

# Choice certainty, consistency, and monotonicity in discrete choice experiments

Matteo Mattmann<sup>a,b</sup>, Ivana Logar<sup>a</sup> and Roy Brouwer<sup>id a,c</sup>

<sup>a</sup>Eawag, Swiss Federal Institute of Aquatic Science and Technology, Dübendorf, Switzerland; <sup>b</sup>Department of Environmental Economics, Institute of Environmental Studies, VU University Amsterdam, Amsterdam, The Netherlands; <sup>c</sup>Department of Economics and The Water Institute, University of Waterloo, Waterloo, Canada

## ABSTRACT

This study investigates choice certainty, choice consistency, and choice monotonicity and their underlying common and idiosyncratic determinants in discrete choice experiments. We test the equality of choice behaviour between respondents who differ with respect to these concepts. Our results suggest that there are significant differences in the choice behaviour between certain and uncertain, as well as consistent and inconsistent, respondents. The hypothesis of equality of choice behaviour between samples with and without a self-reported choice certainty follow-up question cannot be rejected. We identify a variety of idiosyncratic determinants of choice certainty, consistency, and monotonicity, but only the time spent reading informational pages and gender are identified as common drivers. We find that female respondents are less certain about their choices, but display a higher degree of monotonicity and consistency in their choice behaviour.

## ARTICLE HISTORY

Received 19 June 2018

Accepted 20 August 2018

## KEYWORDS

Choice certainty; choice consistency; choice monotonicity; choice complexity; discrete choice experiment

## 1. Introduction

This study investigates choice certainty, choice consistency, and choice monotonicity in the context of discrete choice experiments (DCE). There is heterogeneity in the choice literature on the definition of these terms. Our definitions follow the most common practice in the DCE literature, as identified by a literature review of studies which assess the determinants of certainty, consistency, and monotonicity (see Section 2). ‘Choice certainty’ henceforth refers to ‘stated choice certainty’, as it is based on respondents’ self-reported answers to a question on choice certainty after each choice task. In the broader microeconomic literature, ‘choice consistency’ is often linked to the axiom of transitivity (e.g. Jehle and Reny 2001). However, we follow the definition by Carlsson, Mørkbak, and Olsen (2012), and define choice consistency as ‘making the same choice in two equal choice tasks’. An identical definition of choice consistency is used in the DCE literature on the determinants of consistency in a test-retest setting (Brouwer, Logar, and Sheremet 2016; Mørkbak and Olsen 2015; Rigby, Burton, and Pluske 2016; Schaafsma et al. 2014). Finally, ‘choice monotonicity’ is assumed to hold if a respondent chooses a non-dominated alternative in a choice task that contains a dominated alternative, which is a hypothetical alternative that is worse than at least one other alternative in a choice task with respect to all attributes.

**CONTACT** Matteo Mattmann ✉ [matteo.mattmann@eawag.ch](mailto:matteo.mattmann@eawag.ch) Eawag, Swiss Federal Institute of Aquatic Science and Technology, Dübendorf, Switzerland; Department of Environmental Economics, Institute of Environmental Studies, VU University Amsterdam, Amsterdam, The Netherlands

© 2018 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group  
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

The rationale to investigate these concepts simultaneously is threefold. First, in contrast to the existing literature, an integrated analysis of the three constructs on the basis of an identical sample of respondents allows us to examine how they relate to each other and to identify their common and idiosyncratic drivers. Second, although a number of studies investigate potential determinants of choice certainty and choice consistency separately, we are not aware of any studies in the field of environmental economics that aim to identify the drivers of choice monotonicity in DCEs. The simultaneous analysis allows us to assess whether the drivers of choice monotonicity are common to the determinants of choice certainty and consistency, or whether they are idiosyncratic. Finally, the existence of all three concepts disagrees with the conventional assumptions made in microeconomics that people know their preferences, and that these preferences are stable, coherent, and rationally maximised (e.g. Brown et al. 2008; Rabin 1998). Choice monotonicity in the framework of microeconomics is related to the axiom of consumer choice, which, in its strict formulation, implies that more is always better than less (e.g. Jehle and Reny 2001). Accordingly, in the context of a DCE, a choice alternative should be preferred over another alternative when the first is better on one attribute and at least as good on all the other attributes. The existence of choice uncertainty, inconsistency, and non-monotonicity is hence typically ruled out in DCEs. However, even if evidence for their existence is found, this does not necessarily imply choice irrationality. Preferences are, for example, allowed to shift between choice tasks (Rigby, Burton, and Pluske 2016). In addition, a key assumption in random utility theory (McFadden 1974), which underlies the DCE method, is the existence of an informational asymmetry between the decision maker and the analyst. Thus, it is not always straightforward to identify all the underlying causes of choice behaviour from the choice information collected in a DCE. Choice behaviour may be, for example, driven by attributes which are present in a respondent's utility function but not included in a DCE. In this case, a behavioural violation of conventional assumptions would reflect more a misspecification of the DCE than evidence for choice irrationality. It is nevertheless relevant to detect the degree to which choice uncertainty, inconsistency, and non-monotonicity are present, since they affect the outcomes of choice models. Moreover, identifying their determinants is of interest, as it may help to minimise their causes. The occurrence of uncertainty, inconsistency, and non-monotonicity is especially likely in DCEs in an environmental context, since respondents are often confronted with unfamiliar goods and/or attributes. Even if the characteristics of the good are familiar to a respondent, it may still be cognitively challenging to conceptualise specific attribute levels (Rigby, Burton, and Pluske 2016).

In the first part of this paper, we aim to compare choice behaviour of respondents that differ with respect to choice certainty, consistency, and monotonicity. We first investigate stated choice certainty. The information on choice certainty at the choice-task level enables us to investigate possible dynamic effects of choice certainty over the course of the experiment. Furthermore, we assess the effect of including choice certainty questions in the survey. For this purpose, we use a split-sample approach and present the choice certainty questions to only half of the respondents. The procedural effect of posing a question on certainty has, to the best of our knowledge, only been studied in Brouwer et al. (2010). We add to this study by also investigating the potential effects caused by the presence of certainty questions at a choice-task-specific level. Next, in order to analyze choice consistency, the first choice task is randomly repeated at varying positions in the choice task sequence within the same DCE. This allows us to detect the degree of choice consistency, and to check for locational effects of the repeated task on consistency. Finally, choice monotonicity is assessed by including a choice task that involves a dominated choice alternative in a choice set.

The second part of the analysis comprises models that regress choice certainty, consistency, and monotonicity on possible determinants in order to identify common and unique drivers of these concepts. In doing so, we use a single model to simultaneously regress choice consistency and monotonicity on possible drivers as these concepts are likely to be correlated.

The remainder of this paper is structured as follows: Section 2 provides a literature review on choice certainty, consistency, and monotonicity. The literature review aims to identify the key determinants and hypotheses that will be used in the subsequent probit models. Section 3 describes the

econometric models, and Section 4 the case study and the DCE design. Section 5 presents the results, and there follows a discussion and conclusions in Section 6.

## 2. Choice certainty, consistency, and monotonicity in discrete choice experiments

The literature on choice certainty in the context of DCEs can be divided into two streams, one of which investigates how to account for choice certainty in choice models, while the other focuses on the determinants of choice certainty. Within the first stream of literature, there are studies that consider stated choice certainty in choice models by excluding, recoding, or weighting responses according to their self-reported choice certainty (e.g. Beck, Fifer, and Rose 2016; Beck, Rose, and Hensher 2013; Kosenius 2009; Lundhede et al. 2009). Since treating choice certainty as an explanatory construct in this manner raises the issue of endogeneity, some authors have switched to more complex models which treat stated choice certainty as a latent construct (e.g. Dekker et al. 2016). Other studies focus on inferred rather than on stated choice certainty by assuming that a higher choice certainty is reflected in a lower error variance (Brouwer et al. 2010; Uggeldahl et al. 2016). The second stream of research consists of studies that estimate mainly logit or probit models in order to investigate the determinants of stated choice certainty. These studies, and the significant determinants they detect, are listed in Table 1.

The most common definition of choice consistency in the DCE literature is to make identical choices when faced with identical choice tasks<sup>1</sup>. Therefore, the usual test for choice consistency is

**Table 1.** Studies that regress stated choice certainty on its determinants.

Authors (year of publication)	Model estimated	Significant determinants (direction of the effect) <sup>a</sup>
Brouwer et al. (2010)	Ordered probit	Age (+) Credibility of the alternatives (+) Experience with the DCE's topic (+) Familiarity with the information provided (+) Income (+) Male gender (+) Utility difference <sup>b</sup> (+)
Dekker et al. (2016)	Latent variable model approach	Credibility of the alternatives (+) Male gender (+) Entropy (–) Survey length (–)
Hensher, Rose, and Beck (2012) <sup>c</sup>	Generalised ordered logit	Differences between the attribute levels of the status quo and non-status quo alternatives (+) Environmental attitude (+) Stated number of acceptable alternatives per choice set (+) Income (–) Number of children in household (–)
Kosenius (2009)	Binary logit	Acceptance of scenario elements (+) Male gender (+) Positive attitudes towards the scenarios (+) Time spent at the resource valued (+) University degree (+) Utility difference <sup>d</sup> (+)
Olsen et al. (2011) <sup>e</sup>	Ordered probit	Income (+) Utility difference <sup>b</sup> (+)
Uggeldahl et al. (2016)	Fixed effects regression model	Utility difference (+) Choice of the status quo option (+) Gaze shifts (eye-tracking) (–) Later position of the task (–)

Notes: <sup>a</sup>Some studies also included attributes and scenario-related variables as regressors. These are not listed. <sup>b</sup>Between the chosen and the second-best alternative. <sup>c</sup>Certainty in terms of how certain it is that respondents will choose in a real situation the same alternative as they indicated they would choose in the choice tasks. <sup>d</sup>Between the alternatives with the highest and the lowest utility in a choice task. <sup>e</sup>The results of two case studies are reported. We only report the effects that proved to be stable across both case studies.

to repeat one or several choice tasks using either the same or different samples of respondents. This is implemented between separate choice survey occasions (test-retest) which implies a time-lag (Bliem, Getzner, and Rodiga-Laßnig 2012; Brouwer, Logar, and Sheremet 2016; Liebe, Meyerhoff, and Hartje 2012; Mørkbak and Olsen 2015; Rigby, Burton, and Pluske 2016; Schaafsma et al. 2014) or within the same survey which results in almost no time-lag or a lag in the order of few minutes (Brouwer et al. 2010; Brown et al. 2008; Campbell 2007; Carlsson, Mørkbak, and Olsen 2012; Hanley, Wright, and Koop 2002; Rigby, Burton, and Pluske 2016; Rulleau and Dachary-Bernard 2012; Scarpa, Campbell, and Hutchinson 2007; Soliño et al. 2012). There is, again, a stream of literature that investigates the drivers of choice consistency in DCEs by estimating binomial models, listed in Table 2. Note that all studies except Carlsson, Mørkbak, and Olsen (2012) investigate choice consistency in a test-retest setting.

A considerable amount of literature in health economics includes a test for monotonicity in the form of a choice task, where one alternative is superior to the other alternatives on all choice attributes (e.g. De Bekker-Grob, Ryan, and Gerard 2012; Determann et al. 2017; Lancsar and Louviere 2006; Özdemir et al. 2010; Soliño et al. 2012). Based on the outcome of choice monotonicity tests, a part of the survey participants are typically excluded from further analysis in health studies. Examples of DCEs in environmental economics that include a choice task with a dominant or dominated alternative are: Bateman, Munro, and Poe (2008), Campbell (2007), Foster and Mourato (2002), Hanley, Wright, and Koop (2002), and Scarpa, Campbell, and Hutchinson (2007). However, the literature that investigates the determinants of choice monotonicity is limited and, to the best of our knowledge, non-existent in environmental economics. In health economics, San Miguel, Ryan, and Amaya-Amaya (2005), for example, test for monotonicity as suggested in this study and, in addition, run further monotonicity tests and regress their results on explanatory variables. Their most important finding implies that respondents who state they have great difficulty with a choice task are less likely to pass the monotonicity test, i.e. to choose the theoretically expected alternative. They also detect a positive effect of higher education on choice monotonicity, and a bell-shaped effect of age, indicating that both younger and older respondents have a lower probability of passing the monotonicity test. Finally, an inverted bell-shaped effect of the repeated task position reflects the situation that choice consistency is determined by learning effects for an early position of the repeated choice task, and by fatigue setting in for later positions (Swait and Adamowicz 2001a).

The review of the literature reveals several key determinants of choice certainty, consistency, and monotonicity. The most important regressors identified are included in our analysis in Section 5.3. Variables that proved to be significant in at least three studies include gender, where we expect male respondents to state a higher choice certainty than female respondents. Furthermore, income seems

**Table 2.** Studies that regress choice consistency on its determinants.

Authors (year of publication)	Model estimated	Significant determinants (direction of the effect)
Brouwer, Logar, and Sheremet (2016)	Binary probit	Choice certainty (stated) (+) Credibility of the information provided (+) Differences between the choice task attribute levels (+) Membership in an environmental organisation (+)
Carlsson, Mørkbak, and Olsen (2012)	Binary probit	Choice of the status quo option in the first sequence (+) Later position of the repeated task (+) Stated certainty that the chosen alternative provides the largest utility (+) Utility difference <sup>a</sup> (+) Utility difference <sup>a</sup> (+)
Mørkbak and Olsen (2015)	Binary probit	Income (–) Willingness to take on more complex tasks (+) Entropy (–)
Rigby, Burton, and Pluske (2016)	Binary probit	None
Schaafsma et al. (2014)	Binary probit	None

Note: <sup>a</sup>Between the chosen and the second-best alternative.

to be important, although the direction of the effect is unclear: Brouwer et al. (2010) report a positive effect on choice certainty, whereas Hensher, Rose, and Greene (2012) find the opposite. Mørkbak and Olsen (2015) confirm the latter with respect to choice consistency. The authors argue that this may be due to a lower importance of the price attribute for high-income respondents, which adds more randomness to their choices. A further ambiguous variable is higher education, which is reported to have a positive effect on choice certainty and monotonicity by, respectively, Kosenius (2009) and San Miguel, Ryan, and Amaya-Amaya (2005). On the other hand, Olsen et al. (2011) and Dekker et al. (2016) control for it, but find no effect. Based on the literature review, we also expect that positive (environmental) attitudes positively affect choice certainty and consistency. Furthermore, our analysis aims at assessing the explanatory power of entropy. Entropy originates from information theory (Shannon 1948; Soofi 1994; Swait and Adamowicz 2001b). In the realm of DCEs, it describes the equality of a choice task's alternatives based on the predicted individual specific probabilities of choosing each alternative. We hypothesise that the higher the entropy, and therefore the more similar the choice alternatives in a choice task, the higher is the choice-task complexity. Higher choice-task complexity, in turn, is expected to reduce choice certainty, consistency, and monotonicity. Finally, we include two paradata variables related to time: The time spent by survey respondents reading informational pages that appear before the actual choice tasks, and the time used per choice task. The latter has been shown to outperform eye-tracking information in predicting stated choice certainty (Uggeldahl et al. 2016). We expect the time used per choice task to have a negative impact on stated choice certainty, and the time used for reading informational pages to positively affect certainty, consistency, and monotonicity.

### 3. Econometric approach

The Swait-Louviere test procedure (Swait and Louviere 1993) is applied to assess whether there are differences between the choice behaviour of respondents who: (i) state they are certain or uncertain about their choices; (ii) choose consistently or inconsistently; (iii) were and were not asked a question on choice certainty; and (iv) were subjected to different positions of the repeated choice task in the choice sequence. The test procedure for verifying the equality of choice behaviour between all these groups of respondents consists of two steps. In a first step, the equality of preference parameters between different samples is assessed. If equality cannot be rejected, one is able to proceed to the second step of the test procedure and isolate differences in scale parameters between two samples. Note that scale parameters are inversely related to the error variance. If differences in scale parameters between the two samples cannot be rejected we cannot reject the equality of scale parameters between the two samples.

Two models are used for regressing choice certainty, choice consistency, and choice monotonicity on their determinants. More specifically, we estimate a bivariate probit model to identify the determinants of choice consistency and monotonicity simultaneously, and an ordered probit model to assess the drivers of choice certainty (see, e.g. Cameron and Trivedi (2005) for a detailed econometric specification of these models). The latter is a random-effects panel specification, since choice certainty is analyzed at a choice-task level, whereas the model explaining choice consistency and monotonicity can only be estimated at the respondent level. Amongst other regressors, entropy as a measure of choice-task complexity is included. The procedure to entropy begins with estimating an attribute-only mixed logit (MXL) model based on the following MXL probability distribution for each individual  $n$ , alternative  $i$ , and choice task  $t$ :

$$P_{nit}(\beta) = \int \left( \frac{e^{\beta' x_{nit}}}{\sum_j e^{\beta' x_{njt}}} \right) f(\beta) d\beta. \quad (1)$$

On the basis of equation 1, the MXL model parameters are estimated by applying maximum likelihood estimation on the appropriate log-likelihood function. This model does not include any

sociodemographic covariates, since the probit regression models that follow already control for these.

Entropy was calculated by utilising the predicted choice probabilities  $P$  obtained from the MXL model, and estimates a value of choice-task complexity for each individual  $n$  and each choice task  $t$  as follows (Rigby, Burton, and Pluske 2016; Shannon 1948; Soofi 1994; Swait and Adamowicz 2001b):

$$\text{Entropy}_{nt} = - \sum_{i=1}^I P(i) \ln P(i) \geq 0. \quad (2)$$

For the case of three choice alternatives in a choice task, entropy can reach a maximum value of approximately 1.1, when the probability of choosing an alternative is identical for all three choice options. In contrast, if the probability of choosing an alternative equals 1, entropy reaches its minimum value of 0.

## 4. Case-study description

### 4.1. Discrete choice experiment

Following the nuclear accident in Fukushima, Japan, the Swiss government released a policy paper on the long-term Swiss energy strategy in 2011 ('Energy Strategy 2050' (SFOE 2013)). Our study incorporates two main strategic goals of this energy strategy: The expansion of electricity produced by hydropower, and the phase-out of nuclear power plants. Hydropower and nuclear power constitute by far the most important sources of electricity production in Switzerland, with respective shares in total electricity production of 60% and 34% in 2015 (SFOE 2016). From a policy perspective, an expansion of hydropower renders a phase-out of nuclear power more likely, since the probability and the consequences of a loss in electricity production that such a phase-out may involve could be reduced. Due to this interlinkage of the envisaged policies for the two main sources of electricity in Switzerland, the preferences for hydropower and nuclear power should correspondingly be studied simultaneously. For this reason, we designed a DCE that focuses on a hypothetical expansion of hydropower and takes into account the expected decrease in the provision of nuclear power. The DCE includes three choice alternatives: two unlabelled hydropower expansion scenarios, and one status quo scenario. The scenarios are characterised by four attributes and their associated levels, as summarised in Table 3.

The first attribute describes whether the current hydropower plants would be extended, or whether new facilities would be constructed. It is explained to survey respondents that the construction of new plants is associated with stronger detrimental environmental effects compared with the extension of existing plants. The second and third attributes reflect the changes in the risk of death associated with an expansion of hydropower and a reduction in nuclear power production,

**Table 3.** Attribute and attribute levels in the DCE.

Attribute	Attribute levels in hypothetical alternatives	Attribute levels in status quo alternative
Type of hydropower expansion	Extending existing hydropower plants Construction of new hydropower plants	No hydropower expansion
Lifetime risk of death from a dam breach	20% increase in risk (1 in 750,000 people are expected to die) 40% increase in risk (1 in 650,000 people are expected to die)	Current risk (1 in 900,000 people are expected to die)
Lifetime risk of death from a nuclear accident	60% decrease in risk (1 in 7 million people are expected to die) 30% decrease in risk (1 in 4 million people are expected to die)	Current risk (1 in 3 million people are expected to die)
Increase in household's annual electricity bill (CHF)	100, 200, 300, 400, 500,600	No change in the annual electricity bill

respectively. Both risk attributes can take on three levels, one of which appears only in the status quo alternative. The absolute risk magnitudes of both attributes were carefully constructed on the basis of data from existing studies (Burgherr and Hirschberg 2008, 2014; Hirschberg et al. 2016). The relative changes in risk are based on realistic scenarios and reflect the type of hydropower plants that would be built or extended (only plants with dams or various types of plants) and the number of nuclear power plants in Switzerland that would be phased-out<sup>2</sup>. Care was taken in communicating the risks of death to respondents with the help of risk ladders. The risk ladders put the risks associated with hydropower and nuclear power into context by comparing them with other risks of death. Finally, a price attribute was included which was framed as an increase in a household's yearly electricity bill<sup>3</sup>. The results of two pretest rounds with a total of 570 respondents (see Section 4.3) was used to define and calibrate meaningful attribute levels. Figure 1 presents an example of a choice task. Each respondent received a sequence of eight choice tasks.

#### 4.2. Elicitation of choice certainty, consistency, and monotonicity

Choice certainty was elicited for every choice task. That is, we asked the survey participants the following question after each choice task: 'How certain are you about your choice?'. The question appeared on the same page as the choice task. A Likert-type answer scale with the following five options was presented: not certain at all, not certain, neither certain nor uncertain, certain, very certain. We used a split-sample approach, so that one sample of respondents was asked about their choice certainty (sample 1), while the other sample did not receive this question (sample 2). All other elements of the survey were identical for the two samples. This allows us to investigate the procedural effect of asking respondents about their choice certainty. Choice consistency was elicited by repeating the first choice task once again later in the choice-task sequence. The position of the repeated choice task varied between the 5th, 6th, 7th, and 8th (last) position in the choice-task sequence, and was randomly assigned to each respondent. Finally, we included a test for choice










	A) Expansion	B) Expansion	C) No expansion
Type of hydropower expansion	 New construction	 Extension	 No expansion
Risk of dying from a dam breach	 +40% risk (1 in 650,000 people)	 +20% risk (1 in 750,000 people)	 Current risk (1 in 900,000 people)
Risk of dying from a nuclear accident	 -30% risk (1 in 4,000,000 people)	 -60% risk (1 in 7,000,000 people)	 Current risk (1 in 3,000,000 people)
Increase in your household's yearly electricity bill	+200 CHF/year	+300 CHF/year	+0 CHF/year

Figure 1. Choice task example.



monotonicity by adding a choice task with a dominated alternative. In this choice task, one of the two hypothetical alternatives was worse than the other hypothetical alternative with respect to all attributes. Although this choice task includes a strongly dominated choice alternative, there is no unique dominant alternative identifiable, since the non-dominated choice alternative and the status quo alternative do not dominate each other. This is because in our DCE design the levels of some attributes in the status quo alternative are better and others are worse than their levels in the non-dominated choice alternative. The position of the choice task involving a dominated alternative was fixed at the 4th position in the choice-task sequence for all respondents.

#### 4.3. Sampling procedure and choice experiment design

Three rounds of survey pretests were conducted in order to test the understandability of the survey and of the DCE, as well as to derive prior estimates needed for generating the final choice design. First, 20 respondents were asked to fill in a paper-and-pencil version of the survey, followed by a personal interview. Second, 220 respondents, who were representative for the German-speaking part of Switzerland, were recruited to answer a first online-version of the DCE. In a third pretest, another 350 respondents answered the online survey. The final survey version was completed in June 2016 with a sample consisting of 502 respondents. The response rate was 16.3%. The samples for pretests and for the final survey were recruited by Intervista AG, a market research company with a panel of 50,000 registered individuals throughout Switzerland. The final sample of 502 respondents consisted of two independent samples comprising 248 and 254 respondents. The only distinction between them is the presence/absence of the choice certainty questions. Both samples are representative for the German- and French-speaking population of Switzerland which accounts for roughly 95% of the total Swiss population. The representativeness was safeguarded by independent recruitment which ensured equality of the samples and the general population with respect to gender, age, education, and the geographical distribution of residence within Switzerland. The independent recruitment of samples is also expected to minimise the impact of unobserved respondent heterogeneity (Dekker, Koster, and Brouwer 2014).

We generated a D-efficient DCE design in Ngene 1.1.2. Bayesian priors that are bounded by the results of the second and third pretests, and that follow a normal distribution were used for the design-generating process. Such a specification allows for a random instead of a fixed distribution of priors, and hence takes the approximate character of pretest priors into account. A status quo option was included in the design generation using an alternative-specific constant (ASC) for the status quo alternative. Dominant alternatives were excluded in the design generation procedure, but a choice task which involved a dominated alternative was added to the choice sequence after the design generation process. The final experimental design consisted of two blocks of choice sets, each of which contained eight choice tasks. The two blocks were randomly assigned to the survey participants.

## 5. Results

### 5.1. Descriptive results

Following the criteria of Bateman et al. (2002), we identified 31 protest responses which were excluded from further analysis. This leaves us with a remaining sample population of 235 respondents for sample 1 (which includes a question on choice certainty) and 248 respondents for sample 2 (without a question on choice certainty). No statistically significant differences between the samples with respect to any of the variables that were controlled for at the point of recruiting could be found. This holds even after excluding protest respondents (Table 4).

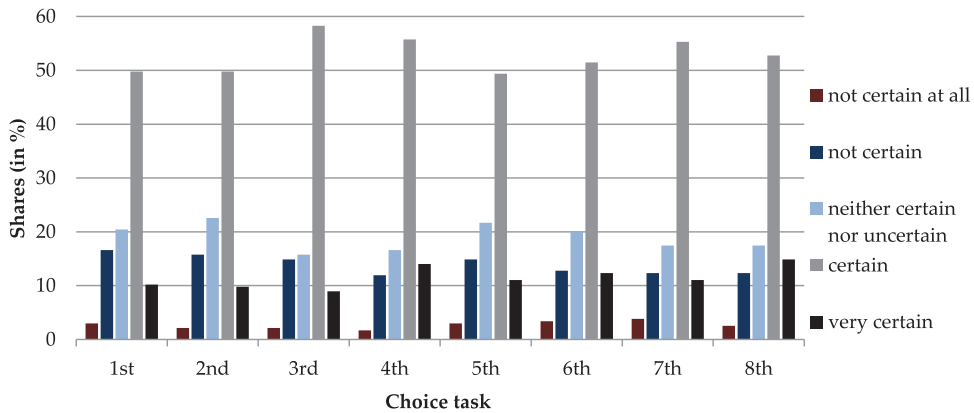
Figure 2 reports the choice-task-specific responses to the question on choice certainty for sample 1. The distribution of certainty values is left-skewed. On average, 64.4% of respondents stated



**Table 4.** Comparison of samples 1 and 2 with respect to sociodemographic characteristics.

Variable	Type	Description	Test	Test-Statistic	<i>p</i> -value
Income	Ordinal	15 income categories	Mann-Whitney	0.772	0.440
Gender	Bivariate	0 = Male; 1 = Female	Chi-squared	0.376	0.540
Age	Continuous	Respondents' age (15–84)	Mann-Whitney	0.069	0.945
Language spoken	Bivariate	0 = German; 1 = French	Chi-squared	0.746	0.389
University degree (Bachelor's, Master's, PhD)	Bivariate	0 = No university degree; 1 = University degree	Chi-squared	0.106	0.744

Note: \*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.1$ .

**Figure 2.** Stated choice certainty across choice tasks.

that they were certain or very certain about all choice tasks. The dynamics of stated choice certainty across the choice tasks is limited: The share of respondents who indicated that they were certain or very certain about their choice slightly increases after the first choice task, but decreases again after the fourth task. The share of respondents who stated that they are either not certain or not certain at all decreases from 19.6% in the first choice task to 14.9% in the last choice task. Note that the 4th choice task includes a dominated alternative. As expected, the share of certain and very certain (not certain at all and not certain) respondents is highest (lowest) for this task. Statistical tests confirm that there is limited dynamics across the choice tasks: The Kruskal-Wallis test statistic for equality of choice certainty across all tasks cannot be rejected at the 10% significance level. The Mann-Whitney test for pairwise comparisons of choice certainty between the tasks rejects equality in 79% of all possible comparisons, and only finds significant differences between the tasks 1 and 4, 1 and 8, 2 and 4, 2 and 8, 4 and 5, and 5 and 8. Finally, the variance of stated choice certainty is lower within than between respondents, since almost 47% of the respondents indicated the same level of choice certainty across all choice tasks.

The inferred choice consistency according to the corresponding position of the repeated choice task is illustrated in Figure 3. Two insights can be gained: First, the respondents in sample 2 seem to be more consistent than the respondents in sample 1, pointing to a possible interaction effect between the presence of a choice certainty question and choice consistency. Second, the choice consistency of respondents who were subjected to the repeated choice task later in the choice sequence seems to be slightly lower, which is especially true for sample 1. Nevertheless, none of these effects is statistically significant. The Mann-Whitney test for the equality in choice consistency between the two samples cannot be rejected ( $p$ -value of 0.14). Within-sample equality between the respondents who differ with respect to the position of the repeated choice task cannot be rejected either, when based on the Kruskal-Wallis test procedure ( $p$ -value of 0.67 and 0.98 for, respectively, samples 1 and 2).

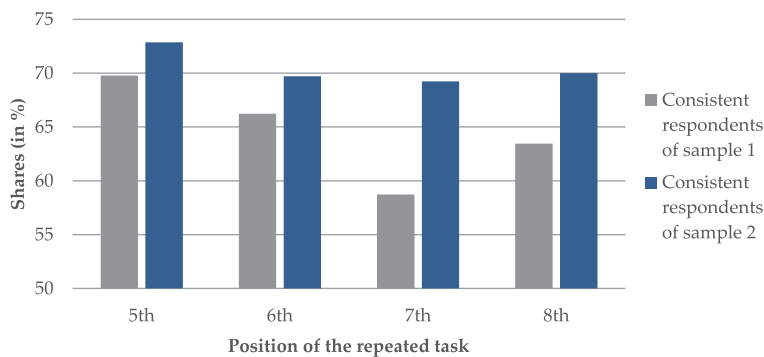


Figure 3. Choice consistency according to the position of the repeated choice task.

The share of monotonic respondents who did not choose the dominated alternative in the choice task that includes such an alternative are similar between the samples, with 88.1% and 86.7% of respondents choosing monotonically in, respectively, sample 1 and 2. These values are higher than the shares of consistent respondents. While choice monotonicity is higher for sample 1, choice consistency is higher for sample 2. The differences in monotonicity and consistency between samples 1 and 2 are, however, not statistically significant.

In order to investigate possible interactions between the three concepts across the two samples, Table 5 shows a cross-tabulation of frequencies of consistent and monotonic respondents. Table 5 reveals that the share of inconsistent respondents who also violate monotonicity (11%) does not differ significantly from the share of consistent respondents violating monotonicity (15.9%). A similar result holds for the share on non-monotonic respondents who behave inconsistently (59%), which is not dissimilar to the share of monotonic respondents who violate consistency (68.7%). Hence, the descriptive results shown here reveal only limited behavioural interactions between choice consistency and monotonicity. This is supported by the low degree of association measured by Goodman and Kruskal's gamma. Nevertheless, it may be the case that average choice certainty levels differ between the four categories of respondents shown in Table 5. However, the descriptive statistics reveal that this is not the case either: The self-reported average choice certainty values range from 2.5 (on the scale from 0 to 4) for non-consistent respondents to 2.6 for non-monotonic participants. Finally, Table 6 reports the pairwise correlation coefficients for choice consistency, monotonicity, and average choice certainty. Again, the low and statistically insignificant correlation values suggest independence of the three behavioural concepts in this specific study.

Table 5. Cross-tabulation of frequencies between monotonic and consistent respondents.

Monotonic respondents (total)	87.4%	Consistent respondents (total)	67.5%
% of which are consistent	31.3%	% of which are monotonic	84.1%
% of which are inconsistent	68.7%	% of which are non-monotonic	15.9%
Non-monotonic respondents (total)	12.6%	Inconsistent respondents (total)	32.5%
% of which are consistent	41.0%	% of which are monotonic	89.0%
% of which are inconsistent	59.0%	% of which are non-monotonic	11.0%

Table 6. Pairwise correlation coefficients for choice consistency, monotonicity, and average certainty.

	Consistency	Monotonicity	Certainty
Consistency	1.00		
Monotonicity	0.07	1.00	
Certainty	0.09	−0.03	1.00

## 5.2. *Swait-Louviere test results*

The Swait-Louviere test procedure as described in Section 3, was run to test the equality of choice behaviour between: (i) respondents who were certain and uncertain about their choices; (ii) consistent and inconsistent respondents; (iii) samples 1 and 2; and (iv), respondents who were subjected to the repeated choice task at different positions in the choice-task sequence. The test results are shown in Tables 7 (i-ii), 8 (iii), and 9 (iv)<sup>4</sup>.

Table 7 reports the results of the Swait-Louviere tests that compare the choice models estimated separately for certain and uncertain respondents, as well as for consistent and inconsistent respondents. The test for equality of choice models estimated for monotonic and non-monotonic respondents could not be performed because the number of respondents categorised as non-monotonic was too low. Certain respondents were defined as respondents who stated that they were either certain or very certain about their choices (on average across all choice tasks). Survey participants who stated that they were neither certain nor uncertain (i.e. who selected the response category 'neither certain nor uncertain') are hence classified as uncertain. This definition results in 46% certain and 54% uncertain respondents. The first step of the Swait-Louviere test procedure convincingly rejects the equality of certain and uncertain, as well as the equality of consistent and inconsistent respondents (column 5 in Table 7). In this case, it is meaningless to proceed with the second step of testing the equality of scale parameters, which for this reason is not reported in Table 7.

In order to compare choice behaviour between samples 1 and 2, Swait-Louviere tests were applied at the sample level and at the individual choice-task level. In the latter case, two choice tasks were merged for each test to ensure a sufficient number of observations for the identification of the models. That is, the choice tasks 1 and 2, 3 and 4, 5 and 6, and 7 and 8 were merged and, as such, compared for the two samples. Two main conclusions can be derived from the results presented in Table 8. First, the hypothesis of equality of preference parameters between the samples 1 and 2 (row 1, column 5), as well as the test for equality of scale parameters (row 1, column 6), cannot be rejected. In other words, the choice behaviour between the respondents in samples 1 and 2 is not significantly different. This result implies that asking the respondents about their choice certainty does not lead to systematic differences in their choices. Second, the test for equality of the samples with respect to specific (pairs of) choice tasks confirms this result, since the equality of preference and scale parameters between the samples cannot be rejected for any task combinations. This means that the presence of a choice certainty question also has no effect on choices while controlling for the position of choice tasks within the sequence.

The results of the Swait-Louviere tests on the equality of choice behaviour between samples that differ with respect to the position of the repeated choice task show that the equality of the preference parameters cannot be rejected for any of the comparisons (column 5 in Table 9). In other words, different positions of the repeated choice task are not associated with procedural biases. This conform to our expectations, since respondents were not informed either about the presence of a repeated task or about its position. Neither can the majority of the tests for the equality of the scale parameters reject equality (column 6 in Table 9). Significant differences in scale are found between the samples that received the repeated task at the 5th and 6th position, and weakly significant differences between the samples that were shown the repeated task at the 6th and 7th position. The results concerning differences in scale parameters suggest that the error variance of the sample with a repeated choice task at the 5th position is generally higher than the error variance of samples that were shown the repeated task at a later position (column 8 in Table 9). A possible explanation is that the 5th choice task appears immediately after the task involving a dominated alternative, which could have caused increased randomness in the following choice task.

**Table 7.** Swait-Louviere test results for certain vs. uncertain respondents (sample 1 only) and consistent vs. inconsistent (sample 1 and 2 merged) respondents.

Respondents' choice characteristics	(1) LL 1 <sup>st</sup> sample	(2) LL 2 <sup>nd</sup> sample	(3) LL pooled ( $\mu_1 \neq \mu_1$ ) <sup>a</sup>	(4) LL pooled ( $\mu_1 = \mu_1$ ) <sup>b</sup>	(5) 1 <sup>st</sup> LR-test: $p$ -value (11df) <sup>c</sup>	(6) 2 <sup>nd</sup> LR-test: $p$ -value (1df) <sup>d</sup>	(7) Rel. scale ( $\mu_1/\mu_2$ )	(8) Rel. variance ( $\sigma_1^2/\sigma_2^2$ )
Certain vs. uncertain	-826.174	-557.462	-1407.469	-1408.279	0.00***	<i>n/a</i>	<i>n/a</i>	<i>n/a</i>
Consistent vs. inconsistent	-1594.651	-1115.740	-2767.317	-2795.189	0.00***	<i>n/a</i>	<i>n/a</i>	<i>n/a</i>
Monotonic vs. non-monotonic	not enough respondents to estimate a model based on non-monotonic respondents only							

Notes: \*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.1$ ; LL: Log-likelihood; df: degrees of freedom. <sup>a</sup>Pooled MXL model allowing the scale parameters to vary between the samples (the scale parameter of the 1st sample is fixed at 1.) <sup>b</sup>Pooled MXL model constraining the scale parameters to be equal between the samples. <sup>c</sup>Test for differences in the preference parameters between the samples. <sup>d</sup>Test for differences in the scale parameters between the samples.

**Table 8.** Swait-Louviere test results for equality of choice behaviour between samples 1 and 2.

Choice samples	(1) LL 1 <sup>st</sup> sample	(2) LL 2 <sup>nd</sup> sample	(3) LL pooled ( $\mu_1 \neq \mu_1$ ) <sup>a</sup>	(4) LL pooled ( $\mu_1 = \mu_1$ ) <sup>b</sup>	(5) 1 <sup>st</sup> LR-test: $p$ -value (11df) <sup>c</sup>	(6) 2 <sup>nd</sup> LR-test: $p$ -value (1df) <sup>d</sup>	(7) Rel. scale ( $\mu_1/\mu_2$ )	(8) Rel. variance ( $\sigma_1^2/\sigma_2^2$ )
All tasks - sample 1 vs. 2	-1408.279	-1380.541	-2794.920	-2795.189	0.349	0.464	0.94	1.13
Tasks 1 & 2 - sample 1 vs. 2	-424.821	-462.513	-893.229	-893.255	0.380	0.819	1.04	0.92
Tasks 3 & 4 - sample 1 vs. 2	-361.595	-382.052	-747.710	-748.311	0.702	0.273	0.76	1.73
Tasks 5 & 6 - sample 1 vs. 2	-427.809	-437.435	-871.809	-871.813	0.285	0.927	1.02	0.96
Tasks 7 & 8 - sample 1 vs. 2	-399.478	-407.668	-811.260	-812.477	0.693	0.119	0.69	2.10

Notes: \*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.1$ ; LL: Log-likelihood; df: degrees of freedom.

**Table 9.** Swait-Louviere test results for equality of choice behaviour between samples that differ in the position of the repeated choice task (samples 1 and 2 merged).

Position of the repeated choice task	(1) LL 1 <sup>st</sup> sample	(2) LL 2 <sup>nd</sup> sample	(3) LL pooled ( $\mu_1 \neq \mu_1$ ) <sup>a</sup>	(4) LL pooled ( $\mu_1 = \mu_1$ ) <sup>b</sup>	(5) 1 <sup>st</sup> LR-test: $p$ -value (11df) <sup>c</sup>	(6) 2 <sup>nd</sup> LR-test: $p$ -value (1df) <sup>d</sup>	(7) Rel. scale ( $\mu_1/\mu_2$ )	(8) Rel. variance ( $\sigma_1^2/\sigma_2^2$ )
5th vs. 6th	-666.055	-828.457	-1495.979	-1495.979	0.992	0.016**	0.74	1.83
5th vs. 7th	-666.055	-659.276	-1327.242	-1327.242	0.975	0.292	0.86	1.35
5th vs. 8th	-666.055	-629.770	-1300.589	-1300.589	0.573	0.177	0.84	1.42
6th vs. 7th	-828.457	-659.276	-1488.794	-1488.794	0.998	0.075*	1.24	0.65
6th vs. 8th	-828.457	-629.770	-1462.262	-1462.262	0.707	0.255	1.16	0.74
7th vs. 8th	-659.276	-629.770	-1292.725	-1292.725	0.769	0.536	0.92	1.18

Notes: \*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.1$ ; LL: Log-likelihood; df: degrees of freedom.

**Table 10.** Probit regression results.

Variable	Choice certainty <sup>a</sup> Coeff. (s.e.)	Choice consistency <sup>b</sup> Coeff. (s.e.)	Choice monotonicity <sup>b</sup> Coeff. (s.e.)
Choice-task-related variables			
Position of choice task in sequence (1 to 8)	−0.134* (0.070)		
Position of choice task in sequence (squared)	0.013* (0.007)		
Time spent per choice task (log)	−0.429*** (0.076)		
Entropy	−2.041*** (0.290)	−0.484 (0.804)	−0.179 (1.051)
Question on choice certainty (1 = yes, 0 = no)		−0.193 (0.121)	0.024 (0.158)
Position of the repeated task (1 = 7./8., 0 = 5./6.)		−0.124 (0.121)	
Individual-specific variables			
Income above sample average (1 = yes, 0 = no)	0.205 (0.270)	0.252* (0.128)	0.211 (0.169)
Female (1 = yes, 0 = no)	−1.014*** (0.264)	0.172 (0.126)	0.579*** (0.169)
Age	−0.001 (0.008)	0.006 (0.004)	0.002 (0.005)
University education (1 = yes, 0 = no)	0.270 (0.281)	0.051 (0.133)	0.262 (0.183)
Recreates in/at waters (1 = yes, 0 = no)	0.884* (0.499)	−0.345 (0.224)	0.213 (0.245)
Member of an env. org. (1 = yes, 0 = no)	0.038 (0.265)	0.155 (0.125)	0.586*** (0.174)
Survey understandability <sup>c</sup>	1.368*** (0.315)	0.139 (0.145)	0.333* (0.174)
Time spent reading informational pages	0.048 (0.031)	0.040** (0.017)	0.056* (0.024)
Model characteristics			
Constant		0.519 (0.844)	−0.139 (1.081)
$\mu_1^d$	−3.706*** (0.768)		
$\mu_2^d$	−2.307*** (0.764)		
$\mu_3^d$	0.839 (0.760)		
$\sigma$	3.442*** (0.422)		
# Observations	1,880	483	483
# Respondents	235	483	483
Log-likelihood	−1434.2	−453.5	−453.5
Mcfaddens's pseudo $R^2$	0.049	0.068	0.068

Notes: \*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.1$ .<sup>a</sup>Random-effects ordered probit regression.<sup>b</sup>Bivariate probit regression.<sup>c</sup>1 = very/rather, 0 = partly/rather not/not at all.<sup>d</sup>Cut-off points of the ordered logistic regression.

### 5.3. Probit regression model results

The results of the ordered and bivariate probit models are shown in Table 10. The model on choice certainty is estimated on sample 1 only, because only the respondents in this sample received the choice certainty questions. The model which simultaneously regresses choice consistency and monotonicity on possible determinants is estimated on the pooled sample of respondents. The effect of the follow-up certainty questions on choice consistency and monotonicity is controlled for by a dummy variable for sample 1 (the sample that contains the certainty question). Apart from theoretical expectations about the determinants of choice certainty, consistency, and monotonicity based on the

existing literature, log-likelihood ratio tests were applied to find the model with the best fit. The correlations of the variables in both regressions are reasonable, with only a few pairs of variables that have correlation coefficients which are larger than 0.2. Excluding these variables does not affect the main conclusions. Note that the ordinal variable for choice certainty was reduced from five to four levels by merging the ‘not certain at all’ and ‘not certain’ response categories, because there were only a few observations in the most uncertain category. No case study-specific attributes are included in the analysis in order to ensure the generalisability of the results. Furthermore, choice task-specific effects are controlled for by choice-task-related variables, e.g. the variables which reflect the position of a choice task in the choice sequence and the measure of complexity (entropy). McFadden’s (McFadden 1974) pseudo- $R^2$  values suggest that the models explain a rather small part of the variation in the data. However, we consider this not to be critical for three reasons: First, many of the pseudo- $R^2$  values found in the literature, if they are reported at all, are of comparable magnitudes (e.g. Colombo, Glenk, and Rocamora-Montiel 2016; Hensher, Rose, and Beck 2012; Kosenius 2009; Mørkbak and Olsen 2015); second, pseudo- $R^2$  values are, in general, considerably lower than the traditional  $R^2$  values obtained in OLS regressions (McFadden 1979); and, third, the pseudo- $R^2$  is of minor interest in the context of this paper, as we are more interested in testing relationships than in building prediction models.

The results of the model on stated choice certainty indicates that the position of the choice task in the choice sequence has an inverted bell-shaped effect on choice certainty. This supports the results of San Miguel, Ryan, and Amaya-Amaya (2005), but, in our case, with respect to choice certainty. Since the choice certainty analysis embraces all choice tasks faced by each respondent, we include the time used per choice task as a regressor. This proves to exert a highly significant and negative effect, meaning that survey participants who used more time to complete a choice task are more likely to report a low choice certainty for that task. On the other hand, the time a respondent spent on reading relevant informational pages of the survey has no statistically significant effect on choice certainty as measured by common significance thresholds. However, the coefficient almost reaches the 10% significance level with a  $p$ -value of 0.12. All these informational pages (seven in total) appeared before the choice tasks, and conveyed information on hydropower, nuclear power, their associated risks, the expansion scenarios of hydropower, and the DCE. None of the informational pages contained survey questions. As expected, the coefficient of entropy is negative, implying that higher entropy, associated with increased choice-task complexity, results in lower choice certainty. The only sociodemographic variable that turns out to have a significant effect on choice certainty is gender. Female respondents seem to be less certain about their choices than male respondents. In contrast with the existing literature, we cannot report any effect of income on choice certainty. Respondents who state the survey to be very or rather understandable are significantly more likely to have a higher choice certainty. Finally, the random effects parameter,  $\sigma$ , is significant, meaning that stated choice certainty is correlated across the choice tasks within the set of an individual’s responses (Olsen et al. 2011). Log-likelihood ratio tests that compare the current model with ordered probit models without panel specifications show that there exists sufficient variability between individuals, and hence that a panel specification performs better than a standard ordered logistic regression<sup>5</sup>.

Turning next to the results of the bivariate probit model on choice consistency and monotonicity, we find that female respondents are significantly more likely to choose monotonically. Note that we find the opposite effect of gender in the case of choice certainty, which shows that female respondents are less certain about their choices. Neither age nor the square root of age has an impact in any of the models. We hence cannot confirm the bell-shaped effect of age which is reported in San Miguel, Ryan, and Amaya-Amaya (2005). As other variables are unaffected by the inclusion of the square root of age, we decided to exclude the square root of age in the final models. In the case of choice consistency and monotonicity, and in contrast to choice certainty, entropy remains insignificant. The time spent reading informational pages in the survey increases both, consistency and monotonicity. Finally, income uniquely affects choice monotonicity.



## 6. Discussion and conclusions

This study simultaneously investigates choice certainty, choice consistency, and choice monotonicity in the context of DCEs. Several conclusions can be drawn from the results of the Swait-Louviere tests. First, choice behaviour significantly differs between certain and uncertain, as well as between consistent and inconsistent, respondents. Second, we do not find a procedural effect caused by the inclusion of a choice certainty question after each choice task on choice certainty, consistency, or monotonicity. The procedural invariance with respect to choice certainty questions confirms the findings of Brouwer et al. (2010). In addition to Brouwer et al. (2010), we also investigate whether the same result holds at a choice-task level, and find that it does. In other words, there exists neither an overall nor a sequence-dependent procedural effect of asking follow-up choice certainty questions. Third, as expected, the equality of choice behaviour between samples that differ by the position of the repeated choice task cannot be rejected, which implies that the position of a repeated choice task in a choice-task sequence does not lead to any systematic changes in choice behaviour.

We identify a number of idiosyncratic determinants for each of the three concepts. However, we only find limited evidence for common drivers of choice certainty, consistency, and monotonicity, suggesting that these are rather separate constructs. The factors that seem to have an effect on all three concepts are the time spent reading informational pages in the survey and gender, although the evidence for the impact of gender on choice consistency is weak. Female participants are less certain about their choices, confirming the findings of Brouwer et al. (2010), Dekker et al. (2016), and Olsen et al. (2011). At the same time, female respondents more often choose monotonically.

Choice uncertainty, inconsistency and non-monotonicity violate several assumptions and axioms underlying DCEs. Therefore, we need to better understand the underlying mechanisms to be able to minimise them and increase the validity and reliability of the DCE results. To this end, we develop a number of recommendations for the DCE literature based on our main findings. Significant differences between the choice behaviour of certain and uncertain and consistent and inconsistent survey participants imply that choice uncertainty and inconsistency can significantly influence the DCE results. It is hence important to control for these factors in the choice models in order to be able to capture their effects. The drivers of choice certainty, consistency, and monotonicity identified in this study could be used as explanatory variables in choice models in order to achieve this goal and possibly mitigate the issue of endogeneity. Moreover, there is evidence that it is necessary to ensure the utmost cognitive ease of a survey, as several results show that this is an important factor in reducing choice uncertainty, inconsistency, and non-monotonicity. Complexity of the choice tasks and the understandability of the survey, both of which are likely to be associated with the cognitive ease of responding to the choice questions, are expected to increase choice consistency. The time spent on choice tasks has a negative effect on choice certainty, and the time spent reading informational pages positively affects choice consistency and monotonicity. This suggests that the information part of the survey plays a crucial role in respondent decision-making and needs to be covered in as much detail as possible by the respondents. The level of information provision in stated preference research has a long history. Ensuring that all respondents fully understand the provided information is a major challenge in social surveys. Possible solutions to ensure respondents have read the information statement in online surveys is to ask them a limited number of follow-up questions, which can only be answered if they have read the preceding information, and not allowing them to proceed without answering these information related questions first. This option might be particularly challenging, however, to keep respondents interested in completing the survey. Procedurally, no reservations against directly asking survey respondents about their choice certainty can be confirmed.

Further research combining the analysis of choice certainty with the analysis of choice consistency and monotonicity could provide additional evidence on common and idiosyncratic drivers of these concepts. The results of the bivariate probit model estimated in this paper suggests that consistency and monotonicity are influenced by idiosyncratic regressors. In addition, a likelihood-ratio test

shows that the independence of the errors terms between choice consistency and monotonicity in the bivariate probit model cannot be rejected. Both of these results suggest that choice consistency and monotonicity are independent constructs. Nevertheless, due to data limitations we could not combine the analysis of choice consistency and monotonicity with an analysis of choice certainty. An instrumental variable (IV) approach would allow this question to be answered more reliably. Since we did not find instruments that exclusively describe any of the three concepts with sufficient predictive power, and, at the same time, are not expected to be correlated with the other concepts, we were not able to conduct such an analysis. If other unique determinants were found, or those identified in this study were confirmed by other studies that analyze all three concepts simultaneously, it may be justified to estimate instrumental variable models which can provide further evidence for (the lack of) possible interrelations between the three concepts.

Finally, the results of this study suggest strong effects of gender on choice certainty and monotonicity. Although a rather extensive literature on gender differences exists in other areas of economics, such as behavioural experiments (e.g. Charness and Gneezy 2012) or marketing (e.g. Meyers-Levy and Loken 2015), there are only a few valuation studies in environmental economics which explicitly focus on gender effects. An example is Ladenburg and Olsen (2008), who report a starting-point bias induced by an instructional choice set for female but not for male respondents. The authors explain their findings by the 'selective hypothesis' of Meyers-Levy (Meyers-Levy 1989; Meyers-Levy and Loken 2015), which states that females have a tendency to process information more comprehensively than males, who are more selective information processors and more often rely on heuristics. Testing the selective hypothesis and, more generally, an in-depth examination of the gender effects in the realm of DCE constitutes another area for further research.

## Notes

1. Some studies test for alternative definitions of choice consistency. Brown et al. (2008), for example, focus on circular triads, that is, on preferences which imply  $A > B > C > A$ .
2. We assume that, even for the case of a complete phase-out of nuclear power in Switzerland, French nuclear plants that are located within 40km of the Swiss border would still be a cause of risk.
3. Note that the average annual electricity bill per household in Switzerland roughly equals 930 CHF (ca. 930 USD) (ElCom 2014).
4. See Appendix for the choice model results underlying the Swait-Louviere analysis (MXL choice models based on sample 1 only, sample 2 only, and the combined sample).
5. The results are available from the authors on request.

## Disclosure statement

No potential conflict of interest was reported by the authors.

## Funding

This study was funded by the Swiss Federal Institute of Aquatic Science and Technology (Eawag), and is part of the Competence Center for Research in Energy, Society and Transition (SCCER CREST).

## ORCID

Roy Brouwer  <http://orcid.org/0000-0002-0525-2050>

## References

- Bateman, I. J., R. T. Carson, B. Day, M. Hanemann, N. Hanley, T. Hett, and M. Jones-Lee, et al. 2002. *Economic Valuation with Stated Preference Techniques*. Cheltenham: Edward Elgar.

- Bateman, I. J., A. Munro, and G. L. Poe. 2008. "Decoy Effects in Choice Experiments and Contingent Valuation: Asymmetric Dominance." *Land Economics* 84: 115–127.
- Beck, M. J., S. Fifer, and J. M. Rose. 2016. "Can you ever be Certain? Reducing Hypothetical Bias in Stated Choice Experiments via Respondent Reported Choice Certainty." *Transportation Research Part B: Methodological* 89: 149–167.
- Beck, M. J., J. M. Rose, and D. A. Hensher. 2013. "Consistently Inconsistent: The Role of Certainty, Acceptability and Scale in Choice." *Transportation Research Part E: Logistics and Transportation Review* 56: 81–93.
- Bliem, M., M. Getzner, and P. Rodiga-Laßnig. 2012. "Temporal Stability of Individual Preferences for River Restoration in Austria using a Choice Experiment." *Journal of Environmental Management* 103: 65–73.
- Brouwer, R., T. Dekker, J. Rolfe, and J. Windle. 2010. "Choice Certainty and Consistency in Repeated Choice Experiments." *Environmental and Resource Economics* 46: 93–109.
- Brouwer, R., I. Logar, and O. Sheremet. 2016. "Choice Consistency and Preference Stability in Test-Retests of Discrete Choice Experiment and Open-Ended Willingness to Pay Elicitation Formats." *Environmental and Resource Economics* 68: 729–751.
- Brown, T. C., D. Kingsley, G. L. Peterson, N. E. Flores, A. Clarke, and A. Birjulin. 2008. "Reliability of Individual Valuations of Public and Private Goods: Choice Consistency, Response Time, and Preference Refinement." *Journal of Public Economics* 92: 1595–1606.
- Burgherr, P., and S. Hirschberg. 2008. "A Comparative Analysis of Accident Risks in Fossil, Hydro, and Nuclear Energy Chains." *Human and Ecological Risk Assessment: An International Journal* 14: 947–973.
- Burgherr, P., and S. Hirschberg. 2014. "Comparative Risk Assessment of Severe Accidents in the Energy Sector." *Energy policy* 74: S45–S56.
- Cameron, A. C., and P. K. Trivedi. 2005. *Microeconometrics: Methods and Applications*. New York, NY: Cambridge University Press.
- Campbell, D. 2007. "Willingness to Pay for Rural Landscape Improvements: Combining Mixed Logit and Random-Effects Model." *Journal of Agricultural Economics* 58: 467–483.
- Carlsson, F., M. R. Mørkbak, and S. B. Olsen. 2012. "The First Time is the Hardest: A Test of Ordering Effects in Choice Experiments." *Journal of Choice Modelling* 5: 19–37.
- Charness, G., and U. Gneezy. 2012. "Strong Evidence for Gender Differences in Risk Taking." *Journal of Economic Behaviour and Organization* 83: 50–58.
- Colombo, S., K. Glenk, and B. Rocamora-Montiel. 2016. "Analysis of Choice Inconsistencies in On-line Choice Experiments: Impact on Welfare Measures." *European Review of Agricultural Economics* 43: 271–302.
- De Bekker-Grob, E. W., M. Ryan, and K. Gerard. 2012. "Discrete Choice Experiments in Health Economics: A Review of the Literature." *Health Economics* 21: 145–172.
- Dekker, T., S. Hess, R. Brouwer, and M. Hofkes. 2016. "Decision Uncertainty in Multi-Attribute Stated Preference Studies." *Resource and Energy Economics* 43: 57–73.
- Dekker, T., P. Koster, and R. Brouwer. 2014. "Changing with the Tide: Semi-Parametric Estimation of Preference Dynamics." *Land Economics* 90: 717–745.
- Determann, D., M. S. Lambooi, E. W. Steyerberg, E. W. De Bekker-Grob, and G. Ardine De Wit. 2017. "Impact of Survey Administration Mode on the Results of a Health-Related Discrete Choice Experiment: Online and Paper Comparison." *Value in Health* (in press).
- ElCom. 2014. Strompreise 2015: Durchschnittlich steigende Tarife für Haushalte und mittlere Betriebe. <https://www.admin.ch/gov/de/start/dokumentation/medienmitteilungen.msg-id-54336.html>.
- Foster, V., and S. Mourato. 2002. "Testing for Consistency in Contingent Ranking Experiments." *Journal of Environmental Economics and Management* 44: 309–328.
- Hanley, N., R. E. Wright, and G. Koop. 2002. "Modelling Recreation Demand Using Choice Experiments: Climbing in Scotland." *Environmental and Resource Economics* 22: 449–466.
- Hensher, D. A., J. M. Rose, and M. J. Beck. 2012. "Are There Specific Design Elements of Choice Experiments and Types of People that Influence Choice Response Certainty?" *Journal of Choice Modelling* 5: 77–97.
- Hensher, D. A., J. M. Rose, and W. H. Greene. 2012. "Inferring Attribute Non-Attendance from Stated Choice Data: Implications for Willingness to Pay Estimates and a Warning for Stated Choice Experiment Design." *Transportation* 39: 235–245.
- Hirschberg, S., C. Bauer, P. Burgherr, E. Cazzoli, T. Heck, M. Spada, and K. Treyer. 2016. "Health Effects of Technologies for Power Generation: Contributions from Normal Operation, Severe Accidents and Terrorist Threat." *Reliability Engineering & System Safety* 145: 373–387.
- Jehle, G., and P. Reny. 2001. *Advanced Microeconomic Theory*. Boston: Addison Wesley.
- Kosenius, A.-K. 2009. "Causes of Response Uncertainty and its Implications for WTP Estimation in Choice Experiment." Discussion Papers No. 29, University of Helsinki.
- Ladenburg, J., and S. B. Olsen. 2008. "Gender-Specific Starting Point Bias in Choice Experiments: Evidence from an Empirical Study." *Journal of Environmental Economics and Management* 56: 275–285.
- Lancsar, E., and J. Louviere. 2006. "Deleting 'Irrational' Responses from Discrete Choice Experiments: A Case of Investigating or Imposing Preferences?" *Health Economics* 15: 797–811.

- Liebe, U., J. Meyerhoff, and V. Hartje. 2012. "Test-Retest Reliability of Choice Experiments in Environmental Valuation." *Environmental and Resource Economics* 53: 389–407.
- Lundhede, T. H., S. B. Olsen, J. B. Jacobsen, and B. J. Thorsen. 2009. "Handling Respondent Uncertainty in Choice Experiments: Evaluating Recoding Approaches Against Explicit Modelling of Uncertainty." *Journal of Choice Modelling* 2: 118–147.
- McFadden, D. 1974. "Conditional Logit Analysis of Qualitative Choice Behaviour." In *Frontiers in Econometrics*, edited by P. Zarembka. New York: Academic Press.
- McFadden, D. 1979. "Quantitative Methods for Analyzing Travel Behaviour on Individuals: Some Recent Developments." In *Behavioural Travel Modelling*, edited by D. A. Hensher and P. R. Stopher, 279–318. Kent: Croom Helm.
- Meyers-Levy, J. 1989. "Gender Differences in Information Processing: A Selective Interpretation." In *Cognitive and Affective Responses to Advertising*, edited by P. Cafferata and A. M. Tybout. Lanham, MA: Lexington.
- Meyers-Levy, J., and B. Loken. 2015. "Revisiting Gender Differences: What We Know and What Lies Ahead." *Journal of Consumer Psychology* 25: 129–149.
- Mørkbak, M.R., and S.B. Olsen. 2015. "A Within-Sample Investigation of Test-Retest Reliability in Choice Experiment Surveys with Real Economic Incentives." *Australian Journal of Agricultural and Resource Economics* 59: 375–392.
- Olsen, S.B., T.H. Lundhede, J.B. Jacobsen, and B.J. Thorsen. 2011. "Tough and Easy Choices: Testing the Influence of Utility Difference on Stated Certainty-in-Choice in Choice Experiments." *Environmental and Resource Economics* 49: 491–510.
- Özdemir, S., A.F. Mohamed, F.R. Johnson, and A.B. Hauber. 2010. "Who pays Attention in Stated-choice Surveys?" *Health Economics* 19: 111–118.
- Rabin, M. 1998. "Psychology and Economics." *Journal of Economic Literature* 36: 11–46.
- Rigby, D., M. Burton, and J. Pluske. 2016. "Preference Stability and Choice Consistency in Discrete Choice Experiments." *Environmental and Resource Economics* 65: 441–461.
- Rulleau, B., and J. Dachary-Bernard. 2012. "Preferences, Rational Choices and Economic Valuation: Some Empirical Tests." *Journal of Socio-Economics* 41: 198–206.
- San Miguel, F., M. Ryan, and M. Amaya-Amaya. 2005. "'Irrational' Stated Preferences: A Quantitative and Qualitative Investigation." *Health Economics* 14: 307–322.
- Scarpa, R., D. Campbell, and W.G. Hutchinson. 2007. "Benefit Estimates for Landscape Improvements : Sequential Bayesian Design and Respondents –Rationality in a Choice Experiment." *Land Economics* 83: 617–634.
- Schaafsma, M., R. Brouwer, I. Liekens, and L. de Nocker. 2014. "Temporal Stability of Preferences and Willingness to Pay for Natural Areas in Choice Experiments: A Test-Retest." *Resource and Energy Economics* 38: 243–260.
- SFOE (Swiss Federal Office of Energy). 2013. Energieperspektiven 2050. Zusammenfassung. <http://www.bfe.admin.ch/energiestrategie2050/index.html?lang=de>.
- SFOE (Swiss Federal Office of Energy). 2016. Schweizerische Elektrizitätsstatistik 2015. BFE. <http://www.bfe.admin.ch/themen/00526/00541/00542/00630/index.html?lang=de>.
- Shannon, C.E. 1948. "A Mathematical Theory of Communication." *The Bell System Technical Journal* 27: 379–423.
- Soliño, M., B.A. Farizo, M.X. Vázquez, and A. Prada. 2012. "Generating Electricity with Forest Biomass: Consistency and Payment Timeframe Effects in Choice Experiments." *Energy policy* 41: 798–806.
- Soofi, E.S. 1994. "Capturing the Intangible Concept of Information." *Journal of the American Statistical Association* 89: 1243–1254.
- Swait, J., and W. Adamowicz. 2001a. "Choice Environment, Market Complexity, and Consumer Behaviour: A Theoretical and Empirical Approach for Incorporating Decision Complexity into Models of Consumer Choice." *Organizational Behaviour and Human Decision Processes* 86: 141–167.
- Swait, J., and W. Adamowicz. 2001b. "The Influence of Task Complexity on Consumer Choice : A Latent Class Model of Decision Strategy Switching." *Journal of Consumers Research* 28: 135–148.
- Swait, J., and J.J. Louviere. 1993. "The Role of the Scale Parameter in the Estimation and Use of Multinomial Logit Models." *Journal of Marketing Research* 30: 305–314.
- Uggeldahl, K., C. Jacobsen, T.H. Lundhede, and S.B. Olsen. 2016. "Choice Certainty in Discrete Choice Experiments: Will Eye Tracking Provide Useful Measures?" *Journal of Choice Modelling* 20: 35–48.

**Appendix. Choice model results (attribute-only models)**

Variable	MXL model (sample 1) Coeff. (s.e.)	MXL model (sample 2) Coeff. (s.e.)	MXL model (sample 1 & 2) Coeff. (s.e.)
Mean estimates			
ASC	−4.812*** (0.402)	−6.047*** (0.624)	−5.306*** (0.343)
Construction of new hydropower hydropower plants	−0.746*** (0.131)	−0.899*** (0.157)	−0.827*** (0.102)
Increase in lifetime risk of death from a dam breach by 40%	−0.444*** (0.101)	−0.330*** (0.099)	−0.406*** (0.071)
Increase in lifetime risk of death from a nuclear accident by 60%	0.740*** (0.098)	0.715*** (0.105)	0.706*** (0.072)
Increase in household's yearly electricity bill	−0.446*** (0.052)	−0.490*** (0.060)	−0.476*** (0.040)
Standard deviations of normally distributed random parameters			
ASC	3.002*** (0.347)	4.712*** (0.518)	3.838*** (0.320)
Construction of new hydropower hydropower plants	1.520*** (0.146)	1.890*** (0.173)	1.723*** (0.114)
Increase in lifetime risk of death from a dam breach by 40%	0.795*** (0.146)	0.625*** (0.174)	0.715*** (0.109)
Increase in lifetime risk of death from a nuclear accident by 60%	0.811*** (0.125)	0.910*** (0.142)	0.859*** (0.091)
Increase in household's yearly electricity bill	0.508*** (0.052)	0.606*** (0.068)	0.543*** (0.040)
Model characteristics			
Number of observations	1880	1984	3864
Log-likelihood (restricted)	−2065.391	−2179.647	−4245.038
Log-likelihood (unrestricted)	−1406.276	−1381.806	−2793.311
AIC/N	1.507	1.403	1.451
McFaddens's pseudo $R^2$	0.319	0.366	0.342

Notes: \*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.1$ .