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
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


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 \([\text{l}/100\text{km}]\)
- Engine power \([\text{bhp}]\)
- Price after discounts and government schemes \([\text{€}]\)
- Reported range (EV) \([\text{km}]\)
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

\[ \text{Car type} = \text{market segment} + \text{fuel type} \]
Choice-set definition

Choice set

\[ \text{Car type} = \text{market segment} + \text{fuel type} \]

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

Fuel type
- Hybrid
- Diesel
- Petrol
- Electric
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

![Box plots showing estimates for various car types and alpha market segment.](image)
Substitution patterns
Market shares before and after a 50% increase in price of alternative 5 (diesel offroad)

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

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

Market segment

- full
- luxury
- medium
- MPV
- offroad
- hybrid
- small

1 8 2 9 3 10 4 11 5 12 7 14 6 15 13
Nesting structure 2

Fuel type
- diesel
  - 1
  - 2
  - 3
  - 4
  - 5
  - 6
  - 7
- electric
  - 15
- petrol
  - 14
  - 13
  - 12
  - 11
  - 10
  - 9
  - 8
Attributes of the car

Income

Parameter estimation: CNL
Parameter estimation: CNL

Socioeconomics

Estimated values for socioeconomics variables:
- PRICE_INC
- CONSO_INC
- MAX_POWER
- RANGE_EV

Dummy variables

Estimated values for dummy variables:
- ASC_fuel
- ASC_luxury
- ASC_medium
- ASC_mpv
- ASC_offroad
- ASC_hybrid
- ASC_electric
- ASC_petrol
Parameter estimation: Logit

Attributes of the car

Income
Parameter estimation: Logit

Socioeconomics

Dummy variables
## Model specification (1/2)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASCfull</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>ASCluxury</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>ASCmedium</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>ASCMPV</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>ASCoffroad</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>ASCpetrol</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>ASCelectric</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$\beta_{\text{inc}_\text{full}}$</td>
<td>income</td>
<td>00000</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$\beta_{\text{inc}_\text{luxury}}$</td>
<td>income</td>
<td>income</td>
<td>income</td>
<td>00000</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$\beta_{\text{inc}_\text{medium}}$</td>
<td>income</td>
<td>00000</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$\beta_{\text{inc}_\text{MPV}}$</td>
<td>income</td>
<td>00000</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$\beta_{\text{inc}_\text{offroad}}$</td>
<td>income</td>
<td>00000</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$\beta_{\text{inc}_\text{hybrid}}$</td>
<td>income</td>
<td>00000</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$\beta_{\text{nbr}_\text{adults_small}}$</td>
<td>income</td>
<td>00000</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$\beta_{\text{nbr}_\text{children_small}}$</td>
<td>income</td>
<td>00000</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$\beta_{\text{nbr}_\text{cars_lux}}$</td>
<td>income</td>
<td>00000</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$\beta_{\text{nbr}_\text{cars_hybrid}}$</td>
<td>income</td>
<td>00000</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$\beta_{\text{university}}$</td>
<td>income</td>
<td>00000</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$\beta_{\text{town}_\text{rural_EV}}$</td>
<td>income</td>
<td>00000</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$\beta_{\text{town}_\text{rural_hybrid}}$</td>
<td>income</td>
<td>00000</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$\beta_{\text{price}_\text{inc}}$</td>
<td>price</td>
<td>10000</td>
<td>price</td>
<td>10000</td>
<td>price</td>
<td>10000</td>
<td>price</td>
<td>10000</td>
</tr>
<tr>
<td>$\beta_{\text{conso}_\text{inc}}$</td>
<td>price</td>
<td>10000</td>
<td>price</td>
<td>10000</td>
<td>price</td>
<td>10000</td>
<td>price</td>
<td>10000</td>
</tr>
<tr>
<td>$\beta_{\text{max}_\text{power}}$</td>
<td>price</td>
<td>10000</td>
<td>price</td>
<td>10000</td>
<td>price</td>
<td>10000</td>
<td>price</td>
<td>10000</td>
</tr>
<tr>
<td>$\beta_{\text{range}_\text{EV}}$</td>
<td>price</td>
<td>10000</td>
<td>price</td>
<td>10000</td>
<td>price</td>
<td>10000</td>
<td>price</td>
<td>10000</td>
</tr>
</tbody>
</table>

A. Fernández Antolín (TRANSP-OR EPFL)
<table>
<thead>
<tr>
<th>Parameter</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASC_full</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>ASC_luxury</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>ASC_medium</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>ASC_MPV</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>ASC_offroad</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>ASC_petrol</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>ASC_electric</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>β_{inc_full}</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>β_{inc_luxury}</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>β_{inc_medium}</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>β_{inc_MPV}</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>β_{inc_offroad}</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>β_{inc_hybrid}</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>β_{nbr_adults_small}</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>β_{nbr_chilren_small}</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>β_{nbr_cars_Lux}</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>β_{nbr_cars_hybrid}</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>β_university</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>β_{town_rural}</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>β_{price_inc}</td>
<td>price_{9,100}</td>
<td>price_{10,100}</td>
<td>price_{11,100}</td>
<td>price_{12,100}</td>
<td>price_{13,100}</td>
<td>price_{14,100}</td>
<td>price_{15,100}</td>
</tr>
<tr>
<td>β_conso_inc</td>
<td>price_{cons9,pp,100}</td>
<td>price_{cons10,pp,100}</td>
<td>price_{cons11,pp,100}</td>
<td>price_{cons12,pp,100}</td>
<td>price_{cons13,pp,100}</td>
<td>price_{cons14,pp,100}</td>
<td>price_{cons15,pp,100}</td>
</tr>
<tr>
<td>β_max_power</td>
<td>max_power_{9}</td>
<td>max_power_{10}</td>
<td>max_power_{11}</td>
<td>max_power_{12}</td>
<td>max_power_{13}</td>
<td>max_power_{14}</td>
<td>max_power_{15}</td>
</tr>
<tr>
<td>β_range_EV</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>