

Uncovering substitution patterns in new car sales using a cross nested logit model

Swiss Transportation Research Conference

Anna Fernández Antolín Matthieu de Lapparent Michel Bierlaire

May 19, 2016



Outline

- 1 State of the art
- 2 Case study
- 3 Results
- 4 Conclusions and future work

Research question

Real research question

Can we model more flexible substitution patterns using Choice Probability Generation Functions (CPGF) based models?

Research question

Real research question

Can we model more flexible substitution patterns using Choice Probability Generation Functions (CPGF) based models?

But before...

- Start by a Cross-Nested Logit
- Add on the car-type ownership literature

Car-type models

Why are they interesting?

- For car manufacturers: valuation of car attributes
- For governments, forecasts of:
 - Tax revenues
 - Energy consumption
 - Emission levels
- Can be used for policy measures



This is preliminary work. Comments and suggestions are more than welcome!

Identifying a vehicle type

Make-model-engine ²

- An alternative: Volvo XC90 2.4
- Over 1000 alternatives
- Sampling of alternatives needed



Market and fuel type ³

- An alternative: Small petrol car
- Between 15 and 30 alternatives
- No sampling of alternatives needed



² Birkeland, M. E. & Jordal-Jorgensen, J. (2001) Energy efficiency of passenger cars. Paper presented at the European Transport Conference 2001, PTRC, Cambridge, UK.

³ Page, M., Whelan, G., & Daly, A. (2000) Modelling the factors which influence new car purchasing. Paper presented at the European Transport Conference 2000, PTRC, Cambridge, UK.

Data: France 2014

Decision makers

- 40,000 observations
- 20,000 contain no NAs

Data: France 2014

Decision makers

- 40,000 observations
- 20,000 contain no NAs

Attributes

- Reported fuel consumption [l/100km]
- Engine power [bhp]
- Price after discounts and government schemes [€]
- Reported range (EV) [km]

Data: France 2014

Decision makers

- 40,000 observations
- 20,000 contain no NAs

Attributes

- Reported fuel consumption [l/100km]
- Engine power [bhp]
- Price after discounts and government schemes [€]
- Reported range (EV) [km]

Socioeconomic variables

- Income
- Number of adults/children in the household
- Residential location (agglomerations vs. rural areas)
- Education level (university vs. no university)

Choice-set definition

Choice set

Car type = market segment + fuel type

Choice-set definition

Choice set

Car type = market segment + fuel type

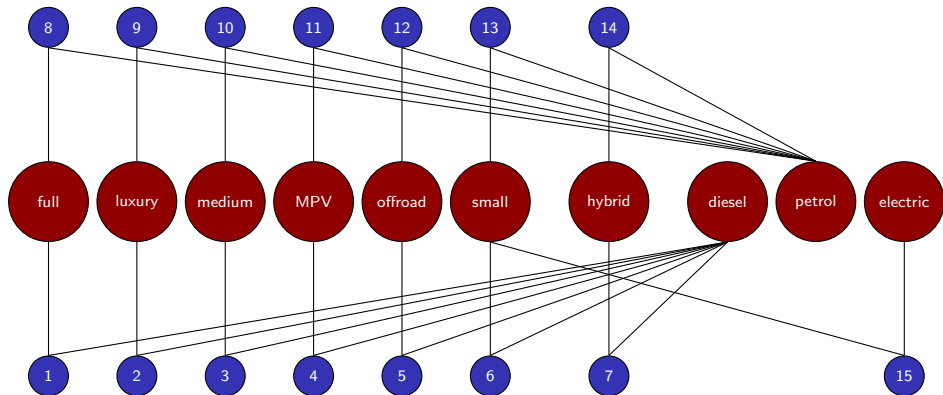
Market segment

- Full
- Luxury
- Medium
- Multi-purpose vehicle (MPV)
- Off-road
- Small

Fuel type

- Hybrid
- Diesel
- Petrol
- Electric

Cross-nesting structure



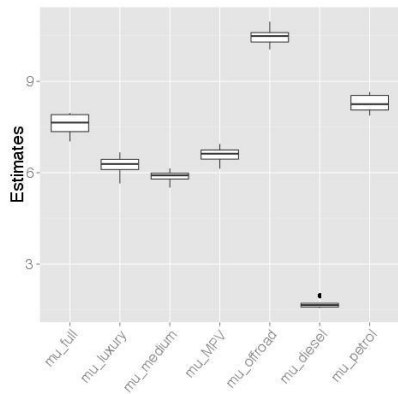
Attributes of non-chosen alternatives

What are the attributes of an off-road diesel car that I didn't choose?

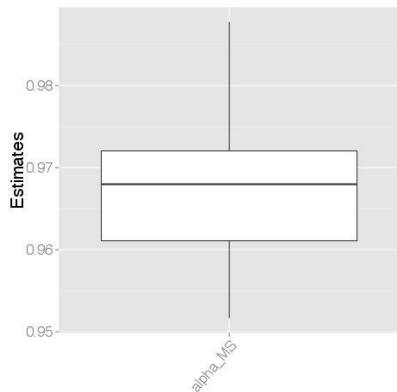
- 1 Draw vectors of attributes from the empirical distribution.
- 2 Define the unchosen alternatives for each respondent.
- 3 Estimate the parameters of the model with this dataset.
- 4 Iterate.

Parameter estimation: CNL

Scale parameters ($\mu_{\text{small}} = \mu_{\text{hybrid}} = \mu_{\text{electric}} = 1$)



Alpha market segment



Substitution patterns

Market shares before and after a 50% increase in price of alternative 5 (diesel offroad)

Alternative	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Observed	1.72	0.95	11.84	7.31	10.87	29.45	0.86	0.36	0.29	3.53	1.65	1.41	26.79	2.64	0.35
Logit	1.30	0.81	9.96	6.06	5.47	36.14	2.34	0.83	0.45	5.79	3.13	4.68	21.42	1.24	0.36
CNL	1.48	0.88	10.74	6.48	7.73	30.49	1.90	0.64	0.39	5.69	2.82	3.88	25.01	1.52	0.35

How does the market share for alternative 12 increase when the market share of alternative 5 decreases?

- Logit:

$$\frac{4.68 - 1.41}{10.87 - 5.47} \cdot 100 = \frac{3.27}{5.40} \cdot 100 = 61\%$$

- CNL:

$$\frac{3.88 - 1.41}{10.87 - 7.73} \cdot 100 = \frac{2.47}{3.14} \cdot 100 = 79\%$$

Conclusions and Future work

Conclusions

- Most results are in line with our expectations and the literature.
- Results seem stable with only 10 draws.
- Substitution patterns seem more intuitive with the CNL.

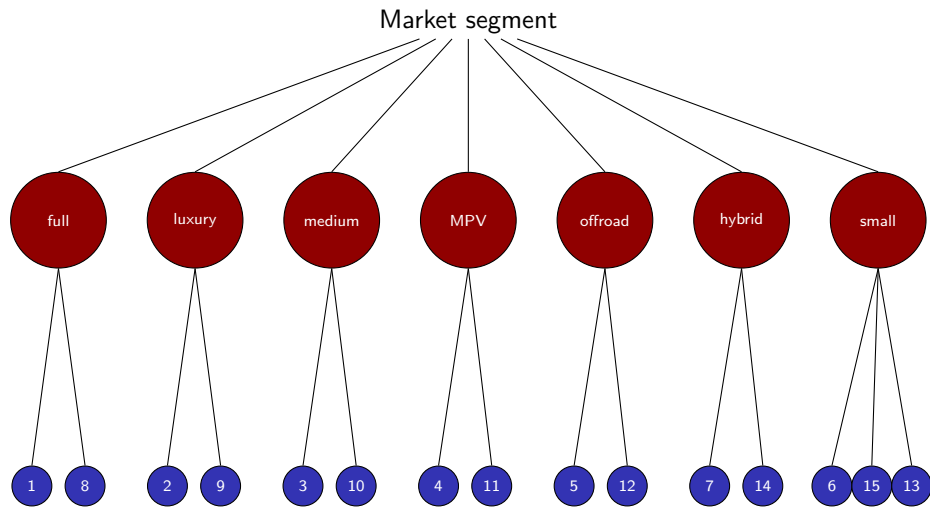
Future work

- Endogeneity of price and fuel consumption
- CPGF-based models

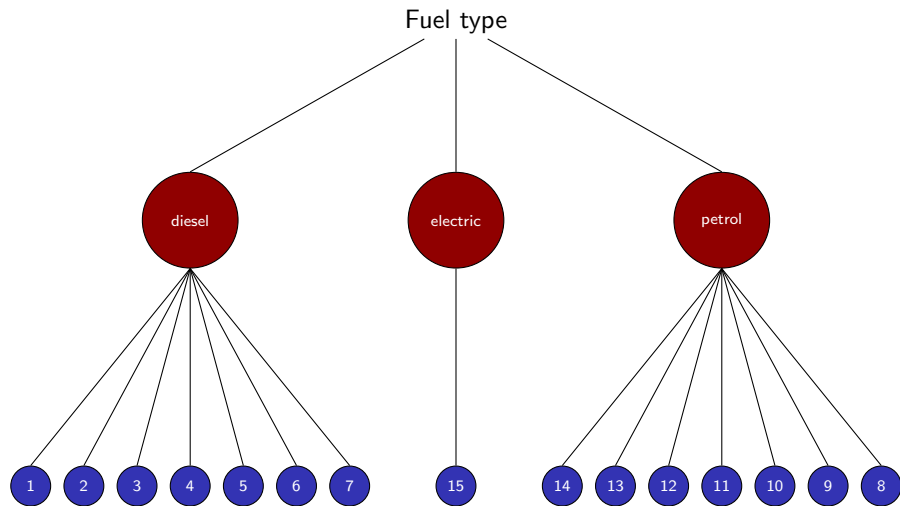
Thank you for your attention!
Questions?

`anna.fernandezantolin@epfl.ch`

Nesting structure 1

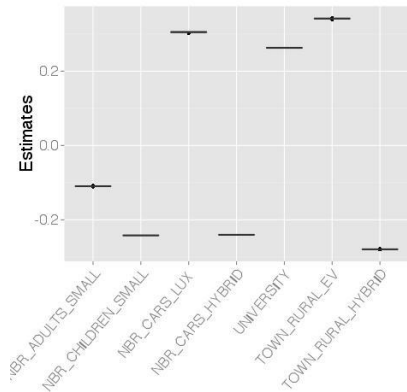


Nesting structure 2

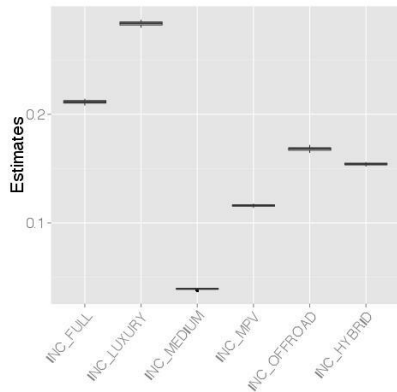


Parameter estimation: CNL

Attributes of the car

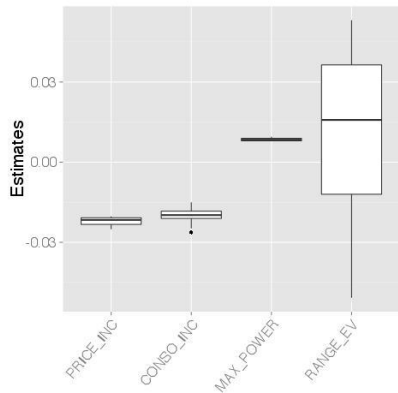


Income

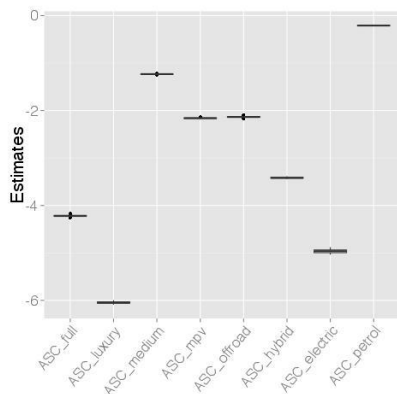


Parameter estimation: CNL

Socioeconomics

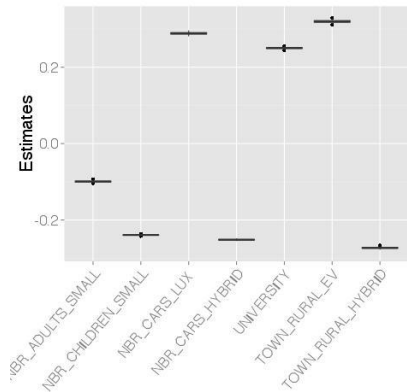


Dummy variables

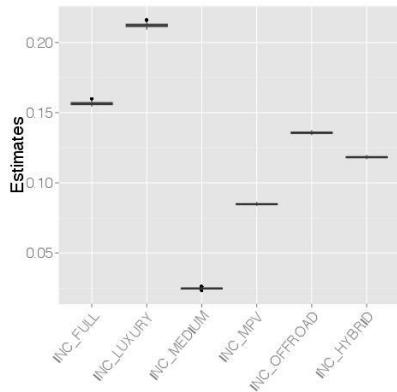


Parameter estimation: Logit

Attributes of the car

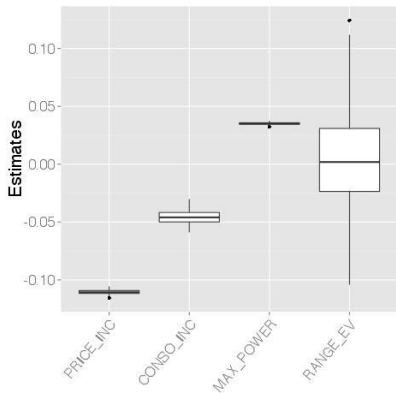


Income

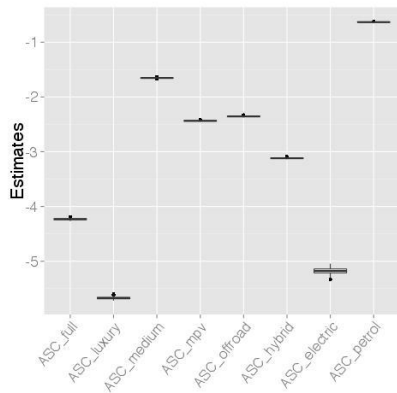


Parameter estimation: Logit

Socioeconomics



Dummy variables



Model specification (1/2)

Parameter	1	2	3	4	5	6	7	8
ASC _{full}	1	0	0	0	0	0	0	1
ASC _{luxury}	0	1	0	0	0	0	0	0
ASC _{medium}	0	0	1	0	0	0	0	0
ASC _{MPV}	0	0	0	1	0	0	0	0
ASC _{offroad}	0	0	0	0	1	0	0	0
ASC _{petrol}	0	0	0	0	0	0	0	1
ASC _{electric}	0	0	0	0	0	0	0	0
β_{inc_full}	$\frac{income}{10000}$	0	0	0	0	0	0	$\frac{income}{10000}$
β_{inc_luxury}	0	$\frac{income}{10000}$	0	0	0	0	0	0
β_{inc_medium}	0	0	$\frac{income}{10000}$	0	0	0	0	0
β_{inc_MPV}	0	0	0	$\frac{income}{10000}$	0	0	0	0
$\beta_{inc_offroad}$	0	0	0	0	$\frac{income}{10000}$	0	0	0
β_{inc_hybrid}	0	0	0	0	0	0	$\frac{income}{10000}$	0
$\beta_{nbr_adults_small}$	0	0	0	0	0	nbr. adults	0	0
$\beta_{nbr_children_small}$	0	0	0	0	0	nbr. child.	0	0
$\beta_{nbr_cars_lux}$	0	nbr. cars	0	0	0	0	0	0
$\beta_{nbr_cars_hybrid}$	0	0	0	0	0	0	nbr. cars	0
$\beta_{university}$	0	0	0	0	0	0	1	0
$\beta_{town_rural_EV}$	0	0	0	0	0	0	0	0
$\beta_{town_rural_hybrid}$	0	0	0	0	0	0	town_rural	0
β_{price_inc}	$\frac{price_1 - 100}{income}$	$\frac{price_2 - 100}{income}$	$\frac{price_3 - 100}{income}$	$\frac{price_4 - 100}{income}$	$\frac{price_5 - 100}{income}$	$\frac{price_6 - 100}{income}$	$\frac{price_7 - 100}{income}$	$\frac{price_8 - 100}{income}$
β_{conso_inc}	$\frac{cons_1 - pd \cdot 100}{income}$	$\frac{cons_2 - pd \cdot 100}{income}$	$\frac{cons_3 - pd \cdot 100}{income}$	$\frac{cons_4 - pd \cdot 100}{income}$	$\frac{cons_5 - pd \cdot 100}{income}$	$\frac{cons_6 - pd \cdot 100}{income}$	$\frac{cons_7 - pd \cdot 100}{income}$	$\frac{cons_8 - pp \cdot 100}{income}$
β_{max_power}	$\frac{max_power_1}{10}$	$\frac{max_power_2}{10}$	$\frac{max_power_3}{10}$	$\frac{max_power_4}{10}$	$\frac{max_power_5}{10}$	$\frac{max_power_6}{10}$	$\frac{max_power_7}{10}$	$\frac{max_power_8}{10}$
β_{range_EV}	0	0	0	0	0	0	0	0

Model specification (2/2)

Parameter	9	10	11	12	13	14	15
ASC _{full}	0	0	0	0	0	0	0
ASC _{luxury}	1	0	0	0	0	0	0
ASC _{medium}	0	1	0	0	0	0	0
ASC _{MPV}	0	0	1	0	0	0	0
ASC _{offroad}	0	0	0	1	0	0	0
ASC _{petrol}	1	1	1	1	1	1	0
ASC _{electric}	0	0	0	0	0	0	1
β_{inc_full}	0	0	0	0	0	0	0
β_{inc_luxury}	$\frac{income}{10000}$	0	0	0	0	0	0
β_{inc_medium}	0	$\frac{income}{10000}$	0	0	0	0	0
β_{inc_MPV}	0	0	$\frac{income}{10000}$	0	0	0	0
$\beta_{inc_offroad}$	0	0	0	$\frac{income}{10000}$	0	0	0
β_{inc_hybrid}	0	0	0	0	0	$\frac{income}{10000}$	0
$\beta_{nbr_adults_small}$	0	0	0	0	nbr. adults	0	0
$\beta_{nbr_children_small}$	0	0	0	0	nbr. child.	0	0
$\beta_{nbr_cars_lux}$	nbr. cars	0	0	0	0	0	0
$\beta_{nbr_cars_hybrid}$	0	0	0	0	0	nbr. cars	0
$\beta_{university}$	0	0	0	0	0	1	1
$\beta_{town_rural_EV}$	0	0	0	0	0	0	town_rural
$\beta_{town_rural_hybrid}$	0	0	0	0	0	town_rural	0
β_{price_inc}	$\frac{price_9 \cdot 100}{income \cdot cons_9 \cdot pp \cdot 100}$	$\frac{price_{10} \cdot 100}{income \cdot cons_{10} \cdot pp \cdot 100}$	$\frac{price_{11} \cdot 100}{income \cdot cons_{11} \cdot pp \cdot 100}$	$\frac{price_{12} \cdot 100}{income \cdot cons_{12} \cdot pp \cdot 100}$	$\frac{price_{13} \cdot 100}{income \cdot cons_{13} \cdot pp \cdot 100}$	$\frac{price_{14} \cdot 100}{income \cdot cons_{14} \cdot pp \cdot 100}$	$\frac{price_{15} \cdot 100}{income}$
β_{conso_inc}	$\frac{income}{max_power_9}$	$\frac{income}{max_power_{10}}$	$\frac{income}{max_power_{11}}$	$\frac{income}{max_power_{12}}$	$\frac{income}{max_power_{13}}$	$\frac{income}{max_power_{14}}$	$\frac{income}{max_power_{15}}$
β_{max_power}	10	10	10	10	10	10	10
β_{range_EV}	0	0	0	0	0	0	$\frac{range_EV}{100}$