

CHOICE MODELING IN R

FINAL PROJECT

Wiktor Budziński

FINAL PROJECT

The final project consists of conducting an empirical choice modeling analysis using the **Apollo package in R** on a dataset of your choice

Your analysis should include:

- A baseline model (Multinomial Logit, MNL)
- At least one model capturing preference heterogeneity:
 - Latent Class Model (LCM),
 - Or Mixed Logit (Random Parameters Logit)
- Exploration of model specification:
 - Different numbers of classes (LCM),
 - Different distributions of random parameters (Mixed Logit)
 - Functional form (linear vs. nonlinear attributes)
- One behavioral extension, such as:
 - Attribute non-attendance
 - Random regret minimization
 - Hybrid choice model
 - This behavioral component should be combined with preference heterogeneity

Models should include a mix of:

- Alternative-specific variables (attributes)
- Individual-specific variables (e.g., socio-demographics)

FINAL PROJECT

The project should take the form of a research analysis addressing a clear question or hypothesis.

This can be:

- Applied:
 - What drives preferences for [product/service/policy]?
 - How do socio-demographic characteristics affect choice behavior?
- Methodological:
 - How does the number of alternatives affect estimated preferences?
 - Which individuals are more likely to follow random regret vs. random utility?

The research question should guide:

- Model specification
- Interpretation of results

FINAL PROJECT

You should submit a **report in a pdf format** with description of the conducted analysis

- Introduction and motivation (Brief)
- Description of the data
- Model specifications
- Results and interpretation (**Main part**)
- Conclusions and key findings (Brief)

The report should be accompanied by an .R file with the codes you used for the analysis

- Use comments to describe briefly what are you doing
- Code should allow to reproduce the results reported in the pdf report

Deadlines:

- 11th of May – please send me what data you want to work on and what do you want to investigate
- 28th of June – submit the final project
- Send them to my email: wiktor_budzinski@uw.edu.pl

FINAL PROJECT

Evaluation will be based on:

- Correct specification and estimation of models
- Code clarity, structure, and reproducibility
- How well the conducted analysis addresses the selected research question
- Quality of interpretation and presentation of results
- Overall quality and clarity of the written report

DATA SOURCES

General repositories

- Mendeley database: <https://data.mendeley.com/>
- ICPSR database: <https://www.icpsr.umich.edu/sites/icpsr/home>
- Harvard dataverse: <https://dataverse.harvard.edu/>
- Search for “stated preference”, “choice experiment” etc.

The Discrete Choice Data Bank

- [GitHub - alvarogutierrez/TheDiscreteChoiceDataBank: Collection of publicly available Discrete Choice Datasets · GitHub](#)

Datasets from Mikolaj Czajkowski's research

- [Mikołaj Czajkowski - Supplementary materials](#)

EPFL Discrete Choice Data Collection

- <https://transp-or.epfl.ch/discretechoice/data.html>