Working Paper in Economics No. 746

Which type of policy instrument do citizens and experts prefer? A choice experiment on Swedish marine and water policy

Claes Ek, Katarina Elofsson, Carl-Johan Lagerkvist

Department of Economics, November 2018

ISSN 1403-2473 (Print)
ISSN 1403-2465 (Online)

UNIVERSITY OF GOTHENBURG
SCHOOL OF BUSINESS, ECONOMICS AND LAW

Which type of policy instrument do citizens and experts prefer? A choice experiment on Swedish marine and water policy

Claes Ek[a,b,1], Katarina Elofsson[c], Carl-Johan Lagerkvist[c]

[a] *Centre for Collective Action Research, University of Gothenburg, P.O. Box 711, 40530 Gothenburg, Sweden*

[b] *Department of Economics, University of Gothenburg, P.O. Box 640, 40530 Gothenburg, Sweden*

[c] *Department of Economics, Swedish University of Agricultural Sciences, P.O. Box 7090, SE-750 07 Uppsala, Sweden*

Abstract

In the choice between alternative environmental policy instruments, economists tend to favor policies capable of attaining cost-efficiency, but other considerations may be important to stakeholders. We perform a choice experiment modeled on Swedish water and marine policy to estimate preferences for different types of environmental policy instruments among citizens and municipal experts. To approximate preferences for each instrument *per se*, choice sets include several attributes that respondents may otherwise view as correlated with instrument type, such as how costs are shared between taxpayers and farmers. In our mixed-logit regressions, both the modal citizen and the modal expert prefer direct regulation and subsidies to nutrient trading. Experts weight taxpayer costs less heavily, implying larger WTP estimates; in particular, nutrient trading is unlikely to deliver sufficiently large cost savings for experts to prefer it to other instrument types. This potentially explains the low takeup of water quality trading outside the US.

Keywords: choice experiments, instrument choice, nutrient trading, water policy

JEL classification: H23, Q53, Q58

☆This work was funded under the framework of the EU's joint Baltic Sea research and development program, through Call no. 2014-12: BONUS Go4Baltic, and under the framework of EU Interreg Central Baltic through the project NutriTrade. The authors have no conflict of interest to declare.

1 Introduction

In the choice between alternative environmental policy instruments, economists tend to favor incentive-based policies capable of attaining a given environmental quality in a cost-effective way (Goulder and Parry, 2008). Within the domain of water quality, and specifically policies to mitigate eutrophication problems, the main candidate is trading in nutrient credits. Fisher-Vanden and Olmstead (2013) list a number of such trading schemes that are now in operation in the United States, Canada, and Australia (there is also a trading system in New Zealand; see Duhon et al., 2015). However, despite water quality problems being widespread, takeup outside of these countries has not occurred. Furthermore, most trading systems are local in scope. What is the cause of the apparent reluctance to adopt nutrient trading?

One possible answer, of course, is that incentive-based policies tend to be less straightforward for water than for air, where trading has been relatively more successful. This is typically due to both physical and institutional constraints. First, while carbon dioxide mixes uniformly in the atmosphere, environmental damages from nutrients may exhibit a great deal of spatial variation, necessitating the introduction of trading ratios across different areas included in trading systems (Montgomery, 1972; Krupnick et al., 1983; Hung and Shaw, 2005). Second, nonpoint emissions (e.g. from agriculture) are inherently more uncertain and difficult to monitor than emissions from point sources (e.g. wastewater treatment plants), which may necessitate a second trading ratio between different source types (Malik et al., 1993; Horan, 2001; Horan and Shortle, 2005). Third, agricultural sources are sometimes exempt from direct regulation, causing institutional design issues regarding how best to involve them in trading, including interactions between nutrient trading and existing policies. For instance, compensatory payments for agricultural nutrient abatement may compete with existing agri-environmental subsidies or regulations.

The above illustrates the fact that it can be difficult to find a suitable design of nutrient trading schemes. However, their use in some countries suggests that other factors (including historical ones) may explain the reluctance to apply this kind of instrument in the European context. The present paper focuses on Swedish water and marine policy and investigates the notion that, within this context, regulators or the general public simply do not like the idea of nutrient trading. More precisely, they may hold preferences over different types of policy instruments, regardless of their impact on private or aggregate outcomes. Our purpose is thus to estimate expert and citizen willingness-to-pay (WTP) for meeting water quality targets through the use of nutrient trading as opposed to direct regulation ('command-and-control') or subsidies.

The idea that some policy instruments may, all else being equal, be preferred to others is not new; there is, for example, a well-documented tendency for a large fraction of lab subjects to vote against proposed 'taxes' that impose marginal costs on certain decisions but are expected to increase individual payoffs as well as efficiency (e.g. Cherry et al., 2012; Kallbekken et al., 2011; Heres et al., 2017; Cherry et al., 2017). These choices do not appear to be caused by confusion: Kallbekken et al. (2011) found that explaining the properties of the tax made experimental subjects more likely to correctly predict its impact, but did not make them significantly more likely to vote in its favor. Moreover, Cherry et al. (2017) found clear evidence that policy aversion is mediated by subjects' cultural worldview, e.g. whether public intervention against private interests is considered legitimate.

While laboratory studies grant researchers a desirable degree of control over choice environments, this may come at the expense of realism. We therefore take the complementary approach of estimating instrument-specific preferences within a stated-preference discrete choice experiment explicitly modeled on actual Swedish marine policy. The respondent pool consists of two groups: (i) Swedish citizens, and (ii) municipal officials ('experts') specializing in issues related to the environment, water quality and sewerage, or local business. Each choice set in the experiment asks respondents to choose between hypothetical policies for meeting Swedish obligations under the 2007 Baltic Sea Action Plan. Each alternative includes an attribute describing whether the policy instrument used is direct regulation, agri-environmental subsidies, or nutrient trading.

Preferences for the instrument-type attribute need to be estimated with care. In particular, respondents may believe that a given policy instrument is correlated with e.g. the probability that an intervention will be effective, and unless that probability is explicitly accounted for within each choice situation, regression estimates may capture perceived correlations rather than a preference for the instrument *per se*.[1] For this reason, our design includes several potentially correlated attributes, including one for the type of 'delivery uncertainty' (Glenk and Colombo, 2011) just described.

We are aware of no other study that attempts to estimate preferences for emissions trading.[2] Furthermore, to our knowledge, only two previous choice experiments have explicitly considered preferences across different types of environmental policy: Johnston and Duke (2007) estimates preferences for alternative land conservation measures (zoning, outright purchase, etc.), while Brännlund and Persson (2012) examines preferences for climate policies. In terms of study design, the former is the most similar to ours: like us, they include additional attributes with the explicit aim of controlling for outcomes that respondents may consider correlated with instrument type.[3] Brännlund and Persson (2012) also include an impressive set of auxiliary attributes in their survey, but do not explicitly code instrument type as an attribute. Instead, they examine the effect of labeling a particular policy alternative either as 'Tax' or 'Other'. In line with the tax-aversion studies mentioned previously, the effect of such a label is found to be negative.

Our paper also builds on a small set of choice experiments that examine differences

[1]Strictly speaking, of course, it is not clear whether such fully intrinsic preferences really exist, or whether (as in the Lancaster model) policy instruments are always preferred based on some perceived set of associated attributes. However, unlike the attributes included in our choice experiment, some characteristics are arguably included in the definition of the instrument: it is difficult to imagine that a preference for emissions trading *per se* could be separated from a preference for a market-based quantity instrument.

[2]Although the lab studies mentioned above do not focus exclusively on taxes, none of them include emissions trading of any kind.

[3]They also provide a useful theoretical framework for thinking about endogeneity due to omitted correlated attributes in choice experiments.

between members of the general public and experts, typically government officials. As in the present paper, these studies typically present both respondent samples with identical or highly similar questionnaires and compare responses. Most find that experts generally have a higher valuation of attributes related to environmental quality (Carlsson et al., 2011; Rogers, 2013; Eggert et al., 2016), though there is at least one example of the opposite (Nordén et al., 2017). All of these studies, however, differ from ours in that they examine preferences or WTP for some environmental policy or good, irrespective of how it is delivered. Thus, the present paper extends the literature by assessing whether citizens and experts differ systematically in their preferred *type* of policy.

The remainder of this paper is structured as follows. Section 2 provides brief background on the issues underlying the choice experiment, i.e. Swedish marine and water policy. Section 3 describes the design of our survey in more detail. Section 4 outlines our econometric strategy, and Section 5 summarizes our results. Section 6 concludes the paper.

2 Swedish water and marine policy

The 2007 Baltic Sea Action Plan set country and basin-specific 'Maximum Allowable Inputs' (MAI) of nitrogen and phosphorus, to be reached by 2021. As per the sixth HELCOM Pollution Load Compilation (Swedish Agency for Marine and Water Management, 2016), Sweden is currently in compliance with all but two of these basin- and nutrient-specific targets. First, the MAI for nitrogen discharges to the Bothnian Bay is set at 17,924 tons/year, but actual discharges were calculated at 19,500 tons/year. Second, the MAI for phosphorus to the Baltic Proper is 308 tons/year, but current discharges were estimated at 780 tons/year.

Of these targets, the second is seen as the more problematic one, for at least three reasons. First, eutrophication is more severe in the Baltic Proper than in the Bothnian Bay (HELCOM, 2014). Second, the limiting nutrient in the Bothnian Bay is not nitrogen, but phosphorus (Swedish Environmental Protection Agency, 2014). Third, the calculations for the sixth Pollution Load Compilation separated total loads into anthropogenic and background loads, and found that the background load alone for phosphorus in the Baltic Proper (370 tons/year) exceeds the Swedish MAI to the same basin. Thus, meeting this target is likely to be very challenging.

Forty percent of the Swedish anthropogenic net phosphorus load to the Baltic Proper arises within the agricultural sector, with wastewater treatment plants being the second largest source of excess phosphorus (22%). However, discharges from large wastewater treatment plants have been substantially reduced since the year 2000, mainly because of bans on using phosphates in detergents (South Baltic Water Authority, 2014a).

Beyond the BSAP targets, Sweden is also obligated by the EU Urban Wastewater Treatment Directive (91/271/EEG) and the EU Water Framework Directive (2000/60/EC) to reduce emissions of nutrients to inland lakes, rivers, and coastal waters. The Water Directive requires all such water bodies to achieve 'good' or 'high' ecological status, including with respect to eutrophication, by 2021 or 2027. These targets are currently relatively far from being met. Within the Baltic Proper catchment area, 28% of all water bodies have yet to be classified, but of those remaining, only 48% currently have good or high ecological status with respect to nutrients.

As for Swedish marine and water policy, it relies heavily on legal requirements and permitting for point sources (wastewater treatment plants and industries), and environmental subsidies for agriculture. Point-source regulation involves permit requirements set based on environmental quality standards ('EQ standards') mainly corresponding to good or high ecological status. Most of these are local in scope and concern lakes, rivers, and Swedish coastal waters. In agriculture, the Swedish Rural development program provides funding (supplemented with information campaigns) for farmers willing to take abatement measures.

While it is clear that Swedish policies have been effective in reducing nutrient loads, economists have found that outcomes have not been cost-effective (e.g. Gren et al., 1997; Elofsson, 2010, 2012). The past few decades has seen increased interest in economic instruments capable of bringing down abatement costs, especially nutrient discharge trading systems, although proposals by the Swedish EPA introduce such systems have ultimately been unsuccessful. An initial proposal (Swedish Environmental Protection Agency, 2008) outlined a trading scheme where a regulated sector (e.g. municipal wastewater treatment plants) are able to meet binding emissions standards by financing compensatory measures within a non-regulated sector (e.g. agriculture), with these transactions handled by the regulating authority.

The proposal was, however, widely criticized by stakeholders for conflicting with existing regulations. First, agri-environmental subsidies may undermine farmers' incentives to supply compensatory measures, and moreover credit payments for measures that are already subsidized likely conflict with additionality requirements within the Rural development program. Second, the point of trading systems is to carry out load reductions (with respect to the Baltic) where they are least expensive. A potential side effect is that measures may be diverted from inland waters subject to EQ standards. While it may be possible to add special provisions to avoid such regulatory conflicts, these auxiliary rules will undermine the cost-efficiency of the trading system.[4] The situation is especially problematic if, as is the case, EQ standards are both abundant and stringent.[5]

A second proposal for nutrient trading was presented in Swedish Environmental Protection Agency (2012). The updated trading system was less ambitious, covering only municipal wastewater treatment plants and not allowing offsets from e.g. agriculture. This resolved conflicts with existing policies to some degree. First, since only municipal wastewater treatment plants were included, there was no obvious risk that trading would undermine subsidies within the Rural development program (or vice versa). Second, trading covered only nitrogen

[4]The EPA concluded that for nutrient trading to work as intended, environmental subsidies within the Rural development program would need to be scrapped and the relationship between trading and EQ standards closely examined at the least. Adopting a fully functional trading system would involve a "fundamental shift in Swedish environmental policy and a regime change from regulation to market-based instruments" (Swedish Environmental Protection Agency, 2010, p. 143).

[5]In fact, the South Baltic Water Authority (2014b) calculated that a program of measures leading to (near) compliance with EQ standards for lakes and rivers in the South Baltic water district would itself imply approximate compliance with the BSAP target for phosphorus discharges to the Baltic Proper.

emissions. EQ standards for water quality largely concern phosphorus, and include nitrogen obligations only for Swedish coastal waters, where significant synergies with BSAP targets are likely. Despite this, the government decided against adoption, stating that the current approach of regulating nitrogen and phosphorus by permitting is "difficult to reconcile with a charge system for these emissions". Subsequent Swedish policy initiatives have again focused mainly on direct regulatory approaches and agri-environmental subsidies (e.g. North Baltic Water Authority, 2016).

3 Survey design

Our aim is to compare expert responses, particularly with respect to the policy-instrument type attribute, to those of the general public. Previous research (e.g. Colombo et al., 2009; Carlsson et al., 2011; Rogers, 2013; Nordén et al., 2017) has demostrated that preferences for environmental policy can differ substantially between citizens and government officials or various stakeholder groups. However, these studies have focused on the stringency of environmental policy itself rather than the choice of instrument in attaining a given target. Our experiment is designed to permit estimation of potential expert/citizen differences in such instrument-type preferences (Johnston and Duke, 2007).

Like previous studies, our sample consists of two groups of respondents. First, in April 2017, the choice experiment was sent out to a representative panel of Swedish citizens aged 18-75. Data collection from this group concluded by May 2017 and yielded 2001 complete responses. Second, also in May 2017, we sent an email to the registry office of all municipalities in Sweden, requesting that an online link to the survey be forwarded to as many municipal officials as possible working within the water quality domain. We specifically asked for experts on (i) environmental issues, particularly with respect to wastewater treatment plants, small-scale wastewater treatment, or local water quality, (ii) water and sewerage, and/or (iii) local business. By August 2017, after several reminders, 146 experts had completed at least part of the survey; 115 respondents had completed the entirety of the choice experiment (though not necessarily the post-experimental questionnaire; see below).[6]

The survey was an online questionnaire consisting of several parts. Respondents first read a general description of the study design, including a cheap-talk script similar to that developed by Carlsson et al. (2005) and shown to often lower WTP significantly. This section also contained a link to optional background information on Swedish marine and water policy, which described the Swedish BSAP target for phosphorus discharges to the Baltic Proper, as well as EU targets for good ecological status in lakes, streams, and coastal waters within the Baltic Proper catchment area.

Following the existing research on the benefits from initiating preference formation (Ben-Akiva and Gershenfeld, 1998; Cohen and Liechty, 2007), the second part of the survey familiarized respondents with each attribute included in the choice experiment. We included this section to make sure that each participant understood the meaning of all attributes and as an ex-ante measure against attribute non-attendance. First, participants read a brief description of an attribute (translated versions of these descriptions are given in Online Appendix A.1), and were asked to select their preferred level for that attribute. After this process had been repeated for each attribute, they then progressed to the choice experiment itself.

Table I lists all attributes and their associated levels. The choice experiment faced by each respondent consisted of 20 choice sets, each of which included an invariant status-quo option reflecting the likely outcome of current Swedish marine and water policy, and two alternative (more ambitious) policy packages. An example choice set is presented in Online Appendix A.2 (along with screenshots in Appendix B). At any point during the choice experiment, participants could review the background information, and could also recall attribute descriptions by hovering the mouse cursor over an attribute. In section 5.4, we check whether the relatively large number of choice sets elicited incoherent preferences, such

[6]All data collection was carried out in collaboration with market research firm GfK, which maintains the general-public panel. Prior to the main data collection phase, we checked the design in a prestudy conducted on a smaller sample of 66 respondents.

as a fatigue effect leading to less considered choices in later tasks (Swait and Adamowicz, 2001). At this point, we simply note that increased survey length need not be detrimental to choice reliability and precision (Hess et al., 2012); indeed, at least up until around 20 choice sets, it might improve reliability through a learning effect (Johnson and Orme, 1996; Carlsson et al., 2012).

We used a heterogeneous design in which each subject faced one out of 50 distinct surveys, each of which consisted of a full 20 choice tasks. This approach increases the variation in attribute levels across the entire sample of respondents and has been shown to provide substantial efficiency improvements, especially when the number of attributes is relatively large, as in our study (Sándor and Wedel, 2005).

Our main attribute of interest is 'Type of policy instrument'. As Johnston and Duke (2007) argue, it is likely that preferences for different types of policy instruments are partly driven by, and thus confounded with, preferences for outcomes believed to be correlated with the use of those instruments. Within the context of marine policy, for example, a preference for environmental subsidies could be driven by a desire to safeguard farmers' competitiveness. Explicitly adding farmer profits as an attribute within the choice sets may alleviate this problem, leading to better estimates of preferences for each instrument *per se*, which was the main objective of the survey.

We therefore include the cost to farmers as a separate attribute, along with two others likely to be seen as correlated with instrument type: (i) compliance with targets for good/high ecological status in domestic inland waters, and (ii) perceived 'delivery uncertainty' as to whether the policy will have the intended effect. Delivery uncertainty can be included in choice sets as a quantitative (Glenk and Colombo, 2011) or a qualitative (Lundhede et al., 2015) attribute; because of the inherent difficulty in quantifying *ex ante* the probability that a policy will be effective, we use a qualitative measure. The status-quo level for the uncertainty attribute was chosen as 'Not applicable' rather than e.g. 'Very certain'. While we cannot rule out the existence of additional omitted attributes, the discussion in section

Table I: *Attributes and levels in the choice experiment*

Attribute	Status quo	Alternative
Target compliance: Baltic Proper	12%	40%; 70%; 100%
Target compliance: lakes, streams, coastal waters	50%	65%; 80%; 100%
Type of policy instrument	N/A	Legislation and permitting; emissions trading; environmental subsidies
Likelihood that policy is effective	N/A	Very certain; rather certain; rather uncertain; very uncertain
Cost to farmers (SEK per year per farmer)	0	+10,000; +20,000; +30,000
Cost to taxpayers (SEK per year per taxpayer)	0	+100; +150; +200; +250; +300

2 suggests that the Swedish debate on instrument choice in marine and water policy is largely framed around these three additional attributes.[7] To mitigate hypothetical bias, our design restricted attribute-level combinations in two ways: neither "Emisions trading" nor "Environmental subsidies" was ever combined with the highest level of "Cost to farmers" (+30,000).[8]

The last stage of the survey was a questionnaire on mainly demographic and socioeconomic characteristics. In addition, we included a two-part item on attribute nonattendance, asking whether respondents ever ignored an attribute while making choices and, if so, which attribute(s) they had ignored. We elicited stated nonattendance in this way — i.e. after the entire choice experiment was complete rather than after each choice task — to avoid priming effects, as suggested e.g. by Scarpa et al. (2010).

In most ways, the design of the choice experiment was identical across the general-public and the expert samples, but since we wanted experts to respond in their capacity as municipal officials rather than as private citizens, we rephrased certain parts of the pre-experimental information slightly. In particular, while citizens were asked to to select the alternative that they would prefer for society to adopt, municipal experts were prompted to select the alternative that 'is the best, given the conditions that apply within your municipality'.

[7] An important caveat is that Lundhede et al. (2015) find that respondents do not simply accept a stated 'objective' level of uncertainty (whether qualitative or not), but base their choice partly on their priors regarding the perceived effect of policy. This suggests that while our approach of including explicit 'control' attributes is probably useful, it may not fully solve the omitted-variables problem with respect to those attributes. We will return to this issue in section 5.

[8] Furthermore, the levels chosen were calibrated to real-world figures. For instance, South Baltic Water Authority (2014b) calculated that a package of measures aimed at attaining compliance with Swedish water quality regulations would cost a total of SEK 4.5 billion, with SEK 2 billion borne by farmers (p. 168-169). This was mostly assumed to involve direct regulatory policies, so we treat these estimates as an upper bound. Dividing the total costs by the number of taxpayers and farmers in the Baltic Proper catchment area produces costs of about SEK 28,000 per farmer and SEK 350 per taxpayer (the details of these calculations are available upon request).

4 Econometric specification

Our empirical analysis is based on a standard random-utility framework where each individual i's utility of choosing an alternative j in choice set t is assumed to be given by

$$U_{ijt} = v(\mathbf{X}_{ijt}) + \epsilon_{ijt} \tag{1}$$

where v is a systematic component and ϵ is a random error term. v is assumed to depend on an observable vector $\mathbf{X}_{ijt}$ which could include attribute levels as well as personal characteristics. We will focus on the case where $\mathbf{X}_{ijt}$ contains only attribute levels and an alternative-specific constant (ASC) associated with choosing a non-status quo policy package. Individual i is taken to choose an alternative $j = A$ over $j = B$ if $U_{iAt} > U_{iBt}$.

We use a random parameter logit model to estimate preferences. The random parameter (or mixed) logit framework specifies the systematic component in equation (1) as $v = \boldsymbol{\beta}_i' \mathbf{X}_{ijt}$. This involves making two assumptions. First, systematic utility is assumed linear in each variable in $\mathbf{X}_{ijt}$. Second, this model allows for heterogeneity in tastes: some or all of the marginal utility parameters $\boldsymbol{\beta}_i$ are assumed to vary across respondents according to some random distribution(s) prespecified by the analyst. Random parameter logit additionally exploits the panel structure of our choice data in that parameters are assumed to be constant within each individual, i.e. across the 20 choice sets faced by a given respondent, making it possible to derive an estimate of each individual-specific β_i (Revelt and Train, 2000).

For estimation we use Nlogit 6. In our regressions, we assume that all parameters (including the ASC) except those related to the 'Cost to taxpayers' attribute are normally distributed across the population. Tax-cost preferences are assumed to be common to all respondents and thus nonstochastic. The likelihood function associated with random parameter logit cannot be evaluated directly, so we estimate the parameters using maximum simulated likelihood with 500 Halton draws; for details, see Revelt and Train (1998) and Hole (2007). The resulting marginal utilities can be used to construct measures of willingness to pay (WTP). For each attribute, mean WTP is calculated as the ratio of the mean estimated coefficient for that attribute divided by the fixed coefficient for tax cost.

5 Results

5.1. Estimated preferences

Table II reports the results from a pair of random parameter logit regressions. Column 1 and 2 estimate preferences for the general-public sample, while column 3 and 4 repeat the analysis for experts.[9] In both cases, mean coefficients for environmental attributes are relatively large and significant; estimated utility is monotonically decreasing in the degree of uncertainty; and coefficients for both cost attributes are negative, with the magnitude of the disutility being much larger for taxpayer cost. Finally, on average, both citizens and experts prefer both environmental subsidies and legislation and permitting to nutrient trading.[10]

Estimated standard deviations are generally large and highly significant within both samples, indicating there was substantial preference heterogeneity among both citizens and experts. This holds particularly true for the instrument-type attribute, where estimated means are typically smaller than corresponding standard deviations. To further explore the distribution of preferences for instrument type, we estimated individual preference parameters conditional on the observed data and estimated population-level parameters (Revelt

[9]Note that since the non-status quo levels of the discretized attributes (e.g. policy-instrument types) never appear as part of the status quo alternative, the ASC will be confounded with these attributes in any regression. The constant thus captures the combined utility effect of (i) rejecting the status quo in favor of some alternative, as well as (ii) moving from the attribute levels associated with the status quo to a set of non-status quo reference levels. This also implies that for dummy coded attributes, we cannot identify utility parameters for all non-status quo levels. For instance, the coefficient on 'Very uncertain' should be interpreted as the utility difference between that level and the omitted non-status quo level for that attribute, which is 'Very certain'.

[10]We also ran regressions where all attributes were dummy coded, yielding separate estimates for each possible attribute level. Overall, results were quite similar. There was one apparent exception: the high-resolution model reveals that expert preferences for tax cost trace an inverted-U shape, with moderately expensive policies preferred to inexpensive ones. While this may seem inconsistent with the relatively substantial negative coefficient estimated in column 3 of Table II, it is not. In fact, fitting a regression line through the points implied by the parameter estimates in the high-resolution model yields a slope coefficient of -1.1875.

Table II: *Random parameter logit estimates*

	Citizens		Experts	
	Mean	Standard deviation	Mean	Standard deviation
ASC, non-status quo	3.538***	5.525***	6,667***	4.226***
	(0.146)	(0.160)	(0.733)	(0.266)
Baltic Proper	0.834***	1.552***	1.519***	1.903***
	(0.052)	(0.062)	(0.225)	(0.266)
Lakes, streams, coastal waters	0.617***	0.137	0.809***	1.129**
	(0.064)	(0.237)	(0.298)	(0.477)
Emissions trading	-0.267***	0.675***	-1.025***	1.218***
	(0.027)	(0.033)	(0.146)	(0.161)
Legislation and permitting	0.050	0.670***	-0.088	1.374***
	(0.036)	(0.046)	(0.183)	(0.215)
Rather certain	-0.270***	0.064	-0.447***	0.201
	(0.025)	(0.054)	(0.112)	(0.222)
Rather uncertain	-1.343***	0.917***	-1.820***	0.953***
	(0.036)	(0.038)	(0.156)	(0.158)
Very uncertain	-1.956***	1.491***	-2.670***	1.392***
	(0.049)	(0.046)	(0.214)	(0.190)
Cost to farmers	-0.041***	0.065***	-0.067***	0.120***
	(0.002)	(0.002)	(0.012)	(0.015)
Cost to taxpayers	-3.326***		-1.224**	
	(0.144)		(0.616)	
Observations	40,020		2,449	
Respondents	2,001		146	
R-squared (constants only)	0.414		0.387	

Standard errors are given within parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Target compliance variables (Baltic Proper; lakes, streams, coastal waters) are coded as proportions, i.e. take values between 0 and 1; these coefficients thus reflect a 100% difference. Cost to farmers and cost to taxpayers expressed in thousands of SEK.

and Train, 2000). Table III presents the corresponding distributions of preference orderings, which are based on Table II.[11]

We confirm, first, that there is substantial heterogeneity within the population, with the modal preference ordering (Legislation $\succ$ Subsidies $\succ$ Trading) applying to less than half of respondents in each sample. Second, there are some qualitative differences between the samples. In particular, as demonstrated in the lower part of the table, a majority of citizens prefer legislation and permitting to all other instrument types, while a majority of experts hold environmental subsidies as their most preferred instrument type. In both groups, however, support for emissions trading is low, with only about 10-15% of respondents preferring it to all other instrument types.[12]

The above discussion suggests there may be systematic differences in preferences across samples; however, the task of performing formal statistical inference with respect to such differences is complicated by the fact that coefficient estimates in random parameter logit models are confounded with a 'scale parameter' representing the amount of variation within each group of respondents. To address this issue, we use the Swait and Louviere (1993) procedure which tests the hypothesis of parameter equality across samples while making no assumptions about the relative scale of the two samples. In practice, it involves first estimating the relative scale factor, and then testing for parameter equality given that estimate. The former step entails finding the relative scale that maximizes the log-likelihood function

[11] The second column in Table III uses all experts data, including incomplete responses. This may imply that some preference orderings are imprecisely estimated; all qualitative results are robust to repeating the analysis without these respondents, however.

[12] The individual-specific estimates converge to the true individual marginal utilities only if the number of choices per respondent rises without bound (Train, 2009). Thus, as an alternative to Table III, we also derived population shares by making 10,000 draws from a bivariate normal distribution for "Emissions trading" and "Legislation and permitting", with population means and standard deviations as given in Table II (covariances of zero were assumed). In these simulations, the modal preference ordering applied to less than 40% of respondents in either sample. Although no instrument type was favored by a majority in the sense of being the most preferred option of more than 50% of respondents, legislation and permitting was the modally most preferred instrument in both groups, at around 45%. Despite this, a narrow majority of experts (53%) preferred subsidies to legislation, with the implication that subsidies would be chosen under any pairwise-majority rule. Among citizens, there was instead a majority (54%) for choosing legislation over subsidies.

Table III: *Distribution of estimated individual preferences for instrument type*

Preference orderings	Respondent share (citizens)	Respondent share (experts)
Subsidies $\succ$ Legislation $\succ$ Trading	21%	38%
Subsidies $\succ$ Trading $\succ$ Legislation	11%	18%
Trading $\succ$ Subsidies $\succ$ Legislation	11%	6%
Trading $\succ$ Legislation $\succ$ Subsidies	5%	1%
Legislation $\succ$ Subsidies $\succ$ Trading	46%	32%
Legislation $\succ$ Trading $\succ$ Subsidies	6%	3%
Sum	100%	100%
Top choice: Subsidy	32%	57%
Top choice: Trading	16%	8%
Top choice: Legislation	53%	36%
Sum	100%	100%

for the pooled (citizen/expert) sample; we perform this grid search using random parameter logit regressions with 25 rather than the full 500 Halton draws. This leads us to strongly reject the hypothesis that parameters are equal across samples ($\chi^2 = 2683.870$).

It is possible that, e.g. as a simplifying heuristic, certain respondents ignored some attributes in the choice sets, and that this may affect the results in Table II. In our end-of-session questionnaire, 32.2% of citizen respondents stated that they ignored at least one attribute at some point during the choice experiment, with instrument type the most commonly ignored attribute. This is comparable to self-stated non-attendance rates reported in other studies (see e.g. Hensher et al., 2005; Nguyen et al., 2015). Attribute nonattendance was considerably higher among experts, where a majority ignored one or more attribute; and close to 30% ignored the instrument-type attribute. Given that instrument type is our main variable of interest, it seems useful to see whether such attribute nonattendance mediates our results. We use the approach proposed by Hensher et al. (2005), in which each attribute parameter is restricted to equal zero exactly for the set of respondents stating that they ignored the attribute. In Online Appendix Table C.I, our main random parameter logit regressions are re-run under these alternative assumptions; however, this produces results that are typically both qualitatively and quantitatively very similar to those in Table II.[13]

5.2. WTP estimates

WTP estimates for citizens and experts are given in Table IV. Within each respondent group, we first present estimates based on Table II. WTP, which we define as each respondent's marginal rate of substitution between an attribute and tax cost, is here calculated by dividing each marginal-utility parameter with the estimated marginal disutility of tax payments. Standard errors are calculated by the delta method.[14] For continuous attributes

[13] We do note that experts' coefficient for taxpayer cost is no longer significant ($p = 0.126$).

[14] For the general-public sample, the Krinsky and Robb (1985) simulation procedure produces practically identical results. For experts however, Krinsky-Robb simulation yields standard errors that are typically almost 100 times larger than those produced using the delta method, with the result that $p > 0.9$ for all attributes. Our best guess as to what is going on is that for experts, some of the random draws of the

(Baltic Proper; Lakes, streams, and coastal waters; Cost to farmers), coefficients should be interpreted as the WTP for a positive or negative one-unit change in each variable, expressed in thousands of SEK. In the first column, for example, respondents are willing to pay an estimated additional 12 SEK to reduce farmer costs by 1000 SEK. For discrete variables, the estimates describe the WTP to move from the reference category to the attribute level corresponding to a given variable.

Clearly, WTP measures for experts are less precisely estimated but also larger in magnitude; recall that the tax-cost attribute was likewise relatively small and imprecisely estimated in Table II. The point estimates on emissions trading suggest that Swedish citizens have a mean willingness to pay of approximately 80 SEK/year for basing national marine and water policy on environmental subsidies rather than emissions trading. The corresponding figure for experts is about an order of magnitude larger.

Given that respondents clearly do not prefer nutrient trading, a natural question is: how much more favorable would a policy involving trading need to be along other dimensions in order to be preferred to another policy based on subsidies or direct regulation? While the number of possible combinations makes it infeasible to provide a full answer to this question, some observations may be made. In Table V, we use the (preference-space) WTP estimates to make the comparison separately for each attribute.[15] For example, a representative citizen choosing between one policy package involving nutrient trading and another involving agri-environmental subsidies would prefer the former if it improves target compliance with respect to the Baltic Proper by at least $-(-0.080/0.251) = 31.9$ percentage points, all else being equal.

The table reveals that the relative strength of the aversion to trading is substantially

(relatively small) estimated taxpayer cost parameter are very close to zero and thus produce extremely large WTP estimates which end up inflating the empirical standard error of mean WTP. The delta method, which uses only means and (co)variance estimates, is not subject to the same issue.

[15] The delivery-uncertainty attribute is not included in Table V because it is coded as discrete variables, making comparisons impractical.

Table IV: *Willingness-to-pay estimates*

	Citizens		Experts	
	Mean WTP	p	Mean WTP	p
Baltic Proper	0.251***	0.000	1.241*	0.056
	(0.019)		(0.648)	
Lakes, streams, coastal waters	0.186***	0.000	0.661	0.111
	(0.021)		(0.415)	
Emissions trading	-0.080***	0.000	-0.838*	0.050
	(0.009)		(0.428)	
Legislation and permitting	0.015	0.168	-0.072	0.640
	(0.011)		(0.154)	
Rather certain	-0.081***	0.000	-0.366*	0.075
	(0.008)		(0.206)	
Rather uncertain	-0.404***	0.000	-1.487**	0.049
	(0.020)		(0.757)	
Very uncertain	-0.588***	0.000	-2.182**	0.049
	(0.028)		(1.107)	
Cost to farmers	-0.012***	0.000	-0.054*	0.060
	(0.001)		(0.029)	

Standard errors calculated by the delta method are given within parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Table is based on Table II. Target compliance variables (Baltic Proper; lakes, streams, coastal waters) are coded as proportions, i.e. take values between 0 and 1; these coefficients thus reflect a 100% difference. Cost to farmers and cost to taxpayers expressed in thousands of SEK.

Table V: *Compensation required for nutrient trading to be preferred to other instruments, by attribute*

	General public		Experts	
Attribute	Subsidy	Legislation	Subsidy	Legislation
Baltic Proper (%)	31.9%	37.8%	67.5%	61.7%
Lakes, streams, coastal waters (%)	43.0%	51.1%	126.8%	115.9%
Cost to farmers (SEK/year)	6,509	7,730	15,404	14,080
Cost to taxpayers (SEK/year)	80	95	838	766

greater among experts than the general public. Experts need at least twice as much compensation, along any dimension, to be willing to switch instrument type. In fact, the required compensating improvement is unattainable for the 'Lakes, streams, and coastal waters' attribute, exceeding 100%; thus, a necessary condition for acceptance among experts is that at least one other attribute improves as well.

Arguably, the same applies to taxpayer costs: the required compensation is on the order of 800 SEK/year, which is much larger than the upper bound of 300 SEK/year used in the experiment. There are only a few empirical estimates of the cost savings from nutrient trading in Sweden, or other regions or countries surrounding the Baltic Sea (e.g. Elofsson, 2010; Swedish Environmental Protection Agency, 2010, p. 86). For our purposes, the most relevant study may be Elofsson (2012), who compared actual Swedish nitrogen and phosphorus reductions over 1995-2005 with a cost-effective solution achieving the same reductions to each Baltic Sea basin. She found that the cost differential was only around 5 MEUR per year, translating into approximately 7 SEK/year and Swedish taxpayer. While attaining BSAP targets (to a greater degree) would entail higher costs and thus is likely to enhance the cost savings from using economic instruments, it seems implausible that nutrient trading could deliver the compensating cost savings required by municipal experts.

5.3. Interaction analysis

It is clear from the previous section that preferences for instrument type remain significant even when a relatively large number of possibly correlated attributes are explicitly included in the choice experiment. There are three possible explanations for this observed pattern: (i) respondents have a preference for instrument types *per se*, (ii) important correlated variables remain uncontrolled for in the design, (iii) the inclusion of correlated attributes failed to elicit *ceteris-paribus* preferences, i.e. respondents did not fully hold the level of other attributes constant when choosing among instrument types. While testing explanation (ii) is not possible, we may look for indirect signs of (iii) by studying interactions between instrument type and other attributes.

For example, suppose respondents are found to be more averse to uncertainty under nutrient trading than under legislation and permitting. Such a pattern would arguably suggest than even after explicitly controlling for uncertainty, subjects view nutrient trading as inherently more uncertain, making added uncertainty under nutrient trading worse than the same amount of added uncertainty under some other policy instrument. Similarly, if our control strategy is effective, there is little reason to expect the effect of e.g. environmental effectiveness or farmer costs to differ across instrument types.

To assess whether each other attribute was valued differently based on the instrument type it was paired with, we perform a sequence of two-way interactions. This analysis is presented in Table VI. Each section of the table corresponds to a separate regression in which we interact instrument type with one other attribute. Although all regressions include coefficients for all attributes, only interaction-relevant parameters are reported in the table. Note that environmental subsidies represent the reference category; hence, the attribute main effect(s) reported at the start of each table section (e.g. Baltic Sea in the top section) should be interpreted as the effect of that attribute under environmental subsidies. Correspondingly, the interaction terms represent the differential effect of each attribute under legislation or trading, as opposed to environmental subsidies.

We now wish to test whether instrument type was a moderator of the marginal utility of other attributes. The null for a given attribute corresponds to the joint hypothesis that all associated interaction parameters are zero, in which case the marginal utility of that attribute exhibited no significant variation across different instrument types. As an example, for the Baltic Sea attribute we test whether mean coefficients in Table VI satisfy

$$\text{Baltic Sea} * \text{Legislation} = \text{Baltic Sea} * \text{Trading} = 0$$

It should be clear that fully analogous expressions exist for all other attributes; for delivery uncertainty, we test whether all six interaction parameters are jointly zero.

Table VI: *Random parameter logit estimates, interaction analysis*

	Citizens		Experts	
	Mean	Standard deviation	Mean	Standard deviation
Interaction: instrument type & Baltic Sea				
Baltic Sea	0.902***	1.545***	1.547***	1.630***
	(0.086)	(0.062)	(0.389)	(0.227)
Legislation	0.181*	0.399***	-0.446	1.785***
	(0.093)	(0.094)	(0.437)	(0.189)
Trading	-0.144	0.461***	-1.553***	1.206***
	(0.088)	(0.087)	(0.413)	(0.171)
Baltic Sea * Legislation	-0.173	0.755***	0.276	0.740**
	(0.124)	(0.094)	(0.571)	(0.338)
Baltic Sea * Trading	-0.179	0.667***	0.477	0.184
	(0.122)	(0.097)	(0.556)	(0.468)
All interactions = 0 (p-value)	0.246		0.689	
Interaction: instrument type & lakes, streams, coastal waters				
Lakes, streams, coastal waters	0.757***	0.255	0.729	1.756***
	(0.136)	(0.232)	(0.645)	(0.518)
Legislation	0.321*	0.667***	-0.728	1.205***
	(0.177)	(0.048)	(0.818)	(0.190)
Trading	-0.208	0.443***	-0.874	0.643*
	(0.173)	(0.072)	(0.788)	(0.344)
Lakes... * Legislation	-0.325	0.067	0.628	0.148
	(0.211)	(0.193)	(0.980)	(0.349)
Lakes... * Trading	-0.072	0.612***	-0.135	1.381***
	(0.209)	(0.090)	(0.956)	(0.351)
All interactions = 0 (p-value)	0.280		0.706	
Interaction: instrument type & uncertainty				
Rather certain	-0.217***	0.036	-0.197	0.081
	(0.052)	(0.051)	(0.242)	(0.260)
Rather uncertain	-1.310***	0.918***	-1.611***	1.008***
	(0.058)	(0.038)	(0.257)	(0.160)
Very uncertain	-1.939***	1.532***	-2.416***	1.666***
	(0.067)	(0.049)	(0.293)	(0.256)
Legislation	0.071	0.625***	0.132	1.806***
	(0.057)	(0.049)	(0.288)	(0.274)
Trading	-0.233***	0.661***	-0.736***	1.029***
	(0.054)	(0.035)	(0.248)	(0.170)

Rather certain * Legislation	-0.104	0.068	-0.391	0.856*
	(0.078)	(0.105)	(0.378)	(0.476)
Rather uncertain * Legislation	-0.055	0.353***	-0.611	0.211
	(0.081)	(0.130)	(0.381)	(0.556)
Very uncertain * Legislation	0.060	0.252	-0.718*	0.262
	(0.083)	(0.160)	(0.389)	(0.533)
Rather certain * Trading	-0.052	0.145	-0.266	0.319
	(0.078)	(0.092)	(0.367)	(0.314)
Rather uncertain * Trading	-0.027	0.200	-0.095	0.183
	(0.079)	(0.122)	(0.352)	(0.651)
Very uncertain * Trading	-0.016	0.330**	-1.128**	1.270***
	(0.085)	(0.165)	(0.452)	(0.418)
All interactions = 0 (*p*-value)	0.593		0.165	
Interaction: instrument type & cost to farmers				
Cost to farmers	-0.038***	0.072***	-0.088***	0.124***
	(0.004)	(0.003)	(0.021)	(0.013)
Legislation	0.343***	0.444***	0.627	0.572***
	(0.117)	(0.078)	(0.537)	(0.193)
Trading	-0.304***	0.672***	-1.561***	1.019***
	(0.104)	(0.033)	(0.487)	(0.203)
Cost to farmers * Legislation	-0.012**	0.017***	-0.027	0.048***
	(0.005)	(0.003)	(0.025)	(0.008)
Cost to farmers * Trading	0.003	0.001	0.036	0.044***
	(0.007)	(0.005)	(0.031)	(0.014)
All interactions = 0	0.007		0.036	
Interaction: instrument type & cost to taxpayers				
Cost to taxpayers	-3.125***		-0.892	
	(0.268)		(1.181)	
Legislation	0.096	0.671***	0.092	1.371***
	(0.087)	(0.046)	(0.396)	(0.216)
Trading	-0.193**	0.676***	-1.007***	1.222***
	(0.083)	(0.033)	(0.381)	(0.162)
Cost to taxpayers * Legislation	-0.234		-0.901	
	(0.397)		(1.766)	
Cost to taxpayers * Trading	-0.372		-0.099	
	(0.393)		(1.749)	
All interactions = 0	0.632		0.859	
Observations	40,020		2,449	
Respondents	2,001		146	

Standard errors are given within parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Target compliance variables (Baltic Proper; lakes, streams, coastal waters) are coded as proportions, i.e. take values between 0 and 1; these coefficients thus reflect a 100% difference. Cost to farmers and cost to taxpayers expressed in thousands of SEK. Each twoway interaction corresponds to a separate regression. All attribute coefficients were included in all regressions (which used 500 Halton draws in estimation), but only those relevant to the interaction in question are reported in the table. Interactions of instrument type and cost to taxpayers are assumed nonstochastic.

Testing each of the resulting five null hypotheses (per respondent group) separately and applying a Bonferroni correction for multiple hypothesis testing (implying that the critical p-value is $0.05/5 = 0.01$), we reject the null hypothesis in one case only: the interaction between instrument type and cost to farmers, for citizens. This reflects a signficant negative interaction between cost to farmers and direct regulation, suggesting that farmer costs of a given magnitude were perceived as a larger disutility under direct regulatory policy than under another instrument type.

Although any interpretation will necessarily be preliminary, this seems consistent with the fact that only direct regulation does not explicitly compensate farmers for abatement measures. As a result, respondents may have found statements about farmer costs under both nutrient trading and environmental subsidies relatively unrealistic, producing the observed significant interaction. However, an alternative (and, for our purposes, less problematic) explanation of the observed pattern is offered by the fact that, as noted in section 2, our design combined high farmer costs (30,000) only with the "Legislation and permitting" instrument type. Because of this, the interaction between instrument type and cost to farmers necessarily entails a somewhat apples-and-oranges comparison. For subsidies and trading, the interaction captures the difference between the levels 10,000 and 20,000; but for legislation and permitting, it fits the entire interval between 10,000 and 30,000.[16] Thus, if the marginal disutility associated with farmer costs is increasing in the level of those costs, we might expect precisely the type of significant interaction that is actually observed.

Moving on, another observation in Table VI is that, although the main effects of legislation and emissions trading generally have the same sign as in Table II, they are not always significant. For each of the regressions in Table VI, we therefore complement the analysis by performing a joint test of the hypothesis that both instrument-type main effects are equal to

[16] The number of observations associated with each level of farmer cost also differed substantially across instrument type. For instance, low farmer costs (10,000) was combined with environmental subsidies in 12,338 choice sets in the citizen data, but with legislation and permitting in only 1,808 sets.

zero (again, with Bonferroni-corrected significance levels). For citizens, this null hypothesis is rejected in all cases except for the 'lakes, streams, coastal waters' attribute ($p = 0.014 > 0.1$), and even there, the marginal utility of legislation is significantly different from the marginal utility of nutrient trading ($p = 0.004$). For experts, the null is rejected in all cases except, again, the 'lakes, streams, coastal waters' attribute ($p = 0.499$), as well as the tax cost attribute, where however the null is very nearly rejected ($p = 0.011$). On the whole, we conclude that (possibly excepting legislation and cost to farmers), our exploratory interaction analysis does not reveal compelling evidence that estimated instrument-type preferences are biased by the other included attributes.

5.4. Order effects

As discussed in Section 3, the questionnaire contained a large number of discrete choice tasks, so order effects reflecting e.g. learning or fatigue are potentially a concern. Table C.II in the Online Appendix reports output from multinomial logit models corresponding to the entire citizen data set, choice sets 1-10, and choice sets 11-20, respectively. Table C.III presents the same analyses for the expert data set.[17]

To check for changes in preferences over time, we follow Carlsson et al. (2012) and again use the Swait and Louviere (1993) procedure to test for differences between the first (sets 1-10) and second half (11-20) of the survey, while allowing for relative scale (and thus variance) to differ between these two halves. As in the above comparison between citizens and experts, the test compares the likelihood of a pooled model — where attribute variables in observations belonging to the second half of the survey are scaled by the proportion yielding the highest likelihood value — with the sum of likelihoods for two models, each of which separately analyzes one half of the data. Since our main analysis considered citizens and experts separately, we likewise perform this test independently for each respondent group.

[17]We report MNL results here because there were some convergence problems with the mixed logit model, making the tests described below difficult to evaluate. However, qualitative results (where applicable) are similar in mixed logit models.

For experts, we can reject neither the hypothesis that parameters are equal when allowing for potential differences in scale ($\chi^2 = 6.081$), nor the hypothesis that the scale parameter is itself equal across early and late choice sets ($\chi^2 = 0.109$). For citizens, however, we reject the null hypothesis of equal preferences and scale across choice set 1-10 and 11-20 ($\chi^2 = 80.234$); thus, in this case preferences did differ significantly over the course of the 20 choice tasks. However, the only major qualitative differences between the two subsamples is that legislation is not significantly preferred to environmental subsidies in the first half of the survey (note that standard errors are generally larger in the subsample regressions). As for quantitative differences, the most apparent one is that the magnitudes of the alternative-specfic constant and the tax cost attribute are larger in late choice sets. The difference in tax-cost magnitude impacts the relative size of WTP estimates, which are generally larger in the first half of the survey (Table C.IV). Despite this, point estimates for instrument-type WTP are remarkably stable across time. Overall, we find little apparent cause for concern in our sample.

6 Concluding remarks

This paper has presented results from a choice experiment designed to estimate citizen and expert preferences for three different types of policy instruments applicable to marine and water policy: (i) agri-environmental public subsidies, (ii) legislation and permitting, and (iii) nutrient trading. Respondents were presented with a hypothetical scenario based on actual Swedish conditions, and were asked to choose between alternative policies for the Baltic Proper catchment area.

Results indicate that both groups clearly and significantly prefer subsidies as well as legislation and licensing to nutrient trading. This applies even when explicitly holding other attributes fixed, including impacts on the Baltic Sea and inland waters, the likelihood that each policy alternative will be effective, and costs to farmers and/or taxpayers. Moreover, in a separate interaction analysis, we find little indication that preferences for instrument type remain dependent on these other attributes, suggesting that respondents are averse to nutrient trading for some reason other than perceived correlations between instrument type and other included factors.

We also find that preferences differ significantly between citizens and experts. While the modal preference ordering among citizens is to prefer legislation and permitting to environmental subsidies, the opposite holds among experts; though it should be noted that the null hypothesis of indifference between these instruments can be rejected in neither group. Perhaps more importantly, experts weight taxpayer costs relatively less heavily. This implies generally higher marginal WTP (in tax terms) for all other attributes, including instrument type. Another implication is that experts require more compensation in terms of improvements to other attributes in order to accept a shift from another policy regime to trading. In particular, we find that compensation, in terms of abatement cost, required by experts is unlikely to be delivered by any real-world trading scheme.

In our view, these findings offer a partial explanation of the reluctance of Swedish policy makers to adopt nutrient trading. What they imply for future attempts to introduce trading schemes is less clear, as this depends on *why* respondents remain skeptical to trading. In principle, it may be possible to design nutrient trading in ways that address remaining concerns; unfortunately however, our data does not permit us to determine what these remaining aversive characteristics are, and they are also not obvious from the policy debate over the last few decades.

Potentially, remaining disutility may be due to fairness concerns. Nutrient trading is essentially described in terms of offsetting in our survey background information, where we emphasize the key innovation that regulated point sources are allowed to buy abatement measures elsewhere. This forms a departure from the 'polluter pays principle' which may activate fairness preferences if regulated sources are deemed to avoid their just responsibilities. We note that although agriculture is likewise a major source of phosphorus leakage, emissions from this sector is currently not subject to regulation, and thus respondents may not view e.g. subsidy payments to farmers as similar fairness violations. If this conjecture is correct, then what respondents dislike is precisely the flexibility that forms the rationale for trading, and it is doubtful whether clever policy design would be sufficient to sway the skeptics. However, further research is certainly required to confirm or reject these ideas.

References

M. Ben-Akiva and S. Gershenfeld. Multi-featured Products and Services: Analysing Pricing and Bundling Strategies. *Journal of Forecasting*, 17(3-4), 1998.

R. Brännlund and L. Persson. To tax, or not to tax: preferences for climate policy attributes. *Climate Policy*, 12(6), 2012.

F. Carlsson, P. Frykblom, and C.J. Lagerkvist. Using cheap talk as a test of validity in choice experiments. *Economics Letters*, 89(2), 2005.

F. Carlsson, M. Kataria, and E. Lampi. Do EPA Administrators Recommend Environmental Policies That Citizens Want? *Land Economics*, 87(1), 2011.

F. Carlsson, M.R. Mørkbak, and S.B. Olsen. The first time is the hardest: A test of ordering effects in choice experiments. *Journal of Choiuce Modelling*, 5(2), 2012.

T.L. Cherry, S. Kallbekken, and S. Kroll. The acceptability of efficiency-enhancing environmental taxes, subsidies and regulation: An experimental investigation. *Environmental Science & Policy*, 16:90–96, 2012.

T.L. Cherry, S. Kallbekken, and S. Kroll. Accepting market failure: Cultural worldviews and the opposition to corrective environmental policies. *Journal of Environmental Economics and Management*, 85:193–204, 2017.

S. Cohen and J. Liechty. Have it Your Way: Menu-based conjoint analysis helps marketers understand mass customization. *Marketing Research*, 19(3), 2007.

S. Colombo, A. Angus, J. Morris, D.J. Parsons, M. Brawn, K. Stacey, and N. Hanley. A comparison of citizen and "expert" preferences using an attribute-based approach to choice. *Ecological Economics*, 68(11), 2009.

M. Duhon, H. McDonald, and S. Kerr. Nitrogen Trading in Lake Taupo. 2015. Motu Working Paper 15-07. Motu Economic and Public Policy Research.

H. Eggert, M. Kataria, and E. Lampi. Difference in Preferences or in Preference Orderings? Comparing Choices of Environmental Bureaucrats, Recreational Anglers, and the Public, 2016. Working Papers in Economics no. 669.

K. Elofsson. Cost-effectiveness of the Baltic Sea Action Plan. *Marine Policy*, 34(5), 2010.

K. Elofsson. Swedish nutrient reduction policies: an evaluation of cost-effectiveness. *Regional Environmental Change*, 12(1), 2012.

K. Fisher-Vanden and S. Olmstead. Moving Pollution Trading from Ait to Water: Potential, Problems, and Prognosis. *Journal of Economic Perspectives*, 27(1), 2013.

K. Glenk and S. Colombo. How Sure Can You Be? A Framework for Considering Delivery Uncertainty in Benefit Assessments Based on Stated Preference Methods. *Journal of Agricultural Economics*, 62(1), 2011.

L.H. Goulder and I.W.H. Parry. Instrument Choice in Environmental Policy. *Review of Environmental Economics and Policy*, 2(2), 2008.

I. Gren, P. Jannke, and K. Elofsson. Cost-Effective Nutrient Reductions to the Baltic Sea. *Environmental and Resource Economics*, 10(4), 1997.

HELCOM. Eutrophication status of the Baltic Sea 2007-2011 — A concise thematic assessment, 2014. Baltic Sea Environment Proceedings No. 143.

D.A. Hensher, J. Rose, and W.H. Greene. The implications on willingness to pay of respondents ignoring specific attributes. *Transportation*, 32(3), 2005.

D.R. Heres, S. Kallbekken, and I. Galarraga. The Role of Budgetary Information in the Preference for Externality-Correcting Subsidies over Taxes: A Lab Experiment on Public Support. *Environmental and Resource Economics*, 66(1), 2017.

S. Hess, D.A. Hensher, and A. Daly. Not bored yet — Revisiting respondent fatigue in stated choice experiments. *Transportation Research Part A: Policy and Practice*, 46(3), 2012.

A.R. Hole. Fitting mixed logit models by using maximum simulated likelihood. *The Stata Journal*, 7(3), 2007.

R.D. Horan. Differences in Social and Public Risk Perceptions and Conflicting Impacts on Point/Nonpoint Trading Ratios. *American Journal of Agricultural Economics*, 83(4), 2001.

R.D. Horan and J.S. Shortle. When Two Wrongs Make a Right: Second-Best Point-Nonpoint Trading Ratios. *American Journal of Agricultural Economics*, 87(2), 2005.

M-F. Hung and D. Shaw. A trading-ratio system for trading water pollution discharge permits. *Journal of Environmental Economics and Management*, 49(1), 2005.

R.M. Johnson and B.K. Orme. How Many Questions Should You Ask In Choice-Based Conjoint Studies?, 1996. American Marketing Association Advanced Research Forum: Beaver Creek, CO, USA.

R.J. Johnston and J.M. Duke. Willingness to Pay for Agricultural Land Preservation and Policy Process Attributes: Does the Method Matter? *American Journal of Agricultural Economics*, 89(4), 2007.

S. Kallbekken, S. Kroll, and T. Cherry. Do you not like Pigou, or do you not understand him? Tax aversion and revenue recycling in the lab. *Journal of Environmental Economics and Management*, 62:53–64, 2011.

I. Krinsky and A.L. Robb. On approximating the statistical properties of elasticities. *Review of Economics and Statistics*, 68(4), 1985.

A.J. Krupnick, W.E. Oates, and E. Van de Verg. On Marketable Air-Pollution Permits: The Case for a System of Pollution Offsets. *Journal of Environmental Economics and Management*, 10(3), 1983.

T. Lundhede, J. Bredahl Jacobsen, N. Hanley, N. Strange, and B. Jellesmark Thorsen. Incorporating Outcome Uncertainty and Prior Outcome Beliefs in Stated Preferences. *Land Economics*, 91(2), 2015.

A.S. Malik, D. Letson, and S.R. Crutchfield. Point/Nonpoint Source Trading of Pollution Abatement: Choosing the Right Trading Ratio. *American Journal of Agricultural Economics*, 75(4), 1993.

W.D. Montgomery. Markets in Licenses and Efficient Pollution Control Programs. *Journal of Economic Theory*, 5(3), 1972.

T.C. Nguyen, J. Robinson, J.A. Whitty, S. Kaneko, and T.C. Nguyen. Attribute non-attendance in discrete choice experiments: A case study in a developing country. *Economic Analysis and Policy*, 47:22–33, 2015.

A. Nordén, J. Coria, A.M. Jönsson, F. Lagergren, and V. Lehsten. Divergence in stakeholders' preferences: Evidence from a choice experiment on forest landscapes preferences in Sweden. *Ecological Economics*, 132, 2017.

North Baltic Water Authority. Åtgärdsprogram 2016-2021. Åtgärder riktade till myndigheter och kommuner samt konsekvensanalys, 2016. Report. In Swedish.

D. Revelt and K. Train. Mixed logit with repeated choices: households' choices of appliance efficiency level. *Review of Economics and Statistics*, 80(4), 1998.

D. Revelt and K. Train. Customer-Specific Taste Parameters and Mixed Logit: Households' Choice of Electricity Supplier, 2000. Working Paper No. E00-274, Department of Economics, University of California at Berkeley.

A.A. Rogers. Public and Expert Preference Divergence: Evidence from a Choice Experiment of Marine Reserves in Australia. *Land Economics*, 89(2), 2013.

R. Scarpa, M. Thiene, and D.A. Hensher. Monitoring Choice Task Attribute Attendance in Nonmarket Valuation of Multiple Park Management Services: Does It Matter? *Land Economics*, 86(4), 2010.

Z. Sándor and M. Wedel. Heterogeneous Conjoint Choice Designs. *Journal of Marketing Research*, 42(2), 2005.

South Baltic Water Authority. Samhällsekonomisk konsekvensanalys av förslag på åtgärdsprogram för Södra Östersjöns vattendistrikt, 2014a. Report. In Swedish.

South Baltic Water Authority. Förslag på åtgärdsprogram 2015-2021 för Södra Östersjöns vattendistrikt, 2014b. Report. In Swedish.

J. Swait and W. Adamowicz. The Influence of Task Complexity on Consumer Choice: A Latent Class Model of Decision Strategy Switching. *Journal of Consumer Research*, 28 (1), 2001.

J. Swait and J. Louviere. The Role of the Scale Parameter in the Estimation and Comparison of Multinomial Logit Models. *Journal of Marketing Research*, 30(3), 1993.

Swedish Agency for Marine and Water Management. Näringsbelastningen på Östersjön och Västerhavet 2014, 2016. Report 2016:12. In Swedish.

Swedish Environmental Protection Agency. Förslag till avgiftssystem för kväve och fosfor, 2008. Report 5913. In Swedish.

Swedish Environmental Protection Agency. Vidareutveckling av förslag till avgiftssystem för kväve och fosfor, 2010. Report 6345. In Swedish.

Swedish Environmental Protection Agency. Styrmedel för ökad rening från kommunala reningsverk, 2012. Report 6521. In Swedish.

Swedish Environmental Protection Agency. Wastewater treatment in Sweden, 2014. Report.

K.E. Train. *Discrete Choice Methods with Simulation.* Cambridge University Press, Cambridge, 2nd edition, 2009.