Conduite de Projet

Méthodologie

Master 1 SE

Youssef El Yaakoubi, Nathalie Picard

The Nested Logit model

What is the purpose of the Nested Logit model ?

The Red bus/Blue bus paradox

- The logit model is based on the **i.i.d assumption**
  $\longrightarrow$ identically and independently distributed error terms
- Mathematically convinient but may lead to counter-intuitive forecasts

- Let's say that a Bus company wants us to analyze the choice of commuters bewteen buses and cars
- The color of their buses is **only blue** but they planned to buy red buses and they want to know if it will make a difference

The Red bus/Blue bus paradox

- Let's assume that individuals face the same travel time by car and bus and it's the only variable in the model :

Model :

$$U_{car} = \beta T + \varepsilon_{car}$$

$$U_{bus} = \beta T + \varepsilon_{bus}$$

Choice probability :

$$P(car|\{car, bus\}) = \frac{e^{\beta T}}{e^{\beta T} + e^{\beta T}} = \frac{1}{2}$$

The Red bus/Blue bus paradox

- Now, let's apply the same model with the purchase of those new red buses :

Model :

$$U_{car} = \beta T + \varepsilon_{car}$$
$$U_{bluebus} = \beta T + \varepsilon_{bluebus}$$
$$U_{redbus} = \beta T + \varepsilon_{redbus}$$

Choice probability :

$$P(car|\{car, bus\}) = \frac{e^{\beta T}}{e^{\beta T} + e^{\beta T} + e^{\beta T}} = \frac{1}{3}$$

- So, if the company introduce red buses and blue buses in the city, **the share of private cars drops from 50% to 33%** $\rightarrow$ obviously wrong result

Why the logit model failed to forecast the right shares ? :

- Only travel time was included in the model
- So, all other attributes are captures by the error terms
- $\varepsilon_{\mathtt{bluebus}}$ and $\varepsilon_{\mathtt{redbus}}$ share common unobserved attributes (type of transport, comfort, fare, etc.)

Logit model :

- Assumes that $\varepsilon_{\mathtt{bluebus}}$ and $\varepsilon_{\mathtt{redbus}}$ are independent $\longrightarrow$ we saw that it is inappropriate here
- The logit model cannot be applied when alternatives share unobserved attributes
- Solution : add some structure with a Nested Logit modelisation

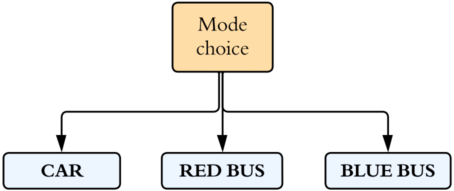


Figure 1: Logit model

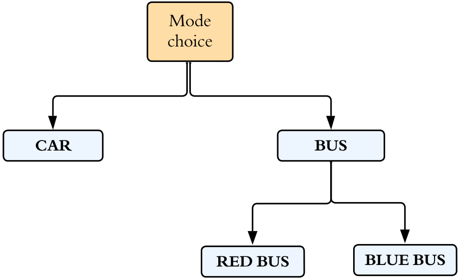


Figure 2: Nested Logit model

Model :

$$U_{car} = \beta T + \varepsilon_{car}$$

$$U_{bluebus} = \beta T + \varepsilon_{bluebus} + \varepsilon_{bus}$$

$$U_{redbus} = \beta T + \varepsilon_{redbus} + \varepsilon_{bus}$$

- ε_{bus} captures all unobserved attributes that are common to the red bus and the blue bus
- while $\varepsilon_{bluebus}$ captures unobserved attributes specific to blue buses (same for red buses)

Within the nest bus :

$$P(\text{blue bus}|\{\text{blue bus}, \text{red bus}\}) = \Pr(\beta T + \varepsilon_{bluebus} + \cancel{\varepsilon_{bus}} \geq \beta T + \varepsilon_{redbus} + \cancel{\varepsilon_{bus}})$$

- We can now assume that $\varepsilon_{bluebus}$ and ε_{redbus} are i.i.d EV

Choice probability :

$$P(\text{blue bus}|\{\text{blue bus}, \text{red bus}\}) = \frac{e^{\beta T}}{e^{\beta T} + e^{\beta T}} = \frac{1}{2}$$

The Red bus/Blue bus paradox

Resulting from the nest bus :

$C \in \{\text{blue bus, red bus}\}$

$\tilde{V}_{bus} = \ln \sum_{j \in C} e^{V_j/\sigma_{bus}}$

Across nests :

$$U_{car} = \beta T + \varepsilon_{car}$$

$$U_{bus} = \tilde{V}_{bus} + \tilde{\varepsilon}_{bus}$$

Choice probability :

$$P(car) = \frac{e^{\beta T/\sigma}}{e^{\beta T/\sigma} + e^{\frac{\sigma_{bus}}{\sigma}\tilde{V}_{bus}}}$$

The Red bus/Blue bus paradox

Usually, one of the scale parameters is normalized to 1 :

$$P(car) = \frac{e^{\beta T}}{e^{\beta T} + e^{\sigma_{bus} \tilde{V}_{bus}}}$$

The choice between bus and car depends on the ratio $\frac{\sigma_{bus}}{\sigma}$

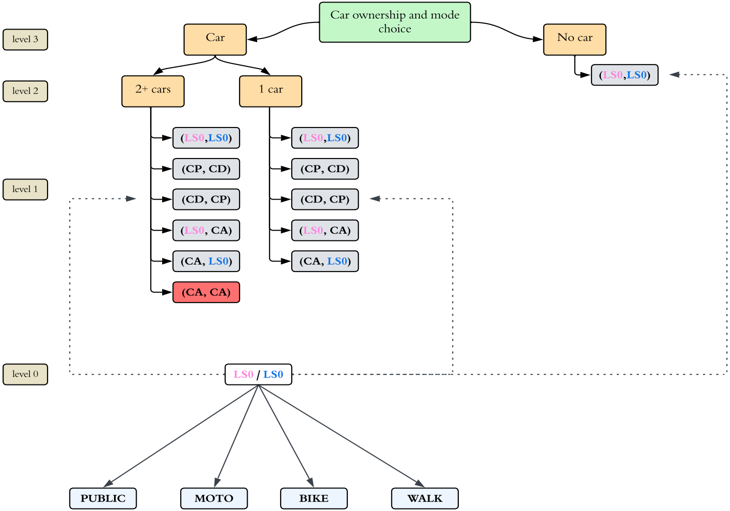
so, if $\frac{\sigma_{bus}}{\sigma} \rightarrow 1$:

$$P(car) = \frac{1}{2}$$

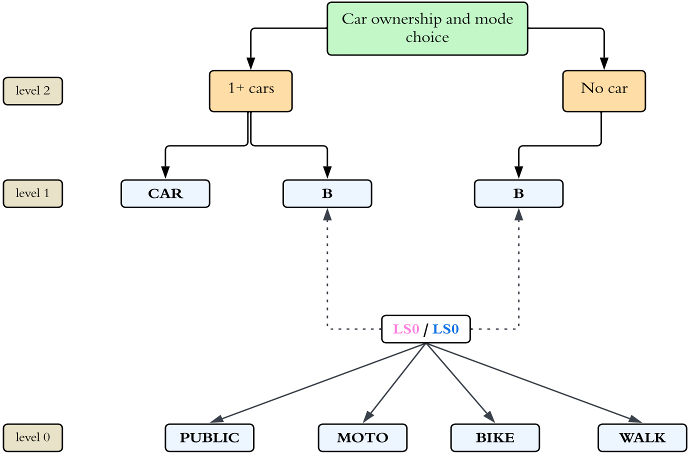
$\rightarrow$ in this case, for the point of view of the decision maker, **red bus and blue bus are exactly the same** : they are fully correlated $\rightarrow$ same as applying a logit

Notes: σ_{bus} and σ are the scale parameters of the bottom and the top respectively and they are estimated from data. One of them is normalized to 1

Study of car ownership and mode choice for commuting to work


level 3
level 2
level 1
level 0
Car ownership and mode choice
Car
No car
2+ cars
1 car
(LS0,LS0)
(CP, CD)
(CD, CP)
(LS0, CA)
(CA, LS0)
(CA, CA)
(LS0,LS0)
(CP, CD)
(CD, CP)
(LS0, CA)
(CA, LS0)
(LS0,LS0)
LS0 / LS0
PUBLIC
MOTO
BIKE
WALK

3-level Nested Logit Model for Singles and Mono-actives



Mode choice (level 0)

Choice between :

$$j \in \{\text{PUBLIC TRANSPORT}, \text{MOTO}, \text{BIKE}, \text{WALK}\}$$

Utility functions :

$$V^j_{i,G} = \alpha^j_G + \beta^j_{i,G}\boldsymbol{X}^j_{i,G} - \left(\exp(\delta^j_{i,G})\right)Z^j_{i,G} + \gamma^j_{i,G}Z^2_{i,G}$$

$$Z^j_{i,G} = \begin{cases} \text{Distance} & \text{if } j \in \{\text{BIKE}, \text{WALK}\} \\ \text{Travel time} & \text{if } j = \text{PUBLIC TRANSPORT} \\ \text{Free-flow travel time} & \text{if } j = \text{MOTO} \end{cases}$$

Probabilities :

$$P_{i,G}(j/K) = \frac{e^{V^j_{i,G}/\sigma_{G,0}}}{\sum_K e^{V^k_{i,G}/\sigma_{G,0}}}$$

Mode choice (level 1) - Individual utilities

$$V_{i,G}^{C^A} = -tt_{i,G}^{C} \cdot VOT_{i,G}^{C^A}$$

$$V_{i,G}^{C^P} = \delta_{\alpha,G}^{C^P} + \boldsymbol{X}_{G}^{C^P} \boldsymbol{\beta}_{G}^{C^P} - VOT_{i,G}^{C^A} \cdot \exp\left(\delta_{VOT,G}^{C^P}\right) \cdot tt_{i,G}^{C}$$

$$V_{i,G}^{C^D} = \delta_{\alpha,G}^{C^D} + \boldsymbol{X}_{G}^{C^D} \boldsymbol{\beta}_{G}^{C^D} - VOT_{i,G}^{C^A} \cdot \exp\left(\delta_{VOT,G}^{C^D}\right) \cdot tt_{i,G'}^{C} - VOT_{i,G}^{C^A} \cdot tt_{i,G'G}^{C}$$

$$V_{i,G}^{B} = \delta_{\alpha,G}^{B} + \boldsymbol{X}_{\mathbf{G}}^{\mathbf{B}} \boldsymbol{\beta}_{\mathbf{G}}^{\mathbf{B}} + \tilde{\boldsymbol{V}}_{\mathbf{G},l}$$

with $l = 0$ we have,

$$\tilde{\boldsymbol{V}}_{\mathbf{G},l} = \sigma_l \cdot \log\left(\sum_{j \in \{PT,MOTO,BIKE,WALK\}} \exp(V_{G,l,j}/\sigma_l)\right)$$

Woman's Pareto weight:

$$\lambda_i = \frac{1}{1+\exp(-\boldsymbol{X}_i^{\lambda}\beta^{\lambda})} \in [0,1]$$

Estimated value of time:

$$\mathrm{VOT}_{i,G}^{j} = \exp(\alpha_{i,G}^{j} + \boldsymbol{X}_{\mathbf{G}}^{\mathbf{VOT,j}}\boldsymbol{\beta}_{\mathbf{G}}^{\mathbf{VOT,j}})$$

Each spouse's utilities:

$$V_{i,G}^{j} = \delta_{\alpha,G}^{j} + \beta_{i,G}^{j}\boldsymbol{X}_{i,G}^{j} - tt_{i,G}^{j}\mathrm{VOT}_{i,G}^{j}$$

Couple utilities:

$$V_i^{j,j'} = \lambda_i V_{i,W}^{j} + (1-\lambda_i)V_{i,M}^{j'}$$

With:

- $j \in \{B, C^A, C^D, C^P\}$
- $tt_{i,G}^{C} = tt_{i,G}^{C^A} = tt_{i,G}^{C^P}$
- $tt_{i,G}^{C^D} = tt_{i,G'}^{C} + tt_{i,G'G}^{C}$
- $\mathrm{VOT}_{i,G}^{j} = \mathrm{VOT}_{i,G}^{C^A}\exp(\delta_{\mathrm{VOT},G}^{j})$

  if the reference being Car Alone (C^A)
- with $\exp(\delta_{\mathrm{VOT},G}^{j}) = \frac{\mathrm{VOT}^{j}}{\mathrm{VOT}^{C^A}}$

Mode choice (level 1) - Probabilities

When #cars ≥ 2

$$P_i(C, C) = \frac{e^{(V_i^{C^A, C^A}/\sigma_{1,2})} + e^{(V_i^{C^D, C^P}/\sigma_{1,2})} + e^{(V_i^{C^P, C^D}/\sigma_{1,2})}}{\Sigma_{1,2}}$$

$$P_i(B, B) = \frac{e^{(V_i^{B,B}/\sigma_{1,2})}}{\Sigma_{1,2}}$$

$$P_i(C^A, B) = \frac{e^{(V_i^{C^A,B}/\sigma_{1,2})}}{\Sigma_{1,2}}$$

$$P_i(B, C^A) = \frac{e^{(V_i^{B,C^A}/\sigma_{1,2})}}{\Sigma_{1,2}}$$

When #cars $= 1$

$$P_i(C, C) = \frac{e^{(V_i^{C^D, C^P}/\sigma_{1,1})} + e^{(V_i^{C^P, C^D}/\sigma_{1,1})}}{\Sigma_{1,1}}$$

$$P_i(B, B) = \frac{e^{(V_i^{B,B}/\sigma_{1,1})}}{\Sigma_{1,1}}$$

$$P_i(C^A, B) = \frac{e^{(V_i^{C^A,B}/\sigma_{1,1})}}{\Sigma_{1,1}}$$

$$P_i(B, C^A) = \frac{e^{(V_i^{B,C^A}/\sigma_{1,1})}}{\Sigma_{1,1}}$$

Choice to buy a second car (level 2)

Resulting from level 1:

$$\tilde{V}_{1,1} + \tilde{\varepsilon}_{1,1} = \sigma_{1,1} \ln \left(e^{\frac{V^{B,B}}{\sigma_{1,1}}} + e^{\frac{V^{B,C^A}}{\sigma_{1,1}}} + e^{\frac{V^{C^A,B}}{\sigma_{1,1}}} + e^{\frac{V^{C^D,C^P}}{\sigma_{1,1}}} + e^{\frac{V^{C^P,C^D}}{\sigma_{1,1}}} \right) + \tilde{\varepsilon}_{1,1}$$

$$\tilde{V}_{1,2} + \tilde{\varepsilon}_{1,2} = \sigma_{1,2} \ln \left(e^{\frac{V^{B,B}}{\sigma_{1,2}}} + e^{\frac{V^{B,C^A}}{\sigma_{1,2}}} + e^{\frac{V^{C^A,B}}{\sigma_{1,2}}} + e^{\frac{V^{C^D,C^P}}{\sigma_{1,2}}} + e^{\frac{V^{C^P,C^D}}{\sigma_{1,2}}} + e^{\frac{V^{C^A,C^A}}{\sigma_{1,2}}} \right) + \tilde{\varepsilon}_{1,2}$$

Couple's utility at level 2:

$$\begin{cases} V_{1car} = \tilde{V}_{1,1} + V_{2,1} + \underbrace{\tilde{\varepsilon}_{1,1} + \varepsilon_{2,1}}_{\sigma_2 \varepsilon'_{2,1}} \\ V_{2cars} = \tilde{V}_{1,2} + V_{2,2} + \underbrace{\tilde{\varepsilon}_{1,2} + \varepsilon_{2,2}}_{\sigma_2 \varepsilon'_{2,2}} \end{cases}$$

$$\frac{\tilde{V}_{1,2} + V_{2,2}}{\sigma_2} - \frac{\tilde{V}_{1,1} + V_{2,1}}{\sigma_2} + \varepsilon'_{2,2} - \varepsilon'_{2,1}$$

Choice to buy the first car (level 3)

Resulting from level 2:

$$\tilde{V}_{2,1+} + \varepsilon'_{2,1+} = \sigma_2 \ln\left(e^{\frac{\tilde{V}_{1,2}+V_{2,2}}{\sigma_2}} + e^{\frac{\tilde{V}_{1,1}+V_{2,1}}{\sigma_2}}\right) + \varepsilon'_{2,1+}$$

Couple's utility at level 3:

$$\begin{cases} V_{NoCar} = V_1(B,B) + \varepsilon'_{1,1} \\ V_{1+car} = \tilde{V}_{2,1} + V_{3,1+} + \underbrace{\varepsilon'_{2,1+} + \varepsilon_{3,1}}_{\sigma_3 \varepsilon'_{3,1+}} \end{cases}$$

The resulting binary logit between the choice to buy a car or not:

$$\frac{\tilde{V}_{2,1} + V_{3,1+}}{\sigma_3} - \frac{V_1(B,B)}{\sigma_3} + \varepsilon'_{3,1+} - \varepsilon'_{1,1}$$

Data

- French population's Census data from **2007 to 2021** (INSEE)
  - Not a Panel-data
  - Rolling sample of households/communes: each year all households in a fraction of small communes and a fraction of households in each large commune
  - In this course we study only the Paris region (île-de-France) through several years
- Data are available here
- Codes are available here

References I

[1] Pierre-André Chiappori. "Rational household labor supply". In: *Econometrica: Journal of the Econometric Society* (1988), pp. 63–90.

[2] de Palma and Picard N. Ziegelmeyer. "Individual and couple decision behavior under risk: evidence on the dynamics of power balance". In: *Theory Decision* (2011), pp. 45–64.

[3] Inoa, Picard, and de Palma. "Effect of an Accessibility Measure in a Model for Choice of Residential Location, Workplace, and Type of Employment". In: *Mathematical Population Studies* (2014), pp. 4–36.

[4] Picard, Dantan, and de Palma. "Mobility decisions within couples". In: *Theory Decision* (2018), pp. 149–180.

[5] Pierre-André Chiappori. "Collective labor supply and welfare". In: *Journal of political Economy* 100.3 (1992), pp. 437–467.

[6] Bhat and Pendyala. "Modeling intra-household interactions and group decision-making". In: *Transportation* 32 (2005), pp. 443–448.

References II

[7] Ho and Mulley. “Intra-household interactions in transport research: a review”. In: *Transport Reviews* 35.1 (2015), pp. 33–55. ISSN: 0144-1647.

[8] Kroesen. “Modeling the behavioral determinants of travel behavior: An application of latent transition analysis”. In: *Transportation Research Part A: Policy and Practice* 65 (2014), pp. 56–67. ISSN: 0965-8564.

[9] Litman and Todd. “Land Use Impacts on Transport How Land Use Factors Affect Travel Behavior”. In: (Jan. 2008).

[10] Borgers and Timmermans. “Transport facilities and residential choice behavior: A model of multi-person choice processes”. In: *Papers in Regional Science* 72.1 (1993), pp. 45–61. ISSN: 1056-8190.

[11] Timmermans et al. “Residential Choice Behaviour of Dual Earner Households: A Decompositional Joint Choice Model”. In: *Environment and Planning A: Economy and Space* 24.4 (1992), pp. 517–533.

[12] Leon Festinger. “A theory of cognitive dissonance”. In: *Row and Peterson* (1957).

References III

[13] Javaudin and de Palma. "METROPOLIS2: Bridging Theory and Simulation in Agent-Based Transport Modeling". In: (2024).

[14] Geoffrey R. Dunbar, Arthur Lewbel, and Krishna Pendakur. "Children's Resources in Collective Households: Identification, Estimation, and an Application to Child Poverty in Malawi". In: *American Economic Review* 103.1 (Feb. 2013), pp. 438–71.