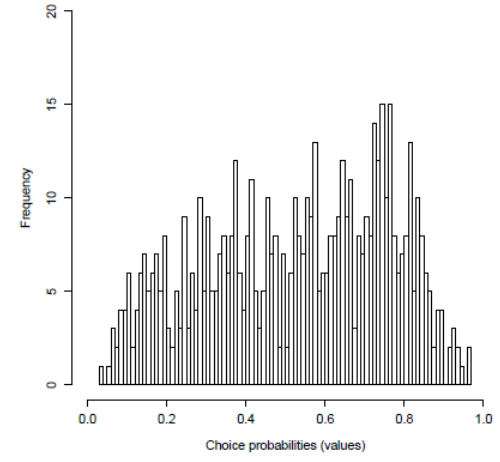
ESTIMATION RESULTS

- $\beta_{AttC} < 0$ and significant:
pro-convenience individuals
less price-sensitive
- $\beta_{AttL} < 0$ and significant:
pro-leasing individuals less
affected by changes in battery
leasing price

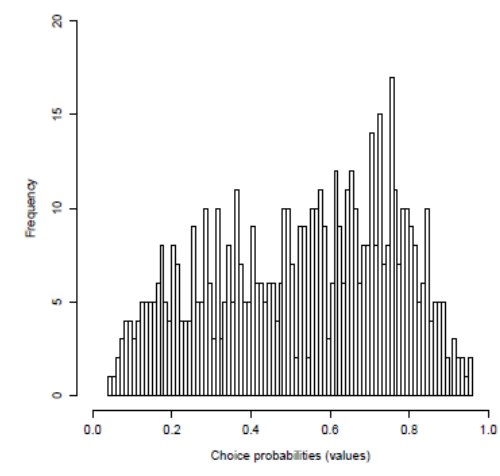
VALIDATION

Histogram of **choice probabilities** predicted by MNL and ICLV (80%/20%)

ICLV



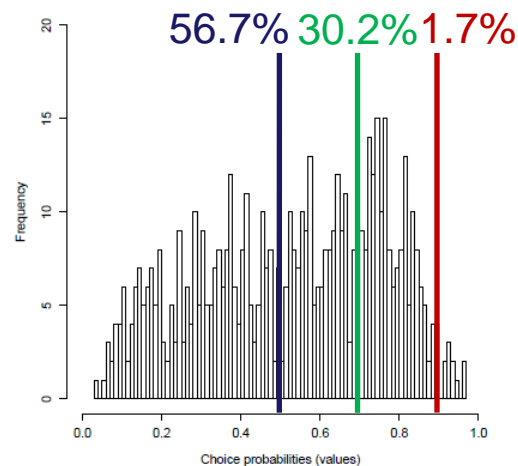
MNL



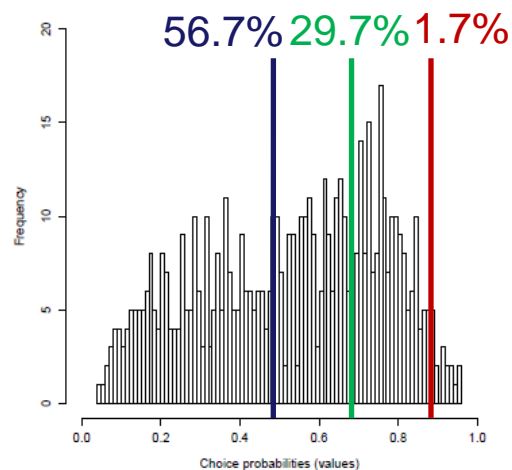
Value

Histogram of **choice probabilities** predicted by MNL and ICLV (80%/20%)

ICLV



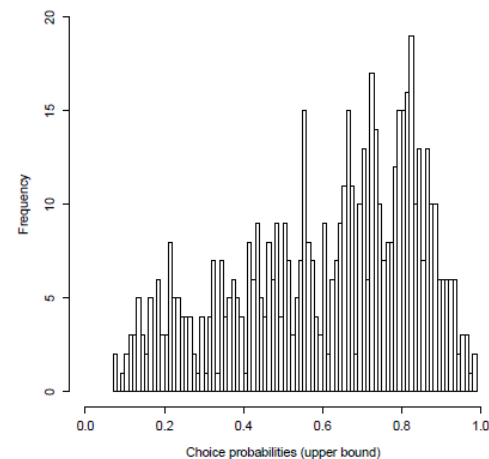
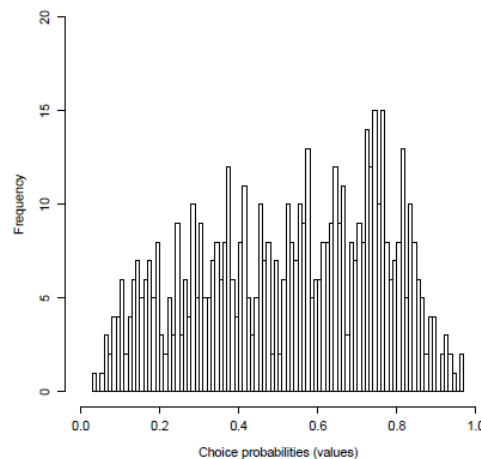
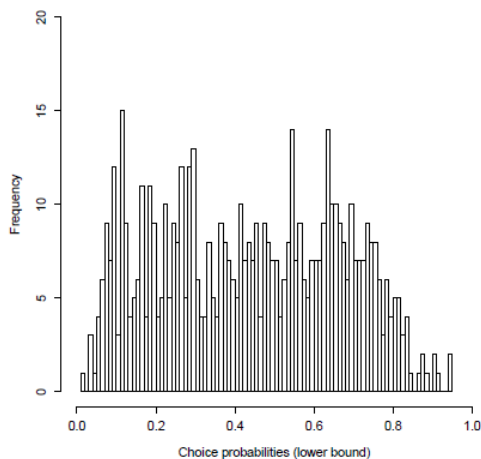
MNL



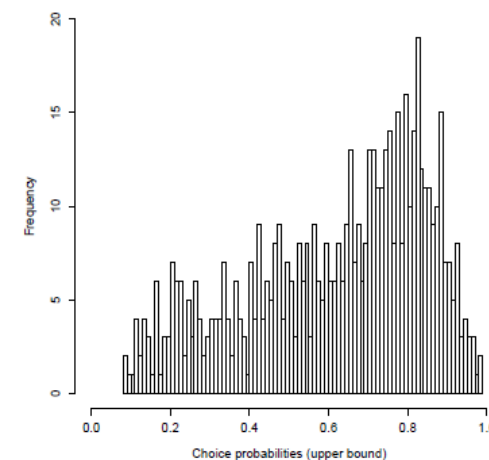
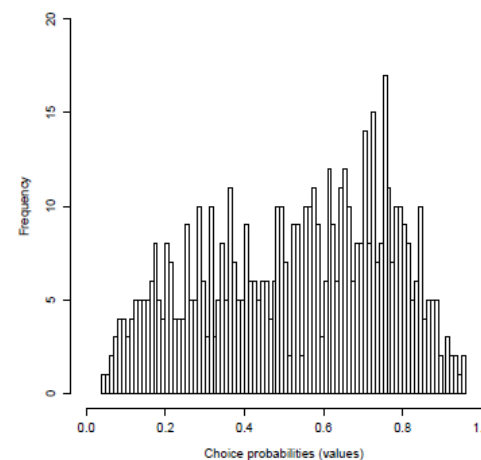
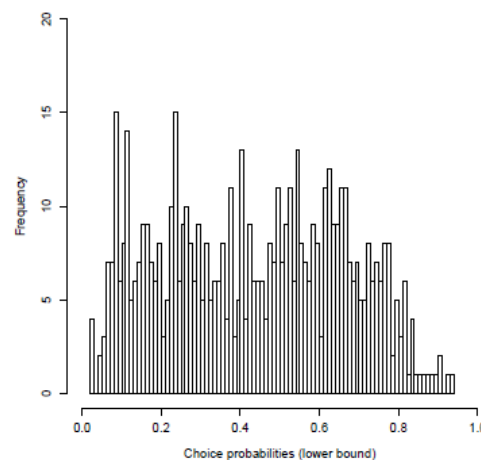
Value

Histogram of choice probabilities predicted by MNL and ICLV (80%/20%)

ICLV



MNL



Lower bound

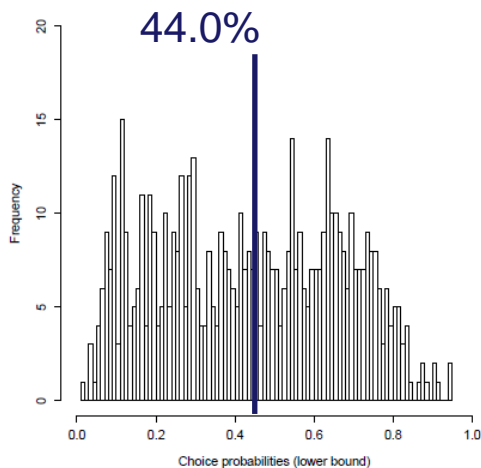
Value

Upper bound

VALIDATION

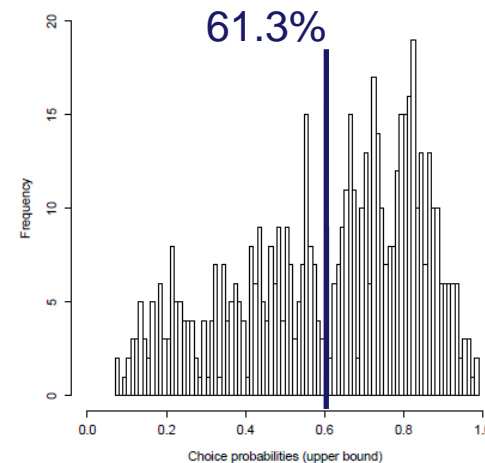
Histogram of **choice probabilities** predicted by MNL and ICLV (80%/20%)

ICLV

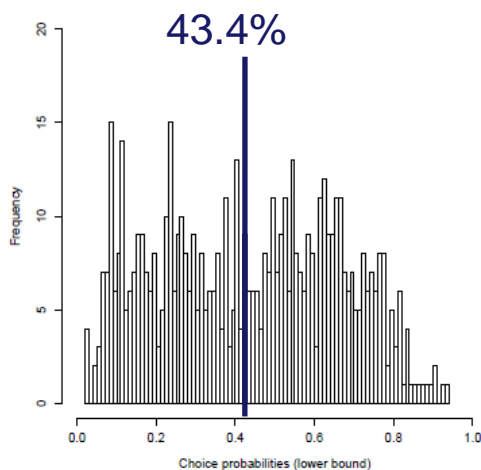


Difference
between average
confidence bounds

17.3%

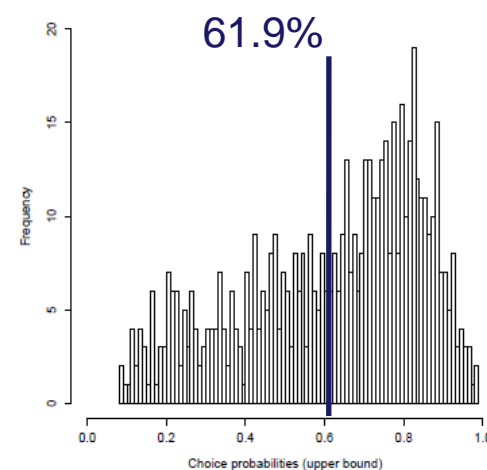


MNL



^

18.5%



Lower
bound

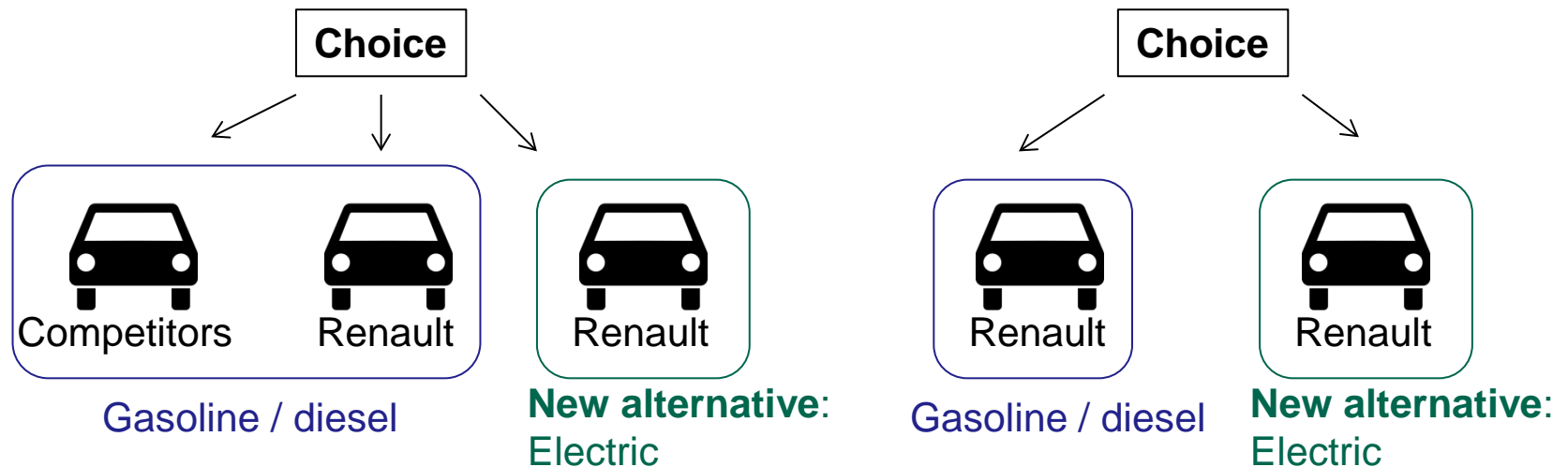
Upper
bound

Several corrections to the SP model are needed before the model can be applied for scenario forecasting:

1. Introduction of an **aggregate alternative** for car models from competitors (using logsum)
2. Correction of constants:
 - Current **ratio of market shares** between Renault and competitors is **preserved**.
 - Estimate potential market share of EV using acceptance rate and Swiss market data.

1. AGGREGATE ALTERNATIVE

Two possible choice situations



Issue:

- Choice is supposed to represent all possible alternatives for decision maker
- Not the case for owners of Renault cars

Solution:

- Impute aggregate alternative of gasoline – competitors for these individuals

1. AGGREGATE ALTERNATIVE

Aggregate alternative imputed for Competitors – Gasoline (CG)

$$V_{CG} = \log \sum_{l \in L} \exp U_{ln}$$

$$U_{ln} = ASC_{CG} + \sum_{s \in S_n} \beta_s \cdot x_s - \exp(\beta_{price_{CG}} + \beta_{AttC} \cdot AttC_n) \cdot price_l \\ + \beta_{UseCostGasoline} \cdot Cost100_l \cdot (Cost100_l \leq 12) + \varepsilon_{ln}$$

Generated from **prices** & **operating costs** of new cars on market
(matching segment of 2 other alternatives in choice situation)

2. CORRECTIONS OF CONSTANTS

Idea:

Use:

- Market data of current alternatives
- SP survey data

} To estimate possible share for new alternative

2. CORRECTIONS OF CONSTANTS

Idea:

Use:

- Market data of current alternatives
- SP survey data

To estimate possible share for new alternative

Evaluation of potential market share (MS) for EV

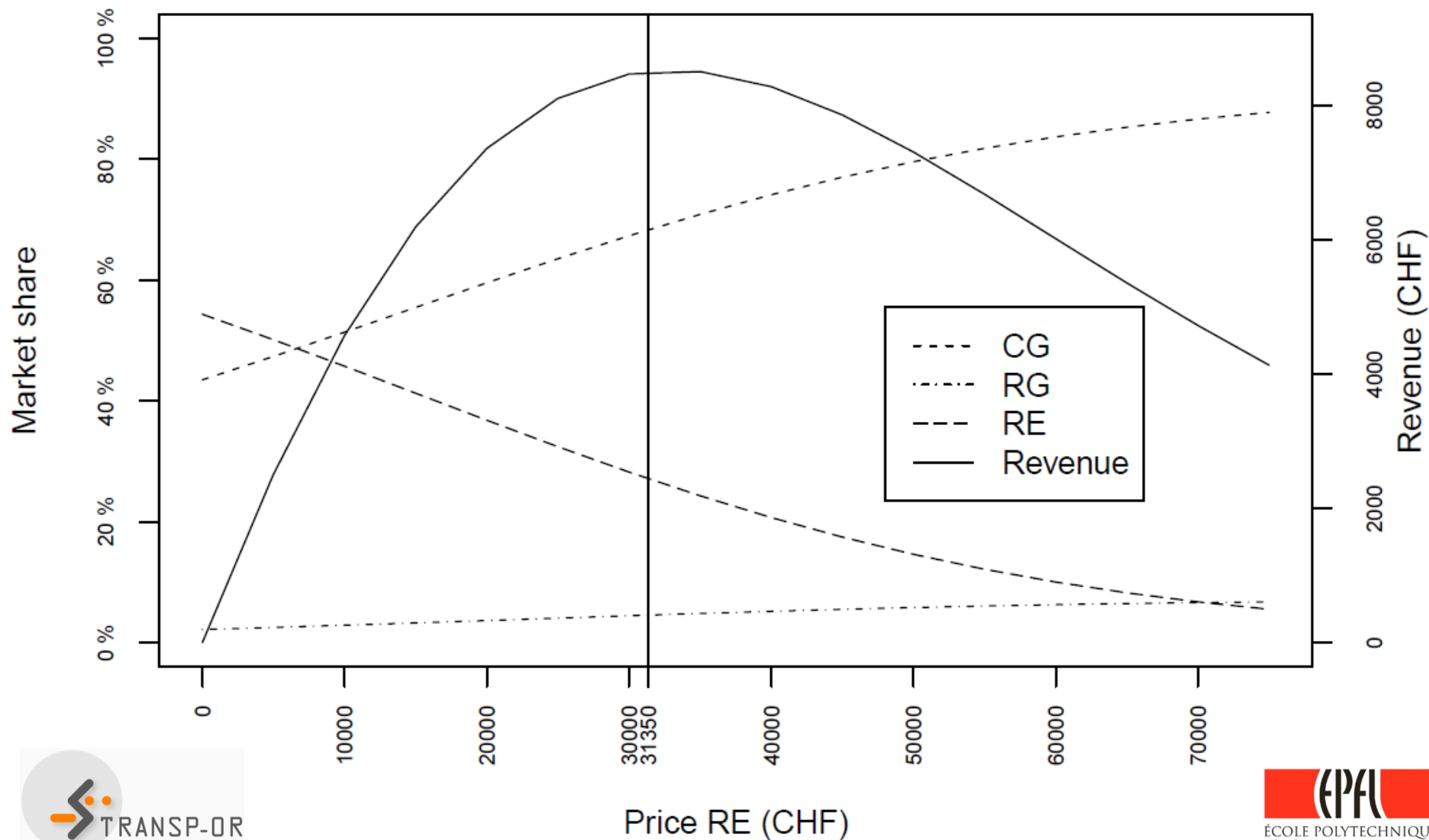
Acceptance rate EV in the questionnaire for CG owners (weighted)

Acceptance rate EV in the questionnaire for RG owners (weighted)

$$MS(RE) = \underbrace{\%(\text{Choice RE} \mid \text{Owns CG})}_{\text{Acceptance rate EV in the questionnaire for CG owners (weighted)}} \cdot \underbrace{94\%}_{\text{Market share of competitors}} + \underbrace{\%(\text{Choice RE} \mid \text{Owns RG})}_{\text{Acceptance rate EV in the questionnaire for RG owners (weighted)}} \cdot \underbrace{6\%}_{\text{Market share of Renault}}$$

$= 27\%$

Example of scenario



Conclusions

- Operational model obtained by the presented procedure: from data collection to model application
- Important to include market data when forecast for a new alternative

Future analyses

- Analyzed the demand for EV for private use, but alternative uses exist (e.g. car sharing)
- Now that EVs are more present on the market, revealed preferences (RP) data can be collected and the model can integrate both.

USING ADJECTIVES TO MEASURE PERCEPTIONS IN HYBRID CHOICE MODELS

Aurélie Glerum
Bilge Atasoy
Michel Bierlaire

Introduction & motivation

The data

Model specification

Estimation results

Validation

Conclusion

Issues related to the integration of latent variables into choice models:

1. **Measurement of latent variable**

⇒ How to obtain the most realistic and accurate measure of a perception?

2. **Integration of the measurement into the choice model**

⇒ How to incorporate this information in the choice modeling framework?

1. Measurement of latent variable:

- Use of **opinion statements**
Five-point Likert scale } Usual way in literature
(Likert, 1932; Bearden and Netemeyer, 1999)

- **Recent technique** developed in **social sciences:**

Respondents report **adjectives** characterizing a variable of interest (Kaufmann et al., 2001; Kaufmann et al., 2010)

Reflects **spontaneous** perceptions of individuals
(≠ survey designer's conception of the perception)

1. Measurement of latent variable:

- Use of **opinion statements**
Five-point Likert scale
- } Usual way in literature
(Likert, 1932; Bearden and Netemeyer, 1999)

- **Recent technique** developed in **social sciences**:

Respondents report **adjectives** characterizing a variable of interest (Kaufmann et al., 2001; Kaufmann et al., 2010)

Reflects **spontaneous** perceptions of individuals
(≠ survey designer's conception of the perception)

➡ **1ST AIM OF THIS RESEARCH: USE THE ADJECTIVES TO MEASURE PERCEPTIONS**

2. Integration of the measurement into the choice model:

- **Structural equation model (SEM)** framework used to characterize latent variable and relate it to its measurement indicators (e.g. Bollen, 1989).
- Latent variable model embedded into DCM \Rightarrow **HCM framework**
- **Integration of measurements into HCM framework:**
 - Well-established for models with opinion statements
 - **Adjectives need to be quantified**

2. Integration of the measurement into the choice model:

- **Structural equation model (SEM)** framework used to characterize latent variable and relate it to its measurement indicators (e.g. Bollen, 1989).
- Latent variable model embedded into DCM \Rightarrow **HCM framework**
- **Integration of measurements into HCM framework:**
 - Well-established for models with opinion statements

- **Adjectives need to be quantified**

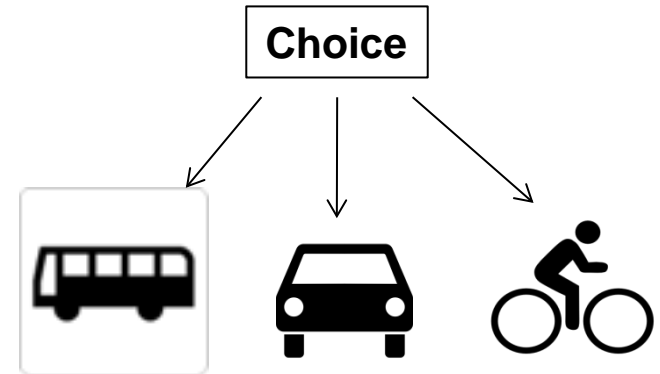
\Rightarrow 2ND AIM OF THIS RESEARCH: QUANTIFY THE ADJECTIVES TO INTEGRATE THEM IN AN HCM

Two surveys:

- Revealed preferences (RP) survey
- Survey with evaluators (adjective quantification survey)

RP survey

- **Mode choice study**
- Conducted between 2009-2010 in low-density areas of Switzerland
- Conducted with PostBus (major bus company in Switzerland, operates in low-density areas)
- Info on **all trips performed by inhabitants in one day**:
 - Transport mode
 - Trip duration
 - Cost of trip
 - Activity at destination
 - Etc.
- **1763 valid questionnaires** collected



Adjective data for perception of transport modes:

For each of the following transport modes, give three adjectives that describe them best according to you.

		Adjective 1	Adjective 2	Adjective 3
1	The car is:			
2	The train is:			
3	The bus, the metro and the tram are:			
4	The post bus is:			
5	The bicycle is:			
6	The walk is:			

Adjective data for perception of transport modes:

For each of the following transport modes, give three adjectives that describe them best according to you.

		Adjective 1	Adjective 2	Adjective 3
1	The car is:	convenient	comfortable	expensive
2	The train is:	relaxing	punctual	restful
3	The bus, the metro and the tram are:	fast	frequent	cheap
4	The post bus is:	punctual	comfortable	cheap
5	The bicycle is:	stimulating	convenient	cheap
6	The walk is:	healthy	relaxing	independent

Extraction of information on perceptions

1. Classification into themes:

- Perception of cost
- Perception of time
- Difficulty of access
- Flexibility
- Comfort, etc.

2. Focused on adjectives related to one theme only and one mode only:

Comfort in public transportation (PT)

Comfort

hardly full

packed

bumpy

comfortable

hard

irritating

tiring

unsuitable with bags

uncomfortable

bad air

...

Extraction of information on perceptions

1. Classification into themes:

- Perception of cost
- Perception of time
- Difficulty of access
- Flexibility
- Comfort, etc.

2. Focused on adjectives related to one theme only and one mode only:

Comfort in public transportation (PT)



**LATENT VARIABLE
WE STUDY**

Comfort

hardly full

packed

bumpy

comfortable

hard

irritating

tiring

unsuitable with bags

uncomfortable

bad air

...

ADJECTIVE QUANTIFICATION SURVEY

SURVEY ON ADJECTIVES

Email address (optional):

Age:

Gender: -- ▾

Nationality: (please select a country) ▾

Native language: -- ▾

Language at work: -- ▾

Language used at home: -- ▾

Highest educational degree: -- ▾

Professional occupation: -- ▾

For each adjective below, we ask you to rate how strongly it characterizes the concept of *comfort* of a transportation mode (car, bus, train, etc.) Please select an integer **between -2 and 2** on the corresponding slider, where a positive number corresponds to adjectives associated with comfort, and a negative number with discomfort. If you do not associate an adjective with the concept of *comfort*, rate it with 0.

Hard	-2	<input type="range"/>	2	<input type="text"/>
Relaxing	-2	<input type="range"/>	2	<input type="text"/>
Stressful	-2	<input type="range"/>	2	<input type="text"/>
Without stress	-2	<input type="range"/>	2	<input type="text"/>

- Asked 25 external evaluators to rate the adjectives on scale of comfort.
- Discrete scale: ratings from -2 to 2.

ADJECTIVE QUANTIFICATION SURVEY

Aims:

- Use adjectives to measure perceptions
- Quantify them to integrate them into the HCM framework

Now:

Ratings from 25 different evaluators

⇒ How reliable is each set of ratings?

Next step:

- Estimate an HCM for each set of ratings (i.e. for each evaluator)
 - LV is perception of comfort in PT
 - LV measured by ratings of one evaluator
- Compare the fit & prediction capabilities of each model

ADJECTIVE QUANTIFICATION SURVEY

Adjectives	Mean	Median	Standard deviation	Central evaluator	Outlying evaluator
bad air	-1.52	-2	0.714	-1	1
bumpy	-1.08	-1	0.862	-1	1
comfortable	1.72	2	0.542	1	2
difficult	-1.00	-1	0.707	-1	-2
empty	0.880	1	0.726	1	1
expensive	-0.680	0	0.988	-2	-1
fast	1.04	1	0.735	2	1
full	-1.00	-1	1.00	-2	2
hard	-0.920	-1	0.640	-1	-1
hardly full	-0.280	0	1.28	1	2
irritating	-1.44	-2	0.870	-1	1
packed	-0.880	-1	1.20	-2	1
relaxing	1.72	2	0.737	1	1

THE DATA

ADJECTIVE QUANTIFICATION SURVEY

Adjectives	Mean	Median	Standard deviation	Central evaluator	Outlying evaluator
bad air	-1.52	-2	0.714	-1	1
bumpy	-1.08	-1	0.862	-1	1
comfortable	1.72	2	0.542	1	2
difficult	-1.00	-1	0.707	-1	-2
empty	0.880	1	0.726	1	1
expensive	-0.680	0	0.988	-2	-1
fast	1.04	1	0.735	2	1
full	-1.00	-1	1.00	-2	2
hard	-0.920	-1	0.640	-1	-1
hardly full	-0.280	0	1.28	1	2
irritating	-1.44	-2	0.870	-1	1
packed	-0.880	-1	1.20	-2	1
relaxing	1.72	2	0.737	1	1
...

Smallest Euclidean distance to other evaluators

ADJECTIVE QUANTIFICATION SURVEY

Adjectives	Mean	Median	Standard deviation	Central evaluator	Outlying evaluator
bad air	-1.52	-2	0.714	-1	1
bumpy	-1.08	-1	0.862	-1	1
comfortable	1.72	2	0.542	1	2
difficult	-1.00	-1	0.707	-1	-2
empty	0.880	1	0.726	1	1
expensive	-0.680	0	0.988	-2	-1
fast	1.04	1	0.735	2	1
full	-1.00	-1	1.00	-2	2
hard	-0.920	-1	0.640	-1	-1
hardly full	-0.280	0	1.28	1	2
irritating	-1.44	-2	0.870	-1	1
packed	-0.880	-1	1.20	-2	1
relaxing	1.72	2	0.737	1	1
...

Furthest away from central evaluator

Hybrid choice model

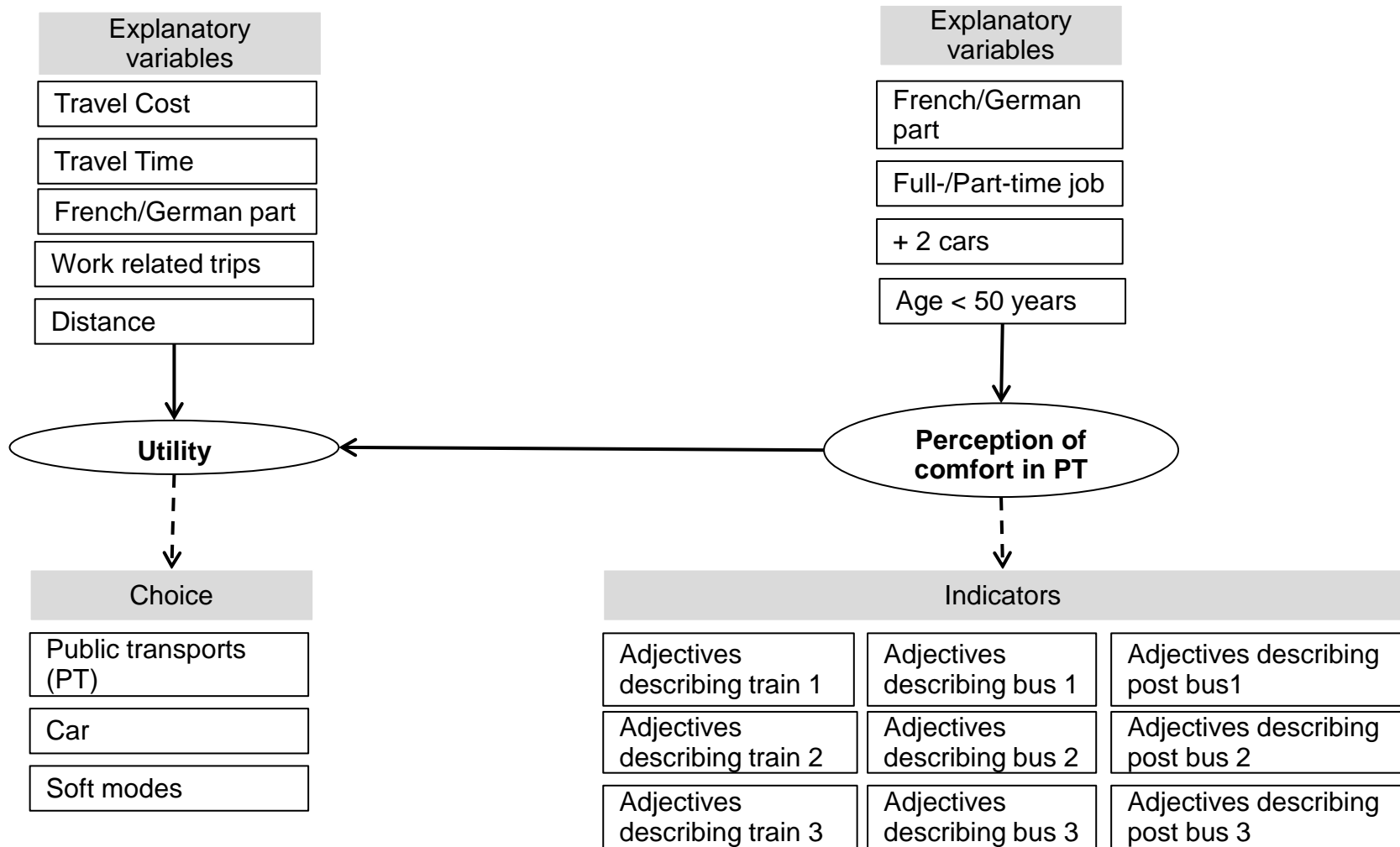


Figure based on Walker and Ben-Akiva, 2002.

Hybrid choice model

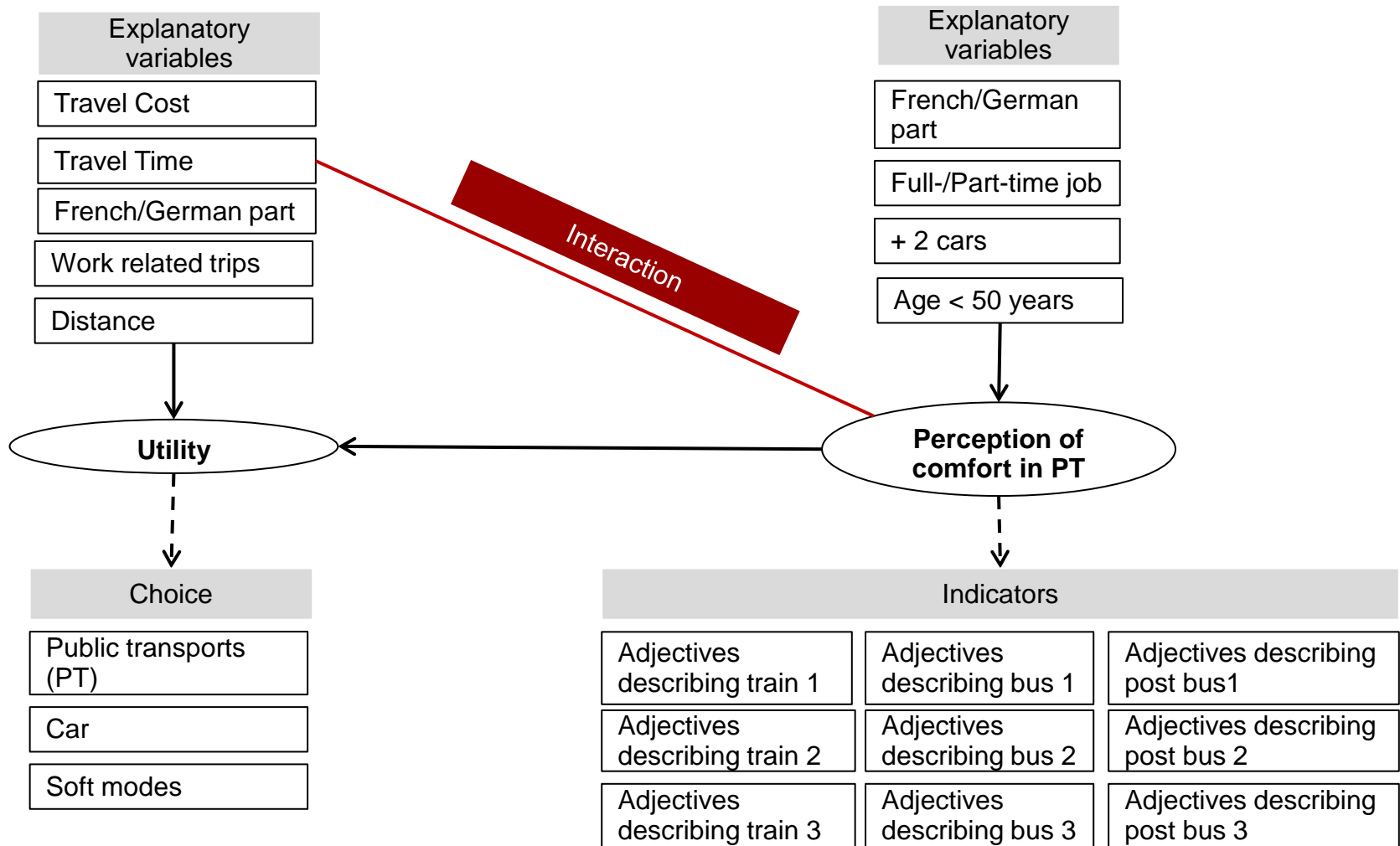


Figure based on Walker and Ben-Akiva, 2002.

- **Model estimated for each evaluator** (\rightarrow 25 models estimated)
- **Fit indices (for the choice model):**

Indicator	Logit	Central evaluator	Outlying evaluator	Median ratings
Loglikelihood	-1153	-1192	-1199	-1190
J	14	11	11	11
$\bar{\rho}^2$	0.443	0.425	0.422	0.427

ESTIMATION RESULTS

Name	Logit			Central evaluator			Outlying evaluator			Median ratings		
	Value	<i>t</i> -test		Value	<i>t</i> -test		Value	<i>t</i> -test		Value	<i>t</i> -test	
ASC _{PT}	-0.178	-0.88	*	-0.155	-0.77	*	-0.132	-0.66	*	-0.155	-0.77	*
ASC _{PMM}	0.423	2.30		0.416	2.26		0.410	2.22		0.419	2.27	
β_{cost}	-0.0658	-8.67		-0.0637	-8.11		-0.0628	-8.02		-0.0653	-8.08	
β_{timePT}	-0.00600	-3.34		-0.0143	-7.71		-0.0293	-4.17		-0.0208	-7.06	
β_{timePMM}	-0.0330	-10.27		-0.0313	-9.53		-0.0312	-9.55		-0.0323	-9.43	
β_{distance}	-0.236	-11.51		-0.233	-11.4		-0.233	-11.38		-0.235	-11.45	
β_{workPT}	0.0987	0.42	*	-0.0602	-0.26	*	-0.0928	-0.40	*	-0.0474	-0.20	*
β_{workPMM}	-0.613	-2.77		-0.572	-2.58		-0.560	-2.53		-0.575	-2.60	
β_{FrenchPT}	-0.228	-0.61	*	-0.073	-0.24	*	-0.113	-0.37	*	-0.0808	-0.26	*
$\beta_{\text{FrenchPMM}}$	0.990	3.64		0.966	3.56		0.969	3.57		0.967	3.56	
β_{comfort}	-	-		1.06	3.46		1.09	2.65		1.33	4.34	
λ_{mean}	-	-		3.33	9.40		15.7	11.46		7.47	9.98	
λ_{French}	1.11	0.44	*	-0.559	-1.80		-0.139	-0.48	*	-0.456	-1.58	*
$\lambda_{\text{age}_{50}}$	1.42	1.25	*	-1.30	-5.53		0.0643	0.30	*	-1.04	-4.62	
λ_{active}	-8.34	-6.77		-1.10	-4.37		-0.582	-2.68		-1.12	-4.62	
λ_{cars}	-7.81	-6.59		-0.730	-3.06		-0.362	-1.58	*	-0.688	-3.04	

$\beta_{\text{comfort}} > 0$: A high perception of comfort of PT increases its utility.

ESTIMATION RESULTS

Name	Logit			Central evaluator			Outlying evaluator			Median ratings		
	Value	<i>t</i> -test		Value	<i>t</i> -test		Value	<i>t</i> -test		Value	<i>t</i> -test	
ASC _{PT}	-0.178	-0.88	*	-0.155	-0.77	*	-0.132	-0.66	*	-0.155	-0.77	*
ASC _{PMM}	0.423	2.30		0.416	2.26		0.410	2.22		0.419	2.27	
β_{cost}	-0.0658	-8.67		-0.0637	-8.11		-0.0628	-8.02		-0.0653	-8.08	
$\beta_{time_{PT}}$	-0.00600	-3.34		-0.0143	-7.71		-0.0293	-4.17		-0.0208	-7.06	
$\beta_{time_{PMM}}$	-0.0330	-10.27		-0.0313	-9.53		-0.0312	-9.55		-0.0323	-9.43	
$\beta_{distance}$	-0.236	-11.51		-0.233	-11.4		-0.233	-11.38		-0.235	-11.45	
$\beta_{work_{PT}}$	0.0987	0.42	*	-0.0602	-0.26	*	-0.0928	-0.40	*	-0.0474	-0.20	*
$\beta_{work_{PMM}}$	-0.613	-2.77		-0.572	-2.58		-0.560	-2.53		-0.575	-2.60	
$\beta_{French_{PT}}$	-0.228	-0.61	*	-0.073	-0.24	*	-0.113	-0.37	*	-0.0808	-0.26	*
$\beta_{French_{PMM}}$	0.990	3.64		0.966	3.56		0.969	3.57		0.967	3.56	
$\beta_{comfort}$	-	-		1.06	3.46		1.09	2.65		1.33	4.34	
λ_{mean}	-	-		3.33	9.40		15.7	11.46		7.47	9.98	
λ_{French}	1.11	0.44	*	-0.559	-1.80		-0.139	-0.48	*	-0.456	-1.58	*
$\lambda_{age_{50}}$	1.42	1.25	*	-1.30	-5.53		0.0643	0.30	*	-1.04	-4.62	
λ_{active}	-8.34	-6.77		-1.10	-4.37		-0.582	-2.68		-1.12	-4.62	
λ_{cars}	-7.81	-6.59		-0.730	-3.06		-0.362	-1.58	*	-0.688	-3.04	

$\beta_{comfort} > 0$ and $\beta_{time_{PT}} < 0$: Travel time sensitivity decrease with an increased perception of comfort of PT.

Estimation on 80% data / Application on 20 % data

Fit indices:

Indicator	Logit	Central evaluator	Outlying evaluator	Median ratings
Loglikelihood	-220	-227	-229	-226
J	14	11	11	11
$\bar{\rho}^2$	0.459	0.449	0.444	0.452
Percentage choice probabilities > 0.5	70.5%	69.7%	70.0%	70.5%

Estimation on 80% data / Application on 20 % data

Fit indices:

Indicator	Logit	Central evaluator	Outlying evaluator	Median ratings
Loglikelihood	-220	-227	-229	-226
J	14	11	11	11
$\bar{\rho}^2$	0.459	0.449	0.444	0.452
Percentage choice probabilities > 0.5	70.5%	69.7%	70.0%	70.5%

- Fit similar across evaluators
- Slightly lower for outlying evaluator

Analysis of demand indicators across evaluators

Example: computation of market shares

		Example of respondents with different probabilities for the chosen alternative						All observations	
Indicator	Mode	Low		Medium		High		Mean	SD
		Mean	SD	Mean	SD	Mean	SD		
Probability of choice / Market share	PT	0.047	0.008	0.543	0.013	0.067	0.005	0.278	0.001
	PMM	0.953	0.008	0.457	0.013	0.933	0.005	0.659	0.001
	SM	0.000	0.000	0.000	0.000	0.000	0.000	0.063	0.000

Analysis of demand indicators across evaluators

Example: computation of market shares

		Example of respondents with different probabilities for the chosen alternative						All observations	
Indicator	Mode	Low		Medium		High		Mean	SD
		Mean	SD	Mean	SD	Mean	SD		
Probability of choice / Market share	PT	0.047	0.008	0.543	0.013	0.067	0.005	0.278	0.001
	PMM	0.953	0.008	0.457	0.013	0.933	0.005	0.659	0.001
	SM	0.000	0.000	0.000	0.000	0.000	0.000	0.063	0.000

Disaggregate indicators → Probabilities to choose each mode.

Analysis of demand indicators across evaluators

Example: computation of market shares

		Example of respondents with different probabilities for the chosen alternative						All observations	
Indicator	Mode	Low		Medium		High		Mean	SD
		Mean	SD	Mean	SD	Mean	SD		
Probability of choice / Market share	PT	0.047	0.008	0.543	0.013	0.067	0.005	0.278	0.001
	PMM	0.953	0.008	0.457	0.013	0.933	0.005	0.659	0.001
	SM	0.000	0.000	0.000	0.000	0.000	0.000	0.063	0.000

Aggregate indicators → market shares for each mode.

Analysis of demand indicators across evaluators

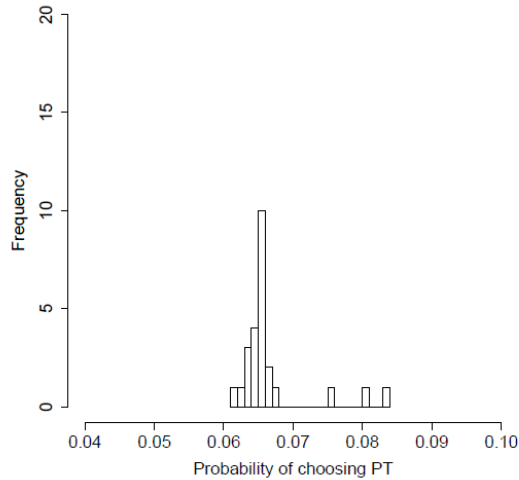
Example: computation of market shares

		Example of respondents with different probabilities for the chosen alternative						All observations	
Indicator	Mode	Low		Medium		High		Mean	SD
		Mean	SD	Mean	SD	Mean	SD		
Probability of choice / Market share	PT	0.047	0.008	0.543	0.013	0.067	0.005	0.278	0.001
	PMM	0.953	0.008	0.457	0.013	0.933	0.005	0.659	0.001
	SM	0.000	0.000	0.000	0.000	0.000	0.000	0.063	0.000

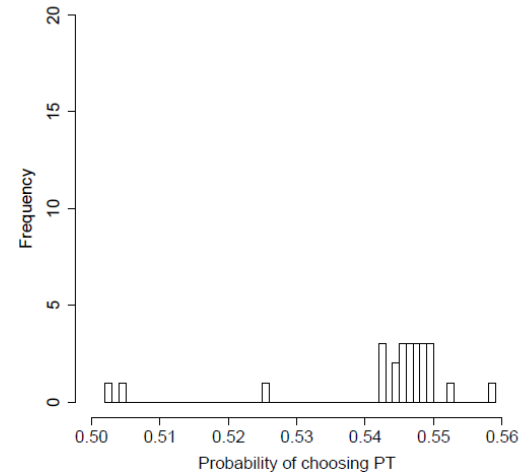
Aggregate indicators have lower standard deviations than disaggregate indicators → more consistency at aggregate level.

Analysis of demand indicators across evaluators

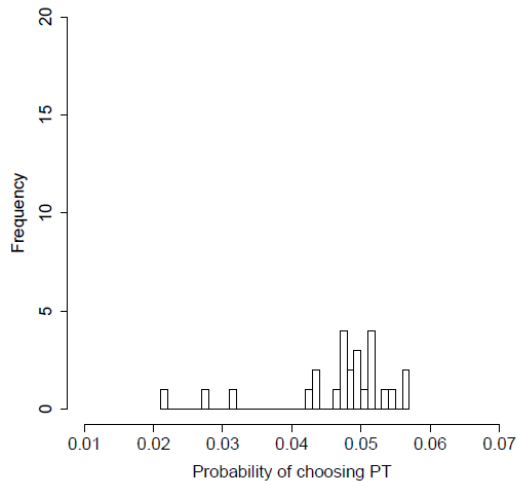
High



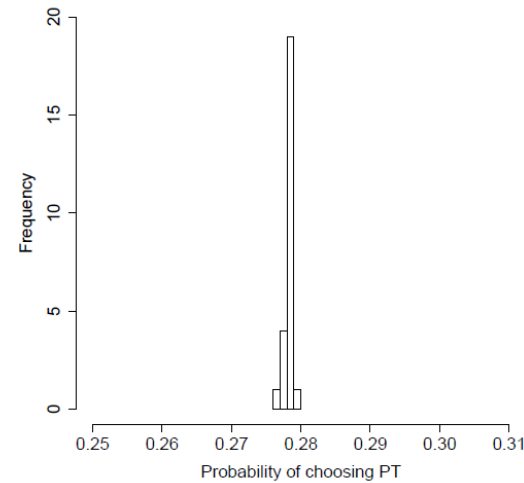
Medium



Low



All observations



Conclusions:

- Adjectives: **alternative** way to **measure** perceptions
- Propose a **methodology to rate adjectives** in order to minimize subjectivity
- Method is **robust** with respect to **poor evaluators**

Further improvements:

- Two surveys in **one step**
- **Comparative approach** between classical **opinion questions** and adjectives
- Comparison between adjective rating on a **discrete** (-2 to 2) and **continuous scale** (-1000 and 1000)
- Investigate the effect of **other themes** (than comfort in PT)

Thank you!