

CHOICE MODELING IN R

CLASS 6

Wiktor Budziński

	Alternative 1	Alternative 2	Alternative 3	Alternative 4
Protection of ecologically valuable forests	 Status quo Passive protection of 50% of the most ecologically valuable forests (1.5% of all forests)	 Status quo Passive protection of 50% of the most ecologically valuable forests (1.5% of all forests)	 Status quo Passive protection of 50% of the most ecologically valuable forests (1.5% of all forests)	 Substantial improvement Passive protection of 100% of the most ecologically valuable forests (3% of all forests, 100% increase)
Litter in forests	 Status quo No change in the amount of litter in the forests	 Partial improvement Decrease the amount of litter in the forests by half (50% reduction)	 Status quo No change in the amount of litter in the forests	 Partial improvement Decrease the amount of litter in the forests by half (50% reduction)
Infrastructure	 Status quo No change in tourist infrastructure	 Status quo No change in tourist infrastructure	 Partial improvement Appropriate tourist infrastructure in an additional 50% of the forests (50% increase)	 Substantial improvement Appropriate tourist infrastructure available in twice as many forests (100% increase)
Cost	0 PLN	10 PLN	25 PLN	100 PLN
Your choice	☐	☐	☐	☐

Read the *Forests.xlsx* data into R

- Transform it into the wide format using codes from the last class

Estimate latent class models with different number of classes

- Select the one with best ICs
- Interpret the results

Add *ForArea* variable to explain class membership probability in the best model

- Area of the forest in the 10x10km grid
- Calculate WTP for the final model

EXERCISE

CHOICE MODELING

ATTRIBUTE NON-ATTENDANCE

ATTRIBUTE NON-ATTENDANCE

In DCE we present to respondents some hypothetical scenarios

They are usually designed beforehand so that attributes included in the scenarios were relevant to the public

- These are chosen based on literature reviews, focus groups and pilot studies

The usual assumption is that individuals will evaluate presented choice sets and make decisions accordingly with their preferences

ANA refers to the situation when some respondents do not take into account one or more of the presented attributes when making choices

- This will show in the model as individuals' coefficients being equal to 0

This is also very likely in RP data

ATTRIBUTE NON-ATTENDANCE

There could be several different reasons for ANA

- Some respondent can simply not care about some attributes and have 0 marginal utility for them, although we would not expect that for the cost
- It could be used as simplifying heuristic if DCE design is too complex
- Respondents could consider some attributes as not realistic and ignore them
- Respondents could consider survey to have little consequences and therefore make little effort in their choices
- There is probably more...

ATTRIBUTE NON-ATTENDANCE

There could be several different reasons for ANA

- Some respondent can simply not care about some attributes and have 0 marginal utility for them, although we would not expect that for the cost
- It could be used as simplifying heuristic if DFE design is too complex
- Respondents could consider some attributes as not realistic and ignore them
- Respondents could consider survey to have little consequences and therefore make little effort in their choice
- There is probably

This is not a problem in general, although we should account for it in the modelling

ATTRIBUTE

These are problematic for estimation of WTP and making inference from the DCE

There could be several different reasons for ANA

- Some respondent can simply not care about some attributes and have 0 marginal utility for them, although we would not expect that for the cost
- It could be used as simplifying heuristic if DCE design is too complex
- Respondents could consider some attributes as not realistic and ignore them
- Respondents could consider survey to have little consequences and therefore make little effort in their choices
- There is probably more...

ATTRIBUTE NON-ATTENDANCE

There is no clear guidance regarding how to differentiate between different reasons for ANA

- We can always ask about it after DCE in the survey

Table 3 Example of follow-up question regarding why respondents ignored specific attributes—in this case for the attribute ‘Catch opportunity’

What was your primary cause for ignoring this attribute? (Tick only one)	Catch opportunity
It is not important to me	☐
It made it easier to choose between the alternatives	☐
The levels for the attribute were unrealistically high/low	☐
I don’t think that this attribute should be weighed against the others	☐
Don’t know	☐

ATTRIBUTE NON-ATTENDANCE

In most basic setting ANA is modelled as LCM in which every parameter can either be equal to some estimated value or to 0

- This gives 2^K classes, where K is the no. of attributes
- We can either assume that probability of ANA for each attribute is independent (which leads to K additional coefficients)
 - Discrete mixture (DM)
- Or we can assume that they are correlated (which leads to $2^K - 1$ additional coefficients)

ATTRIBUTE NON-ATTENDANCE

In most basic setting ANA is modelled as LCM in which every parameter can either be equal to some estimated value or to 0

- This gives 2^K classes, where K is the no. of attributes
- We can either assume that probability of ANA for each attribute is independent (which leads to K additional coefficients)
- Or we can assume that they are correlated (which leads to $2^K - 1$ additional coefficients)

Let's go to



ATTRIBUTE NON-ATTENDANCE

Potential problems with ANA:

- If preference heterogeneity is not accounted for in any way, then it is likely that we will capture preference heterogeneity instead of ANA
- Even if we add random parameters then it is possible that part of the ANA effect will be driven by assumptions regarding preference distribution. It is advised to compare the results across different specifications
- ANA can be also confounded with low preferences for a given attribute
 - For example, if the lowest bid in DCE is 5 PLN then we may have problem with identifying WTP of 0.5 PLN, and count it as 0

Choice Situation 1.	Alternative 1	Alternative 2	Alternative 3
Method of sorting in household	Into 5 categories	Into 2 categories	None
Frequency of collection	Once every 4 weeks	Once every 2 weeks	Once every week
Monthly cost for your household	75 PLN	50 PLN	100 PLN
Your choice:	☐	☐	☐

Estimate ANA models on recycling data

- LCM specification
- DM specification (simplified LCM)
- DM with preference heterogeneity

EXERCISE

Let's go to

