Journal of Choice Modelling 50 (2024) 100465

Contents lists available at ScienceDirect

Journal of Choice Modelling

journal homepage: www.elsevier.com/locate/jocm

On the impact of decision rule assumptions in experimental designs on preference recovery: An application to climate change adaptation measures

Sander van Cranenburgh [a], Jürgen Meyerhoff [b], Katrin Rehdanz [c], Andrea Wunsch [d,*]

[a] Engineering Systems and Services, Delft University of Technology, Jaffalaan 5, 2628 BX, Delft, the Netherlands
[b] Hochschule für Wirtschaft und Recht Berlin (HWR), Badensche Straße 52, 10825, Berlin, Germany
[c] Kiel University, Institute for Environmental, Resource and Spatial Economics, Olshausenstraße 40-60, 24098, Kiel, Germany
[d] Kiel Institute for the World Economy, Kiellinie 66, 24105, Kiel, Germany

ARTICLE INFO

Keywords:
Coastal adaptation
Climate change
Experimental design theory
Decision rules
Random regret minimisation

ABSTRACT

Efficient experimental designs aim to maximise the information obtained from stated choice data to estimate discrete choice models' parameters statistically efficiently. Almost without exception efficient experimental designs assume that decision-makers use a Random Utility Maximisation (RUM) decision rule. When using such designs, researchers (implicitly) assume that the decision rule used to generate the design has no impact on respondents' choice behaviour. This study investigates whether the decision rule assumption underlying an experimental design affects respondents' choice behaviour. We use four stated choice experiments on coastal adaptation to climate change: Two are based on experimental designs optimised for utility maximisation and two are based on experimental designs optimised for a mixture of RUM and Random Regret Minimisation (RRM). Generally, we find that respondents place value on adaptation measures (e.g., dykes and beach nourishments). We evaluate the models' fits and investigate whether some choice tasks particularly invoke RUM or RRM decision rules. For the latter, we develop a new sampling-based approach that avoids the confounding between preference and decision rule heterogeneity. We find no evidence that RUM-optimised designs invoke RUM-consistent choice behaviour. However, we find a relationship between some of the attributes and decision rules, and compelling evidence that some choice tasks invoke RUM consistent behaviour while others invoke RRM consistent behaviour. This implies that respondents' choice behaviour and choice modelling outcomes are not exogenous to the choice tasks, which can be particularly critical when information on preferences is used to inform actual decision-making on a sensitive issue of common interest as climate change.

1. Introduction

Compelling evidence in the choice modelling field (and beyond) shows that decision-makers employ a wide variety of decision rules. In the choice modelling field, in particular Random Regret Minimisation (RRM) models have emerged as a popular alternative to traditional Random Utility Maximisation (RUM) based models (Chorus et al., 2014). RRM models postulate that decision-makers

* Corresponding author.
E-mail address: andrea.wunsch@ifw-kiel.de (A. Wunsch).

https://doi.org/10.1016/j.jocm.2023.100465
Received 14 April 2023; Received in revised form 28 November 2023; Accepted 18 December 2023
Available online 3 January 2024

minimise their regrets, which are experienced when a competing option outperforms the chosen option on a specific attribute (Chorus, 2010). Since its inception, RRM models have been applied to both stated choice (SC) and revealed choice data sets, across a diverse range of fields, including health (Buckell et al., 2021; Paul et al., 2018), environment (Long et al., 2022; Mao et al., 2020), food (Biondi et al., 2019), transportation (Zhu et al., 2021), and tourism (Masiero et al., 2019). The compelling evidence in support of RRM and other non-RUM decision rules urges choice modellers against uncritically assuming that decision-makers always maximise their utility.

Choice modellers widely use Stated Choice (SC) experiments to collect choice data. Whereas traditionally choice modellers have used orthogonal experimental designs, in recent years so-called efficient designs have gained in popularity. Efficient designs aim to maximise the information gained from the data by carefully selecting the choice tasks presented to respondents while minimising the number of choice observations required to estimate the parameters of interest. Generating an efficient design requires the researcher to specify the choice model he/she intends to estimate, including its decision rule and expected parameters (priors). Almost without exception efficient experimental designs have been optimised for the assumption that decision-makers use a RUM decision rule. In doing so, researchers (implicitly) assume that the decision rule used to create the efficient design has no impact on respondents' choice behaviour.

However, the validity of the assumption that the decision rule used to create efficient designs has no impact on the decision rule used by respondents has been called into question. In 2018, efficient design theory is extended to RRM models, making it possible to generate efficient designs assuming decision-makers minimise regret (Van Cranenburgh et al., 2018). In their study, the authors use their extended design theory to investigate whether the decision rule assumption taken to generate the experimental design impacts the decision rule used by decision-makers. Based on a small sample ($N = 106$), the authors find preliminary evidence that the decision rule used to generate the design (RUM or RRM) might invoke that particular decision rule. To date, it remains unclear whether the recovery of one or the other decision rule as being most likely depends on the experimental design, or, possibly, on the inclusion of particular choice tasks in the experimental design. If such a relation exists, it will imply that the modelling outcomes and, particularly, preferences recovered based on efficient designs are not exogenous to the experimental design itself.

This paper further scrutinises the assumption that the decision rule used to create the experimental design does not impact the recovered preferences and/or decision rules. To do so, we generate four independent SC experiments: two using RUM-optimised designs and the other two optimised for a mixture of RUM and RRM. The SC experiments concern climate change adaptation measures in coastal areas of both the Baltic Sea and the North Sea in Germany. Coastal areas are especially vulnerable to climate change due to rising sea levels and an increasing frequency of severe events such as storms. As decisions about coastal adaptation measures require costly trade-offs (Johnston et al., 2018), a choice experiment was considered an adequate tool for evaluating their magnitude. Its design closely follows a study by Meyerhoff et al. (2021), which was conducted only for coastal adaptation at the Baltic Sea. A choice set with three alternatives was used, two hypothetical, and one that would maintain today's measures. Adaptation measures included beach nourishment, dyke heightening and realignment of dykes and dunes, among others, following expert consultations and matched to the two coasts. Given the potentially severe consequences of poorly chosen adaptation measures in coastal areas, we expected that regret minimisation behaviour might explain our choice data well. This applies, in particular, to the adaptation measure dyke heightening. It refers to a traditional, widely used hard defence structure that we expected respondents would regret if it had a lower height. Further, especially when information on preferences derived from a SC experiment is intended to guide policy-making on sensitive issues of common interest with severe consequences, such as climate change, it is crucial to be aware of a potential relationship between the type of attributes and decision rules applied.

For each of the four resulting data sets, we look at whether RUM models fit comparatively better on the design optimised for RUM and RRM models or whether they fit better on the design optimised for a mixture of RUM and RRM. Additionally, we develop a novel sampling-based approach to investigate whether some choice tasks particularly invoke RUM or RRM decisions rules (and thus are comparatively better explained by RUM or RRM). Importantly, our sampling-based approach avoids the confounding between preference and decision rule heterogeneity that thwarts commonly used latent class-based approaches (Hess et al., 2012; Boeri and Longo, 2017; Charoniti et al., 2020; Nielsen and Jacobsen 2020; Buckell et al., 2021).

The remainder of this paper is organised as follows. Section 2 describes the data collection, including the experimental designs. Section 3 presents the methodology, including the sampling-based approach to investigate whether particular choice tasks invoke decision rules. Section 4 reports the (estimation) results. Finally, section 5 provides the conclusions and a discussion.

2. Data collection

2.1. Treatments

The overall objective of the SC experiments used in this study was to elicit the preferences of the German population for adaptation to climate change on both coasts in the northern part of Germany, the North Sea coast in the West and the Baltic Sea coast in the East. Our SC experiment comprises of four overall treatments (Table 1). The treatments vary across coast lines (Baltic Sea vs. North Sea) and decision rules to create the experimental design (RUM vs. a mixture of RUM and RRM). As both coasts are not directly connected but separated by the German mainland and differ slightly in their topographical characteristics, the treatment was adapted for each coast. Furthermore, we used two experimental designs, a RUM-optimised design and a design optimised for a mixture of RUM and RRM.

Table 1
Treatments.

		Coastline	
		Baltic Sea	North Sea
Experimental design optimised for:	RUM only	T1	T3
	Mixture of RUM and RRM	T2	T4

Importantly, since the number of attributes and the levels are (almost) equal across coastlines, we have used the same experimental designs across the North Sea and the Baltic Sea.

Respondents were (semi-) randomly assigned to one of the four treatments and subsequently randomly assigned to one of the four blocks of that design.[1]

2.2. Attributes and experimental design

The adaptation measures are described by six attributes: beach nourishment, dyke heightening, planting of dykes, shoreline conditions/cliff protection, realignment of dykes and dunes, and individual payment (Table 2). Since the Baltic Sea coast entails large stretches of cliffs, in contrast to the North Sea, we adapted one attribute to match the coast. In the SC experiment relating to the Baltic Sea (T1 and T2), we have used the attribute "cliff protection", and in the SC experiment relating to the North Sea (T3 and T4), we have used the attribute "shoreline conditions". A related minor difference is that cliff protection comprises four levels, while shoreline conditions comprise three levels. An example of a choice task, in this case from a North Sea treatment, is given in Fig. 1. A detailed description of the attributes is given in Table A1 in the appendix. Importantly, in the remainder of this paper we treat all attribute levels as continuous in order to be able to estimate regret minimisation models (Hess et al., 2014).

Ngene software (ChoiceMetrics 2021) is used to generate the experimental designs, allocating the attribute levels across the two hypothetical alternatives. Since version 1.3, it has the capability to generate designs that are simultaneously efficient for RUM and RRM models (van Cranenburgh and Collins, 2019). Two different designs were created (for completeness we report the full designs Tables A3 and A4 in the appendix). Both designs were generated as Bayesian efficient designs (Rose et al., 2008; Rose and Bliemer, 2009) with weak uniform priors for the attribute parameters. For optimisation, the D-efficiency criterion for the MNL model was selected. To account for the uncertainty regarding the value of the priors, we used 1000 Sobol draws from the uniform distribution for each parameter. Dominated alternatives were excluded. We let the algorithm ran for a couple of hours until no meaningful improvements in minimising the D-error took place. The final designs comprised 48 choice tasks that were assigned to four blocks. Each respondent thus faced 12 choice tasks in a randomised order. The first design assumes that decision-makers use RUM as a decision rule; while the second design is optimised assuming that decision-makers use RUM and/or RRM as a decision rule. In other words, the created design is robust towards uncertainty relating to the decision rule (RUM or RRM). In the latter case, the researcher needs to specify the importance of each decision rule to the optimisation (van Cranenburgh and Collins, 2019). As we were not expecting that one of the decision rules would dominate, we weighted both the RUM and the RRM design by 0.5.

The choice tasks (Fig. 1 gives an example) involved two hypothetical alternatives (*Adaptation measure A and B*) and a status quo alternative. The hypothetical alternatives offered a change in at least one of the attribute levels compared to the zero-priced status quo alternative. Apart from the one attribute that was different for the two coasts, the survey and questionnaire design were identical. This comprises both the information respondents received about the potential consequences of climate change for the coasts and specific survey features such as an opt-out reminder included in each choice task (Ladenburg and Olsen, 2014). It informed respondents to choose the zero-price alternative if the costs of the hypothetical alternatives were higher than the amount of money they would be willing to pay.

2.3. Survey implementation

The online survey was implemented in June and July 2021 by a survey company. Since coastal protection in Germany is financed both by the coastal states and from the federal budget (Federal Ministry of Food and Agriculture, 2021), all people in Germany of age 18 and older were defined as the target population. To mitigate protest responses, respondents who lived in one of the coastal states were assigned to a treatment involving the sea adjacent to their state of residence (Baltic Sea: Mecklenburg-Western Pomerania, Schleswig Holstein; North Sea: Lower Saxony, Hamburg, Bremen, Schleswig-Holstein; residents of Schleswig-Holstein randomly assigned to one of the two coasts). We assumed that individuals in a coastal state would rather pay for coastal protection in their state, but not in another coastal state. However, all respondents living in a state adjacent to the sea were randomly assigned to a RUM design (T1 or T3) or to a mixture design (T2 or T4). All remaining respondents were randomly assigned to one of the four treatments (T1 to T4). Given that both the Baltic Sea coast and the North Sea coast are popular holiday destinations across all of Germany, no prior designation was made. The survey was designed to take about 20 min, respondents were recruited from the online-access-panel Gapfish, and both response time and responses to specific follow-up questions were monitored to control for quality.

[1] Residents of a coastal state were however always assigned to "their" respective coast. For further details see Section 2.3 below.

Table 2
Attributes and levels: Coastal adaptation.

Beach nourishment	20 m, *40 m*, 60 m
Dyke heightening	*50 cm*, 100 cm, 150 cm
Planting of dykes	*5 km*, 30 km, 60 km
Cliff protection[1]	*15 km*, 30 km, 45 km
Shoreline conditions[2]	*0 km*, 70 km, 140 km
Realignment of dykes and dunes	*1 spot*, 3 spots, 6 spots, 9 spots
Individual payment (cost)	*0 €*, 8 €, 20 €, 35 €, 70 €, 110 €, 190 €

Note: Attribute levels of the status quo alternative are displayed in italics. The prior values along with the Ngene syntax used are reported in the Appendix (Table A2).
1) The attribute *cliff protection* is only included in the Baltic Sea treatments (T1 and T2).
2) The attribute *shoreline conditions* is only included in the North Sea treatments (T3 and T4).

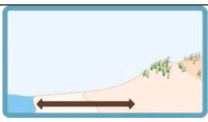
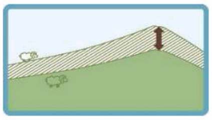
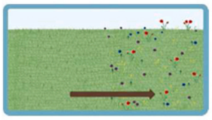
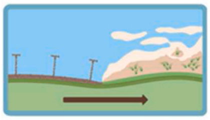
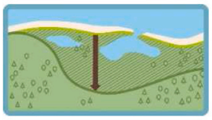
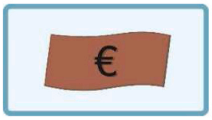
		Adaptation measure A	Adaptation measure B	Status quo
Beach nourishment (60 km total)		20 m width	60 m width	40 m width
Dyke heightening (950 km total)		50 cm height	100 cm height	50 cm height
Planting of dykes		60 km	30 km	5 km
Shoreline conditions		70 km	140 km	0 km
Realignment of dykes and dunes		1 spot	3 spots	1 spot
Individual payment (cost)		20 €	35 €	0 €
I choose		☐	☐	☐

Fig. 1. Example choice task North Sea (T3 and T4).

2.4. Sample characteristics

The number of respondents across the four treatments (Table 3) is not balanced as participants who live in a coastal state were assigned to treatment with the sea adjacent to their state. As more states are adjunct to the North Sea, and Lower Saxony, as one of them, has the largest population among them, the samples of the treatments T3 and T4 are larger. The effect of this type of allocation on socio-demographics is small. They differ slightly concerning some of the characteristics such as gender or share of people with higher education. Mean age, income, and people per household, however, are very similar across treatments. The share of people who have visited the coast at the Baltic Sea or the North Sea suggests, as expected, that more participants of the Baltic Sea treatments have visited the Baltic Sea and vice versa. Overall, given the distribution of socio-demographics, we do not expect that differences across samples are driving potential differences in choice patterns across treatments.

3. Methodology

Section 3.1 describes the models we use to investigate the impact of the decision rule assumption in efficient experimental designs

Table 3
Descriptive statistics by split treatment.

	Baltic Sea		North Sea	
	T1: RUM	T2: RUM-RRM	T3: RUM	T4: RUM-RRM
Characteristics	mean (sd)	mean (sd)	mean (sd)	mean (sd)
Number or respondents	830	800	959	990
Number of choice observations	9 960	9 600	11 508	11 880
Resident in a coastal state (1 = yes)	0.38	0.39	0.44	0.45
Age (years)	46.3 (15.7)	47.7 (15.5)	47.5 (15.8)	47.4 (15.9)
Female (1 = yes)	0.55	0.51	0.49	0.48
Higher education[a] (1 = yes)	0.28	0.32	0.27	0.29
Household income[b] (€ per month)	2 709 (1 627)	2 686 (1 568)	2 671 (1 521)	2 722 (1 617)
People per household (number)	2.4 (1.3)	2.3 (1.2)	2.3 (1.2)	2.4 (1.2)
Visit at North Sea (1 = yes)	0.55	0.57	0.68	0.65
Visit at Baltic Sea (1 = yes)	0.64	0.69	0.62	0.58

Note: 1) Higher education is defined as general qualification for university entrance; 2) Due to missing values the mean income was calculated each time for a subset of respondents.

on choice behaviour. Section 3.2 describes the approach to investigate whether particular choice tasks invoke a certain decision rule, and thus are comparatively better explained by RUM or RRM. Thus, whereas the models presented in section 3.1 apply to data sets as a whole, the approach presented in section 3.2 focusses on individual choice tasks.

3.1. Discrete choice models

To investigate the impact of the decision rule assumption in efficient experimental designs on choice behaviour, we estimate RUM and RRM models on each of the four data sets (i.e., T1 to T4) and compare their results. Specifically, we compare the results in terms of model fit (log-likelihood) and preferences (marginal utilities and regrets). Regarding the models, we estimate two models: a linear-additive RUM model and a μRRM model with attribute specific μ parameters (Van Cranenburgh et al., 2015). The μRRM model is a generalisation of the RRM model proposed by Chorus (2010). The μ parameter governs the degree of regret minimisation. It determines the extent to which losses loom larger than equivalently sized gains. When μ is estimated to be small (i.e., close to zero), the model postulates considerable regret, while when μ is estimated to be large (typically >5), the model postulates losses loom equally large as gains. In that case, the model is observationally equivalent to a linear-additive RUM model. Below, we briefly revisit the linear-additive RUM and μRRM models. In all models we estimate in this paper, we assume the unobserved part of utility/regret is i.i.d. extreme value type I distributed. Hence, all models are estimated in multinomial logit form.

3.1.1. Random Utility Maximisation model

Equation (1) presents the canonical linear-additive RUM model (McFadden 1974). U_{in} denotes the total utility a decision-maker derives from alternative i. The total utility is composed of an observed part, V_{in}, and an unobserved part, ε_{in}. The observed part of utility is a function of attributes the analyst observes. In line with standard practice, we assume utility is linear with the attributes, denoted m, and additive in utility space. Under this specification, β_m denotes the utility of a unit change in attribute m. The unobserved part of utility, ε_{in}, accounts for everything that is not observed by the analyst but still matters to the decision of the decision-maker n.

$$U_{in}^{RUM} = V_{in} + \varepsilon_{in} \text{ , } where \ \varepsilon_{in} \sim i.i.d.\ EV(0,1), V_{in} = \sum_m \beta_m x_{imn} \tag{1}$$

3.1.2. Random Regret Minimisation models

Equation (2) presents the μRRM model. RRM models postulate that when choosing, decision-makers want to minimise regret, which is experienced when one or more non-chosen alternatives outperform the chosen one in terms of one or more attributes m. The total regret is conceived to be the sum of all binary regrets associated with the binary comparisons between a considered alternative i and its competitor alternatives j for all attributes m. Through μ, the μRRM can capture different degrees of regret minimisation behaviour. The parameter μ can be estimated for each attribute individually, for all attributes generically, or for a combination of the two. In practice, estimating μ's for all attributes individually is challenging. Therefore, in this paper, we estimate μRRM models with attribute specific μ's for two attributes of particular interest, namely 'Realignment of dykes and dunes' and 'Individual payment', and a generic μ for the remaining attributes.

$$RR_{in}^{\mu RRM} = R_{in}^{\mu RRM} + \varepsilon_{in}, where - \varepsilon_{in} \sim i.i.d.EV(0,1), R_{in}^{\mu RMM} = \sum_{j \neq i} \sum_m \mu_m \ ln\left(1 + exp\left(\frac{\beta_m}{\mu_m}\left[x_{jmn} - x_{imn}\right]\right)\right) \tag{2}$$

3.2. Sampling-based approach

To investigate whether some choice tasks particularly invoke RUM or RRM decision rules, we propose a sampling-based approach.

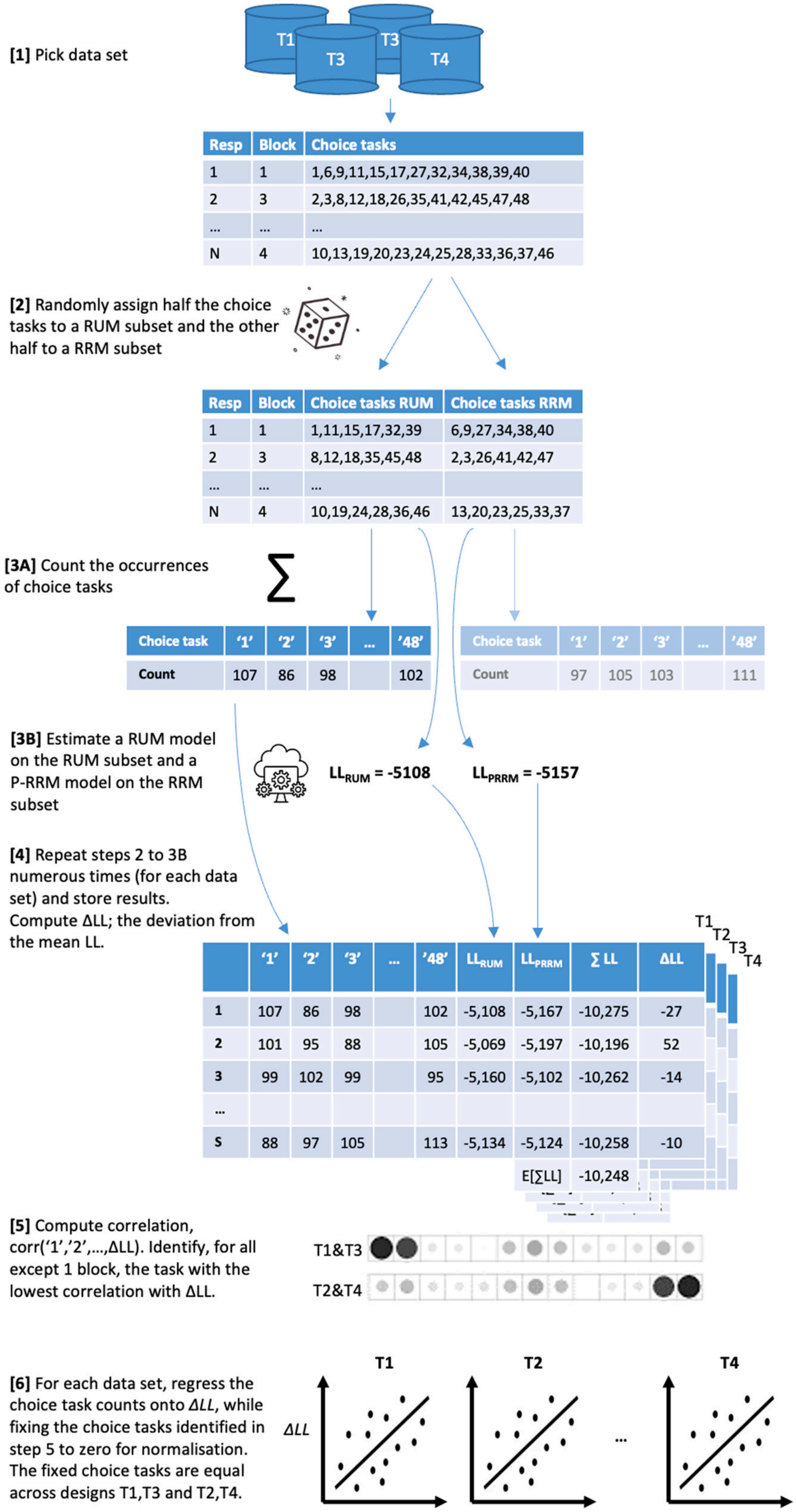


Fig. 2. Procedure to investigate decision rule heterogeneity.

The main reason to develop this new approach is because previous approaches to investigate decision rule heterogeneity, which are primarily based on latent class modelling, suffer from the confounding of taste and decision rule heterogeneity (Hess et al., 2012; Boeri and Longo, 2017; Charoniti et al., 2020).

Below, we describe the six steps of this approach; Fig. 2 visualises the steps.

1. Select a choice data set (i.e., one treatment).
2. For each respondent in the data set, randomly assign half of the choice tasks for estimation using RUM and the other half of the choice tasks for estimation using RRM. Hence, the two newly created subsets contain the same number of observations and the same respondents.
3. [A] For each (sub) data set, count how often each of the (48) choice tasks is present in the data set.[B] Estimate a RUM model on the RUM-subset and a P-RRM model on the RRM-subset and sum their log-likelihoods (LL). The P-RRM model is a special case of the μRRM model. Specifically, the P-RRM model is a μRRM model in which all μ parameters approach zero (from above). As such, the P-RRM model postulates stronger regret minimisation behaviour than the μRRM model (with nonzero μ parameters). See van Cranenburgh et al. (2015) for more details about this model. For the sampling method, we use the P-RRM model instead of the μRRM model to magnify the difference between the decision rules (RUM vs RRM).
4. Repeat steps 2 to 3 numerous times. In our application, we repeat steps 2 and 3 25k times.
5. Compute the correlation coefficients to identify the choice task that correlates least with the model fit for each block in our experimental design. Specifically, we compute Pearson correlation coefficients between the counts and the deviation of the log-likelihood from the mean log-likelihood (ΔLL).
6. Estimate a regression model in which the response variables are the counts of the choice tasks (for each manifestation of the data), and the independent variable is the deviation of the log-likelihood from the mean log-likelihood, ΔLL. The choice tasks identified under step 5 are fixed to zero to identify the model. The fixed tasks are equal across the treatments T1 and T2 and treatments T3 and T4. In case a particular choice task invokes RUM (or RRM) behaviour, we will find a significant regression coefficient.

Table 4
Estimation results.

		Baltic Sea				North Sea				Parameter differences	
	Design treatment	T1 (RUM-only)		T2 (RUM-RRM)		T3 (RUM-only)		T4 (RUM-RRM)		T1 – T2	T3 – T4
	Number of individuals	830		800		959		990			
		Coef.	|t-value|	Coef.	|t-value|	Coef.	|t-value|	Coef.	|t-value|	|t-value|	|t-value|
RUM-MNL	ASC Adaptation A	−0.051	1.93	0.003	0.13	−0.036	1.48	−0.090	3.83		
	ASC status quo	−0.319	5.28	0.026	0.42	−0.303	5.30	−0.331	5.00	3.99	
	Beach nourishment	0.320	4.04	0.496	6.06	0.325	4.50	0.192	2.64		
	Dyke heightening	0.118	2.37	0.269	4.84	0.187	3.99	0.216	4.41	2.02	
	Planting of dykes	0.298	4.48	0.137	1.73	0.247	3.86	0.185	2.57		
	Cliff protection[1]	−0.226	1.46	−0.081	0.60						
	Shoreline conditions[2]					−0.041	1.25	−0.190	4.64		2.84
	Realignment of dykes and dunes	3.723	7.60	3.184	5.53	2.717	6.02	3.729	7.37		
	Individual payment	−0.098	25.01	−0.071	17.78	−0.093	25.47	−0.073	20.49	4.83	3.52
	Final Log-likelihood	−10.183		−10.145		−11.932		−12.435			
	Number of parameters/Rho square	8/0.069		8/0.038		8/0.056		8/0.047			
μRRM-MNL	ASC Adaptation A	−0.052	1.95	−0.001	0.05	−0.037	1.52	−0.092	3.91		
	ASC status quo	−0.315	5.12	0.094	1.32	−0.310	5.21	−0.297	4.29	4.34	
	Beach nourishment	0.225	4.22	0.373	6.74	0.225	4.61	0.131	2.65		
	Dyke heightening	0.061	1.84	0.214	5.63	0.115	3.56	0.154	4.53	3.03	
	Planting of dykes	0.204	3.86	0.129	1.88	0.156	3.19	0.122	2.26		
	Cliff protection[1]	−0.147	1.53	0.134	1.45					2.11	
	Shoreline conditions[2]					−0.027	1.34	−0.119	3.96		2.54
	Realignment of dykes and dunes	0.032	6.58	0.032	6.40	0.021	3.83	0.028	5.98		
	Individual payment	−0.066	24.86	−0.051	18.21	−0.063	25.15	−0.049	19.35	3.89	3.93
	μ_dyke_heightening	0.121	1.63	0.039	19.33	0.118	1.31	0.042	24.77		
	μ_individual_payment	10.000	–fixed	1.117	0.17	10.000	–fixed	10.000	–fixed		
	μ[3]	0.064	23.60	0.007	84.20	0.078	11.84	0.112	7.77		
	Final Log-likelihood	−10.181		−10.137		11.932		12.434			
	Number of parameters/Rho square	10/0.070		11/0.039		10/0.056		10/0.047			

Note: 1) only at Baltic Sea coast; 2) only at North Sea coast; 3) *t*-test against 1.

The reason why this approach avoids confounding between preference and decision rule heterogeneity is that it does not try to capture heterogeneity. Instead, it is based on the notion that when a choice set invokes a particular decision rule, say RUM, then it should result in a noticeable (and statistically significant) improvement in model fit of the RUM model (compared to the expected log-likelihood, i.e., the situations where choice tasks are randomly put in the RUM basket).

4. Results

4.1. Model fit and preferences for coastal adaptation

Starting with model fit, a comparison across the four treatments reveals the following (Table 4). Firstly, the results do not provide evidence for (or against) the conjecture that the decision rule used to create the efficient design invokes that decision rule. Across all designs we see a modest model fit improvement of the μRRM over the RUM model. However, we do not see larger improvement in the RUM-RRM mixture design than in the RUM-only design. Looking at the estimation results for the μRRM model, we see that the μs for dyke heightening are always found to be small (<1) (indicating behaviour consistent with regret minimisation behaviour), while the μ's for individual payment are found to be large (>10), except for T2 (indicating behaviour consistent with utility maximisation behaviour). Overall, the comparison of the model fits across the four treatments does not suggest a clear association between invoked decision rules and the underlying type of experimental design.

Next, we turn to the parameter estimates for the coastal adaptation attributes (Table 4). We first compare across treatments for each type of model (RUM versus μRRM) before comparing across models and all four designs. Starting with the estimates from the RUM model, individuals prefer wider beaches as well as higher dykes compared to the situation today. Furthermore, they prefer to change the planting of dykes to increase the food supply for insects and are in favour of a higher number of locations with realignments of dykes and dunes in the future. In contrast, respondents are less likely to select an alternative when costs increase. Moving to the attributes that differ across the two seas, cliff protection at the Baltic Sea is not significant in both treatments T1 & T2, while for shoreline conditions at the North Sea we find that the sign is negative but the parameter is only significant in T4. The parameter estimates of the cost attributes are clearly statistically significant. The ASC for the status quo alternative has a negative sign and is statistically significant in all treatments except T2. In both models for treatment T2 this ASC has a positive sign and is not significant. We observe the same pattern of signs for the ASC for the alternative *Adaptation A*, but in this case the parameter is significant at the 5%-level only in T4 and at 10% in T1, revealing a mixed impact.

A similar picture emerges for the parameter estimates of the μRRM models across all four treatments. Again, the signs of the parameter estimates are persistent for the attributes beach nourishment, dyke heightening, planting of dykes, realignment of dykes and dunes and cost. All estimates are also highly statistically significant across treatments. This differs again when we look at the two attributes cliff protection (Baltic Sea) and shoreline conditions (North Sea). The former is not statistically significant, the later only in T4 (mixed design). Regret would decrease if shoreline conditions were changed at a longer stretch along the coast. The same pattern of signs and significances is also observed for the ASC parameter.

In addition, Table 4 shows *t*-test statistics for the differences in estimates across the designs for the same sea. In case the experimental design has no impact on a recovered parameter, we expect an absolute value of the t-statistic smaller than 1.96, while in case the experimental design has impact on a recovered parameter, we expect an absolute value of the t-statistic larger than 1.96. Recall that respondents were randomly assigned to one of the design treatments. Table 4 shows that for the Baltic Sea the estimates of three parameters are statistically different from one another. For the North Sea, the estimates of two parameters are statistically different from each other. This observation goes for both the RUM and μRRM model estimates. Also, it consistently concerns the same attributes, namely: 'ASC status quo', 'Dyke heightening', 'Shoreline conditions' and 'Individual payment'. This observation is unsettling and raises concerns about the overall reliability of results obtained from stated choice experiments. It highlights that even a minor alteration in the experimental design can lead to significant changes in the behavioural findings.

For adaptation to climate change in coastal areas, the results across all four treatments clearly indicate that both protection against the sea and promotion of nature are preferred as components of the adaptation process. All models show that increased beach nourishment and higher dykes are preferred, in each case compared to the current situation. Both are measures at different sections of the German coast to protect the hinterland against rising sea levels and extreme events such as storms. Not only higher protection levels, however, but also nature conservation is preferred when implementing adaptation measures. Changing the way the surfaces of dykes are planted on longer stretches along the coast was a highly significant measure in all but one model. While this represents only a minor intervention in coastal protection compared to the other measures, it increases the food supply for insects on a significant scale given that dykes stretch for several hundred kilometres along both coasts. A stronger change of the coastline would come along with an increase of the number of spots where dykes and dunes are moved to the hinterland. One effect of this adaptation measure is that habitats for typical coastal animals and plants are provided. At the same time, the preferred higher number of spots with realignments could at specific locations increase safety by giving more space to the sea and reducing construction and maintenance costs. In contrast, the protection of cliffs at the Baltic Sea coast was not preferred following the MNL models. The same partially applied to shoreline conditions at the North Sea coast, where the softening of shoreline conditions was valued negatively. While the latter would also create habitats for animals and plants in the transition zone between land and sea, respondents might perceive this measure as a potential weakening of the level of protection against the sea.

4.2. Impact of decision rule assumptions in experimental designs at choice task level

To visualise the regression results, stem plots are used (Fig. 3). Specifically, on the x-axis, the estimates for the choice task counts are sorted based on their magnitudes, from smallest to largest. Estimates that are found to be significant are flagged with a red fill. A negative estimate indicates that modelling this choice task as RUM deteriorates the overall model fit (relative to the mean when choice tasks are randomly split). Likewise, a positive estimate indicates that modelling this choice task as RUM improves the overall model fit. Importantly, the numbering of choice tasks is consistent across data sets T1 and T3 and across data sets T2 and T4. In other words, choice task 30 in data set T1 is the same choice task as choice task 30 in data set T3, but not the same as choice task 30 in data set T2 (or T4). Finally, to ease interpretation, we have rescaled the *y*-axis, by normalising the counts. As a result of this, the magnitudes of the estimates reflect the expected change in log-likelihood when the choice task is estimated as RUM (as compared to when the choice task would be randomly allocated). This rescaling of the explanatory variables does not impact model fit or the statistical significance of estimates.

Fig. 3 provides solid statistical evidence that choice tasks can provoke decision rules. Specifically, in each data set, we see that some estimates are statistically significant. Moreover, *F*-statistics and their associated *p*-values (both reported in the annotations in the north-west corners of the plots) indicate that all regression models are significant. This means they explain ΔLL statistically better than a model without explanatory variables (i.e., a constant only model). But we also see that many choice tasks do not statistically significantly impact ΔLL. In addition, the (very) low coefficients of determination (i.e., R^2) tell us that the extent to which choice tasks provoke decision rules is modest. Or, at least, their impact on the model fit is modest. After all, a low R^2 indicates the model is hardly able to explain the variability in ΔLL.

Choice tasks invoking a particular decision rule do not seem to have particular characteristics. We inspected these choice tasks and searched for common features, i.e., compromise attribute levels for certain attributes as well as attribute level allocation across the identified choice tasks. However, no pattern has been found. Future research, based on larger and more diverse data sets, may shed light on the salient features of choice tasks that invoke the decision rules. In conclusion, we find solid evidence that choice tasks can provoke a particular choice behaviour that is statistically better explained by a RUM than RRM decision rule (or vice versa), its impacts on model fit seem modest, however. We have not been able to associate features of choice tasks that invoke particular decision rules.

To validate this finding, we pose two conjectures:

1. The impact of choice tasks in invoking RUM (or RRM) decision rules should be stable across data sets. In other words, if a certain choice task invokes RUM (or RRM) in data set T1, it should also do so in data set T3. The same goes for data sets T2 and T4.
2. When all choice tasks invoking RUM are assigned to one subset A and all choice tasks invoking RRM are assigned to another subset B and we estimate a RUM model on subset A and a P-RRM model on subset B, then we expect $LL_{subsetA} + LL_{subsetB}$ to be (much) higher than if all choice tasks are randomly assigned to either subsets A or B.

To shed light on Conjecture 1, Fig. 4 shows two scatter plots. The scatter plot on the left shows the results for the RUM experimental design; the scatter plot on the right shows the results for the RUM-RRM experimental design. Each data point is a choice task. Its *x*-coordinate is the magnitude of the regression coefficient obtained from the Baltic Sea data sets; its y-coordinate is the magnitude of the regression coefficient obtained from the North Sea data sets. A data point is filled in red when the regression coefficient is statistically significant in both the Baltic Sea and North Sea data sets. In case the choice task has a similar effect across the two data sets, we expect

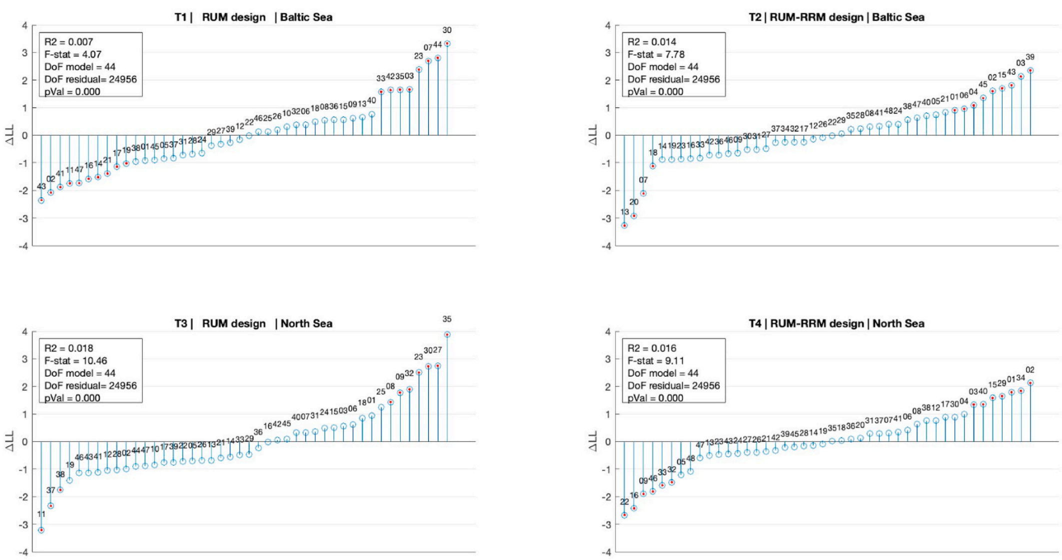


Fig. 3. Differences in LL-values of RUM and PRRM model by treatment.

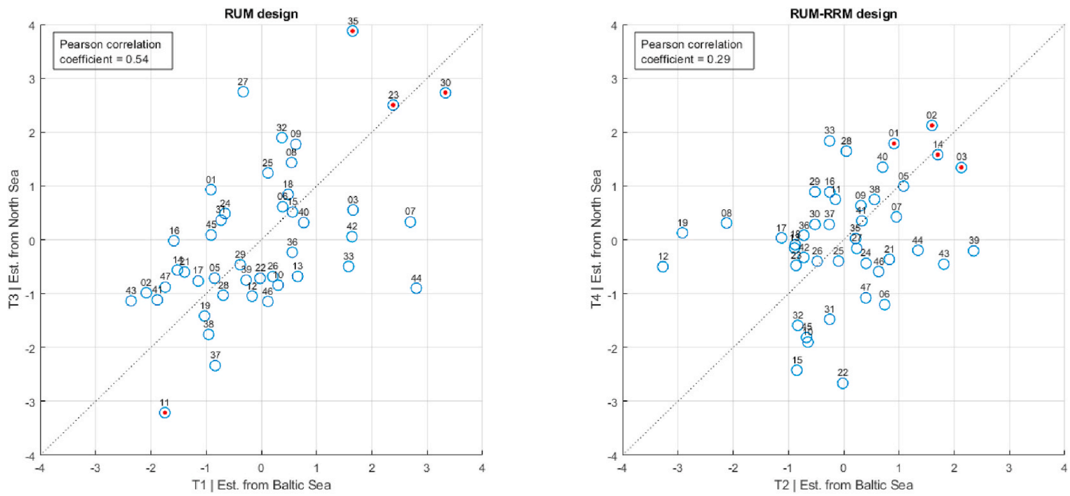


Fig. 4. Comparison of Baltic and North Sea treatments by design type.

the data points to be on the diagonal.

Based on Fig. 4, we make two observations. First, in line with expectations, the RUM design (left-hand side plot) displays a pattern in which the data points are scattered around the diagonal. The Pearson correlation coefficient for the RUM design equals 0.54 ($p = 0.000$). Based on this, Conjecture 1 can be accepted. Moreover, we see that four choice tasks are significant in both data sets. Choice tasks 23, 30 and 35 are found to be positive. This means that modelling these choice tasks as RUM positively impacts on the model fit. Choice task 11 is found to be negative. This means modelling this choice task as RUM negatively impacts the model fit. In other words, choice 11 is best modelled using RRM. Second, the RUM-RRM design (right-hand side plot) displays a much weaker pattern than the RUM design. The Pearson correlation coefficient is considerable: 0.29 ($p = 0.052$). This tentatively suggests that RUM-RRM designs produce choice tasks that are less pronounced in terms of the decision rule they invoke. Nonetheless, also for the RUM-RRM design we find four statistically significant choice tasks, namely 01, 02, 03 and 14. All these choice tasks are positive, meaning that modelling these as RUM positively impacts the model fit.

To test Conjecture 2, for T1 and T3 we place choice tasks 23, 30 and 35 in subset A and choice task 11 in subset B. The remaining choice tasks are randomly allocated to either subset A or subset B, while ensuring that for each respondent exactly six choice tasks are in subset A and six choice tasks are in subset B. We take the same approach for T2 and T4. That is, we place choice tasks 01, 02, 03 and 14 in subset A. The remaining choice tasks are randomly allocated to either subset A or B, while ensuring six choice tasks of each

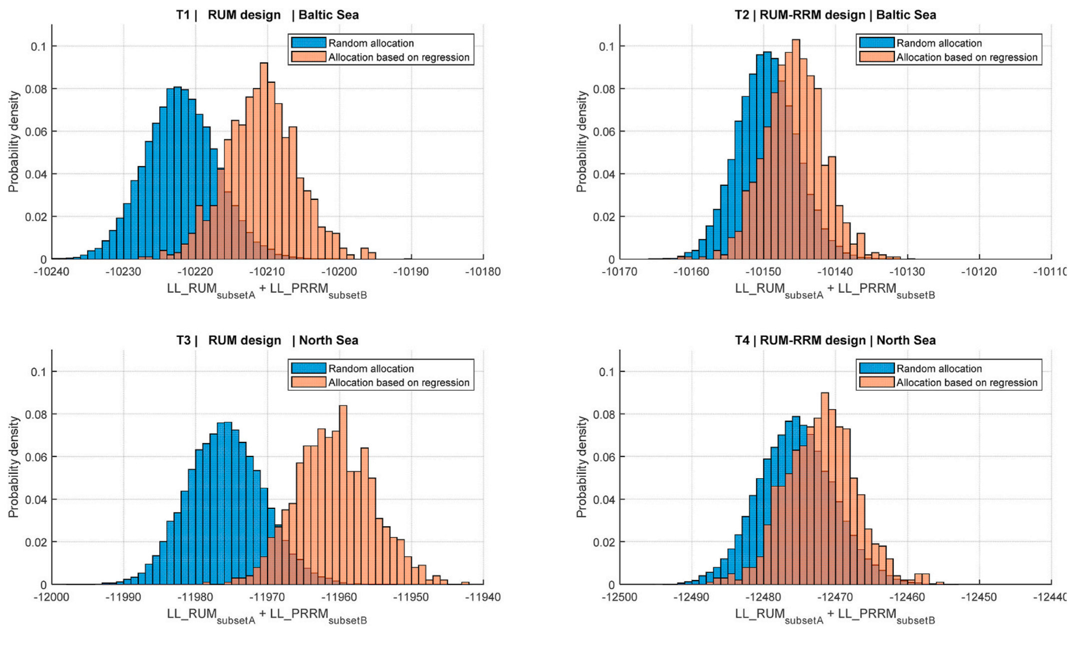


Fig. 5. Distribution of the sum of the log-likelihoods for random and non-randomly allocated choice tasks.

individual are in both subsets. After creating the eight subsets, we estimate a RUM model on all subsets A and a P-RRM model on all subsets B. Finally, we take the sum of the log-likelihoods of both estimations, for each data set. To avoid presenting results of a particular manifestation of the data (related to the random allocation of choice tasks to the subsets), we show the results across 1 000 repetitions of these steps.

Fig. 5 shows two histograms per data set. The blue-coloured histograms show the distribution of the sum of the log-likelihoods when, for each respondent, randomly six choice tasks are allocated to subset A and the remaining six choice tasks are allocated to subset B. The orange-coloured histograms show the distribution of the sum of the log-likelihoods when the choice tasks are allocated to the subsets based on the sign and significance level from the regression results (Fig. 3). Based on Fig. 5 Conjecture 2 can be accepted. In each plot, we see that the orange-coloured histograms are shifted to the right, relative to the blue histograms. Two-sample t-tests confirm that means *LL* for the randomly allocated choice tasks are statistically different from the means *LL* for the non-randomly allocated choice tasks ($p = 0.000$ for all data sets). This means that, in line with our conjecture, the model fit improves by allocating the choice tasks to decision rules, relative to a random allocation of choice tasks to decision rules (blue-coloured histograms). A final noteworthy observation is that difference in log-likelihood is considerably larger in the RUM designs (on avg. 13 LL points) than in the RUM-RRM design (on avg. 3.5 LL points). More research is needed to determine whether this difference can be attributed to the type of experimental design.

5. Conclusion and discussion

This study scrutinised the assumption that the decision rule used to create an efficient experimental design does not impact the recovered preferences and/or decision rules. We have conducted four independent SC experiments: two using RUM-optimised designs and the other two optimised for a mixture of RUM and RRM. Our SC experiments were concerned with climate change adaptation measures in coastal areas. Coastal areas are especially vulnerable to climate change due to rising sea levels and an increasing frequency of severe events such as storms. This paper has made a methodological and a substantive contribution. Methodologically, we have shed light on the extent to which choice behaviour is exogenous to the decision used to create the experimental design and individual choice tasks. Substantively, we have contributed to a better understanding of people's preferences over climate change adaptation measures in coastal areas.

Starting with our substantive conclusions, we find that people in Germany prefer to increase the height of dykes and broaden the width of beaches to increase protection against sea level rise and an expected higher frequency of severe events. People also place value to nature protection. Our results show support for changing the currently dominating planting of dykes, which is not supportive to insects, and to increase the number of locations where dykes and dunes are realigned. This would create new natural bays and wetlands that provide habitats for typical coastal animal and plant species. For other aspects such as protecting cliffs at the Baltic Sea coast and changing shoreline conditions at the North Sea coast, the mean preferences do not show a clear trend. A limitation, due to the focus of this study, is that we have confined the analysis to MNL models not capturing observed or unobserved taste heterogeneity. For example, a previous study by Meyerhoff et al. (2021), conducted only for the Baltic Sea coast, showed that cliff protection was on average also not statistically significant but a latent class model accounting for taste heterogeneity showed that a subgroup of respondents strongly supported cliff protection. Future analyses, not focusing primarily on the impact of the underlying experimental design, should therefore investigate the extent of taste heterogeneity for all adaptation measures along both the Baltic and the North Sea coasts.

Regarding a relationship between the type of attributes and decision rules, we find evidence that the dyke heightening attribute triggered choice behaviour that is best explained by the regret minimisation model, while the individual payment attribute triggered behaviour that is best explained by the utility maximisation model. The former finding is in line with our expectations. Dykes are a well-known measure in public to secure the hinterland against the sea, and heightening them is to many an obvious measure to prepare for future rising sea-levels and more frequent extreme events. We therefore expected respondents to experience regret from having a lower dyke (than competitor alternatives), which would not easily be compensated by a better performance on another attribute (e.g., payment).

Regarding our methodological conclusions, we find no clear evidence that an efficient design optimised for RUM invokes RUM decision rules, or vice versa. However, we find strong statistical evidence that some choice tasks invoke RUM consistent behaviour while others invoke RRM consistent behaviour. Furthermore, our results suggest that RUM-RRM designs produce choice tasks that are less pronounced in terms of the decision rule they invoke. Follow-up research should be conducted to more unequivocally draw this conclusion. The finding that choice tasks can invoke certain choice behaviour is, in itself, not surprising. After all, it is widely known that the choice set composition can be used to lure consumers into buying particular products (Guevara and Fukushi, 2016). Nonetheless, the observation that experimental designs produce such choice tasks implies that respondents' choice behaviour, and thus choice modelling outcomes, are not exogenous. This suggests that, for instance, the findings of numerous studies that have used stated choice data to investigate decision rule heterogeneity are likely partially the result of experimental design choices by the researcher, rather than being inherent to the decision-making behaviour of the respondents. For issues of common interest and with potentially severe consequences, such as adaptation to climate change, it will then be especially important to account for potentially invoked choice behaviour and non-exogenous choice modelling outcomes in future research. In addition, our investigation reveals that even a minor adjustment in the experimental design, specifically regarding the decision rule for which it is optimised, leads to statistically significant outcomes within the same model. This discovery encourages a critical examination of the reliability and robustness of behavioural insights derived from stated choice experiments.

Furthermore, we believe that the proposed sampling-based method provides a useful new tool to investigate the impact of individual choice tasks on choice behaviour. In this study we employed it to investigate which choice tasks invoke RUM behaviour and which choice tasks invoke RRM behaviour. Future research could explore using this method for other purposes, such as identifying choice tasks that lead to high or low willingness to pay estimates or violations of RUM properties.

Finally, future studies should also account for a couple of limitations to our study. We based the comparative analysis solely on RUM-only and RUM-RRM mixture designs. Future research on the relation between choice tasks and experimental designs should additionally include RRM-only optimised designs. Also, an orthogonal design can be added to a comparative study. Moreover, the point of departure of our research was that a decision maker either uses a RUM or a RRM data generation process (at the choice task level). Although we found strong empirical support for both RUM and RRM models, there is a possibility that the results are driven by a third, yet not considered decision rule, which might even be invariant of the choice task composition.

Funding

The data were collected in the context of the research project GoCoase, funded by the Federal Ministry of Education and Research (BMBF), Germany, with grant number 1LA1812 A/C, which is gratefully acknowledged.

CRediT authorship contribution statement

Sander van Cranenburgh: Conceptualization, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Jürgen Meyerhoff:** Conceptualization, Formal analysis, Funding acquisition, Investigation, Methodology, Validation, Writing – original draft, Writing – review & editing. **Katrin Rehdanz:** Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Validation, Writing – original draft, Writing – review & editing. **Andrea Wunsch:** Conceptualization, Investigation, Methodology, Project administration, Validation, Writing – original draft, Writing – review & editing.

Declaration of competing interest

None.

Appendix A. Supplementary data

Supplementary data to this article can be found online at odoi:mmcdoin

References

Biondi, B., Van der Lans, I.A., Mazzocchi, M., Fischer, A.R.H., Van Trijp, H.C.M., Camanzi, L., 2019. Modelling consumer choice through the random regret minimization model: an application in the food domain. Food Qual. Prefer. 73, 97–109. https://doi.org/10.1016/j.foodqual.2018.12.008.

Boeri, M., Longo, A., 2017. The importance of regret minimization in the choice for renewable energy programmes: evidence from a discrete choice experiment. Energy Econ. 63, 253–260. https://doi.org/10.1016/j.eneco.2017.03.005.

Buckell, J., Vasavada, V., Wordsworth, S., Regier, D.A., Quaife, M., 2021. Utility maximization versus regret minimization in health choice behavior: evidence from four datasets. Health Econ. 31 (2), 363–381. https://doi.org/10.1002/hec.4455.

Charoniti, E., Kim, J., Rasouli, S., Timmermans, H.J.P., 2020. Intrapersonal heterogeneity in car-sharing decision-making processes by activity-travel contexts: a context-dependent latent class random utility – random regret model. Int. J. Sustain.Transport. 15, 501–511. https://doi.org/10.1080/15568318.2020.1768608.

ChoiceMetrics, 2021. Ngene 1.3 User Manual & Reference Guide. Sydney, Australia.

Chorus, C.G., van Cranenburgh, S., Dekker, T., 2014. Random regret minimization for consumer choice modeling: assessment of empirical evidence. J. Bus. Res. 67 (11), 2428–2436. https://doi.org/10.1016/j.jbusres.2014.02.010.

Chorus, C.G., 2010. A new model of random regret minimization. Eur. J. Transport Infrastruct. Res. 10 (2), 181–196. https://doi.org/10.18757/ejtir.2010.10.2.2881.

Federal Ministry of Food and Agriculture, 2021. Rahmenplan der Gemeinschaftsaufgabe "Verbesserung der Agrarstruktur und des Küstenschutzes" 2021-2024. Bonn (in German).

Guevara, C.A., Fukushi, M., 2016. Modelling the decoy effect with context-RUM Models: diagrammatic analysis and empirical evidence from route choice SP and mode choice RP case studies. Transp. Res. Part B Methodol. 93, 318–337. https://doi.org/10.1016/j.trb.2016.07.012. Part A.

Hess, S., Beck, M.J., Chorus, C.G., 2014. Contrasts between utility maximisation and regret minimisation in the presence of opt out alternatives. Transport. Res. Pol. Pract. 66, 1–12. https://doi.org/10.1016/j.tra.2014.04.004.

Hess, S., Stathopoulos, A., Daly, A., 2012. Allowing for heterogeneous decision rules in discrete choice models: an approach and four case studies. Transportation 39 (3), 565–591. https://doi.org/10.1007/s11116-011-9365-6.

Johnston, R.J., Makriyannis, C., Whelchel, A.W., 2018. Using ecosystem service values to evaluate tradeoffs in coastal hazard adaptation. Coast. Manag. 46 (4), 259–277. https://doi.org/10.1080/08920753.2018.1474067.

Ladenburg, J., Olsen, S.B., 2014. Augmenting short cheap talk scripts with a repeated opt-out reminder in choice experiment surveys. Resour. Energy Econ. 37, 39–63. https://doi.org/10.1016/j.reseneeco.2014.05.002.

Long, Y., Mao, B., Xu, L., Zheng, D., Ao, C., 2022. Public choices and welfare estimate under wetland improvement context: utility maximization, regret minimization or both? Wetlands 43 (3), 1–14. https://doi.org/10.1007/s13157-022-01645-4.

Mao, B., Ao, C., Wang, J., Sun, B., Xu, L., 2020. Does regret matter in public choices for air quality improvement policies? A comparison of regret-based and utility-based discrete choice modelling. J. Clean. Prod. 254 https://doi.org/10.1016/j.jclepro.2020.120052.

Masiero, L., Yang, Y., Qiu, R.T.R., 2019. Understanding hotel location preference of customers: comparing random utility and random regret decision rules. Tourism Manag. 73, 83–93. https://doi.org/10.1016/j.tourman.2018.12.002.

McFadden, D., 1974. Conditional logit analysis of qualitative choice behaviour. In: Zarembka, P. (Ed.), Frontiers in Econometrics, vols. 105–142. Academic Press, New York.

Meyerhoff, J., Rehdanz, K., Wunsch, A., 2021. Preferences for coastal adaptation to climate change: evidence from a choice experiment. J. Environ.Econ. Pol. 10 (4), 374–390. https://doi.org/10.1080/21606544.2021.1894990.

Nielsen, M.R., Jacobsen, J.B., 2020. Effect of decision rules in choice experiments on hunting and bushmeat trade. Conserv. Biol. 34 (6), 1393–1403. https://10.1111/cobi.13628.

Paul, P., Berlin, C., Maessen, M., Valtonen, H., 2018. A comparison of regret-based and utility-based discrete choice modelling – an empirical illustration with hospital bed choice. Appl. Econ. 50, 4295–4305. https://doi.org/10.1080/00036846.2018.1444260.

Rose, J.M., Bliemer, M.C.J., Hensher, D.A., Collins, A.T., 2008. Designing efficient stated choice experiments in the presence of reference alternatives. Transp. Res. Part B Methodol. 42 (4), 395–406. https://doi.org/10.1016/j.trb.2007.09.002.

Rose, J.M., Bliemer, M.C., 2009. Constructing efficient stated choice experimental designs. Transport Rev. 29 (5), 587–617. https://doi.org/10.1080/01441640902827623.

van Cranenburgh, S., Collins, A.T., 2019. New software tools for creating stated choice experimental designs efficient for regret minimisation and utility maximisation decision rules. J. Choice Model. 31, 104–123. https://doi.org/10.1016/j.jocm.2019.04.002.

van Cranenburgh, S., Rose, J.M., Chorus, C.G., 2018. On the robustness of efficient experimental designs towards the underlying decision rule. Transport. Res. Pol. Pract. 109, 50–64. https://doi.org/10.1016/j.tra.2018.01.001.

van Cranenburgh, S., Guevara, C.A., Chorus, C.G., 2015. New insights on random regret minimization models. Transport. Res. Pol. Pract. 74, 91–109. https://doi.org/10.1016/j.tra.2015.01.008.

Zhu, D., Sze, N.N., Feng, Z., 2021. The trade-off between safety and time in the red light running behaviors of pedestrians: a random regret minimization approach. Accid. Anal. Prev. 158, 106214 https://doi.org/10.1016/j.aap.2021.106214.