

## 2.3 Mode choice - Choice modelling (Due on 17.11.2025)

Suggested section length: ~ 5.

In this exercise you will estimate mode choice models with the "apollo" package. The package is well maintained, has a manual, a FAQ and a forum where virtually all questions on modelling are answered. Check: <http://www.apollochoicemodelling.com>

You have to include the following alternatives in the model:

1. Walk
2. Bicycle
3. Public transport (PT)
4. Car

Your utility functions should at least include:

- Alternative-specific constants: You have to define a reference alternative yourself.
- Alternative-specific travel time parameters.
- Alternative-specific parameters for the rest of public transport related variables
- A generic cost parameter common for all alternatives.

We prepared a dataset containing all variables for chosen and non-chosen mode alternative trips in the microcensus with over 200m of crowfly distance. This is the starting point for the estimation. But note that not all alternatives are available for each trip!

Mode	Variable name	description
Car/individual motorized travel	totalTravelTime_car	travel time (min)
	cost_car	variable fuel travel costs (CHF)
Public transport: <i>The public transport variables are the median values of all possible connections within a 2h time frame starting from the departure time.</i>	access_time	total time spent accessing public transport (min)
	wait_time	total time spent waiting for first connection (min)
	transfers_pt	number of transfers
	transferWaitingTime_pt	total transfer time (min)
	frequency_pt	headway of possible connections (min)
	totalTravelTime_pt	sum of access, wait, transfer, egress time and the in-vehicle time (min)
	cost_pt	Estimation of PT ticket cost. This value already accounts for the subscription available.
Cycling	totalTravelTime_bike	cycling travel time (accounts for topography) in min
	conditions/icon	weather conditions at time of travel
Walking	totalTravelTime_walk	walking travel time (min)
	conditions/icon	weather conditions at time of travel

Further we generated the following availability variables: avail\_pt, avail\_car, avail\_walk and avail\_bike and a column "CHOICE" with the revealed mode choice from the MZMV where car=1, pt=2, bike=3, walk=4.

You have to do the following things:

1. Read the data and split the data into 80% for estimation and 20% testing
2. Define a first model that explains mode choice
  - stepwise you can make it more and more complex
  - Eventually include sociodemographic information and possibly interact those with supply

attributes of the modes.

3. Test the model's prediction performance with train and test datasets
4. Calculate the value of travel time (VTT) for car and public transport alternatives.
5. Calculate the direct elasticities for all cost and travel time parameters as well as at least two cross elasticities of your choice.
6. Lastly, apply your model to the results from the gravity model (trip distribution section 2.2) in order to find the modal split for commuting trips. Therefore, make sure that you take the average of the probabilities calculated for each alternative per trip for each origin-destination pair of zones.

Describe all your modelling steps in detail. Include an output table for the final MNL model, discuss important coefficients of your model. Here again: One model is no model, so develop your model and show your model development steps. So for example, you want to test sociodemographic parameters and their interactions with attributes of the alternatives. A good reference how to structure your code can be found here:

<https://cran.r-project.org/web/packages/apollo/vignettes/apollofirstexample.R>