Supplementary Material for “Assessing Student Preferences for Quantitative Methods Courses Using Discrete Choice Experiments”

# STEPS OF CREATING THE DISCRETE CHOICE EXPERIMENT

The following is a detailed description of the different steps to create a discrete choice experiment. The different steps are discussed in terms of determining a suitable discrete choice experiment to assess sociology students’ quantitative methods course preferences.

### A) CHOOSING THE ATTRIBUTES FOR EXPERIMENTAL MANIPULATION

As with any research design, the first step involves matching the research question at hand to a suitable design. In the present case, we are interested in the question of what makes quantitative methods courses more attractive to sociology students, which helps addressing statistical anxiety or low engagement in such courses. Against the background outlined in the article and since we shouldn’t (cognitively) overwhelm our respondents by manipulating too many characteristics (Auspurg & Hinz, 2015), the following six aspects were chosen for experimental manipulation: The number of credit points students get, the prerequisites for the course, its schedule, the form of assessment, the teaching methods, and whether it is a course that can be taken by Bachelor’s and Master’s students alike.

Since students will be presented with hypothetical courses (vignettes), a valid design requires that they have sufficient information to make a decision. Otherwise, they might resort to unobserved heuristics which could, in the worst case, bias the results (Zangger & Becker, 2019). To ensure that students base their decision only on the information presented to them, they were informed that the hypothetical course descriptions in each choice set represent different arrangements of the same course and that the alternatives do not differ other than by the dimensions specified in the vignettes.

### B) CHOOSING THE LEVELS OF THE ATTRIBUTES

In a second step, we need to specify the different levels (values) of all the attributes in the experiment. To this end, we consider a qualitative and a quantitative aspect. Starting with the latter, we have to decide how many levels we need for each of the attributes. Foremost, this question is driven by the underlying research interest. However, there are good reasons not to include too many different levels for a dimension since the full factorial (universe of all possible treatment combinations) increases exponentially with the number of levels, making it harder to find an efficient experimental design (Auspurg & Hinz, 2015). If the dimension to manipulate can be conceived as a metric variable, for example, the credit points students get for a course, an ex-ante consideration of the functional form of the effect can guide the decision. If we assume linear effects, two levels suffice, three allow for the estimation of (inversely) U-shaped effects, four for an S-shaped relationship with the outcome. Finally, ensuring that all dimensions in the experiment have a similar number of levels or multiples of each other results in more efficient designs (Auspurg & Hinz, 2015).

From a qualitative point of view, we must consider the exact values of the experimentally manipulated attributes. They should introduce enough variability to detect treatment effects while simultaneously not being too extreme such that they would dominate all other aspects of interest (Johnson et al., 2013). For example, using the three values “1 ECTS”, “2 ECTS”, and “8 ECTS” for the credit points, the discriminatory power between the first two categories is likely too small whereas the third level could become a dominating factor that could drive students’ decision, irrespective of other dimensions that are manipulated in the experiment.

In the current study, the levels summarized in Table 1 in the main article were chosen for the six dimensions of interest. The design is rather minimalistic and only two of the 6 attributes (prerequisites and mode of assessment) have three levels. Together, they yield a full factorial of different treatment combinations. For the validity of the experiment, it is further crucial to exclude illogical or implausible combinations of the experimentally altered characteristics during the construction of the design. For a detailed discussion of this topic, see Auspurg & Hinz (2015) and Kuhfeld (2010). In the present case, no combinations among the characteristics in Table 1 were considered illogical or implausible.

144 possible treatment combinations are still too many for the application in even large classrooms, since each vignette (arrangement of attributes) should be evaluated by at least five respondents (Auspurg & Hinz, 2015). Thus, we need to find a design that works with a fraction of the vignette universe and that still allows for the unbiased estimation of treatment effects.

### C) FINDING A D-EFFICIENT DESIGN

Algorithmic searches are the most flexible way to find a suitable fractional factorial (Kuhfeld, 1997). In this regard, we must consider two questions:

* How large should the fractional factorial be?
* What kind of interaction effects do we expect among the experimentally manipulated dimensions?

To address the two questions, we use the SAS-macros developed by Warren Kuhfeld (2010).[[1]](#footnote-2) All the code can be found in Appendix A; the data are available at [blinded for review]. In a first step, we determine suitable sizes for a fractional factorial using the *%mktruns* macro. It requires the number of levels of all the dimensions that we want to manipulate (here: *%mktruns(2 3 2 3 2 2)*). The smallest possible design is called *saturated design*. It is defined by the degrees of freedom, accounting for one reference group in each of the dimensions. In the present case, the saturated design would comprise 9 out of the 144 different treatment combinations (1+2+1+2+1+1 and an additional parameter for the intercept).

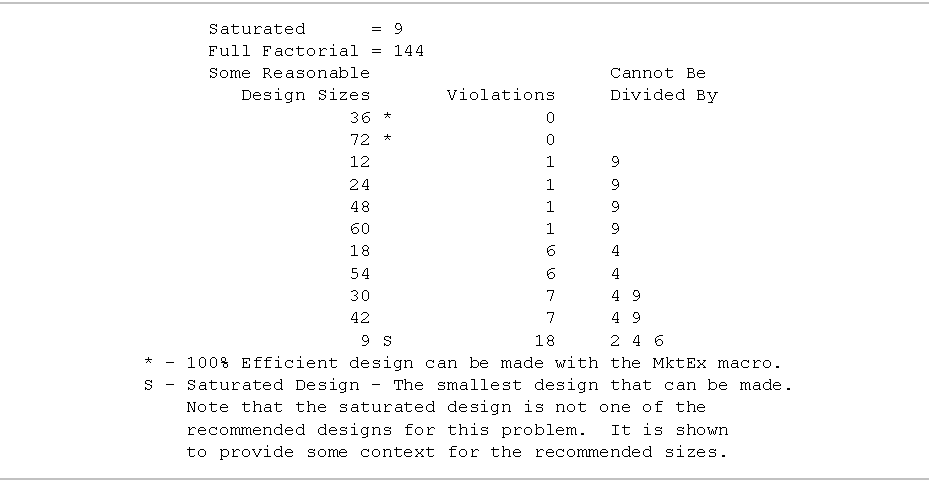


Figure S1: Output from %mktruns

From Figure S1 we infer that, in the present context, a sample size of 36 or 72 out of the 144 total vignettes would result in fully efficient, orthogonal designs. However, this number is still too large for many settings (e.g., when assessing the needs and preferences of seminar attendees or if we want to dry-run our intended course design with a small sample of students). In the context of the present study, we opt for a fractional factorial with only 12 different vignettes (treatment combinations) to ensure that we get at least 5 judgements of each vignette and choice set from different students.



Figure S2: D-efficiency using %mktex

In a next step, we evaluate the efficiency of a fractional factorial with 12 vignettes using the *%mktex* macro. Its main arguments are the experimentally altered dimensions and we also need to specify any interaction effects that we want to estimate among the dimensions. Moreover, a random seed ensures reproducibility. In the present case, we opt for a main effects-only design. Doing so results in a more efficient design, but at the cost that the estimation of interaction effects among the dimensions in the experiment might be confounded. The following statement in SAS is used to find a suitable design:

*%mktex(2 3 2 3 2 2, n=12, seed=32719);*

A white square with blue and white text

Description automatically generatedFigure S2 depicts the result of this command. The fractional factorial with 12 different vignettes has a *D*-efficiency of 98.58, which is very close to the maximum value of 100 for a fully efficient design. To check the design in more detail, we use *%mkteval*, which requires only the name of the design it should evaluate (named “design” by default). The output in Figure S3 gives the correlations among the 6 attributes in the experiment. In a fully efficient design, all the off-diagonal elements in Figure S3 should be 0 (orthogonality). In the present case, there is a correlation of 0.25 between the second (prerequisites) and fourth dimension (mode of assessment).

Figure S3: Results from %mkteval

In a factorial survey experiment, we could directly use this design by presenting one or several vignettes to respondents. Meanwhile, in a choice experiment, we first need to efficiently allocate the 12 vignettes to different choice sets. The macro *%choiceff* will do this for us. It requires the fractional factorial as main argument (again named “design” by default). Moreover, we need to specify the model (here: a standard orthogonal design with main effects only, using x1, x2, etc. as placeholders for the experimentally varied dimensions), the number of choice sets to be constructed, and the number of alternatives per choice set (“flags”). Finally, we specify a random seed, set a maximum of 30 searches and ask for a relative measure of efficiency using the following code:

*%choiceff(data=design,*

*model=class(x1 x2 x3 x4 x5 x6 / standorth),*

*nsets=12, flags=3,*

*seed=23843, maxiter=30, options=relative, beta=zero);*

Figure S4 tells us that, from the 30 designs that the algorithm considered, design number 3 was the most efficient one. In contrast to the *D*-efficiency of the fractional factorial in Figure 3, the relative *D*-efficiency in Figure 5 cannot be interpreted in absolute terms but only in relation to other designs. Here, we construct 12 different choice sets, each with three alternatives. However, since the fractional factorial that we specified as an input comprises only 12 different vignettes, it would suffice to construct 4 different choice sets with three alternatives each. Doing so results in a less efficient design with a relative *D*-efficiency of 77.89 and variances for all the attributes that are about three times larger than the one in Figure 5. Although this loss in efficiency could potentially be offset by a larger number of evaluations of each choice set, we nevertheless proceed with the design in Figure 5 to present one last step: the blocking of choice sets to different decks which are then presented to students.

A screenshot of a computer

Description automatically generatedSince rating all 12 choice sets likely overburdens students, the experiment is split into two. The *%mktblock* macro optimally allocates the choice sets to different decks. It requires the design that was chosen by *%choiceff* (which is called “best” if not specified otherwise), the number of alternatives, and the attributes (“factors”). Moreover, we specify the number of blocks (“nblocks”), that is, the number of decks to which the choice sets are blocked. The following code blocks the design into two different decks. The resulting designs are named *blocked* and *blockedr*, the latter comprising randomly reordered choice sets within decks:

Figure S4: Result from %choiceff

*%mktblock(data=best, nalts=3, nblocks=2, factors=x1-x6,*

*out=blocked, outr=blockedr, seed=472);*

Having now found an efficient experimental design, we finally export the (randomized) blocked design and implement it in an online questionnaire.

### D) IMPLEMENTING THE EXPERIMENTAL DESIGN

Just like factorial survey experiments, discrete choice experiments can be either implemented in a paper and pencil (PAPI) or online questionnaire. The process is discussed in detail in the book by Auspurg and Hinz (2015) and comprises labelling the different levels of the experimentally manipulated dimensions, adding an ID variable for later merging of the collected data with the experimental design, and exporting the labelled design. The code for the first part of these steps can be found in Appendix A.

Implementing a discrete choice experiment in an online questionnaire is, however, somewhat more demanding. Since most platforms for online surveys were not designed to incorporate choice experiments, the corresponding question type often needs to be constructed by hand. To this end, the survey platform should allow for adding HTML elements. In addition to the steps briefly summarized before, we need to implement the graphical representation of the choice experiment in a tabular form (see Figure 1 in the main article) and we have to specify all the place-holders for our experimentally altered characteristics in the HTML code. Appendix B contains the HTML code used in the present study. With minor modifications, it can be adapted to other choice experiments. Depending on the online-survey platform used, an alternative strategy would be to let the survey platform itself assign the choice sets to respondents randomly (when using Qualtrics, for example, this is an often-chosen way of implementing choice experiments).

# APPENDIX A

The code below was used to construct a D-efficient experimental design for the choice experiment using the free-to-use service of SAS on demand (<https://welcome.oda.sas.com/login>). The export from SAS is then processed further in Stata for implementation in an online or paper-and-pencil survey. The corresponding Stata code can also be found below.

SAS Code:

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

\*\*\*\* Course Choice \*\*\*\*

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

\*\*\*\*\*\*\*\*\*\*\*\*

\*\* Dimensions (generic design)

x1: ECTS (2 levels)

x2: prerequisites (3 levels)

x3: schedule (2 levels)

x4: assessments (3 levels)

x5: organization (2 levels)

x6: pre-/post graduate mixed (2 levels);

\* 1. Find suitable design sizes

(specify the no levels of the six variables above);

%mktruns(2 3 2 3 2 2);

\* 2. Construct a linear arrangement (vignettes);

%mktex(2 3 2 3 2 2, n=12, seed=32719);

\* 3. Evaluate the design;

/\*Have a look at the vignettes\*/

proc print data=design; run;

%mkteval(data=design);

\* 4. Construct the choice sets;

/\*Minimal design allocating the 12 vignettes to 4 coice sets, each with 3 alternatives\*/

%choiceff(data=design,

model=class(x1 x2 x3 x4 x5 x6 / standorth),

nsets=4, flags=3,

seed=23843, maxiter=30, options=relative, beta=zero);

/\*Allocating the 12 vignettes to 12 coice sets, each with 3 alternatives\*/

%choiceff(data=design,

model=class(x1 x2 x3 x4 x5 x6 / standorth),

nsets=12, flags=3,

seed=23843, maxiter=30, options=relative, beta=zero);

\* 5. Block to decks;

/\*Blocking the 12 choice sets to 2 decks\*/

%mktblock(data=best, nalts=3, nblocks=2, factors=x1-x6,

out=blocked, outr=blockedr, seed=472);

/\*Have a look at the choice sets\*/

\*6. Export for further use (in Stata, etc.);

proc print data=blocked; run;

PROC EXPORT DATA=blocked

OUTFILE= "/home/u47334465/courses.dta"

DBMS=STATA REPLACE;

RUN;

Stata Code:

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

\*\*\* Course choice \*\*\*

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

set seed 14582

\*1. Load the data from SAS

clear

cd "…"

use "courses.dta"

\*2. Rename the vignette dimensions

rename (x1 x2 x3 x4 x5 x6) (ects prerequisites schedule assessments organization mixedgroup)

\*3. Label values (levels of the dimensions)

label def ects\_l 1 "3 ECTS" 2 "6 ECTS"

label values ects ects\_l

label def prerequisites\_l ///

1 "none" ///

2 "introductory statistics" ///

3 "multivariate statistics"

label values prerequisites prerequisites\_l

label def schedule\_l 1 "two blocks, FR-SA, 9 a.m. - 5 p.m." 2 "weekly, Monday, 2 - 4 p.m."

label values schedule schedule\_l

label def assessments\_l ///

1 "written exam, end of semester" ///

2 "oral presentation & term paper" ///

3 "graded exercises & small project"

label values assessments assessments\_l

label def organization\_l 1 "mostly lectures/presentations" 2 "applied, working on problems"

label values organization organization\_l

label def mixedgroup\_l 1 "homogenous (either BA or MA)" 2 "mixed (BA & MA)"

label values mixedgroup mixedgroup\_l

\*4. Reshape to wide format, first for the three alternatives per choice set, then by choice set

\* Alternatives

reshape wide ects prerequisites schedule assessments organization mixedgroup, i(Block Set) j(Alt)

\* Choice sets (such that each individual will rate 6)

reshape wide ects\* prerequisites\* schedule\* assessments\* organization\* mixedgroup\*, i(Block) j(Set)

\*5. Expand and reshuffle the choice sets, make an ID for later matching of results with exp. design

expand 60

gen r = runiform()

sort r

drop r

gen ID = char(runiformint(65,90)) + ///

string(runiformint(0,9)) + ///

char(runiformint(65,90)) + ///

char(runiformint(65,90)) + ///

string(runiformint(0,9))

fre ID

order ID, before(Block)

export excel using "course choice.xls", firstrow(variables) replace

save "courses\_labeled.dta", replace

# APPENDIX B

The following HTML code was used to construct the tabular environment in which the choice experiment was presented to students. Note that c\_0001, c\_0002, etc. are the placeholders (variables) for the experimentally altered dimensions.

HTML code:

<html>

<head>

<title>CE</title>

<style type="text/css">

<!--

td {

text-align: center;

table-layout:fixed;

border-collapse: collapse;

}

.center {

text-align:center;

height: 50px;

}

.borders {

border:thin solid silver;

}

th {

text-align: left;

height: 50px;

border:thin solid silver;

table-layout:fixed;

border-collapse: collapse;

}

.tablebr {

width:900px;

table-layout:fixed;

border-collapse: collapse;

}

-->

</style>

</head>

<body>

<table class="tablebr" cellpadding="2em">

<colgroup>

<col width="34%">

<col width="22%">

<col width="22%">

<col width="22%">

</colgroup>

<tr>

<th></th>

<th bgcolor= "#E3E3E3" class="center">Course A1</th>

<th class="center">Course A2</th>

<th bgcolor= "#E3E3E3" class="center">Course A3</th>

</tr>

<tr>

<th>ECTS</th>

<td bgcolor= "#E3E3E3" class="borders">#c\_0001#</td>

<td class="borders">#c\_0007#</td>

<td bgcolor= "#E3E3E3" class="borders">#c\_0013# </td>

<td>&nbsp;</td>

</tr>

<tr>

<th>Prerequisites</th>

<td bgcolor= "#E3E3E3" class="borders">#c\_0002#</td>

<td class="borders">#c\_0008#</td>

<td bgcolor= "#E3E3E3" class="borders">#c\_0014#</td>

<td>&nbsp;</td>

</tr>

<tr>

<th>Schedule</th>

<td bgcolor= "#E3E3E3" class="borders">#c\_0003#</td>

<td class="borders">#c\_0009#</td>

<td bgcolor= "#E3E3E3" class="borders">#c\_0015#</td>

<td>&nbsp;</td>

</tr>

<tr>

<th>Assessment</th>

<td bgcolor= "#E3E3E3" class="borders">#c\_0004#</td>

<td class="borders">#c\_0010#</td>

<td bgcolor= "#E3E3E3" class="borders">#c\_0016#</td>

<td>&nbsp;</td>

</tr>

<tr>

<th>Mode of teaching</th>

<td bgcolor= "#E3E3E3" class="borders">#c\_0005#</td>

<td class="borders">#c\_0011#</td>

<td bgcolor= "#E3E3E3" class="borders">#c\_0017#</td>

<td>&nbsp;</td>

</tr>

<tr>

<th>Student body</th>

<td bgcolor= "#E3E3E3" class="borders">#c\_0006#</td>

<td class="borders">#c\_0012#</td>

<td bgcolor= "#E3E3E3" class="borders">#c\_0018#</td>

<td>&nbsp;</td>

</tr>

</table>

</body>

</html>

# APPENDIX C

The following Stata Code was used to analyse the data from the choice experiment.

Stata code:

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

\*\*\* Course Choice: Analysis \*\*\*\*

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

clear

set more off

cd "$data"

import excel "data.xlsx", sheet("Export 1.1") firstrow

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

\*\*\*\* A) Prepare and reshape data \*\*\*\*

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

\*\*\*A.1) Drop unwanted vars

drop rts\* external\_lfdn tester quality lastpage v\_14 v\_31 v\_33 v\_35 v\_37 v\_39 referer device\_type quota\* ///

page\_history hflip vflip output\_mode javascript flash session\_id language cleaned ats ///

date\_of\_last\_access date\_of\_first\_mail

\*\*\*A.2) Rename

rename (c\_0001 c\_0002 c\_0003 c\_0004 c\_0005 c\_0006 c\_0007 c\_0008 c\_0009 c\_0010 c\_0011 c\_0012 c\_0013 ///

c\_0014 c\_0015 c\_0016 c\_0017 c\_0018 c\_0019 c\_0020 c\_0021 c\_0022 c\_0023 c\_0024 c\_0025 c\_0026 ///

c\_0027 c\_0028 c\_0029 c\_0030 c\_0031 c\_0032 c\_0033 c\_0034 c\_0035 c\_0036 c\_0037 c\_0038 c\_0039 ///

c\_0040 c\_0041 c\_0042 c\_0043 c\_0044 c\_0045 c\_0046 c\_0047 c\_0048 c\_0049 c\_0050 c\_0051 c\_0052 ///

c\_0053 c\_0054 c\_0055 c\_0056 c\_0057 c\_0058 c\_0059 c\_0060 c\_0061 c\_0062 c\_0063 c\_0064 c\_0065 ///

c\_0066 c\_0067 c\_0068 c\_0069 c\_0070 c\_0071 c\_0072 c\_0073 c\_0074 c\_0075 c\_0076 c\_0077 c\_0078 ///

c\_0079 c\_0080 c\_0081 c\_0082 c\_0083 c\_0084 c\_0085 c\_0086 c\_0087 c\_0088 c\_0089 c\_0090 c\_0091 ///

c\_0092 c\_0093 c\_0094 c\_0095 c\_0096 c\_0097 c\_0098 c\_0099 c\_0100 c\_0101 c\_0102 c\_0103 c\_0104 ///

c\_0105 c\_0106 c\_0107 c\_0108) ///

(ects1\_1 prerequisites1\_1 schedule1\_1 assessments1\_1 organization1\_1 mixedgroup1\_1 ects2\_1 ///

prerequisites2\_1 schedule2\_1 assessments2\_1 organization2\_1 mixedgroup2\_1 ects3\_1 prerequisites3\_1 ///

schedule3\_1 assessments3\_1 organization3\_1 mixedgroup3\_1 ects1\_2 prerequisites1\_2 schedule1\_2 ///

assessments1\_2 organization1\_2 mixedgroup1\_2 ects2\_2 prerequisites2\_2 schedule2\_2 assessments2\_2 ///

organization2\_2 mixedgroup2\_2 ects3\_2 prerequisites3\_2 schedule3\_2 assessments3\_2 organization3\_2 ///

mixedgroup3\_2 ects1\_3 prerequisites1\_3 schedule1\_3 assessments1\_3 organization1\_3 mixedgroup1\_3 ///

ects2\_3 prerequisites2\_3 schedule2\_3 assessments2\_3 organization2\_3 mixedgroup2\_3 ects3\_3 ///

prerequisites3\_3 schedule3\_3 assessments3\_3 organization3\_3 mixedgroup3\_3 ects1\_4 prerequisites1\_4 ///

schedule1\_4 assessments1\_4 organization1\_4 mixedgroup1\_4 ects2\_4 prerequisites2\_4 schedule2\_4 ///

assessments2\_4 organization2\_4 mixedgroup2\_4 ects3\_4 prerequisites3\_4 schedule3\_4 assessments3\_4 ///

organization3\_4 mixedgroup3\_4 ects1\_5 prerequisites1\_5 schedule1\_5 assessments1\_5 organization1\_5 ///

mixedgroup1\_5 ects2\_5 prerequisites2\_5 schedule2\_5 assessments2\_5 organization2\_5 mixedgroup2\_5 ///

ects3\_5 prerequisites3\_5 schedule3\_5 assessments3\_5 organization3\_5 mixedgroup3\_5 ects1\_6 ///

prerequisites1\_6 schedule1\_6 assessments1\_6 organization1\_6 mixedgroup1\_6 ects2\_6 prerequisites2\_6 ///

schedule2\_6 assessments2\_6 organization2\_6 mixedgroup2\_6 ects3\_6 prerequisites3\_6 schedule3\_6 ///

assessments3\_6 organization3\_6 mixedgroup3\_6)

rename (v\_24 v\_32 v\_34 v\_36 v\_38 v\_40) ///

(choice\_1 choice\_2 choice\_3 choice\_4 choice\_5 choice\_6)

rename (v\_1 v\_2 v\_3 v\_4 v\_5 v\_6 v\_7 v\_8 v\_46) ///

(field field\_other program program\_other N\_courses sat\_variety sat\_teaching sat\_career comments)

rename lfdn ID

\*\*\*A.3) Reshape to long, for the 6 choice sets each respondent answered

reshape long ects1\_ prerequisites1\_ schedule1\_ assessments1\_ organization1\_ mixedgroup1\_ ects2\_ ///

prerequisites2\_ schedule2\_ assessments2\_ organization2\_ mixedgroup2\_ ects3\_ prerequisites3\_ ///

schedule3\_ assessments3\_ organization3\_ mixedgroup3\_ choice\_, i(ID) j(ChoiceSet)

\*Add the deck information

gen deck=2

replace deck=1 if ects2\_=="6 ECTS"

fre deck

rename (\*\_) (\*)

\*\*\*A.4) Reshape another time to long format, this time for the different alternatives of each choice set

\*Identify unique choice situations for reshape

egen ChoiceSituation=concat(ID ChoiceSet)

destring ChoiceSituation, replace

reshape long ects prerequisites schedule assessments organization mixedgroup, i(ChoiceSituation) j(alternative)

\*Binary choice variable: Has alternative been chosen or not (dep var)

gen chosen=choice==alternative

fre choice

\*Encode string variables to numeric

foreach var of varlist ects prerequisites schedule assessments organization mixedgroup {

encode `var', gen(`var'\_n)

drop `var'

rename `var'\_n `var'

}

\*\*\* A5) Recode missing values (-99)

recode field program N\_courses sat\_variety sat\_teaching sat\_career (-99=.)

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

\*\*\* B: Some descriptives \*\*\*

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

\*Number of times each vignette and choice set was evaluated

sort ects prerequisites schedule assessments organization mixedgroup

quietly by ects prerequisites schedule assessments organization mixedgroup: gen dup = cond(\_N==1,0,\_N)

tab dup

//each vignette evaluated between 22 and 46 times

drop dup

preserve

keep ChoiceSituation alternative ects prerequisites schedule assessments organization mixedgroup

reshape wide ects prerequisites schedule assessments organization mixedgroup, i(ChoiceSituation) j(alternative)

sort ects1 prerequisites1 schedule1 assessments1 organization1 mixedgroup1 ///

ects2 prerequisites2 schedule2 assessments2 organization2 mixedgroup2 ///

ects3 prerequisites3 schedule3 assessments3 organization3 mixedgroup3

quietly by ects1 prerequisites1 schedule1 assessments1 organization1 mixedgroup1 ///

ects2 prerequisites2 schedule2 assessments2 organization2 mixedgroup2 ///

ects3 prerequisites3 schedule3 assessments3 organization3 mixedgroup3: gen dup = cond(\_N==1,0,\_N)

tab dup

//choice sets were evaluated between 10 and 22 times

restore

\*Check: Correlation with deck induced by design!

bysort deck: fre choice //3rd alternative dominates in deck 2 --> also see corr matrix exp. design

label var ects "ECTS"

label var prerequisites "Prerequisites"

label var schedule "Schedule"

label var assessments "Assessment"

label var organization "Methods"

label var mixedgroup "Composition"

label var deck "Deck"

label var chosen "Outcome"

corr ects prerequisites schedule assessments ///

organization mixedgroup deck chosen

matrix correlations = r(C)

capture graph drop \_all

heatplot correlations, values(format(%4.3f) size(medium)) color(tab Gray Warm, intensity(0.5)) ///

lower nodiagonal xlabel(,alternate) label legend(off) graphregion(color(white)) xscale(range(1/8)) name(g0)

graph export "$graphs\correlations\_all.eps", name(g0) replace

graph export "$graphs\correlations\_all.png", name(g0) width(2000) replace

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

\*\*\* C: Conditional logit models \*\*\*

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

\*Control for deck since correlated with vignette dimension(s)!

cmset ID ChoiceSet alternative

sort ID ChoiceSet alternative

list ID ChoiceSet alternative \_\* in 1/20

cmclogit chosen i.ects i.prerequisites i.schedule i.assessments i.organization i.mixedgroup, ///

casevars(i.deck) vce(cluster ID)

\*equivalent

clogit chosen i.alternative i.ects i.prerequisites i.schedule i.assessments i.organization i.mixedgroup ///

i.deck i.alternative#i.deck, ///

group(ChoiceSituation) vce(cluster ID)

eststo clogit

\*Add a respondent-level variable (interaction)

clogit chosen i.alternative i.ects i.prerequisites i.schedule i.assessments i.organization i.mixedgroup ///

i.deck i.alternative#i.deck c.sat\_teaching c.sat\_teaching#i.prerequisites, ///

group(ChoiceSituation) vce(cluster ID)

eststo clogit\_rl

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

\*\*\*\* D: Mixed logit models \*\*\*\*\*

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

\*Panel

cmxtmixlogit chosen i.ects i.prerequisites i.schedule i.assessments i.organization i.mixedgroup, casevars(i.deck)

\*Add random effect for prerequisites

cmxtmixlogit chosen i.ects i.schedule i.assessments i.organization i.mixedgroup, ///

random(i.prerequisites) casevars(i.deck)

eststo mixlogit

esttab clogit clogit\_rl mixlogit, pr2 b(%8.2f) se(%8.2f) star(+ 0.10 \* 0.05 \*\* 0.01 \*\*\* 0.001)

esttab clogit clogit\_rl mixlogit using "$tables/models.rtf", pr2 b(%8.3f) se(%8.3f) ///

star(+ 0.10 \* 0.05 \*\* 0.01 \*\*\* 0.001) onecell replace

# REFERENCES

Auspurg, K., & Hinz, T. (2015). *Factorial Survey Experiments*. Sage.

Johnson, F. R., Lancsar, E., Marshall, D., Kilambi, V., Mühlbacher, A., Regier, D. A., Bresnahan, B. W., Kanninen, B., & Bridges, J. F. P. (2013). Constructing Experimental Designs for Discrete-Choice Experiments: Report of the ISPOR Conjoint Analysis Experimental Design Good Research Practices Task Force. *Value in Health*, *16*(1), 3–13.

Kuhfeld, W. F. (1997). *Efficient Experimental Designs Using Computerized Searches* (Research Paper Series). Sawtooth Software Inc.

Kuhfeld, W. F. (2010). *Marketing research methods in SAS*. SAS Institute Inc. http://support.sas.com/techsup/technote/mr2010.pdf

Zangger, C., & Becker, R. (2019). Experiments in the sociology of education: Causal inference and estimating causal effects in sociological research on education. In R. Becker (Ed.), *Research Handbook on the Sociology of Education* (pp. 153–171). Edward Elgar.

1. SAS can be used for free by signing up to SAS OnDemand: https://welcome.oda.sas.com/login [↑](#footnote-ref-2)