Economic Analysis and Policy 73 (2022) 725–736

Contents lists available at ScienceDirect

Economic Analysis and Policy

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

Analyses of topical policy issues

A test–retest analysis of stated preferences in uncertain times

Andrea Wunsch [a,b,*], Jürgen Meyerhoff [a,c], Katrin Rehdanz [a]

[a] *Kiel University, Institute for Environmental, Resource and Spatial Economics, Olshausenstraße 40-60, 24098 Kiel, Germany*
[b] *Kiel Institute for the World Economy, Kiellinie 66, 24105 Kiel, Germany*
[c] *Technische Universität Berlin, Institute for Landscape Architecture and Environmental Planning, EB 4-2, Straße des 17. Juni 145, 10623 Berlin, Germany*

ARTICLE INFO

Article history:
Received 21 October 2021
Received in revised form 17 December 2021
Accepted 28 December 2021
Available online 4 January 2022

JEL classification:
Q51
Q54
Q57

Keywords:
Adaptation
Choice experiment
Climate change
Coastal protection
Covid-19 pandemic
Test re-test

ABSTRACT

In environmental valuation, the issue of the temporal stability of stated preferences to changes in environmental (dis)amenities is important because their results can be employed to inform decision-making. This includes cost–benefit analysis for large infrastructure projects such as coastal protection. A couple of studies have investigated stability of stated preferences over varying time periods. However, less evidence is available for temporal stability of stated preferences for (dis)amenities in uncertain times, i.e., times that are characterized by larger degrees of uncertainty regarding the (near) future. Using a choice experiment on coastal adaptation to climate change, this paper examines the test–retest reliability of individual preferences and resulting welfare estimates over the course of the Covid-19 pandemic. We do so by surveying the same respondents at two points in time five months apart during the ongoing pandemic. Using a latent class model, we find similar preference heterogeneity patterns but different class sizes. While the welfare measure of an adaptation scenario that focuses on safety increases across survey waves, scenarios that centre on recreation or nature have decreasing welfare effects. This suggests that individuals set other priorities in uncertain times.

1. Introduction

In environmental valuation, the temporal stability of preferences for changes in environmental (dis)amenities is vital as results are used for decision making, i.e., directly incorporated in cost–benefit analyses, or employed for benefit transfer. When the decisions affect large infrastructure projects such as motorways, railway networks or coastal protection, they can have far-reaching consequences. Therefore, the preferences recorded should reflect what people prefer at the time when the preferences are measured. They should also be stable over time, because once projects are implemented they will remain for long periods of time. The issue of temporal stability is particularly relevant for stated preferences recorded on hypothetical markets, which may depend on the context in which they were captured.

First studies on the stability of preferences over time focused on the contingent valuation (CV) approach (see, e.g., the review by McConnell et al. (1998)). For the use of the other stated preference approaches, in particular choice experiments (CE), empirical evidence is much scarcer.[1] However, we also find here test–retest studies that analyse temporal stability

* Corresponding author at: Kiel Institute for the World Economy, Kiellinie 66, 24105 Kiel, Germany.
E-mail address: andrea.wunsch@ifw-kiel.de (A. Wunsch).

[1] Studies related to health economics are not included in our review. For further information, see e.g. Ryan et al. (2006), Brouwer (2006) or Gamper et al. (2018).

https://doi.org/10.1016/j.eap.2021.12.021

either using subsamples of the original samples (Liebe et al., 2012; Matthews et al., 2017) or different samples from the same population (e.g., Bliem et al., 2012; Lee and Wallmo, 2017). The distinct periods investigated vary between six months (Rigby et al., 2016) and 17 months (Lew and Wallmo, 2017), and cover a variety of environmental goods and services. To our knowledge, none of these CE studies has investigated temporal preference stability in uncertain times, characterized by insecure employment opportunities and income prospects. Against the backdrop of the current Covid-19 pandemic, associated with these insecurities, the focus of our paper is to do so.

Reviewing this literature, Bliem et al. (2012) and Liebe et al. (2012) were the first to conduct a test–retest CE study. The former found remarkable temporal stability for two surveys within a time-interval of one year regarding river restoration programs in Austria. In contrast to them, Liebe et al. (2012) invited respondents of the first survey to the retest survey. A time interval of eleven months between the test and the retest is used on the valuation of landscape externalities from onshore wind power in Germany. They find a test–retest reliability of about 60% and insignificant differences between all WTP estimates except for one of the choice attributes. A couple of further studies used this framework of intrapersonal choices with varying test–retest periods. Based on a test–retest period of one year for land use change scenarios in Flanders, Schaafsma et al. (2014) confirm temporal stability of choices in 57% of the choice occasions but significantly different WTP values. Using a six month delay, Czajkowski et al. (2016) report significantly different WTP parameters for public forest management of the Bialowieza Forest in Poland. Similarly, Matthews et al. (2017) found significant differences in WTP values for beach erosion management of beach visitors to the Coromandel peninsula, New Zealand, using two retests over a six-month period. The existing literature finds mixed evidence regarding temporal stability of preferences, yet indicates a high degree of stability overall, during more secure times.

In contrast to the above test–retest studies, we investigate stability in uncertain times. We are not the first to do this. There are a few studies, that were conducted in times of great economic uncertainty.[2] Lew and Wallmo (2017) and Loureiro and Loomis (2017) both are related to the global financial crisis. Based on a framework of different samples from the same population and a test–retest period of 17 months during the financial crisis (spring 2009 and autumn 2010), Lew and Wallmo (2017) find temporal stability in preferences and WTP values associated with threatened and endangered marine species protection in the United States. Loureiro and Loomis (2017) compare temporal stability of WTP estimates of a CV-study linked to the environmental damages caused by the Prestige oil spill before (2006) and after (2009) the recession. They note a significant decline in median WTP estimates for Spanish society as a whole, attributing this to changing priorities during the recession.

To our knowledge, the only other study conducting a test–retest study related to the Covid-19 pandemic is Hynes et al. (2021). They compare WTP values associated with alternative environmental management plans on the high seas before (late 2019) and during the pandemic (May 2020). Using different samples from Canada, Scotland, and Norway their results suggest that both preferences and WTP values remain relatively stable. Our CE test–retest study uses a sample of respondents from Germany who participated in two survey waves during the Covid-19 pandemic. During the first survey wave in April 2020 Covid-19 case numbers rose and began to fall in May/June of the same year. When the second survey wave was conducted five months later, in early September, incidences were low but the arrival of a second wave was anticipated. Instead of comparing choices on the individual level, as done for example by Liebe et al. (2012), we compare them on the level of samples, as was done, for example, by Bliem et al. (2012). Focusing on preference stability in uncertain times, we expect mean WTP estimates to generally decline in the second wave, as respondents are likely to feel a higher degree of insecurity about their own future during this period.

2. Method

In a CE, respondents are requested to choose their most preferred alternative from a set of mutual exclusive alternatives. Each alternative is described by a number of attributes, whose associated levels are systematically varied due to an underlying experimental design (Mariel et al., 2021). Our CE focuses on coastal adaptation options to climate change on the German Baltic Sea coast of Mecklenburg-Western Pomerania (MWP). As climate continues to change and impacts increase, increasing investment is needed to keep up with the changes. This also applies to coastal protection. In Germany, unlike in other countries, coastal protection is a national task, delegated to the coastal states and limited to the protection of settlements. The expenses are mainly borne by the federal government and to a much lesser extent by the coastal state (StALU, 2009; BMEL, 2020). Therefore, preferences of people who contribute to coastal protection but live far away from it or do not use it need to be considered as well.

In the following, we briefly describe the CE. For a comprehensive overview of the study area, the survey design, and implementation see Meyerhoff et al. (2021) focusing on the first wave. Our CE included the following choice situations: two adaptation alternatives *(Adaptation A and B)*, an alternative offering a continuation of present measures *(Present measures)*, and an alternative which requires no additional payments compared to today *(Present budget)*. The latter alternative *Present budget* reflects that individuals may not be willing to pay extra for future adaptation. This implies that it would not be possible to maintain the current level of measures as costs will increase in the future. The third

[2] More generally on the effect of significant changes in information, Morrison and Hatfield-Dodds (2011) used longitudinal surveys to examine the effects of increased media coverage of climate change on public willingness to support greenhouse policies in Australia. The results suggest that support for policies decreases at low and medium costs, while it increases slightly at higher costs.

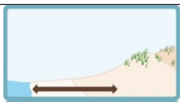
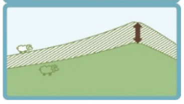
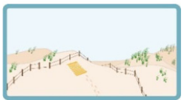
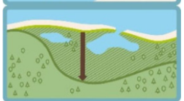
		Adaptation A	Adaptation B	Present measures	Present budget
Beach nourishment (44 km total)		20 m width	60 m width	40 m width	20 m width
Dyke heightening (227 km total)		50 cm height	25 cm height	50 cm height	25 cm height
Access to dunes		Yes. Over a distance of 28 km	Yes. Over a distance of 15 km	No	No
Protection of cliffs		45 km distance	15 km distance	15 km distance	15 km distance
Realignment of dykes and dunes		1 spot (15 km^2)	3 spots (45 km^2)	1 spot (15 km^2)	1 spot (15 km^2)
My payment		20 €	35 €	15 €	No payment
I choose		☐	☐	☐	☐

Fig. 1. Example choice task (Wave 1). Note: The attribute *realignment of dykes and dunes* always has three levels, but the number of spots differs between the waves. In Wave 1 1, 2 and 3 spots were chosen and in Wave 2 1, 4 and 6 spots.

alternative, *Present measures*, reflects that individuals may prefer to maintain current measures and are willing to pay more for them than people in Germany pay on average for coastal protection today. The corresponding payment attribute could only take values from the lower half of the cost vector to reflect that the cost of maintaining the present measures would be lower than the cost of the changes described by the two adaptation alternatives.

The final set of attributes compromised adaptation measures which directly related to coastal protection (*beach nourishment* and *dyke heightening*), present recreational opportunities (*access to dunes* and *protection of cliffs*) and the measure *realignment of dykes and dunes*. A coastal protection levy was used as the cost attribute. The set of attributes including a detailed description, associated pictograms presented to respondents and attribute levels is presented in Table A.1 in the Appendix. A Bayesian efficient design with uniform priors was used as experimental design to assign attribute levels to the different alternatives. D-efficiency as design criterion was optimized for an MNL model to allow for uncertainty in the prior values, 1000 Sobol draws were taken for each parameter prior. As detailed in Table A.1, certain cost attribute levels are only assigned to certain alternatives. In the experimental design, the levels 5/ 10/ 15 Euro are assigned only to the alternative "*Present measures*" via the experimental design. This is indicated in Table A.1 by the addition "(PM)" after these three levels. The other cost levels are assigned to the hypothetical alternatives A or B by the design, indicated by the addition "(A_B)". Overall, 48 choice tasks were created and allocated to four blocks. Respondents were randomly assigned to a block, and within a block the order of presentation was randomized showing respondents different sequences of choice tasks. An example choice task of the test survey is presented in Fig. 1.

The finalized version of the survey instrument was implemented for the first nationwide wave ($N = 1878$) in spring 2020 after we administered a pilot study to about 50 individuals. Five months later, the retest survey was implemented with a smaller sample size due to budget constraints ($N = 880$). In addition, half of the sample of the second wave consists of persons who had already participated in the first wave and half of newly invited respondents. A test–retest study can thus only be conducted with a maximum of 440 respondents from the second wave. For both waves, the same survey instrument was used, except that the levels of the attribute *realignment of dykes and dunes* were changed to find out whether the WTP for this attribute was sensitive to the number of sites. Unfortunately, the survey company was not able to provide the same block of choice sets to respondents in the first wave. This has two implications: First, we had to apply a matching procedure to identify the respondents who had participated in both waves with a high level of certainty.[3] Secondly, we can only compare mean estimates across waves as single choices are not comparable due to the different choice sets for each respondent.

[3] We used the two Stata commands *nnmatch* (nearest neighbour) and *psmatch* (propensity score) for matching. Only those respondents that were identified by both approaches were selected for our analysis.

The matching approach used the information provided by the survey company about a respondent's participation in Wave 1 and Wave 2. In addition, we used socio-demographics (gender, age, and education), the residence (Federal State and postcode), and information about whether the German Sea coast had been visited by the respondent before the pandemic for matching. This resulted in 266 matches out of the 440 possible matches. This matching procedure is rather restrictive, as postcode information is not provided for all respondents and those who moved to another state or had a birthday between the first and second wave were excluded.

3. Econometric modelling

Our analysis of choice data is based on the random utility model by McFadden (1974). Its assumption that complete information regarding an individual *n*'s preferences is not available to the researcher is captured by modelling an individual *n*'s utility as the sum of a systematic and a random part:

$$U_{ni} = V_{ni}\left(x_{ni}\beta\right) + \varepsilon_{ni} \tag{1}$$

U_{ni} is the true but unobservable utility which an individual derives from choosing an alternative i out of a set of several available alternatives j. The deterministic part V_{ni} is a function of the attributes (x_{ni}) and a vector (β) of unknown coefficients representing the desirability of the attributes. ϵ_{ni} denotes the unknown random part.

Different assumptions regarding the distribution of the error term lead to different models. A conditional logit model is obtained when assuming that the error terms are distributed independently and identically (IID) and follow a type 1 extreme value distribution. In this case the choice probability of individual n choosing alternative i is:

$$P_{ni} = \frac{\exp(\mu V_{ni})}{\sum_{j\epsilon C}\exp(V_{nj})} \tag{2}$$

To be able to capture unobserved heterogeneity different models are needed. We therefore apply a latent class (LC) model (Hess, 2014). It assumes that an unknown number of classes exist that each possess their own taste coefficient β. So, instead of estimating one single β for the whole population, as done in the conditional logit model, a taste coefficient is calculated separately for each class. Each individual in the population is assumed to belong to each of the classes with a certain probability. The number of classes is a priori unknown and is selected by comparing information criteria (such as the Bayesian Information Criteria (BIC)) while increasing the number of classes sequentially. Moreover, we use likelihood ratio tests to determine whether non-linear specifications of the attributes outperform linear specifications (Mariel et al., 2021).

Including the monetary attribute "*My payment*" allows to calculate the marginal willingness to pay (*mWTP*) for each of the non-monetary attributes by calculating the ratio: $mWTP = \beta_{s\_attribute}/\beta_{s\_money}$. Based on the LC model results, this is done for each class separately. Using those marginal estimates, we also provide an overall marginal WTP estimate weighted by class size. This weighted estimate is employed for the compensating variation for three coastal adaptation scenarios that we have additionally calculated. Each scenario has a certain focus: safety, recreation, and nature. Following Hanemann (1984) we calculate the welfare measure by applying the following formula:

$$CS_n = -\frac{1}{\beta_{\cos tn}}\left[ln\sum_{n} expV_n^1 - ln\sum_{n} expV_n^0\right] \tag{3}$$

where CS is the compensating surplus welfare measure, β_{cost} refers to the marginal utility of income while V_n^0 and V_n^1 are the n*th* individuals' indirect utility functions before and after the change respectively. To calculate the welfare measures we used only the cost parameter that reflects the cost sensitivity associated with the two hypothetical adaptation alternatives.

4. Results

4.1. Sample characteristics

The characteristics of the full sample of respondents in Wave 1 and those in the matched sample are reported in Table 1. To test for differences in socio-demographic characteristics and familiarity of the German coast between the full sample and our smaller sample of 266 individuals, we used a logistic regression (Table A.2). We find no significant differences between participants in Wave 2 and our matched sample with respect to higher education, household income, household size, state of residence and previous visit to at least one of the German coasts. Significant differences were only found for the variables age and gender. Overall, this points to the fact that the smaller sample of 266 matched respondents appropriately represents the full sample.

Table 1
Descriptive statistics.

Characteristics	Wave1 (N = 1878)[a]		Matched respondents (N = 266)[b]	
	Mean	Std. dev	Mean	Std. dev
Age (years)	48.86	16.46	52.36	15.57
Female (%)	50.85		55.64	
Higher education[c] (%)	36.94		30.07	
Household income[d] (€ per month)	2533.78	1454.10	2489.64	1241.43
People per household (number)	2.23	1.28	2.15	1.13
Resident in a coastal state (%)	17.25		19.55	
Recreational visits to coast in Germany				
only to the North Sea (% yes)	15.10		12.78	
only to the Baltic Sea (% yes)	15.02		13.16	
to the North and Baltic Sea (% yes)	51.60		58.65	

[a]Note that of the 1878 individuals, 440 participated in Wave 2. For further details see section 2.
[b]As expected, there are no significant differences between the characteristics of the two waves, such that only values from Wave 1 are reported here.
[c]Higher education here is defined as education level that is required to study at a college/university.
[d]Due to missing disclosures and some implausible responses to this question the statistics for household income are based on a smaller sample, within the range of € 450 to € 10,000.

Table 2
Observed and expected choices across survey waves.

	Wave 1		Wave 2	
	Frequencies	Percent	Frequencies	Percent
Alternative A	673/ *647.5*	21.1%	622/ *647.5*	19.5%
Alternative B	676/ *651.0*	21.2%	626/ *651.0*	19.6%
Present measures	1023/ *1078.0*	32.1%	1133/ *1078.0*	35.5%
Present budget	820/ *815.5*	25.7%	811/ *815.5*	25.4%
Total	3192	100.0%	3192	100.0%

Note: N = 266 and 3192 choice observations; expected frequencies are in italics.

4.2. Descriptive choice analysis

Proceeding with a descriptive analysis of the recorded choices of the matched sample (Table 2), we find the alternative *Present measures* accounted for most of the choices in both waves (32.1% and 35.5%, respectively). While the share of choices for this alternative increased from the test to the retest survey, the share for the two hypothetical alternatives decreased. In uncertain times, like the Covid-19 crisis, more respondents chose to maintain the current level of protection and express this by choosing *Present measures* at the expense of the two hypothetical and more costly alternatives. The share of choices for the alternative *Present budget* instead remained quite stable across the two survey waves. Thus, we cannot observe a shift of choices towards the no-cost alternative *Present budget*. A Pearson chi-squared test allows to reject the null hypothesis that the observed frequencies in Wave 1 and Wave 2 are from the same distribution (Pearson chi2(3) = 9.591, pr = 0.022).

4.3. Modelling preferences for coastal adaptation

Table 3 presents the results of the LC models. For both waves, the comparison of information criteria statistics suggested that a 3-class segmentation best represents the underlying unobserved taste heterogeneity. The 3-class segmentation is in line with results of a previous study that analysed the entire first wave of the survey, including respondents who were not surveyed a second time (Meyerhoff et al., 2021). The labels assigned to the three classes are still meaningful for the present analysis as well, which also indicates a great stability of the previous results.

In class 1 (named *dykes* group), for both samples the parameter for *dyke heightening* is positive and significant. Respondents likely to be in this class prefer higher dykes compared to the reference alternative *Present measures.* The parameter estimate of ASC_PB, representing the alternative *Present budget*, is also significant in both surveys and has a negative sign. Respondents prefer the reference alternative *Present measures* against the alternative *Present budget*. In the test survey the attribute parameter for a decrease of beach width from 40 to 20 metres is also significantly negative. So, while a decrease in beach width systematically influences respondents' choices in the test survey, this is not the case in the retest survey. The parameter estimates for all other non-monetary attributes are not significant at least at the 5% level of statistical significance in both survey waves and thus do not influence choices of respondents in this class.

Respondents who are likely to belong to the second class, named *improvement* group, do not only prefer higher dykes. Compared to the reference alternative *Present measure,* they prefer more extensive changes. For both waves, this applies to the attributes *access to dunes* and the *realignment of dykes and dunes* at further spots.[4] In the test survey (Wave 1),

4 As noted above, the attribute levels for realignment were changed from the test to the retest survey.

Table 3
Estimates from latent class (LC) logit models.

	LC-Wave 1 (N=266)						LC-Wave 2 (N=266)					
Label Sample size (in%)	Class1 *Dykes* 30.0		Class2 *Improvement* 45.6		Class3 *No payment* 24.4		Class1 *Dykes* 40.6		Class2 *Improvement* 32.1		Class3 *No payment* 27.3	
	Coeff.	| z-val.|	Coeff.	| z-val.|	Coeff.	| z-val.|	Coeff.	| z-val.|	Coeff.	| z-val.|	Coeff.	| z-val.|
Beach nourishment												
40 => 20 metres	-0.978^{**}	2.42	-0.260^{**}	1.99	0.558	0.55	−0.034	0.12	0.005	0.03	−1.097	1.52
40 => 60 metres	-0.735^{*}	1.95	0.027	0.20	−0.565	0.64	−0.011	0.04	0.169	1.04	-1.081^{*}	1.78
Dyke heightening (per centimetre)	0.014^{**}	2.57	0.007^{***}	3.95	-0.035^{**}	2.34	0.012^{***}	3.53	0.007^{***}	3.26	−0.009	0.79
Access to dunes (yes, per kilometre)	−0.000	0.02	0.014^{***}	3.44	−0.040	1.26	-0.015^{*}	1.73	0.015^{***}	3.04	−0.004	0.17
Cliff protection (per kilometre)	0.001	0.05	0.012^{***}	3.58	−0.036	1.47	−0.005	0.79	0.002	0.45	0.010	0.51
Realignment of dikes and dunes (per spot)[a]												
1 spot => 2/4 spots	0.095	0.28	0.310^{***}	2.61	1.320^{*}	1.78	0.096	0.41	0.412^{***}	2.91	0.083	0.13
1 spot => 3/6 spots	0.277	0.84	0.536^{***}	5.30	−0.086	0.10	−0.082	0.37	0.494^{***}	4.14	−0.068	0.10
Payment_PM (in €)	-0.122^{***}	4.59	0.049^{***}	2.91	-0.183^{***}	3.97	−0.018	0.81	0.066^{***}	2.93	-0.209^{***}	6.25
Payment_A_B (in €)	-0.117^{***}	5.74	-0.011^{***}	9.66	-0.073^{***}	3.42	-0.045^{***}	8.01	-0.009^{***}	6.42	-0.042^{***}	3.54
ASC_A	0.190	0.40	1.212^{***}	5.32	−1.164	0.14	0.144	0.38	1.718^{***}	5.90	-1.427^{*}	1.70
ASC_B	0.567	1.17	1.169^{***}	5.15	−0.506	0.45	0.238	0.62	1.633^{***}	5.64	-1.407^{*}	1.65
ASC_PM			Reference						Reference			
ASC_PB	-1.856^{***}	3.76	-2.555^{***}	6.35	−0.026	0.21	-2.652^{***}	7.05	-2.274^{***}	4.71	1.147	1.30

Note: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

[a]The attribute *realignment of dykes and dunes* always has three levels, but the number of spots differs between the waves. In Wave 1 1, 2 and 3 spots were chosen and in Wave 2 1, 4 and 6 spots.

respondents also rate changes in the attribute *cliff protection* positively, whereas the protection of cliffs does not influence respondent's choices in the retest survey. Another difference between the two survey waves is that the attribute parameter for a decrease in beach width from 40 to 20 metres is significantly negative for the test survey while it is insignificant for the retest survey. For both survey waves highly positive significant parameters for ASC_A and ASC_B can be observed, and a highly negative significant parameter estimate for ASC_PB. This underlines the high appreciation of respondents in the second class for far-reaching changes of coastal adaptation to climate change.

Turning to respondents likely to be in the third class (named *no payment* group), the only parameter being significant at the 5% level is the *dyke heightening* in the test survey, but not in the retest survey. Respondents in the test survey exhibit a negative valuation for an extension of existing dykes in height. Choices of respondents are not significantly influenced by any other of the non-monetary attributes in both waves.

Investigating class sizes across the two survey waves, the *no payment* group exhibits the smallest class size with a quite stable share of respondents across the two survey waves (24% and 27%, respectively). This does not hold for the two other groups. While the *dyke* group shows a class size of 30% in the test survey, its size increases to 41% in the retest survey. This is reversed for the *improvement* group. Class size decreases from 46% in the test survey to 32% in the retest survey. With regard to class size, the group *improvement* and the group *dykes* switched places.

4.4. Willingness to pay for coastal adaptation

The marginal WTP estimates, which are reported by class and as a weighted average across classes (Table 4), mirror the underlying preference patterns of both waves. For comparison, we focus on the weighted means, each of which considers only the statistically significant marginal WTP estimates. In Fig. 2 we present the weighted mean WTP estimates and their 95% confidence intervals for different changes in the attributes compared across both waves. As the underlying units differ (cm, km etc.) we used non-marginal changes for some of the attributes to better visualize the results. In particular for dykes we estimated an extension of 20 cm, for access to dunes and protection of cliffs we used a length of 10 km for each. As Fig. 2 reveals, the mean values for most attributes differ across waves while they are very close for the attribute *access to dunes*. Among the most obvious differences, the WTP estimates for reduced *beach nourishments* and *cliff protection* are no longer significant in the second wave. Thus, we cannot compute marginal WTP estimates and assume that the WTP is zero in those cases. In contrast, the marginal WTP estimate for the attribute dyke extension has increased in Wave 2. This applies also for the marginal WTP estimate for the lower level of *the realignment of dykes and dunes* (i.e., four spots instead of one). For the upper level of this attribute, we instead observe a decline from Wave 1 to Wave 2. It is important to note here that the levels of this attribute have changed. In the second wave, *realignment of dykes and dunes* has the same number of levels, but the level values are higher (Table A.1). Thus, a direct comparison of the WTP estimates is not possible. Still, we can conclude that more spots are valued more although the increase from one level to the next is rather non-linear. The marginal value for additional sites is significantly lower in Wave 2 compared to Wave 1.

For both waves also non-marginal welfare measures (compensating surplus; CS) for various policy scenarios are calculated. The scenarios combine certain levels of the attributes, thereby emphasizing different aspects of coastal adaptation to climate change (Table 5) and differ from the current situation (*Present measures*) in at least the level of one attribute. In terms of a possible aggregation of the non-marginal welfare measures, it should be recalled that only a relatively small sample of the 266 matched respondents serves as basis here. Scenario 1 (*Safety*) focuses on better protection of the hinterland. Beach width is increased by 20 metres and dyke height by 25 to overall 75 cm. Scenario 2 (*Recreation*) stresses aspects related to the recreational use of the Baltic Sea in MWP. Beach width is increased to 60

Table 4
Marginal WTP estimates (in €/ year/ household) by class and as a weighted average.

	Wave 1 (March 2020)				Wave 2 (September 2020)			
Label Sample size (in%)	Class 1 Dykes 30.0	Class 2 Improvement 45.6	Class 3 No payment 24.4	Weighted average[a]	Class 1 Dykes 40.6	Class 2 Improvement 32.1	Class 3 No payment 27.3	Weighted average[a]
Beach nourishment								
40 metre => 20 metre	−8.34	−24.54	7.61	−13.79	−0.76	0.54	−26.21	n.s.
	(−15.27/ −1.41)	(−47.34/ −1.75)	(−20.53/ 35.76)	(−24.51/ −3.08)	(−12.92/ 11.41)	(−34.32/ 35.39)	(−57.86/ 5.43)	
40 metre => 60 metre	−6.27	2.56	−7.71	n.s.	−0.25	19.78	−25.84	n.s.
	(−12.95/ 0.40)	(−22.64/27.77)	(−31.75/ 16.33)		(−11.35/ 10.85)	(−20.66/ 60.22)	(−54.34/ 2.65)	
Dyke heightening	0.11	0.69	−0.48	0.24	0.27	0.84	−0.21	0.38
(per centimetre)	(0.03/ 0.20)	(0.32/ 1.06)	(−1.01/ 0.44)	(0.02/ 0.45)	(0.13/ 0.42)	(0.26/ 1.42)	(−0.77/ 0.35)	(0.19/ 0.57)
Access to dunes	−0.00	1.29	−0.55	0.59	−0.32	1.71	−0.10	0.55
(per kilometre)	(−0.21/ 0.21)	(0.52/ 2.07)	(−1.55/ 0.45)	(0.24/ 0.95)	(−0.72/ 0.08)	(0.45/ 2.96)	(−1.30/ 1.07)	(0.14/ 0.95)
Cliff protection	−0.00	1.11	−0.50	0.51	−0.11	0.21	0.23	n.s.
(per kilometre)	(−0.15/ 0.16)	(0.45/ 1.78)	(−1.26/ 0.27)	(0.21/ 0.82)	(−0.40/ 0.18)	(−0.70/ 1.10)	(−0.63/ 1.09)	
Realignment of dikes and dunes[b]								
1 => 2 spots (W1)	0.81	29.23	18.00	13.44	2.15	48.13	1.98	15.40
1 => 4 spots (W2)	(−4.78/ 6.40)	(7.56/ 50.89)	(−1.61/ 37.62)	(3.48/ 23.41)	(−7.97/ 12.26)	(13.96/ 82.30)	(−27.81/ 31.78)	(4.46/ 26.34)
1 => 3 spots (W1)	2.36	50.53	−1.17	23.24	−1.82	57.69	−1.61	18.46
1 => 6 spots (W2)	(−2.97/ 7.70)	(28.94/ 72.10)	(−23.35/ 21.01)	(13.31/ 33.17)	(−11.61/ 7.97)	(23.70/ 91.72)	(−32.18/ 28.94)	(7.57/ 29.35)

Note: Confidence intervals are reported in parenthesis.

[a] We use n.s. to indicate that the weighted averages are not significantly different from zero.

[b] The attribute realignment of dykes and dunes always has three levels, but the number of spots differs between the waves. In Wave 1 1, 2 and 3 spots were chosen and in Wave 2 1, 4 and 6 spots.

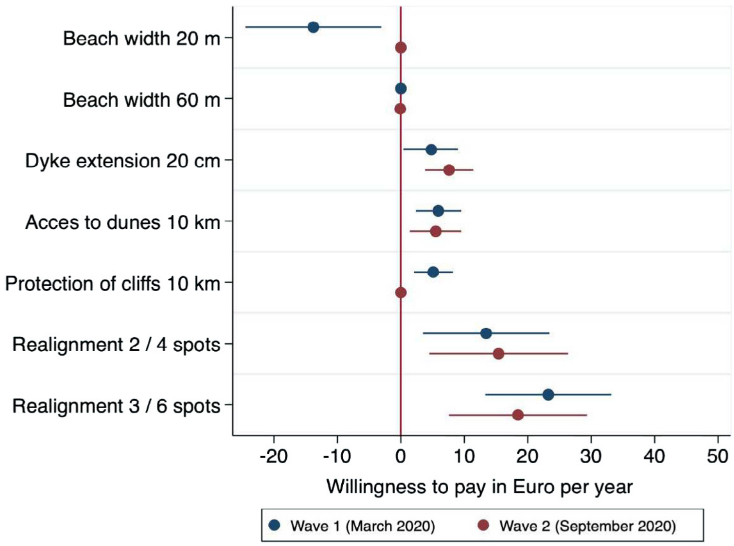


Fig. 2. Weighted mean WTP estimates by wave.

metres as in Scenario 1, access to dunes is allowed on stretches of overall 28 kilometres length, and cliffs are protected on another 30 kilometres of coastline. Finally, Scenario 3 (*Nature*) addresses nature protection. New inter-tidal areas that would serve the conservation of species typical for these habitats would be created through realigning additional spots with dykes and dunes. Recalling the marginal WTP estimates for both waves, one difference is related to the attribute *dyke heightening*. The increase in the CS value for the policy scenario *Safety* is due to the increase in the valuation for this attribute. The decrease in the CS value for the policy scenario *Recreation* can be explained by the insignificance of the attribute cliff protection in Wave 2. The CS for the policy scenario *Nature* decreases because the WTP for realignment sites has gone down.

The mean values of the CS measures clearly differ between waves and would result in different policy conclusions if the aggregated values were incorporated in cost–benefit analysis. Testing whether these measures or the marginal WTP estimates are significantly different is ultimately less informative, in our view, as decision makers generally cannot test for those differences. Locking at the aggregate values in Table 5 for each wave, this is very likely the case. In all three scenarios the values aggregated across the target population differ by more than € 150 million between the waves. In the case of the scenario *Recreation*, the aggregate measure even halved in Wave 2.

5. Discussion and conclusion

Knowledge about the stability of preferences over time is essential when they are used to inform decision-making. This is especially critical for preferences for public goods such as (dis)amenities that are stated on hypothetical markets. These

Table 5
Welfare measure for policy scenarios (in €/ year/household).

	As today	Scenario 1 *Safety*	Scenario 2 *Recreation*	Scenario 3 *Nature*
Beach nourishment (per metre)	40 m	**60 m**	**60 m**	40 m
Dyke heightening (per centimetre)	50 cm	**75 cm**	50 cm	50 cm
Access to dunes (for access)	no	no	**yes (28 km)**	no
Cliff protection (per kilometre)	15 km	15 km	**45 km**	15 km
Realignment of dikes and dunes (spots)	1 spot	1 spot	1 spot	**3 (6) spots**
Wave 1 CS in € per year per household[a]		5.90 (0.56/ 11.26)	32.01 (16.09/ 47.93)	23.24 (13.31/ 33.17)
CS aggregated in mio €[b]		244.9	1328.4	964.5
Wave 2 CS in € per year per household[a]		9.52 (4.70/ 14.34)	15.29 (4.04/ 26.54)	18.46 (7.57/ 29.35)
CS aggregated in mio €[b]		395.1	634.5	766.1

Note: Welfare measures were calculated without incorporating marginal WTP estimates for the ASCs. In bold are attribute levels that have changed compared to "As today". CS refers to compensation surplus.
[a]The 95% confidence intervals are shown in brackets.
[b]The number of households at the time of the survey was 41.5 million (https://www.destatis.de; retrieved September 15, 2020).

preferences might be more susceptible to the valuation context, which might change over time, especially in uncertain times like a pandemic. While test–retest studies of stated preferences exist, this study is one of only a small number of studies to assess preference stability in face of uncertain times such as the current pandemic.

We use a CE focusing on adaptation to climate change along the coast of MWP, Germany. While climate change is more and more advancing, adaptation is becoming increasingly crucial for coastal areas. Literature on people's preferences for coastal adaptation options is limited, so is literature that looks at preference stability. This paper adds to both, but the latter forms the focal point. Along with the Covid-19 pandemic developing significantly in Germany in March and April 2019, the first wave of our survey had to go online. A few months later, the possibility arose to test for preference stability. Applying a matching procedure, we identified 266 respondents who participated in both survey waves. Firstly, the descriptive choice analysis revealed that the share of choices for the alternative *Present measures* increased at the expense of less choices devoted to the two hypothetical, generally more costly alternatives *Alternative A* and *Alternative B*. Secondly, a latent class model with three classes had proven to perform best to model existing preference heterogeneity in both waves. One group of respondents, labelled *dykes*, is willing to pay for an increase in dyke height but is not ready to give up money for a positive change in any other attribute. Respondents in the class *improvement* prefer extensive changes also for attributes such as *access to dunes* and *realignment* in addition to the heightening of dykes. To the contrary, respondents in the class *no payment* are generally not willing to cover any additional expenses required for adapting the German Baltic Sea coast to climate change.

When comparing the results of the two waves in more detail, we find mixed evidence. For two attributes - *beach nourishment* and *cliff protection* - the weighted average WTP values became insignificant in the second wave. For the attribute *dyke extension*, the weighted average WTP value increased while for the attribute *access to dunes* the value rather slightly decreased. A direct comparison of the marginal WTP for the attribute *realignment* is aggravated because the number of potential realignment sites was increased for the second wave. However, we found a lower mean WTP estimate in the second wave for this attribute although the number of realignment sites had increased. This indicates that the mean marginal WTP per relocation site has become smaller. Turning to non-marginal welfare effects, we found a significantly higher effect for the adaptation scenario *Safety* which is due to the attribute dyke. That the value of the policy scenario *Recreation* decreased is instead driven by the attributes *beach nourishment* and *cliff protection* which became insignificant in Wave 2. The decrease in the value for the scenario *Nature* again is caused by the attribute *realignment*.

Despite the differences between the two waves, a similar preference heterogeneity pattern was revealed by the latent class models. Both times we identify a class of people who focused on dyke heightening, on overall improvements, and who are not willing to pay for adaptation measures. The important difference relates to the change in class sizes. This change indicates that preferences have changed and respondents did choose differently in the second wave. While we observe an increase in the size of the *dykes* labelled class from about 30% to 41%, the size of the *improvement* labelled class decreased from 46% to 32%. For the third class, gathering people who are not willing to pay for adaptation measures, class size increased slightly by 3%. Thus, although the underlying preference heterogeneity pattern remain stable, we observe changes in the size of the latent classes: Less respondents are allocated to the class with preferences for overall improvements at the expense of more respondents being assigned to the class only exhibiting preferences for the heightening of dykes. This points, on average, to preference changes towards the most basic and fundamental adaptation measure, dyke heightening. Doing so seems comprehensible, as dykes as hard-engineered constructions may be perceived as safest and thus most important option and stronghold against rising sea-level and flooding events by respondents. Additionally, focusing solely on a single adaptation measure incurs a smaller payment compared with the expenses associated with more extensive adaptation options respondents in the *improvement* labelled class are willing to cover. In this regard, these findings are in line with the results of the descriptive choice analysis that detected a shift of choices from the generally more expensive *Alternative A and B* towards the generally cheaper *Present measures* alternative.

The welfare measures aggregated across the target population result in clearly different values for both waves. For the scenario *Safety*, this estimate has increased by around € 150 million in the second wave. In contrast, for the scenarios *Recreation* and *Nature* there are huge decreases on the order of several million euros. Thus, a cost–benefit analysis using these estimates might reach different conclusions, and decision makers with only the results of the first or the second survey wave available might have concluded differently with respect to investments in adaptation measures. One conclusion from this finding is that in uncertain times it might be necessary to monitor the preferences of the public more closely over time when large-scale investments are involved.

As analysing preference stability in uncertain times, to the best of our knowledge, is novel, a comparison with results from existing literature is rather difficult. However, it can be summarized, that already in secure times, which most existing test–retest studies are embedded in, temporal preference stability is not unambiguous, though a tendency towards stability can be observed. We broadly characterized uncertain times as providing insecure employment opportunities and income perspectives to people. The only other test–retest study that considers the Covid-19 pandemic as setting is by Hynes et al. (2021). However, it is not comparable to our study because it differs regarding the timing of the survey and content wise. Hynes et al. (2021) contrast preferences elicited prior to the pandemic with preferences during the pandemic for environmental management plans on the high seas. Instead, our analysis provides insights on preference stability concerning coastal adaptation at two points in time during the pandemic, i.e. at the advent of the first and second Covid-19 wave in Germany.

The rather vague term of *uncertain times* as used by us calls for future work on a classification of uncertain times. It is conceivable that, for example, uncertainty during a financial crisis is perceived differently in particular in terms of severity than uncertainty during a pandemic. While the current pandemic is causing financial impairments, its level of uncertainty may nevertheless be perceived lower than, for instance, during a financial crisis. With respect to future research, it would also be interesting to direct investigations towards preference stability in light of an extreme event (e.g. storm flood) representing an external shock which is directly related to the environmental good at consideration as pointed out earlier by Bliem et al. (2012). Applied to coastal adaptation options, it would be interesting to check how preferences especially of local people differ subsequent to a larger storm flood.

A limitation of our study is that we could not test preference stability at the individual level and in particular not test choice stability. This is due to the fact that, because of restrictions of the survey company, it was not possible to assign respondents in the retest survey the choice cards they faced in the test survey. Consequently, we can only report results on the aggregated level of the sample, but not on an individual level.

Overall, our analysis provides insights on preference stability during uncertain times, i.e., preferences captured at the advent of the first and second Covid-19 wave in Germany. That more respondents focus on the heightening of dykes in uncertain times seems plausible. Along with the similar preference heterogeneity pattern found in both survey waves, with differing class sizes, however, this constitutes a valuable piece of information for policy makers in directing future adaptation efforts at coastal areas at the German Baltic Sea coast.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

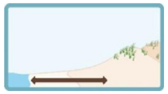
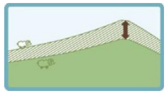
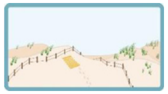
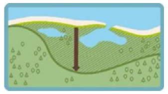
Acknowledgements

We would like to thank K. Sommermeier and L. Tiepolt from Staatliches Amt für Umwelt und Natur - Abteilung Küste - Ministerium für Landwirtschaft, Umwelt und Verbraucherschutz Mecklenburg -Vorpommern (StALU) for providing information about the coastal defence system in MWP and valuable feedback on the survey. Finally, we would like to thank the GoCoase project group for contributions, in particular J. Tiede, J. Visscher and T. Schlurmann, Ludwig-Franzius-Institute, Leibniz Universität Hannover (LUH); A. Dehnhardt (IÖW, Berlin) and N. Stybel (EUCC Deutschland, Rostock). This work stems from the joint research project GoCoase, funded by the Federal Ministry of Education and Research (BMBF), Germany, grant number 1LA1812A/C.

Appendix

See Tables A.1 and A.2.

Table A.1
Attributes and attribute levels.

Attribute	Pictogram	Description	Attribute levels
Beach nourishment (44 km total)		Wider beaches and higher dunes offer better protection for the land behind. Despite the use of groynes, sand is currently being washed up along the coast on 13 sections with a total length of 44 km. To maintain this condition, this process must be repeated every few years. In the future, beach nourishments may also be necessary on further sections. Usually the beach is about 40 m wide after a nourishment. However, less or more sand could also be used for nourishment.	Beach width of: 20 m **40 m** 60 m
Dyke heightening (227 km total)		Dykes protect the land behind them from flooding. In order to meet future requirements for coastal protection, dykes are currently being heightened by 50 cm during scheduled maintenance. However, this increase could also be lower or higher.	Height of: 25 cm **50 cm** 75 cm
Access to dunes		The stability of dunes is increased by planting, especially with beach grass. It also acts as a natural sand catcher. Walking on dunes can impair these functions and is therefore not allowed. In selected sections where dunes are of less importance for coastal protection and where this is compatible with nature conservation, access could be allowed in the future. There, dunes would then have to be maintained more comprehensively. This would involve additional costs.	Length of: 0 km 6 km 15 km 28 km
Protection of cliffs		The sea continuously erodes parts of the cliffs and shifts the coastline. Breakwaters and seawalls can slow down this process. They are currently used in 10 sections where settlements are endangered by the demolition of cliffs. As a result. 15 km of the cliffs are currently fortified. Further sections could be fortified, for example to preserve coastal cycle or hiking paths for a longer period.	Length of: **15 km** 30 km 45 km
Realignment of dykes and dunes[a]		At selected spots along the coast, the protection line (dykes and dunes) could be moved to the hinterland. There, natural bays and wetlands are created which provide habitats for typical coastal animal and plant species. Access to the areas would be restricted and existing cycle and hiking paths would have to be relocated. The safety of settlements would not be endangered at any time. So far, one major realignment has been implemented on the Zingst peninsula (about 15 km^2 or 1500 ha). Similar projects would be conceivable in other parts of the coast.	**1 spot (15 km^2)** 2 spots (30 km^2) 3 spots (45 km^2)
Payment		The existing coastal protection in Germany is financed jointly by the federal government and the coastal states from taxes. For further measures in Mecklenburg-Western Pomerania, a new coastal protection levy would have to be charged per household nationwide from 2021 on. It would have to be paid annually for the next 10 years. After this period, the measures would be reassessed, and a decision would be made on further financing.	**0 €**, 5/ 10/ 15 € (PM) 8/ 20/ 35/ 70/ 110/ 190 € (A_B)

Note: Attribute levels in bold indicate current levels.
[a]In Wave 2 the attribute levels for *realignment of dykes and dunes* were increased from 2 and 3 spots to 4 and 6 spots.

Table A.2
Logit model participation second wave.

	Coeff.	| z-val.|
Female (=1)	0.314**	2.23
Age	0.014	3.07
State (categorial)		
Baden-Wuerttemberg	Reference	
Bavaria	−0.138	−0.51
Berlin	−0.184	−0.46
Brandenburg	0.298	0.76
Bremen	0.931	1.48
Hamburg	0.068	0.14
Hesse	−0.122	−0.38
Mecklenburg Western	−0.318	−0.56
Pomerania	0.136	0.48
Lower Saxony	0.024	0.10
North Rhine-Westphalia	−0.658	−0.18
Rhineland Palatinate	−0.390	−0.60
Saarland	0.122	0.35
Saxony	−0.234	−0.49
Saxony-Anhalt	0.010	0.02
Schleswig Holstein	0.333	0.79
Thuringia		
Higher education (= 1)	−0.210	−1.26
Household size (number)	−0.022	−0.31
Income (categories)		
>900€	0.240	0.68
900€–1.300€	−0.197	−0.55
1.300€–1.500€	Reference	
1.500€–2.000€	0.065	0.19
2.000€–2.600€	0.145	0.43
2.600€–3.200€	0.329	0.97
3.200€–4.500€	0.243	0.73
4.500€–6.000€	−0.248	−0.54
> 9000€	0.516	1.05
not given	−0.020	−0.06
Visitor at North or Baltic Sea coast (=1)	0.148	0.75
Constant	−2.829***	−5.96
Number of observations	1,810	

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

References

Bliem, M., Getzner, M., Rodiga-Laznig, P., 2012. Temporal stability of individual preferences for river restoration in Austria using a choice experiment. J. Environ. Manage. 103, 65–73. http://dx.doi.org/10.1016/J.Jenvman.2012.02.029.

BMEL, 2020. Rahmenplan der Gemeinschaftsaufgabe Verbesserung der Agrarstruktur und des Küstenschutzes 2020–2023. https://www.bmel-statistik.de/fileadmin/daten/GAB-0002000-2020.pdf.

Brouwer, R., 2006. Do stated preference methods stand the test of time? A test of the stability of contingent values and models for health risks when facing an extreme event. Ecol. Econom. 60 (2), 399–406. http://dx.doi.org/10.1016/j.ecolecon.2006.04.001.

Czajkowski, N., Bartczak, A., Budziński, W., Giergiczny, M., Hanley, N., 2016. Preference and WTP stability for public forest management. Forest Policy Econom. 71, 11–22. http://dx.doi.org/10.1016/j.forpol.2016.06.027.

Gamper, E., Holzner, B., King, M.T., Norman, R., Viney, R., Nerich, V., Kemmler, G., 2018. Test-retest reliability of discrete choice experiment for valuations of QLU-C10D health states. Value Health 21 (8), 958–966. http://dx.doi.org/10.1016/j.jval.2017.11.012.

Hanemann, W.M., 1984. Welfare evaluations in contingent valuation experiment with discrete responses. Amer. J. Agricult. Econom. 66, 332–341. http://dx.doi.org/10.2307/1240800.

Hess, S., 2014. Latent class structures: taste heterogeneity and beyond. In: Hess, S., Daly, A. (Eds.), Handbook of Choice Modelling. Cheltenham, pp. 311–329. http://dx.doi.org/10.4337/9781781003152.00021.

Hynes, S., Armstrong, C.W., Xuan, B.B., Ankamah-Yeboah, I., Simpson, K., Tinch, R., Ressurreiçao, A., 2021. Have environmental preferences and willingness to pay remained stable before and during the global Covid-19 shock? Ecol. Econom. 189, http://dx.doi.org/10.1016/j.ecolecon.2021.107142.

Lew, D.K., Wallmo, K., 2017. Temporal stability of stated preferences for endangered species protection from choice experiments. Ecol. Econom. 131, 87–97. http://dx.doi.org/10.1016/j.ecolecon.2016.08.009.

Liebe, U., Meyerhoff, J., Hartje, V., 2012. Test–retest reliability of choice experiments in environmental valuation. Environ. Resour. Econom. 53, 389–407. http://dx.doi.org/10.1007/s10640-012-9567-1.

Loureiro, M., Loomis, J., 2017. How sensitive are environmental valuations to economic downturns? Ecol. Econom. 140, 235–240. http://dx.doi.org/10.1016/j.ecolecon.2017.05.008.

Mariel, P., Hoyle, H., Meyerhoff, J., Czajkowski, M., Dekker, T., Glenk, K., Jacobsen, J.B., Liebe, U., Olsen, S.B., Sagebiel, J., Thiene, M., 2021. Environmental valuation with discrete choice experiments. In: Guidance on Design, Implementation and Data Analysis. Springer.
Matthews, Y., Scarpa, R., Marsh, D., 2017. Stability of willingness-to-pay for coastal management: A choice experiment across three time periods. Ecol. Econom. 138, 64–73. http://dx.doi.org/10.1016/j.ecoleon.2017.03.031.
McConnell, K.E., Strand, I.E., Valdés, S., 1998. Testing temporal reliability and carry-over effect: The role of correlated responses in test-retest reliability studies. Environ. Resour. Econom. 12, 357–374. http://dx.doi.org/10.1023/a:1008264922331.
McFadden, D., 1974. Conditional logit analysis of qualitative choice behaviour. In: Zarembka, P. (Ed.), Frontiers in Econometrics. Academic Press, New York, pp. 105–142.
Meyerhoff, Jürgen, Rehdanz, Katrin, Wunsch, Andrea, 2021. Preferences for coastal adaptation to climate change: evidence from a choice experiment. J. Environ. Econom. Policy http://dx.doi.org/10.1080/21606544.2021.1894990.
Morrison, M., Hatfield-Dodds, S., 2011. The success and failure of an inconvenient trust and the stern report in influencing Australian public support for greenhouse policy. Econ. Rec. 87 (277), 269–281. http://dx.doi.org/10.1111/j.1475-4932.2010.00706.x.
Rigby, D., Burton, M., Pluske, J., 2016. Preference stability and choice consistency in discrete choice experiments. Environ. Resour. Econom. 65, 441–461. http://dx.doi.org/10.1007/s10640-015-9913-1.
Ryan, M., Netten, A., Skåtun, D., Smith, P., 2006. Using discrete choice experiments to estimate a preference-based measure of outcome – an application to social care for older people. J. Health Econom. 25 (5), 927–944. http://dx.doi.org/10.1016/j.jhealeco.2006.01.001.
Schaafsma, M., Brouwer, R., Liekens, I., Nocker, L., 2014. Temporal stability of preferences and willingness to pay for natural areas in choice experiments: a test–retest. Resour. Energy Econ. 38, 243–260. http://dx.doi.org/10.1016/j.reseneeco.2014.09.001.
StALU, 2009. Regelwerk Küstenschutz Mecklenburg–Vorpommern – Übersichtsheft - Grundlage, Grundsätze, Standortbestimmung und Ausblick. Rostock, Germany: Staatliches Amt für Umwelt und Natur - Abteilung Küste - Ministerium für Landwirtschaft, Umwelt und Verbraucherschutz Mecklenburg –Vorpommern.