

available at www.sciencedirect.comwww.elsevier.com/locate/ecolecon

ANALYSIS

Theoretical incentive properties of contingent valuation questions: Do they matter in the field?

Felix Schläpfer^{a,*}, Ingo Bräuer^b

^aInstitut für Umweltwissenschaften, Universität Zürich, Winterthurerstrasse 190, 8057 Zürich, Switzerland

^bEconomics Department, UFZ Center for Environmental Research Leipzig–Halle, Leipzig, Germany

ARTICLE INFO

Article history:

Received 13 October 2005

Received in revised form

15 June 2006

Accepted 6 July 2006

Available online 17 August 2006

Keywords:

Contingent valuation

Beliefs

Incentives

Majority rule

Public goods

Stated preferences

ABSTRACT

Green et al. (1998) [Green, D., Jacowitz, K.E., Kahneman, D., McFadden, D., 1998. Referendum contingent valuation, anchoring, and willingness to pay for public goods. *Resource and Energy Economics* 20 (2), 85–116] show theoretically that stated preference questions about public services can be framed in such ways that if the subjects accept the frame the payoff-maximizing behavior will be to answer truthfully. One key element of such a theoretically incentive-compatible framing is that the (hypothetical) decision rule specified in the survey instrument is understood to be a majority rule rather than the efficiency rule typically used in cost-benefit analysis. We conducted field experiments in Germany and Switzerland to test if a referendum framing as suggested by Green et al. is effective in reducing strategic misrepresentation in a contingent valuation setting. We did not find the expected effects of the framing treatments on stated willingness to pay or on individuals' (stated) beliefs about the social choice context. The results do not support hopes that a theoretically incentive compatible framing could be purposefully used to invoke the specific beliefs about the linkage between responses and policy implementation that would make stated preference questions incentive compatible.

© 2006 Elsevier B.V. All rights reserved.

1. Introduction

Unresolved incentive compatibility issues are an important weakness in the theoretical underpinnings of the contingent valuation method. In 1993, the NOAA Panel on Contingent Valuation recommended the use of a hypothetical referendum question format in surveys designed to estimate losses associated with natural resource damage (Arrow et al., 1993). A hypothetical referendum question asks, for example, "If your personal tax cost for program A were \$D per year, would you vote for the program?" The panel recommended the hypothetical referendum format, because, "[...] as far as strategic reasons go, a respondent who would not be willing to pay D dollars has no reason to answer 'Yes', and a respondent who would be

willing to pay D dollars has no reason to answer 'No'" (p. 4606). Contingent valuation (CV) questions posed in a referendum format with a tax as the (hypothetical) payment mechanism are widely perceived as incentive compatible by CV researchers today (Hoehn and Randall, 1987; Mitchell and Carson, 1989; Arrow et al., 1993; Hanemann, 1994; Carson et al., 2000; Taylor et al., 2001; Champ et al., 2002; Horowitz and McConnell, 2002). Only few authors have noted the panel's mistake (e.g. Diamond and Hausman, 1994; Fisher, 1996; Green et al., 1998; Posovac, 1998; Burton et al., 2003). Fisher (1996), for instance, notes that "[...] the case for closed-ended CV responses being free of strategic bias has not been made either in theory or by empirical findings" (p.26). Green et al. (1998) find "[...] that patterns of response to open-ended and referendum questions that are

* Corresponding author. Tel.: +411 635 4747; fax: +411 635 5711.

E-mail address: schlaepf@uwinst.unizh.ch (F. Schläpfer).

often taken as evidence of incentive-induced misrepresentation are in fact present in situations where there are no economic incentives” (p. 95). Unfortunately, the CV studies comparing responses to open-ended vs. dichotomous choice questions do not directly address the incentive compatibility issue, since potential effects of the question format on incentives cannot be distinguished from other potential effects of the question format (e.g. Brown et al., 1998).

Previous designed experiments on public good provision and incentives have been conducted in laboratory settings (Andreoni, 1988; Andreoni and McGuire, 1993; Cummings et al., 1995; Burton et al., 2003). Experiments have demonstrated that in early rounds, subjects do not free-ride to the extent predicted by theory. In later rounds, free-riding becomes prevalent. These laboratory studies used private goods or (group) donations and, therefore, did not need to worry about the issues that come into play when stated preferences for collectively provided public services are concerned. The incentive properties in private good choices and also in group donation decisions (Cummings et al., 1997) do not depend on the complex linkages between the responses, outcomes and division of costs which complicate individual decision making in a social choice context. In particular, stated preference elicitation in a social choice context must take into account that respondents may not always accept (believe) the social choice framework as posited in a survey.

Green et al. (1998) find that – if a CV experiment can be set up in such ways that an economically rational subject believes there is a positive probability of being decisive – it is possible to frame the elicitation such that if the subject accepts the frame the pay-off maximizing response will be to answer truthfully. However, the authors emphasize that subjects may hold beliefs about the social choice context that induce misrepresentation, because of a particular framing or in spite of it. The paper by Green et al. (1998) is to our knowledge the first to fully identify the elements of the posited social choice framework – or more precisely, the respondent *beliefs* concerning these elements – which jointly determine the incentive properties of preference elicitation for collectively provided public services¹. These elements concern the *decisiveness of the response*, the *decision rule*, the *payment vehicle* and the *uncertainty about actual costs*, as detailed in the following section. The aim of the present study is to test the empirical relevance of Green et al.’s theoretical results regarding the role of these elements for the incentive properties of preference elicitation in a field setting. Our strategy is to design a field experiment in which the important elements of the survey frame identified by Green et al. (1998) can be manipulated individually. We then

test how these manipulations affect both stated beliefs and willingness-to-pay (WTP) responses.

We asked the following specific questions: (1) How do specifications of the survey frame concerning (i) majority vs. mean rule, (ii) cost uncertainty, and (iii) decisiveness of the survey affect (a) beliefs about the social choice context and (b) stated WTP. Since cultural differences may affect the attitudes to valuation surveys (e.g. Loomis et al., 2002) we conducted these tests in two locations with different political institutions and cultures.

The remaining paper proceeds as follows: Section 2 presents the conceptual framework. Sections 3 and 4 contain the method and results. Sections 5 and 6 present the discussion and conclusions.

2. Conceptual framework

Whether CV respondents think and act strategically is an empirical question (Cummings et al., 1995, 1997). There are several reasons why strategic misrepresentation could be quantitatively unimportant in standard referendum CV surveys applied to collectively provided public goods (Smith, 1979). First, subjects confronted with referendum-format questions may not perceive their strategic opportunities. Second, subjects may be influenced by norms that prevent them from misrepresentation (Green et al., 1998, p. 91). Third, individuals may simply believe that their response is inconsequential. However, based on theoretical grounds, more general claims that the CV referendum protocol is incentive compatible are misleading (Green et al., 1998; Riera, 2003; Burton et al., 2003).

In the following, we will call a social choice framework posited in the survey protocol (hereafter for simplicity also called the “frame”) “theoretically incentive compatible” (TIC) if the payoff-maximizing response of a rational respondent who accepts this frame is to answer truthfully.² Following Green et al. (1998, p. 92) a TIC frame would comprise the following elements:

- (1) *decisiveness*: the response affects the probability of implementation;
- (2) *decision rule*: the probability of implementation is proportional to the survey plurality for the policy, given a pre-specified distribution of the costs;
- (3) *payment vehicle*: the payment vehicle is ‘decoupled’ implying that the costs will be distributed across tax payers by a formula that does not depend on the response;
- (4) *cost uncertainty*: the actual individual costs burden if the policy is implemented could be above or below their true WTP for the policy.

Concerning element (2), Baron (1996, p. 153) suggests how this can be implemented in a survey: “Specifically, we can tell [the respondents] that their WTP will be compared to their fair share of the cost of the good. If more than half of the

¹ Riera (2003), in turn, is the first to present a systematic theoretical treatment of incentive properties of various stated preference methods and question formats, although (as Riera admits) the study does not consider important issues concerning respondents’ perception of the relevant social choice framework. Riera’s perhaps most important theoretical result is that, if a set of general assumptions about the social choice framework hold, the median-rule open-ended CV format is incentive compatible. This result is consistent with the conclusions of Green et al. (1998).

² One could also use the term theoretically demand revealing“. We use TIC to remain close to the terminology used by Green et al.

respondents are willing to pay at least their fair share, then the good will be provided, and otherwise it will not". Concerning element (4), If this is not the case, then the set of payoff-maximizing responses in the majority-rule format may comprise an entire range of bid amounts (Green et al., 1998, p. 92). Although there are then no strategic gains to be made, the question is not incentive compatible, since the respondent can state any WTP that lies within the set of payoff-maximizing responses.

However, incentive compatibility of preference elicitation in surveys cannot be defined without taking into account whether the respondents accept the posited choice framework. The additional requirement for a preference elicitation to be *incentive compatible* (IC) is that the respondent accepts this frame. In other words, to accomplish an incentive compatible elicitation, the surveyor must persuade the (rational) respondents that their responses will be counted as a 'Yes' if they exceed the individual share of the costs given a pre-specified (e.g. tax) schedule, that the probability of the implementation is linked to survey plurality (and not mean WTP) and that the respondents do not know if their share of the actual costs of the policy are above or below their WTP. Green et al. (1998) emphasize that a CV experiment can be set up to be what we call TIC *regardless* of the (open-ended vs. closed ended) question format.

For experimental purposes, the specific merit of the TIC property is that it is a property of the survey protocol alone: A survey protocol is TIC if it contains explicit specifications to prompt beliefs (1)–(4) above. These specifications can be experimentally manipulated and their effects on WTP response assessed. Currently standard referendum CV survey protocols are not TIC. Indeed, to our knowledge, not a single CV study has used a TIC frame to date. Since the TIC frame *can* support the beliefs required to make an elicitation incentive compatible, an important empirical issue is therefore whether the TIC framing actually affects these beliefs and is successful in reducing strategic misrepresentation. If responses turn out to be sensitive to the theoretical incentive properties of the preference question, we must conclude that incentives do matter. Accordingly, the null hypothesis we test is that the TIC framing does not affect responses compared with standard non-TIC frames.

In the present study, the beliefs concerning a decoupled payment vehicle (3) are not critical for strategic misrepresentation because, with the proposed public policy, a non-decoupled payment vehicle is hardly conceivable. In contrast, each of the specifications made to prompt beliefs (1), (2) and (4) may be hypothesized to influence the respondents' perception of the linkages between response and policy outcome. We thus devise an experiment in which we manipulate specifically the information to prompt these beliefs (1), (2) and (4). We then examine the effects of these manipulations on stated beliefs and on strategic misrepresentation of WTP. In order to examine strategic misrepresentation in a public policy setting where "true" preferences are unknown, we pursue the following simple strategy. We conduct a survey to evaluate a nature conservation project among a sample of outdoor recreationists. We thus obtain a sample of respondents with a generally great demand for nature conservation policies relative to the entire population. Strategic overbidding should be

particularly frequent among our respondent groups. We argue that any decrease of stated WTP due to a TIC framing could thus be quite plausibly interpreted as evidence for decreased strategic overbidding.

3. Methods

3.1. Experimental design

We conducted two field surveys with outdoor recreationists, one in Germany and one in Switzerland. The valuation scenarios were structured around proposed species re-introduction programs. We used experimental design techniques to investigate individual and interactive effects of different pieces of information in the survey protocol that specify the relevant social choice framework. In a short section explaining the social choice context (preceding the valuation question) we factorially crossed three two-level treatment factors. The treatments were different versions of the information to prompt beliefs (1), (2) and (4): the decision rule (DR; majority vs. mean rule), the decisiveness of the survey (DV; emphasis vs. no special emphasis), and the extent of cost uncertainty of the project (CU; high vs. low) (Table 1). With each information treatment factor, we compare a non-TIC control treatment (coded 0) with a TIC treatment (coded 1; see Table 2). Apart from this section, the eight (2³) questionnaire versions (treatment combinations) were identical.

For a credible implementation of this experiment, the good to be valued had to fulfil a number of criteria: It had to be realistically provided by a public body, amenable to description in a written survey, of appropriate size to be funded by the study population, credibly provided through a tax mechanism, and of a nature where it may be credibly claimed that costs are unknown and yet to be determined. A hypothetical project to re-introduce the Eurasian otter (*Lutra lutra* L.) respectively in the Harz Region of Lower Saxony, Germany, or in the Canton of Zurich, Switzerland, fulfilled the above criteria. Otters are extinct in the Harz Region and in the canton of Zurich (and all over Switzerland). Since other attractive animals such as the lynx had been re-introduced in both areas in recent years, the otter project appeared to be a credible proposition. In fact, there had been discussions about such programs among NGOs in both study locations.

Table 1 – Experimental design

Decision rule (DR)	Cost uncertainty (CU)	Decisiveness (DV)	Harz	Zurich
majority rule	high	emphasized	41	58
		no emphasis	40	60
	low	emphasized	45	60
mean rule	high	no emphasis	34	55
		emphasized	36	58
	low	no emphasis	40	58
		emphasized	39	54
		no emphasis	38	56

Note: Numbers in treatment cells are sample sizes (usable responses).

Table 2 – Levels (wording) of the three information treatment factors “survey decisiveness” (DV), “decision rule” (DR) and “cost uncertainty” (CU)

Factor	Control	TIC_treatment
CU	cost uncertainty low (CU=0): “The costs of the measures for the re-introduction are presently only approximately known . Depending on the results of precursory assessments they could be higher or lower. The project would increase the expenditures of [lower Saxony/the canton of Zurich] for nature conservation by an expected 10 to 20 percent during one year.”	cost uncertainty high (CU=1): “The costs of the measures for the re-introduction are not yet known . Depending on the results of precursory assessments they could be relatively low or quite high. The project would increase the expenditures of [lower Saxony/the canton of Zurich] for nature conservation by between between 1 and 50 percent during one year.”
DR	mean rule (DR=0): “The results of this survey will be used to estimate if the people’s joint willingness to pay is higher or lower than the costs (which are yet to be determined).”	majority rule (DR=1): “The results of this survey will only be used to assess, if a majority of the citizens would approve of the project . Every person has 1 vote. You cannot overproportionately influence the result by overstating or understating your willingness to pay”
DV	no emphasis on decisiveness (DV=0): [no information]	emphasis on decisiveness (DV=1): “The results of this survey can have some influence on whether the re-introduction of the otter will be further pursued and realized by the public institutions.”

Note: Words printed in bold only in this table (not in the original questionnaire).

3.2. Field sites

As explained above, our goal was to obtain respondent samples that included a large proportion of individuals with a particularly supportive attitude towards our hypothetical project. This goal was pursued by conducting the survey among outdoor recreationists at field sites nearby nature reserves. Each survey was conducted in two field sites which were selected to sample slightly different segments of the ‘population’ of outdoor recreationists.

The Harz National Park covers about 15'000 ha. There are several creeks and river systems suitable for a reintroduction in the lower parts of the mountain range, although the Region has also a high touristic potential. The field site ‘Oderbruck’ was a parking lot in the center of the National Park. It is a starting point for many walks in the area. The field site ‘Torfhaus’ was a parking lot on top of the Torfhaus Mountain surrounded by various restaurants and small kiosks. The mountain is one of the main tourist attractions in the Harz. It is famous for its views to the Brocken which is the highest elevation of the Harz.

The canton of Zurich has several smaller nature reserves in which a re-introduction of the otter could be feasible, including the Sihlwald near Zurich’s Uetliberg mountain and the protected areas of the Greifensee lake. The field site ‘Uetliberg’ was the terrace (with a restaurant and playground) on top of the Uetliberg. The place can be reached by train plus a ten minute walk, and is a starting point for walks. The field site ‘Greifensee’ was a passenger-boat landing near the city of Uster. The shoreline is a popular outdoor recreation and hiking area.

The rationale for conducting the survey in both Germany and Switzerland is that institutional differences may influence the perception of the framing. In particular, Swiss respondents are habituated to voting about regional (cantonal) policy issues. On one hand, this might increase the plausibility of the voting scenario. On the other hand, given that constitutional provisions allow actual referendum decisions, the Swiss citizens might be less inclined to perceive surveys as a valid and credible instrument to inform public policy. While it is difficult

to predict how the institutional and cultural differences might affect the perception and effects of alternative frames, conducting the study both in Germany and in Switzerland would provide an indication of whether any observed effects may be sensitive to the cultural and institutional setting.

3.3. Survey procedures

Recreationists aged 18 and over, were intercepted by field assistants (students and PhD students carrying a tag with name and university affiliation) who gave a short description of the survey. Volunteers then immediately completed a written questionnaire at standing-tables provided for this purpose (in the Harz study, writing pads were additionally offered). On the tables, a small brochure with additional information about the biology and ecology of the otter was available for interested respondents. The interviewers were present at all times and ensured that questionnaires were completed independently by each respondent. Ballot boxes were put up to ensure anonymity of the completed questionnaires. As a way to express our appreciation and also to increase participation we offered a beverage coupon from a nearby restaurant or kiosk to each volunteer completing the questionnaire. Social norms and procedural benefits have been found to play an important role in motivating carefully considered survey responses (Hidano et al., 2005a, b). Specifically, we expected that the explicit university affiliation of the surveyors should activate feelings of duty to carefully complete the survey. With the beverage coupons, we aimed at supporting a generally favourable perception of the survey process.

The Zurich surveys were conducted in a period of nice weather on the 10th to 13th (Greifensee) and on the 14th to 17th (Uetliberg) of October 2003 on weekday afternoons and on a Saturday. The Harz surveys were completed on a single weekend (18th and 19th of October) in both field sites simultaneously. A second sampling on the following weekend had to be terminated unsuccessfully because of bad weather. Increasingly low temperatures thereafter prevented us from reaching our target of 220 responses for each field site of the Harz experiment.

3.4. Questionnaire

The survey included four sections: Section one provided background information about the species, its local extinction history, and its habitat requirements with regard to a possible re-introduction. In Section 2 the respondents were first asked about their attitude concerning an increase or decrease of the state (cantonal) expenditures for nature conservation:

Should [the canton of Zurich] [Lower Saxony] spend more or less tax money for nature protection?

Response options were “more”, “somewhat more”, “the same amount as today”, “somewhat less”, “less” and “don’t know”. Furthermore, they were asked whether they would welcome a small local population of otters if these could be re-introduced without any additional nature protection measures.

The following text, comprising the open-ended WTP question, was:

In view of the costs of re-introducing the otter we now ask you if you would be prepared to back this nature conservation project with tax money.³ Consider that this money would then be unavailable for other purposes! Please consider also the following items:

- The project would be realized through the [state of Lower Saxony (State Office for Ecology)][Canton of Zurich (Office of Landscape and Nature)] and would have a direct impact on your personal tax bill.
- [treatment factor CU; Table 2, first row]
- [treatment factor DR; Table 2, second row]
- [treatment factor DV; Table 2, third row]

Up to which amount of personal additional taxes (state income and property taxes) would you be prepared to vote for a successful re-introduction program [in a cantonal referendum]? (Note: if you are married and receive a joint tax bill, divide the amount by two.) ... [EURO] [Francs]. Please think and respond carefully and realistically!

Section 3 asked five questions about socio-economic characteristics (gender, year of birth, education, type of job, and income. Finally, Section 4 asked two “debriefing” questions about the respondents’ beliefs concerning the social choice context. The questions about beliefs allow to examine whether the information treatments were effective in manipulating those beliefs which, based on the theoretical results of Green et al. (1998), would be expected to determine the incentive compatibility of the elicitation. The questions were placed at the end of the

questionnaire to minimise potential effects on the WTP responses (see Pouta, 2004). In a first question, the respondents were asked about their perception of the *decisiveness* of the survey (see belief 1 and the treatment factor DV above):

How strongly do you agree with the following statement: ‘The results of this survey can have an impact on public policy’?

Answering options were from 1 for “fully agree” to 5 for “do not agree at all”. The second question concerned the issue of cost *uncertainty* and was devised to assess to what extent respondents were unsure whether the actual costs would be lower or higher than their WTP (see belief 4 above and the treatment factor CU in Table 2):

How strongly do you agree with the following statement: ‘I expect that it will be possible to re-introduce the otter at agreeable costs’?

Answering options were again from 1 for “fully agree” to 5 for “do not agree at all”. This last section also provided space for comments and offered a contact address for further information.

4. Results

4.1. Descriptive statistics

The Zurich and the Harz surveys yielded a total of 772 usable responses. Means of socioeconomic characteristics and response variables for the two survey locations are listed in Table 3. The responses to the attitude questions suggested that the Zurich respondents were on average more supportive of increased expenditures for conservation than the Harz respondents, and they also stated higher mean WTP for the proposed project (Table 3). The observed mean attitude value of about 4 suggests that the outdoor recreationist sample was strongly biased towards individuals with a generally positive attitude towards nature conservation policies, as required by our empirical strategy (see end of Section 2). The proportion of “zero WTP” responses was higher in the Harz sample.

In the following, we first report the binary probit models of treatment, demographic and attitude effects on stated beliefs about the social choice context. We then present the treatment, demographic and attitude effects on stated WTP. The data from the two field sites (subsamples) of each location were pooled. A dummy variable for the subsamples did not affect the pattern of effects and was therefore dropped in the models presented below.

4.2. Effects on stated beliefs

The effects of the questionnaire treatments (majority vs. mean decision rule (DR), high vs. low cost uncertainty (CU), and emphasis vs. no emphasis of decisiveness (DV)) and of the covariables on stated beliefs about the social choice context were analyzed using binary probit models as implemented in the software package LIMDEP 7.0 (Greene, 1998). We used a coding of the dependent variables CU_belief and DV_belief as defined in Table 3 (last two rows).

³ If an unequal distribution of costs, such as in an income tax schedule, is specified, then the majority-rule open-ended format is not incentive compatible unless an additional assumption is made. This assumption is, in the case of an income tax vehicle, that respondents truthfully report their taxable income. In the present experiment, an income tax vehicle was the only credible option. Theoretical incentive compatibility of our majority-rule format therefore requires the assumption that respondents did not misreport their income.

Table 3 – Means of socioeconomic characteristics and response variables for each sub-sample of the two survey locations^a

Variable name	Description	Location			
		Harz		Zurich	
		Oderbruck	Torfhaus	Greifensee	Uetliberg
nobs	Sample size	159	154	221	238
income	Categorical variable: Harz: 1=0–1000; 2=1001–2000; ...6=>5000 Euros. Zurich: 1=0–2000; 2=2001–4000; ...6=>10000 SFR	2.30 (1.14) 141	2.35 (1.20) 139	3.14 (1.32) 205	2.85 (1.42) 215
female	Binary variable: 1=female; 0=male	0.42 (0.50) 159	0.39 (0.49) 153	0.51 (0.50) 220	0.61 (0.49) 238
age	Age in years	45.94 (12.80) 157	46.44 (13.21) 151	47.23 (16.15) 217	37.61 (18.15) 238
education	Highest educational degree: ranging from 1=mandatory schooling to 5=university degree	3.18 (1.48) 154	2.96 (1.53) 148	2.97 (1.36) 215	2.95 (1.29) 235
attitude	Opinion about public expenditures for nature conservation (see methods): ranging from 1=should spend less to 5=should spend more	3.86 (0.82) 146	3.86 (0.95) 126	4.07 (0.93) 214	4.12 (0.83) 223
wtp	Stated WTP in, respectively, Euro or SFR	25.92 (37.32) 133	30.70 (58.96) 124	132.23 (433.25) 193	93.29 (195.68) 187
wtp0	Stated WTP in, respectively, Euro or SFR: item non-response recoded to zero (see methods)	21.68 (35.44) 159	24.72 (54.25) 154	115.48 (407.14) 221	73.30 (177.54) 238
CU_belief	Stated belief about cost uncertainty: 1 for response options 2–4; 0 otherwise ^b	0.58 (0.50) 159	0.56 (0.50) 154	0.47 (0.50) 221	0.49 (0.50) 238
DV_belief	Stated belief about decisiveness: 1 for response option 1 or 2; 0 otherwise ^b	0.55 (0.50) 159	0.49 (0.50) 154	0.57 (0.50) 221	0.59 (0.49) 238

^a Standard deviations are in parentheses and number of observations in italics.

^b See text (Section 3.4).

Table 4 – Binary probit models of the respondents' stated beliefs about the social choice context

Variable	Dependent variable in the model							
	CU_belief (Zurich)		DV_belief (Zurich)		CU_belief (Harz)		DV_belief (Harz)	
	Coeff.	t-ratio	Coeff.	t-ratio	Coeff.	t-ratio	Coeff.	t-ratio
Constant	-0.130	-0.403	0.370	1.133	0.486	1.154	-0.520	-1.219
DECRULE (DR)	0.092	0.375	-0.198	-0.803	-0.668	-2.064**	0.392	1.208
COSTUNC (CU)	0.062	0.249	-0.073	-0.292	-0.333	-1.082	-0.212	-0.685
DECISIVE (DV)	-0.405	-1.546	-0.351	-1.358	-0.133	-0.418	0.154	0.486
CU × DR	-0.205	-0.599	0.249	0.718	0.781	1.767*	0.043	0.097
CU × DV	0.284	0.789	0.460	1.281	0.307	0.682	0.224	0.495
DR × DV	0.213	0.592	0.617	1.718*	0.562	1.266	-0.380	-0.852
CU × DR × DV	0.278	0.557	-0.789	-1.575	-0.910	-1.463	-0.036	-0.057
age	0.002	0.614	0.005	1.358	-0.005	-0.871	0.013	1.974**
gender	-0.092	-0.704	-0.050	-0.381	-0.027	-0.165	0.102	0.612
education	0.045	0.861	-0.059	-1.120	0.116	2.037**	-0.156	-2.701***
income	-0.034	-0.674	-0.058	-1.155	-0.095	-1.253	0.171	2.181**
Log-likelihood		-281.7		-278.9		-183.3		-179.2
Log-l. restricted		-286.7		-283.7		-189.0		-188.5
χ^2		9.8		9.5		11.4		18.5
Sig. level		0.55		0.57		0.41		0.07
N		414		414		274		274

*** = significant at $p < 0.01$;

** = significant at $p < 0.05$;

* = significant at $p < 0.1$.

Table 5 – Means and standard deviations (SD) of stated willingness to pay (WTP) for the 8 treatment combinations (questionnaire versions)

Decision rule (DR)	Cost uncertainty (CU)	Decisiveness (DV)	Zurich (values in SFR)				Harz (values in Euro)			
			WTP		WTP (missing=0)		WTP		WTP (missing=0)	
			Mean	SD	Mean	SD	Mean	SD	Mean	SD
majority	high	emphasized	146	312	123	291	36	95	25	80
		no emphasis	81	143	70	136	17	19	15	19
	low	emphasized	131	215	105	199	31	54	27	52
mean	high	no emphasis	63	81	51	77	21	26	20	26
		emphasized	63	58	53	58	19	28	13	25
	low	no emphasis	91	224	81	213	27	32	22	31
		emphasized	75	65	61	65	36	47	29	45
		no emphasis	101	242	73	211	36	51	34	50

The specifications of the social choice context as implemented in our treatments did not have any consistent effects on the respondents' stated beliefs about this social choice context in the Zurich or in the Harz sample (Table 4). The statement concerning a high (*vs.* low) uncertainty about the actual costs of the project (treatment factor CU) did not increase perceived cost uncertainty (CU_belief). The statement concerning the decisiveness of the survey results (treatment factor DV) did not affect the stated beliefs about the survey decisiveness (DV_belief).

In the Harz sample we found some significant associations of the respondents' reported beliefs about the social choice context with their socioeconomic characteristics. Education was positively related to uncertainty about whether the project could be implemented at 'agreeable' costs and negatively to the perceived "decisiveness" of the survey results (Table 4). These effects suggest that the more educated individuals may be less inclined to accept a social choice context as posited in the survey. We did not have any clear expectations regarding the effects of age and income on the perceived decisiveness, which were positive in the Harz sample.

4.3. Effects on stated willingness to pay

Mean values of stated WTP for the eight versions of the survey questionnaire and the two locations are listed in Table 5. The theoretically incentive compatible version of the questionnaire (DR=majority, CU=high, DV=emphasized; see Table 4) yielded the highest mean WTP in the Zurich survey and one of the highest in the Harz survey, although standard deviations were also high with this questionnaire version.

The effects of the framing treatments on stated WTP were analysed using ordinary least squares regression. Corresponding with our aim to analyze strategic overbidding, we excluded only very extreme outliers among the stated WTP values in the WTP regressions. In the Harz sample there were no bids above 500 Euro, and in the Zurich sample we deleted two observations above 2000 SFR (2500 and 5000). Following the standard procedure, observations with missing values in any of the variables of the model were excluded from the analysis, which resulted in the sample sizes given in the regression tables. Several alternative definitions of the WTP variable were applied: WTP as given by the respondents (with non-

Table 6 – OLS regression models of the WTP responses (Zurich dataset)

Variable	Dependent variable in the model							
	WTP		WTP (missing=0)		log WTP (missing=0)		log WTP (only >0)	
	Coeff.	t-ratio	Coeff.	t-ratio	Coeff.	t-ratio	Coeff.	t-ratio
Constant	-27.677	-0.327	-67.962	-0.994	-1.151	-1.740*	1.485	2.585
DECRULE (DR)	-51.729	-1.128	-18.145	-0.486	0.669	1.853*	0.000	0.001
COSTUNC (CU)	-10.091	-0.225	24.847	0.656	0.815	2.224***	0.276	0.920
DECISIVE (DV)	-32.737	-0.682	-7.405	-0.188	0.600	1.576	0.384	1.218
CU×DR	21.581	0.354	-16.192	-0.310	-0.893	-1.769*	-0.293	-0.717
CU×DV	-12.837	-0.202	-37.019	-0.679	-0.469	-0.889	-0.518	-1.230
DR×DV	110.384	1.720*	75.452	1.396	-0.323	-0.617	0.068	0.161
CU×DR×DV	-0.339	-0.003	31.630	0.421	0.295	0.406	0.241	0.417
age	2.321	3.429***	1.748	3.052***	0.005	0.819	0.016	3.456***
gender	3.273	0.144	1.830	0.092	0.037	0.195	-0.005	-0.030
education	-13.524	-1.551	-9.502	-1.221	0.064	0.850	-0.059	-1.022
income	11.107	1.291*	10.552	1.391**	0.187	2.542**	0.181	3.137***
attitude	10.978	0.849	14.657	1.340	0.657	6.209***	0.268	2.925***
R ²		0.078		0.063		0.132		0.118
N		336		393		393		318

For significance levels see Table 4.

Table 7 – OLS Regression models of the WTP responses (Harz dataset)

Variable	Dependent variable in the model							
	WTP		WTP (missing=0)		logWTP (missing=0)		logWTP (only >0)	
	Coeff.	t-ratio	Coeff.	t-ratio	Coeff.	t-ratio	Coeff.	t-ratio
Constant	-40.168	-1.490	-28.581	-1.230	0.344	0.451	1.153	1.340
DECRULE (DR)	-13.926	-0.952	-13.948	-1.028	-0.257	-0.577	-0.188	-0.410
COSTUNC (CU)	-9.147	-0.610	-13.898	-1.041	-0.466	-1.064	-0.163	-0.357
DECISIVE (DV)	1.419	0.094	-3.402	-0.248	-0.258	-0.573	-0.021	-0.045
CU × DR	2.473	0.119	1.379	0.074	-0.157	-0.258	-0.119	-0.185
CU × DV	-10.901	-0.497	-9.447	-0.496	-0.655	-1.047	0.066	0.094
DR × DV	5.068	0.251	6.094	0.332	-0.115	-0.191	0.280	0.436
CU × DR × DV	18.168	0.615	16.715	0.648	0.565	0.668	-0.432	-0.460
age	0.377	1.206	0.143	0.560	-0.013	-1.526	0.020	1.953*
gender	5.238	0.671	4.111	0.606	0.179	0.801	-0.038	-0.156
education	-0.400	-0.145	0.709	0.304	0.076	0.997	-0.101	-1.173
income	9.357	2.575**	9.117	2.967***	0.265	2.625***	0.150	1.270
attitude	10.018	2.303**	9.459	2.547**	0.531	4.350***	0.242	1.670*
R ²		0.089		0.091		0.149		0.085
N		205		238		238		165

For significance levels see Table 4.

responses treated as “missings”); WTP with “missings” recoded to zeros (see e.g. Carson et al., 1998); log-transformed WTP (after recoding zeros to 1’s); and finally log-transformed positive (non-zero) WTP values (to retain only those WTP responses which were potentially inflated due to strategic incentives). Log-transformation substantially improved the normality of the WTP distribution.

Using these definitions of the dependent variable, we did not find a consistent pattern of treatment effects on stated willingness to pay (Tables 6 and 7). In other words, statements concerning the decision rule (DR), the degree of cost uncertainty (CU) and the “decisiveness” of the survey responses (DV) did not consistently affect the ways in which respondents answered the willingness to pay question. Only in one model (third model in Table 6) did we find some evidence for a (positive) main effect of the majority decision rule and of increased cost uncertainty on stated WTP, while the interaction of the two factors was negative. However, a similar effect is not found in any of the other WTP models of the Zurich dataset (Table 6). Some of the covariates for demographic characteristics and attitudes, however, were quite consistently related to the WTP responses. In particular, the coefficients on the covariables ‘income’ and ‘attitude’ always had the expected signs and were significant in most models (Tables 6 and 7). Hence, the survey passed the – admittedly weak – internal validity criterion of a significant income effect which is frequently reported in contingent valuation surveys.

We also estimated models in which we fitted dummy variables for each questionnaire version, except for the theoretically incentive compatible (TIC) version (DR=majority, CU=high, DV=emphasized; see Table 4) which was the base case, to contrast each of these treatment combinations with the TIC version. In the Zurich dataset, the version 2, 3 and 6 yielded significantly lower stated WTP values both in models with and without the covariables. In the Harz dataset, however, none of the questionnaire versions yielded significantly different stated WTP than the TIC version, although the covariables for income and attitude were significant.

5. Discussion

From the theoretical perspective of Green et al. (1998), the incentive compatibility of stated preference elicitation about public goods depends on whether the survey frame is successful in supporting a specific set of beliefs about the social choice context. One might therefore expect to find effects of information treatments aimed at supporting these beliefs on stated WTP for public goods. Contrary to this expectation, the present study found no consistent effects of such information treatments on either beliefs about the social choice context or on stated WTP values. This result suggests that individuals’ responses were quite immune to claims about the social choice context of which the respondents, after all, did not know whether or not they were true. Unfortunately, we cannot say whether this is because strategic issues were generally not important or whether our treatments were simply not successful in modifying the respondents’ perceptions of their strategic opportunities. The former would be likely if the respondents did not know their WTP with any precision or if only a minority of the respondents engaged in the kind of analytical thinking that would allow strategic behavior, as has been found in studies using verbal protocol methods (Schkade and Payne, 1994). The latter would be likely if the respondents did not “accept” the frames. A final possible reason for the immunity of responses would be, of course, that Green et al.’s – compelling – proposition regarding incentive compatibility in CV surveys is incorrect.

Two problems of our empirical approach merit special attention. First, as noted by Green et al. (1998), it is conceivable that if subjects who accept some, but not all, of the information about the social choice context, the incentives for misrepresentation may be increased rather than reduced. However, as the theoretically incentive compatible (TIC) questionnaire version yielded the highest WTP responses, the pairwise comparisons of the responses to the TIC questionnaire version with the responses to each of the other questionnaire versions did not support this particular idea of Green and his co-authors.

Second, it appeared to be difficult to provide the relatively complex information about the implementation rule (majority vs. mean rule) in an appropriate way. To characterize the majority rule we had to convey something like: “We are doing this survey so that once we will know the exact cost of the good, we can determine whether enough people would vote for the proposition” (see e.g. Baron, 1996, p. 153; Green et al., 1998, p. 90). From verbal protocol analyses it is well known that CV respondents do not always use the heuristics desired by the surveyors (e.g. Chilton and Hutchinson, 2003; Schkade and Payne, 1994). Burton et al. (2003) who examined the effects of cost uncertainty in a laboratory referendum setting with induced preferences likewise found that cost uncertainty did not affect the choices in expected ways. We are not convinced that our respondents to the majority–rule format fully understood the consequences of this implementation rule for their strategic opportunities. More explicit explanations would perhaps be required to secure this understanding.

What directions could future research take, given the identified challenges in investigating incentive properties in field surveys on tax–financed public goods? Targeted verbal protocol analysis may be one promising avenue. However, as suggested by Green et al. (1998), strategic issues may be a relatively minor problem with respondents who do not even know themselves which response would be in line with their interests and values. The first step would thus be to devise new survey approaches which assist the respondents in reliably “constructing” their preferences in spite of cognitive limitations. Such an approach, involving the provision of “voter recommendations” from familiar parties and interest groups, has recently been proposed by Schläpfer and Schmitt (in revision). Once the respondents “know” their true preference responses, preventing strategic responses should be expected to become a main unresolved issue. Experimental manipulations of the survey frame in the spirit of Green et al. (1998) and the present experiment may then yield more conclusive results about the potential merits of a theoretically incentive compatible framing of preference questions.

6. Conclusion

From the failure of our treatments to consistently affect stated beliefs about the social choice context and stated willingness to pay we conclude that the potential for reducing strategic behavior through theoretically incentive compatible survey protocols may be very limited. Our results are consonant with the conclusion of Green et al. (1998) who find that problems following from cognitive limitations of survey respondents may deserve more attention than the incentive properties of survey questions. Nevertheless, the incentive properties of stated preference elicitation for public goods remain a key unresolved issue in the theoretical underpinnings of stated preference methods. The present study represents one of the first empirical attempts towards its clarification. As such, it also highlights the empirical difficulties involved in investigating the role of incentive properties in preference questions about collectively provided public goods. Further empirical research is needed to understand the factors that drive the perception of the social choice context and how this perception may determine the

incentive properties of stated preference questions about public goods and services. Sharper tests and more conclusive results than those presented here may be sought in the context of novel survey approaches in which the more fundamental problem that respondents may not even know their true preferences with any precision is resolved in a satisfactory way.

Acknowledgements

We thank George Hutchinson, Ricardo Scarpa, Jonathan Baron and two anonymous reviewers for very valuable comments on previous versions of this manuscript. Financial support was provided by the University of Zurich (Research Grant 57230101 to F. Schläpfer).

REFERENCES

- Andreoni, J., 1988. Why free ride? Strategies and learning in public goods experiments. *Journal of Public Economics* 37 (3), 291–304.
- Andreoni, J., McGuire, M.C., 1993. Identifying the free riders: a simple algorithm for determining who will contribute to a public good. *Journal of Public Economics* 51 (3), 447–454.
- Arrow, K.R., Portney, P.R., Leamer, E.E., Radner, R., Schuman, H., Solow, R., 1993. Report of the NOAA panel on contingent valuation. *Federal Register*, vol. 58 (10), pp. 4601–4614.
- Baron, J., 1996. Rationality and invariance: Response to Schuman. In: Bjornstad, D.J., Kahn, J.R. (Eds.), *The Contingent Valuation of Environmental Resources. Methodological Issues and Research Needs*. Edward Elgar, Cheltenham, U.K, pp. 145–163.
- Brown, T.C., Champ, P.A., Bishop, R.C., McCollum, D.W., 1998. Which response format reveals the truth about donations to a public good? *Land Economics* 72 (2), 152–166.
- Burton, A.C., Carson, K.S., Chilton, S.M., Hutchinson, W.G., 2003. An experimental investigation of explanations for inconsistencies to second offers in double referenda. *Journal of Environmental Economics and Management* 36 (3), 472–489.
- Carson, R.T., Hanemann, W.M., Kopp, R.J., Krosnick, J.A., Mitchell, R.C., Presser, S., 1998. Referendum design and contingent valuation: the NOAA panel’s no–vote recommendation. *Review of Economics and Statistics* 80 (3), 484–487.
- Carson, R.T., Groves, T., Machina, M., 2000. Incentive and informational properties of preference questions. Working Paper. Department of Economics, University of California, San Diego, CA.
- Champ, P.A., Flores, N.E., Brown, T.C., Chivers, J., 2002. Contingent valuation and incentives. *Land Economics* 78 (4), 591–604.
- Chilton, S.M., Hutchinson, W.G., 2003. A qualitative examination of how respondents in a contingent valuation study rationalise their WTP responses to an increase in the quantity of the environmental good. *Journal of Economic Psychology* 24 (1), 65–75.
- Cummings, R.G., Harrison, G., Rutström, E.E., 1995. Homegrown values and hypothetical surveys: Is the dichotomous choice approach incentive compatible? *American Economic Review* 85 (1), 260–266.
- Cummings, R., Elliott, S., Harrison, G.W., Murphy, J., 1997. Are hypothetical referenda incentive compatible? *Journal of Political Economy* 105 (3), 609–621.
- Diamond, P.A., Hausman, J.A., 1994. Contingent valuation: is some number better than no number? *Journal of Economic Perspectives* 8 (4), 45–64.
- Fisher, A.C., 1996. The conceptual underpinnings of the contingent valuation method. In: Bjornstad, D.J., Kahn, J.R. (Eds.), *The Contingent Valuation of Environmental Resources*.

- Methodological Issues and Research Needs. Edward Elgar, Cheltenham, UK, pp. 19–37.
- Green, D., Jacowitz, K.E., Kahneman, D., McFadden, D., 1998. Referendum contingent valuation, anchoring, and willingness to pay for public goods. *Resource and Energy Economics* 20 (2), 85–116.
- Greene, W., 1998. LIMDEP Version 7.0. User's Manual. Econometrics Software Inc, Plainview, NY.
- Hanemann, W.M., 1994. Valuing the environment through contingent valuation. *Journal of Economic Perspectives* 8 (4), 19–43.
- Hidano, N., Kato, T., Izumi, K., 2005a. Reciprocity, consequentiality and willingness-to-pay in contingent valuation: an experimental panel analysis on climate changes. Paper presented at the 14th Annual Conference of the European Association of Environmental and Resource Economists, Budapest, Hungary.
- Hidano, N., Kato, T., Aritomi, M., 2005b. Benefits of participating in contingent valuation mail surveys and their effects on respondent behavior: a panel analysis. *Ecological Economics* 52 (1), 63–80.
- Hoehn, J., Randall, A., 1987. A satisfactory benefit–cost indicator for contingent valuation. *Journal of Environmental Economics and Management* 17 (3), 226–247.
- Horowitz, J.K., McConnell, K.E., 2002. A review of WTA/WTP studies. *Journal of Environmental Economics and Management* 44 (3), 426–447.
- Loomis, J.B., Blair, L.S., Gonzalez–Caban, A., 2002. Language–related differences in a contingent valuation study: English versus Spanish. *American Journal of Agricultural Economics* 84 (4), 1091–1102.
- Mitchell, R.C., Carson, R.T., 1989. *Using Surveys to Value Public Goods: The Contingent Valuation Method*. John Hopkins University Press, Baltimore.
- Pouta, E., 2004. Attitude and belief questions as a source of context effect in a contingent valuation survey. *Journal of Economic Psychology* 25 (2), 229–242.
- Posovac, S., 1998. Strategic overbidding in contingent valuation: Stated economic value of public goods varies according to consumers expectations of funding source. *Journal of Economic Psychology* 19 (2), 205–214.
- Riera, P., 2003. Incentive compatibility in stated preference valuation methods: some positive results. Paper presented at the 12th Annual Conference of the Association of Environmental and Resource Economists, Bilbao, Spain.
- Schkade, D.A., Payne, J.W., 1994. How people respond to contingent valuation questions: a verbal protocol analysis of willingness–to–pay for an environmental regulation. *Journal of Environmental Economics and Management* 26 (1), 88–109.
- Schläpfer, F., Schmitt, M. Anchors, endorsements, and preferences: a field experiment. *Resource and Energy Economics*, in revision.
- Smith, V., 1979. An experimental comparison of three public good decision mechanisms. *Scandinavian Journal of Economics* 81, 198–215.
- Taylor, L.O., McKee, M., Laury, S.K., Cummings, R.G., 2001. Induced–value tests of the referendum voting mechanism. *Economics Letters* 71 (1), 61–65.