Article:
Martin-Ortega, J, Azahara Mesa-Jurado, M and Berbel, J (2015) Revisiting the Impact of Order Effects on Sensitivity to Scope: A Contingent Valuation of a Common-Pool Resource. Journal of Agricultural Economics, 66 (3). 705 - 726. ISSN 0021-857X

https://doi.org/10.1111/1477-9552.12105

Revisiting the impact of order effects on sensitivity to scope: a contingent valuation of a common-pool resource

Julia Martin-Ortega, M. Azahara Mesa-Jurado and Julio Berbel[a]

[Original submitted April, 2014, Revision received September 2014, Accepted November 2014]

ABSTRACT

While many studies have employed stated preferences methods to estimate the value of non-market ecosystem goods and services contingent valuation (CV) still generates a significant amount of criticism. Besides ethical concerns, two of the key criticisms relate to insensitivity to scope and order effects. We examine the expectation that the presence of order effects in stepwise disclosure procedures affects the degree of scope sensitivity. We use data from a CV exercise asking farmers in Southern Spain to value two different levels of guarantee of water supply for irrigation in a context of water scarcity. We find that despite order effects being present, they do not affect the existence or the degree of sensitivity to scope. We conclude that, in the light of the mixed evidence found in the literature and the results of our study, it does not seem justified to ascribe order effects and their connection with sensitivity to scope to study design alone (e.g. step-wise versus advanced disclosure, top-down versus bottom-up). The nature of the environmental good under valuation also matters. Our study of irrigation water as a common-pool resource suggests that, when clear private benefits also exist, these appear to override any 'good cause dumping effect' that might arise from the public good component.

Keywords: willingness to pay, contingent valuation, order effects, scope sensitivity, 'good cause dumping effect', irrigation water, warm glow

JEL Classifications: (Please add from here.)

[a] Azahara Mesa-Jurado is the contact author, azaecosur@gmail.com, and is at El Colegio de la Frontera Sur, ECOSUR, Mexico. Julia Martin-Ortega is with the The James Hutton Institute, United Kingdom, and Julio Berbel is at the University of Cordoba, Spain. The survey used in this research was financed by the Spanish Ministry of Science and Innovation (Programa de Formación al Personal Investigador "Análisis prospectivo de la sostenibilidad de los sistemas agrarios nacionales en el marco de la PAC" Ref: AGL2006-05587-c04-02). The work on this manuscript was partly funded by the Scottish Government Rural Affairs and the Environment Portfolio Strategic Research Programme 2011-2016, Theme 1 (Environmental Change: Ecosystem Services and Biodiversity) and by FOMIX (Tabasco Government and CONACYT) through the project "El valor económico del agua en el estado de Tabasco" and El Colegio de la Frontera Sur (Mexico). The authors are very grateful to Sebastiaan Hess for his invaluable comments to a previous version of this manuscript and to Dominic Duckett for inspiring discussions that helped developing ideas in this paper. Thanks are also due to the anonymous reviewers and the editors for valuable comments on earlier drafts.

1. Introduction

In the last two decades, many studies have employed stated preference methods to estimate the value of non-market ecosystem goods and services, especially using the contingent valuation method (CV) (see Carson (2011) for a historical review). Best practice guidelines have been drawn up and methodological improvements continue to be put forward (Arrow et al., 1993; Venkatachalam, 2004; Bateman et al., 2008; Schläpfer, 2008; Carson, 2012). However, despite its widespread adoption and methodological innovations, CV still suffers significant criticism. Besides ethical concerns associated with attaching monetary values to natural resources (Azqueta and Delacamara, 2006; Spangenbergh and Settele, 2010; Gómez-Baggethun et al., 2010), two of the remaining key criticisms relate to insensitivity to scope (Kahneman and Knestch, 1992; Desvousges et al., 1993; Diamond and Hausman, 1994; Desvousges et al., 2012; Hausman, 2012), and the presence of order effects (Mitchell and Carson, 1989; Bateman and Langford, 1996; Payne et al., 2000; Powe and Bateman, 2003; Bateman et al., 2004; Clark and Friesen, 2008).

In principle, the larger the scope of the environmental improvement under valuation, the greater the welfare increases and, hence, the greater the willingness to pay (WTP) should be. In apparent conflict with theory, a considerable number of studies have shown scope-insensitive results (Carson and Mitchell, 1993; Boyle et al., 1993; Diamond et al., 1993; Halvorsen, 1996; Hammitt et al., 1999; Heberlein et al., 2005). These inconsistencies with economic theory have made critics question the overall validity of stated preference techniques (Diamond and Hausman, 1994; Desvousges et al., 1993; Kahneman and Knetsch, 1992; Kahneman et al., 1999; Desvousges et al., 2012; Hausman, 2012).

Order effects appear when changing the order of WTP questions for nested goods (i.e. goods which either encompass or are encompassed by other goods) affects the WTP values for these goods. This is inconsistent with the basic principle of procedural invariance of the rational theory of choice, which states that preference should not depend on the method of elicitation (Boyle et al., 1993).[b] It has been proposed that order effects indicate that

[b] A distinction needs to be made between inclusive and exclusive lists of goods. In inclusive lists, each subsequent good includes the previously valued good(s), while in exclusive lists goods are presented as alternatives. Bateman and Langford (1996) distinguish between theoretically expected 'sequence effects'

preferences are constructed while going through the questionnaire and are not given, as assumed in economic theory (Slovic, 1995).

A key question in this context is whether order effects affect sensitivity to scope. Although some efforts have been devoted to this, the literature on this particular topic is still scarce and results have so far been inconclusive. Annex 1 (on-line) presents a compilation of 9 studies on order effects in CV of exclusive lists of goods. Five of these specifically discuss their impact on sensitivity to scope. A landmark is the study by Bateman et al. (2004) who offered a systematic examination of order effects on observed scope sensitivity. They distinguished between 'advanced disclosure' and 'stepwise disclosure'. Advanced disclosure refers to a situation where the respondent is told at the outset that different goods are going to be valued, while in stepwise designs the respondent is first asked to value one good, and only after that valuation, is told about other goods to be valued. The authors argued that in the stepwise procedure there is an element of surprise that may strategically affect responses to the CV questions. They examined the relationship between advanced and stepwise disclosure and scope sensitivity in field and laboratory trials in the United Kingdom, and found that with the stepwise disclosure procedure the observed degree of scope sensitivity is substantially and significantly affected by the order of the questions, but this is not observed in the advanced disclosure design. In their conclusions Bateman et al. (2004) favoured the advanced disclosure format, arguing that strategic incentives may still apply, but are constant throughout all the valuation exercises. However, the authors were cautious not to state that values from advanced disclosure designs are unbiased, and called for further investigation. In fact, more recent studies have also found order effects in advanced disclosure procedures. Nielsen and Kjær (2011) explored the issue in a large-scale web-based stated preference survey in Denmark, in which split-samples were asked to value two independent gains of life expectancy caused by a change in air pollution. They used advanced disclosure and found order effects that influenced the degree of scope sensitivity. Andersson and Svensson (2013) used advanced disclosure in an experiment on reduced mortality risk in Sweden and could not reject scope insensitivity.

from inclusive lists due to income and substitution effects, and 'order effects' in exclusive lists that are not consistent with theoretical predictions. In this paper we focus on order effects in exclusive lists.

Clark and Friesen (2008) explored the causes of order effects in CV, using a laboratory experiment in which participants were requested to value sequences of nested goods from an exclusive list using an incentive-compatible mechanism in a stepwise disclosure procedure. They tested different types of goods (induced value goods, familiar private goods and the same private goods to be donated to charity). While according to Bateman et al.'s (2004) findings, order effects might be expected due to the stepwise disclosure design, the interesting result in Clark and Friesen's study is that they were found for familiar private goods intended for personal use, and not when the same private goods were intended for donation.

These findings demonstrate that the issue of order effects and their influence on sensitivity to scope is unresolved. This paper explores the effect that order of WTP questions for nested goods in CV surveys have on the existence and the degree of sensitivity to scope. Specifically, we revisit the question of whether the presence of order effects in stepwise disclosure procedures affects sensitivity to scope. We use the results of a CV exercise in which farmers in Southern Spain were asked to value two different levels in the guarantee of water supply for irrigation in a context of water scarcity. While we agree with Bateman et al. (2004) that there is not a general theoretical expectation for the existence of sensitivity to scope, we might expect it for this particular good, because irrigation water is both scarce and a major input in agriculture in the area, so the volume of applied irrigation greatly affects productivity.

In light of the interpretations that have been given to order effects and sensitivity to scope, we discuss our results in relation to procedural invariance and the economic nature of the valued good. The 'common-pool resource' nature of irrigation water, as opposed to fully private and fully public goods, provides a new perspective to this discussion. Irrigation water flowing through the distribution system of an irrigation district is hard to make excludable although yield is subtractable (rival), particularly in contexts of scarcity (Gardner et al., 1990; Ostrom and Gardner, 1993; Hanemann, 2006). By testing order effects on sensitivity to scope in a neither fully private nor fully public good, we reflect on whether reasons of warm glow or 'good cause dump effect' (i.e. by which respondents 'dump' their entire 'good cause' account on the first valuation question they encounter)

associated with the 'common good' element of common-pool resources are affected or mitigated when private benefits also exist.

2. Data

We use data from a CV survey among farmers in the Guadalbullon River Basin in southern Spain that was first published in Mesa-Jurado et al. (2012).[c] This river basin is typically Mediterranean, subject to water scarcity and water supply uncertainty. Almost all farmers belong to 'irrigator communities', in which they manage the water under a democratic co-operative principle, and the water is perceived as a common-pool resource by the farmers.

A sample of olive grove farmers was asked their WTP for increased guarantees of water supply for irrigation. Trained interviewers conducted 151 face-to-face interviews in July 2009 throughout the Guadalbullon River Basin, targeting a random sample of irrigated olive farm owners, belonging to eighteen different irrigator communities.

The questionnaire was organized in four main parts: i) farm characterization (e.g., farm size, crop density); ii) respondents' perception of the current state of the water supply for irrigation in the area (e.g. certainty about the amount of water they will receive at the beginning of the season, satisfaction with the amount of water they receive); iii) respondents' preferences and WTP for increased water supply guarantees; and iv) respondents' socio-economic characteristics.

The valuation scenario consisted of a description of the baseline situation (status-quo) in which the current official water allocation – 'water right' or concession – of 1,500 m^3/ha is rarely met. Farmers were offered the possibility to pay for an increased guarantee of the water supply by contributing to the construction of infrastructure such as reservoirs or irrigation ponds.[d] Pictograms helped explain the questions. Respondents were also reminded to consider their budget constraints and were assured that their money would only be used for the indicated purpose. For those who stated they were not willing to pay, a follow-up question was asked to differentiate legitimate zero bidders from protest answers. Those farmers who stated they were willing to pay in principle were then asked their

[c] The Mesa-Jurado et al. (2012) paper does not include any analysis or discussion on the issue of order effects.
[d] Currently, there is no such infrastructure in the area and irrigators are fully dependent on flow levels (Mesa-Jurado et al., 2012).

maximum WTP to secure a certain water supply: *'What is the maximum amount of money* that you would be willing to pay in addition to your yearly Irrigators Community annual payment to ensure 10,000 litres per olive tree (...)*?'*. As will be explained in the next section, two different levels of increased guarantee were presented to respondents in a stepwise design, in which the order of the levels was alternated in a split sample.

In the pre-test, an increase in the annual payment to the irrigator community (expressed in euros per olive-tree) was identified as the best payment vehicle due to its familiarity. The elicitation format was a payment card (Ready et al., 2001; Kallas et al., 2007). Based on information from pre-test interviews and focus group discussions, payment amounts ranging from €0 to more than €8 per olive tree were included in the payment card, as shown in Annex 2 (on-line).

3. Method

3.1. Split sample design

Two levels of water supply guarantee were defined in order to test for sensitivity to scope: 50 and 90% guarantee.[e] In the pre-test it was observed that farmers had problems with probability expressed in percentages. It was therefore presented as the number of years in a given time frame that they would receive a specified quantity of water with certainty, i.e. WTP to ensure 10,000 litres per olive tree 'in five out of the next ten years' and 'in nine out of the next ten years'. This way of expressing probabilities has been used elsewhere in the literature (Hensher et al., 2005; Hatton McDonald et al., 2010; Martin-Ortega et al., 2011). The two valuation scenarios form an exclusive list, since they were presented as independent and not additional.

To test for order effects, a split sample approach was used. Half the sample was first asked their WTP for a 50% guarantee followed by the 90% guarantee. For the other half the order was reversed. In the remainder of this text we refer to the first sub-sample as 'bottom-up' and to the second as 'top-down', following Carson and Mitchell's terminology (1995). As

[e] The two levels represent an increase on the scope of the service provided (i.e. increased level of guarantee). During the pre-test it was observed that a guarantee of 100% was not a credible scenario, because farmers are aware of the variable climatic conditions and the structural water scarcity in the region, as demonstrated in the preparatory focus groups.

mentioned, we followed a pure stepwise disclosure procedure, i.e. respondents were not warned that there would be two WTP questions.

3.1. Statistical Testing

The analysis was carried out through a series of parametric and non-parametric statistical tests, and univariate and multivariate analysis. The methodological sequence applied in this study is as follows.

First, we examined whether the two split samples are significantly different to ensure that no bias was introduced by the split sample procedure. For that we compared means across the split samples of those variables identified in Mesa-Jurado et al. (2012) as relevant for WTP.

Second, we looked for the existence of order effects, testing for the hypotheses whether WTP for the two levels of water supply guarantee are significantly different in the bottom-up and top-down approaches (H_1, H_2). For that, we followed the literature in using non-parametric Mann-Whitney U tests to compare medians[f] and t-test to compare mean WTP values.

Third, we explored sensitivity to scope in several ways. Following the approach proposed by Nielsen and Kjær (2011), we first looked for respondents who are insensitive to scope, i.e. those farmers who stated identical WTP values for 50% and 90% levels of water supply guarantee. A panel data structure was used, in which the two observations per respondent were identified by a variable taking the value "0" for the 50% level and "1" for the 90% level.[g] In order to analyse the differences between scope sensitive and insensitive respondents, a Probit regression was run, in which the dependent variable was coded "1" for respondents willing to pay more for 90 than for 50%. A variable accounting for the order of the WTP questions was included in the model to study the influence of question-order on the probability of respondents showing insensitivity to scope.

[f] Strictly speaking, the Mann-Whitney U-test is not a comparison of medians. However, when one has similar distributions in shape (spread), as is our case, it is possible to use Mann-Whitney to a test of the medians (Hart, 2001).

[g] The potential correlation between error terms in the answers from the same individual was checked as suggested by Longo et al. (2009), and rejected in Mesa-Jurado et al. (2012).

Sensitivity to scope was further tested in an internal scope test. Using Wilcoxon-signed-rank tests and t-tests, we checked whether median and mean WTP values for 50% and 90% are significantly different for the overall sample (H_3)[h]. Such internal scope tests have been criticized, arguing that respondents may simply be trying to give "internally consistent" answers (Carson et al., 2001). This is why we also used an external scope test, in which WTP values for the first question of the bottom-up sample were compared with those for the first question of the top-down sample: i.e. we tested the hypothesis of WTP values for the 50% guarantee from the first sample being equal to WTP values for the 90% guarantee from the second sample (H_4).

Finally, the effect of order on sensitivity to scope was analysed. For that we looked at two aspects: the effects on the existence and on the degree of sensitivity to scope. This distinction is relevant because, as Bateman et al. (2004) argue, while the existence of sensitivity to scope might not be generally expected theoretically (and is dependent upon the characteristics of each case), there is a clear economic-theoretical expectation that the observed degree of scope sensitivity should not vary by treatment.

To test whether question order affected the existence of sensitivity to scope, we tested whether WTP values for the 50% are equal to WTP values for 90% water supply guarantee in the bottom-up (H_5) and top-down samples (H_6). Following Clark and Friesen (2008) we also tested for order and scope effects while controlling for respondents' characteristics. For this, a Tobit model was used, which allows left censoring of WTP values at zero. A dummy variable was added to control for the order of the WTP questions.

To test the effects of question order on the degree of sensitivity to scope, we followed Bateman et al. (2004) and Clark and Friesen (2008) and compared the difference between WTP for the larger and smaller goods for the two subsamples (H_7). We looked at both mean and median of the difference through t-tests and Mann-Whitney tests, respectively.

4. Results

[h] A paired test was used for the internal scope test.

Farmers were first asked whether they were willing to pay in principle to improve the guarantee of water supply for irrigation, without mentioning any guarantee level. 29% were unwilling to pay. Follow-up questions allowed us to separate protest answers from legititmate zeros. 79% of the negative answers were classified as legitimate zeros and the remaining 21% (6% of the total sample) were judged to be protest responses. These figures are within the limits of acceptable levels of protest according to Carson et al. (2003). A high proportion of the protest votes (89%, i.e. 5.3% of the total sample) involved affirmative response to 'the State should pay for it' (the rest of protesters did not believe there would be enough water to secure such guarantee levels). Following common practice in the literature, protest responses were excluded from the analysis, while the legitimate zeros were retained (Dziegielewksa and Mendelsohn, 2007). Two outliers were identified and also excluded from the analysis.

A detailed descriptive analysis of the survey data can be found in Mesa-Jurado et al. (2012). Here we focus only on those results of interest for the study of order effects and sensitivity to scope.

4.1. Test for split-sample differences

T-test results shown in Table 1 indicate that the two samples are not significantly different on the 29 relevant variables for WTP (as reported in Mesa-Jurado et al., 2012), except for one: the number of olive trees per hectare. We take this to indicate that the two samples are sufficiently similar to rule out any sample selection bias in our study.

Table 1. Test of equality of variables influencing WTP values across the split samples

Variable	Description	Test value[a]	Sig. (2-tailed)
Perceived water allocation	Volume of allocated water as perceived by the farmer (litres/tree)	1.409	0.161
Perceived annual payment	Irrigators Community annual payment declared by farmers (€/tree)	0.952	0.343
Age	Farmer's age	-0.006	0.995
Income	Gross household income (in €) per year per hectare	0.423	0.673
Household size	Number of persons in the household dependent on farm income	0.653	0.514
Olive trees per hectare	Traditional planting frame (fewer than 90 olive trees/ha) Semi-Intensive planting frame (100-150 olive trees/ha) Intensive planting frame (more than 150 olive trees/ha)	9.244	0.015**[b]
Education level	No formal education Elementary school	4.455	0.216

	Secondary education Higher education (university)		
Agricultural training	Identifies if the farmer has received some kind of specific agricultural training	0.750	0.387

[a]t value for interval variables and Pearson Chi^2 value for categorical variables (olive trees per hectare, education level and agricultural training).
[b]Statistical significance levels: *** 1%, ** 5%, *10%

4.2.Testing for order effects

Table 2 shows the mean and median WTP values for the 50% and 90% levels of water supply guarantee for the two sub-samples and the results of all hypotheses tested. The results of the tests for independent samples show that mean and median WTP values for the smaller good are significantly different between the sub-samples: when the 50% question was asked first, farmers stated a higher WTP than when it came after the 90% question (H_1). On the contrary, stated WTP for the larger good is not significantly different between the sub-samples (H_2).

These results are consistent with previous findings in the literature that order effects influence WTP estimates for the smaller good but not for the larger one (Powe and Bateman, 2003; Bateman et al., 2004; Veisten et al., 2004; Longo et al., 2009; Nielsen and Kjær, 2011).

Table 2. WTP for 50 and 90% guarantee (€ cents/olive tree) and summary of statistical test

Sample descriptors	**Bottom-up**		**Top-down**	
	50%	**90%**	**50%**	**90%**
Mean	46	77	50	72
Median	50	60	20	60
Std. error	4.8	7.7	4.0	7.4
Confidence interval	(37 - 56)	(61 – 92)	(24 - 40)	(57 - 86)
Number of observations	70		71	
Hypothesis (H)	**Difference in means (p-value)**		**Difference in medians[a] (Asymptotic p-value)**	
Test of order effects				
(H_1) WTP50 BU = WTP50 TD	0.026**		0.053**	
Number of observations		141		
(H_2) WTP90 BU = WTP 90 TD	0.628		0.737	
Number of observations		141		
Internal test of scope sensitivity				
(H_3) WTP50 pooled = WTP90 pooled	0.000***		0.000***	
Number of observations		141		
External test of scope sensitivity by first response				

(H_4) WTP50 BU = WTP90 TD	0.005***		0.020**
Number of observations		282	
Influence of order on scope sensitivity			
(H_5) WTP50 BU = WTP90 BU	0.001***		0.011**
Number of observations		142	
(H_6) WTP50 TD = WTP90 TD	0.000***		0.000***
Number of observations		140	
(H_7) (WTP90 BU – WTP50 BU) = (WTP90 TD – WTP50 TD)	0.208		0.940
Number of observations		141	

Statistical significance levels: *** 1%, ** 5%, *10%

[a] For the external scope tests and order effect tests this is a Mann-Whitney U test. For the internal scope test this is performed through a Wilconxon-signed-rank test.

4.3. Sensitivity to scope

Factors determining insensitivity to scope

We found 18 respondents who showed insensitivity to scope, i.e. who gave exactly the same valuation for the 50 and 90% scenarios. This represents 12% of the sample and is substantially smaller than the proportion found by Nielsen and Kjær (2011), who report between 40 and 50% of scope-insensitive respondents in their sample, and Jones-Lee et al. (1985), who found that 42 to 47% of respondents were completely scope-insensitive.[i]

Results from the Probit model (Table 3) indicate that age, education level, tree density and farm income influence the likelihood of being sensitive to scope. Regarding age, Nielsen and Kjær (2011) and Andersson and Svensson (2013) also found that younger respondents are more likely to be sensitive to scope. The education level was also significant in Nielsen and Kjær (2011), but contrary to our results, they found that higher levels of education positively affect the likelihood of sensitivity to scope. Regarding income, Powe and Bateman (2003) similarly found positive significant effects, meaning higher income respondents are more likely to be sensitive to scope.

Contrary to the Nielsen and Kjær (2011) and Andersson and Svensson (2013) findings, the order variable (Sub-sample) was not significant. This is a first indication that the order of

[i] Twelve percent refers to our total sample. If we exclude respondents who are not willing to pay in principle, the percentage rises to 18%, still substantially below the figures reported in the cited studies.

the WTP questions does not affect the existence of sensitivity to scope.[j] Further tests of the influence of order effects on sensitivity to scope are presented below (Section 4.4).

Table 3. Probit regression of variables determining sensitivity to scope (dependent variable = 1 if respondent is willing to pay more for 90 than for 50% water supply guarantee)

	Description	Coef. (Standard Error)	z	P > z[a]	(95% Conf. Interval)
Sub-sample	Bottom-up=0 Top-down=1	-0.055 (0.328)	-0.17	0.866	(-0.699, 0.588)
Olive trees per hectare	< 90 olive trees/ha=1 100-150 olive trees/ha=2 > 150 olive trees/ha=3	0.975 (0.399)	2.44	0.015**	(0.193, 1.756)
Perceived Annual Payment	Irrigators Community annual payment declared by farmers (€/tree)	-0.017 (0.141)	-0.12	0.906	(-0.294, 0.260)
Income	Gross household income (€/year/ha)	$3.35 \cdot 10^{-5}$ (0.000)	2.54	0.011**	$(7.62 \cdot 10^{-6}, 5.95 \cdot 10^{-5})$
Respondent's attitude during survey	As judged by enumerator: good=1; neutral=2; uninterested=3	-0.655 (0.401)	-1.63	0.103	(-1.442, 0.131)
Age	Age of the farmer	-0.039 (0.015)	-2.66	0.008***	(-0.068, -0.010)
Education level	No formal education=1 Elementary school=2 Secondary education=3 Higher education (university)=4	-0.641(0.295)	-2.17	0.030**	(-1.221, -0.062)
Model constant		0.804 (1.689)	0.48	0.634	(-2.506, 4.114)
Number of Obs	148				

[a]Statistical significance levels: *** 1%, ** 5%, *10%

Results of scope sensitivity tests

Results of the internal scope test (H_3) are presented in Table 2. These results indicate that the null hypothesis of equality is rejected, showing that the overall sample of farmers is willing to pay significantly more for the 90% guarantee than for the 50% one, as expected. The external scope test (H_4) also confirms sensitivity to scope (also in Table 2). The ratio between the WTP for the two goods valued (i.e. WTP90/WTP50) for the overall sample is 1.89, which is close to the ratio of the two goods being offered (i.e. 90/50 = 1.8). In the bottom-up survey split this ratio is 1.67. The ratio is higher (2.25) in the top-down approach, indicating a higher degree of sensitivity in this approach. Moreover, this ratio is

[j] This finding proved robust across all model specifications that were tested.

higher than the ratio between the two guarantees being offered (90/50=1.8). This is probably explained by the high number of zero bids in the valuation of the smaller good in the top-down approach (24 or 34.3%), which suggests that the lower level of guarantee of water supply (50%) becomes unattractive for many after having been offered the 90% (this is also suggested by the fact that the median WTP for the 50% water supply guarantee is substantially smaller than the mean, see Table 2).

The question now is whether sensitivity to scope is affected by question order, which we check in the next section.

4.4. Scope consistency check

Order effects and the existence of sensitivity to scope

Mean and median tests show that WTP for the 50 and 90% guarantee is significantly different in both the bottom-up (H_5) and top-down (H_6) samples, as shown in Table 2. This means that despite the existence of order effects (i.e. the smaller good is more valued when asked first) sensitivity to scope holds.

A Tobit model was estimated to test sensitivity to scope while controlling for respondents' characteristics. Results in Table 4 show that when controlling for the sub-sample (top-down or bottom-up), sensitivity to scope is present: the coefficient of the variable Sub-sample is not statistically significant, while the variable Level of water supply guarantee has a positive and statistically significant coefficient.

Table 4. Tobit regression of variables determining WTP when controlling for sub-sample

	Description	Coef. (Standard Error)	T	P > t	(95% Conf. Interval)
Sub-sample	Bottom-up = 0 Top-down = 1	-0.044 (0.089)	-0.50	0.619	(-0.222, 0.132)
Level of water supply guarantee	50%=0 90%=1	0.391 (0.088)	4.42	0.000***	(0.217, 0.565)
Cluster ID	Respondents ID variable controlling for panel data effects[a]	-0.001 (0.001)	-0.34	0.737	(-0.002, 0.002)
Olive trees per hectare	< 90 olive trees/ha=1 100-150 olive trees/ha=2 > 150 olive trees/ha=3	-0.003 (0.002)	-1.84	0.067*	(-0.007, 0.001)

Perceived water allocation	Volume of water received, as perceived by the farmer (litres/tree)	$-4.21 \cdot 10^{-5}$ (0.001)	-2.44	0.015**	$(-7.90 \cdot 10^{-5}, -8.44 \cdot 10^{-6})$
Income	Gross household income (€/year/ha)	0.223 (0.121)	1.85	0.066*	(-0.014, 0.461)
Age	Age of the farmer	-0.009 (0.004)	-2.39	0.017**	(-0.016, -0.002)
Agricultural Training	Identifies if the farmer has received agricultural training: no=0; yes=1	0.192 (0.115)	0.1.68	0.095*	(-0.033, 0.418)
Model constant		1.154 (0.388)	2.68	0.003	(0.389, 1.917)
Sigma		0.644 (0.037)			(0.571, 0.718)
Number of Obs[a]	232				

[a] Data in this model are treated as panel data, i.e. two observations per respondents are recorded for the two levels of guarantee. The variable Cluster ID controls for the potential correlation between the two observations per respondent.
Note: Statistical significance levels: *** 1%, ** 5%, *10%.

Order effects and the degree of sensitivity to scope

Results from a Mann-Whitney test show that the median of the differences between WTP90 and WTP50 is not statistically significantly across the two sub-samples either (asymptotic p-value = 0.940). Question order therefore did not affect the observed degree of scope sensitivity (H_7).

5. Discussion

We have looked for the existence of order effects and how these affect the existence and the degree sensitivity to scope in the valuation of increased guarantee in the supply of water for irrigation. Two levels of guarantee (50% and 90%), presented in the form of exclusive goods, were valued by a sample of farmers in southern Spain using a stepwise disclosure procedure. Tests of means and medians have shown order effects for the smaller good (50% guarantee) but not for the larger one (90% guarantee), which is in line with previous findings in the literature. However, these order effects do not have a significant impact on the existence of sensitivity to scope. This was corroborated by a multivariate analysis that showed a statistically significant scope effect (level of guarantee) when controlling for question order. Moreover, and contrary to previous findings for stepwise procedures by Bateman et al. (2004), we find that the degree of scope sensitivity does not vary by sub-

sample either. In this respect, our results are closer to those of Clark and Friesen (2008), who also found order effects for private goods, but no influence on scope sensitivity.

Order effects and insensitivity to scope have been related in the literature to the economic nature of the good being valued, with public goods supposedly being affected by the so-called warm glow effect (the positive emotional feeling obtained from 'giving'; Andreoni, 1990) and the 'purchase of moral satisfaction' (Kahneman and Knetsch, 1992), not present for private goods.[k] Powe and Bateman (2003) made the link between the warm glow hypothesis and the 'anchoring and adjustment' heuristic as interpreted by Kahneman et al. (1982), which suggests that an initial valuation provides an anchor for subsequent responses which typically are not adjusted much from that level. Powe and Bateman (2003) argued that the nature of public goods and their modes of funding might affect the (in)dependence of valuation tasks. They argued that if respondents' first valuation response tends to be a 'dumping ground' for moral satisfaction then this might provide an anchor. As Clark and Friesen (2008) further explained, a possible explanation for the finding that goods of smaller scope are valued more when valued first is the 'Good Cause Dump hypothesis', in which respondents 'dump' their entire 'good cause' account on the first valuation question they encounter (Harrison, 1992). Longo et al. (2009) found that order effects disappeared when respondents were given the opportunity to review their answers, and compared this situation (the possibility of revisiting payments) to that of a supermarket where private goods are purchased. It should be noted, though, that Longo et al. (2009) did not test for the effects of sensitivity to scope, and Nielsen and Kjær (2011), who did test for it, found that these adjustments did not mitigate the scope insensitivity results that they identified.

The present study provides a new perspective to the debate by testing order effects on sensitivity to scope in a good which is neither fully private nor fully public, but a common-pool resource. Irrigation water is generally non-excludable, but is rival, particularly in the context of scarcity as is the case in our study area. Warm glow in terms of purchasing moral

[k] A different interpretation to this discourse is that provided by Bateman et al. (2004) who attribute the effect to the 'surprise' element inherent in the stepwise procedure and not to the warm glow effect. However, as discussed in the introduction, more recent experiments have shown that order effects also affect advanced disclosure formats.

satisfaction might not directly apply in this case. However, when farmers pay for the construction of water storage infrastructure, they are in effect contributing not only to their own (private) benefit, but to a more general common good: increased water availability in the region from which other users benefit as well. In our study area these investments are also explicitly planned to also improve ecological flows for the 'benefit of the environment' (Berbel et al., 2012), in compliance with the legislative framework set up by the European Water Framework Directive (2000/60/EC).

Order effects occurred in our data, but this did not affect either the existence or the degree of sensitivity to scope. One possible interpretation is that the 'the common good element' is not large enough to produce the 'good cause dumping effect', i.e. the public element of the common-pool resource is dominated by the private element, where sensitivity to scope still applies in a significant degree due to the increased (private) benefit that the farmer obtains. This interpretation could also apply to Andersson and Svensson's (2013) results, which also show order effects but still some differences in the degree of sensitivity to scope in the valuation of decreased mortality risk from bus accidents. They refer to it as a non-pure private good in which values might contain an element of paternalistic altruism, but that is likely to be less strong than the private benefit of reducing one's own personal risk of dying in a bus accident.

Our results can also be compared with those of Clark and Friesen (2008), who have a private good operating as such and the same private good operating as a public good (i.e. same good intended for donation), while we have a good that shares elements with private and public goods. They found that order affects the degree of sensitivity to scope in the valuation of private goods, but not when those same goods were intended for donation. As Clark and Friesen, we do not claim to infer from our results that a warm glow effect cannot be the cause of order effects' induced insensitivity to scope in other CV studies. On the contrary, our results could indicate that the warm glow effect is mitigated (or 'switched off') when private benefits also exist, but it could still be affecting sensitivity to scope for pure environmental public goods in which a greater public benefit prevails over the private benefit.

Finally, Clark and Friesen speculate that the familiarity of their private goods may have reduced the order effects to a point where they did not affect sensitivity to scope. Although they do not make clear why this may be the case, it is noteable that we found the same results for irrigation water, which is a tangible and easily recognizable (familiar) good to our group of respondents. It should be noted that in their case, as well as in ours, small sample sizes might be seen as a limitation[1], and replications of this study with larger samples are needed.

6. Conclusions

Despite extended application of the contingent valuation method, the issue of order effects remains unresolved. A key question in this debate is whether order effects influence sensitivity to scope. The question has generated a small body of literature which suggests that order effects might affect sensitivity to scope, although results have so far been inconclusive. We have revisited the question on whether the presence of order effects in stepwise disclosure procedures affects sensitivity to scope. We have found that, while order effects were present, they did not affect the existence nor the degree of sensitivity to scope amongst our sample. This is comparable to other 'puzzling' results found in previous studies (notably by Clark and Friesen (2009) and Andersson and Sevensson (2013)).

In the case of our study of irrigation water as a common-pool resource, we suggest that a 'contribution to the common good' when clear private benefits also exist, is not enough to produce the 'good cause dumping effect', i.e. the public element of the common-pool resource is overruled by the private element and hence sensitivity to scope still applies due to the increased (private) benefit that the farmer obtains.

[1] Our sample is within the range of those used in some of the key studies having addressed this issue: as shown in the Annex, Bateman et al. (2004) use a sample of 149 interviews in their lab trial. For their test, they make two sample splits, so the number of observations in the sub-samples ranges between 36 and 43. Anderson and Svensson (2013) use a slightly larger sample of 200 individuals, but split the sample twice (two treatments), so they deal with sample sizes from 46 to 96 observations. Clark and Friesen (2008) have a larger sample (235), but they split the samples for three different treatments, meaning that the deal with samples between 20 and 48 observations. Other studies using in person interviews do have larger samples, from about 600 to 1,000 interviews (i.e. Powe and Bateman, 2003; Veisten et al. 2004 and Bateman et al. 2006).

More studies (preferably with larger samples) would be necessary to confirm such proposition for the case of common-pool resources. Nevertheless, in the light of the mixed evidence found in the literature and the results of our study, it does not seem justified to ascribe order effects and their connection with sensitivity to scope to study design alone (e.g. step-wise versus advanced disclosure, top-down versus bottom-up). The nature of the environmental good under valuation also appears to matter. Further research might also compare systematically stepwise disclosure with advance disclosure in the case of common-pool resources. However, it seems unlikely that further empirical results alone will be able to resolve the questions of order and sensitivity to scope. We also need theoretical analyses of the possible relationships between private and public (altruistic) valuations of common-pool resources with anchoring and order effects, especially since many environmental public goods also provide considerable private benefits (particularly to suppliers).

References

Andersson, H. and Svensson, M. 'Scale sensitivity and question order in the contingent valuation method', Journal of Environmental Planning and Management, (2013) DOI: 10.1080/09640568.2013.839442

Andreoni, J. 'Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving', Economic Journal, Vol. 100, (1990) pp. 464–477.

Arrow, K., Solow, R., Porteny, P. R., Leamer, E. E., Radner, R. and Schuman, H. 'Report of the NOAA Panel on Contingent Valuation' Federal Register, 58 (1993) pp. 4601-4614.

Azqueta, D. and Delacámara, G. 'Ethics, Economics and Environmental Management', Ecological Economics, Vol. 56, (2006) pp. 524-533.

Bateman, I. J., Cooper, P., Georgiou, S. and Poe, G. L. 'Visible choice sets and scope sensitivity: an experimental and field test of study design effects upon nested contingent values' CSERGE Global Environmental Change Working Paper EDM 01-01, Centre for

Social and Economic Research on the Global Environment, University of East Anglia, (2001).

Bateman, I. J., Cole, M., Cooper, P., Georgiou, S., Hadley, D. and Poe, G. L. 'Visible choice sets and scope sensitivity', Journal of Environmental Economics and Management, Vol. 47, (2004) pp. 71–93.

Bateman, I. J., Cole M. A., Georgiou, S. and Hadley, D. 'Comparing contingent valuation and contingent ranking: A case study considering the benefits of urban river water quality improvements', Journal of Environmental Management, Vol. 79, (2006) pp. 221–231.

Bateman, I. J., Cole M. A., Georgiou, S., Navrud, S., Poe, G., Ready, R., Riera, P., Ryan, M. and Vossler, C. 'Economic valuation of policies for managing acidity in remote mountain lakes: examining validity through scope sensitivity testing', Aquatic Sciences, Vol. 67, (2005) pp. 274–291.

Bateman, I. J., Burgess, D., Hutchinson, W.G. and Matthews, D. I. 'Learning design contingent valuation (LDCV): NOAA guidelines, preference learning and coherent arbitrariness', Journal of Environmental Economics and Management, Vol. 55, (2008) pp. 127–141.

Bateman, I. J. and Langford, I. H. 'Budget-constraint, temporal, and question-ordering effects in contingent valuation studies', Environmental and Planning A, Vol. 28, (1996) pp 1215–1228.

Berbel, J., Kolberg, S., and Martin-Ortega, J. 'Assessment of the draft Hydrological Basin Plan of the Guadalquivir river basin (Spain)', International Journal of Water Resources Development, Vol. 28, (2012) pp. 43-55

Boyle, K., Welsh, M. and Bishop, R. 'The role of question order and respondent experience in contingent valuation studies', Journal of Environmental Economics and Management, Vol. 25, (1993) pp. S80–S99.

Carson, R.T. 'Contingent valuation: A practical alternative when prices aren't available', Journal of Economic Perspectives, Vol. 26, (2012) pp. 27- 42

Carson, R.T. Contingent Valuation: A Comprehensive Bibliography and History. (Cheltenham, UK: Edward Elgar Publishing, 2011).

Carson, R. T., Flores, N. E. and Meade, N. F. 'Contingent Valuation: Controversies and Evidence', Environmental and Resource Economics, Vol. 19, (2001) pp.173–210.

Carson, R.T., Mitchell, R.C., Hanemann, M., Kopp, R.J., Presser, S., Ruud, P.A. 'Contingent valuation and lost passive use: damages from the Exxon Valdez oilspill', Environmental and Resources .Economics, Vol. 25, (2003) pp. 257-286.

Carson, R. T. and Mitchell, R. C. 'Sequencing and nesting in contingent valuation surveys' Journal of Environmental Economics and Management, Vol. 28, (1995) pp. 155-173.

Carson, R. T., and Mitchell, R. C. 'The issue of scope in contingent valuation studies', American Journal of Agricultural Economics, Vol. 75, (1993) pp. 1263-1267.

Clark, J. E. and Friesen, L. 'The Causes of Order Effects in Contingent Valuation Surveys: An Experimental Investigation', Journal of Environmental Economics and Management, Vol. 56, (2008) pp. 195–206.

Desvousges, W. H., Johnson, F. R., Dunford, R. W., Hudson, S. P. and Wilson, K. N. 'Measuring natural resource damage with contingent valuation: a test of validity and reliability', In: Hausman, J. (ed) Contingent valuation: a critical assessment (Amsterdam: North Holland Publishers, 1993).

Desvousges, W., Mathews, K. and Train, K. 'Adequate responsiveness to scope in contingent valuation', Ecological Economics, Vol. 84, (2012) pp. 121–128.

Diamond, P.A. and Hausman, J.A. 'Contingent valuation: Is some number better than no number?', Journal of Economic Perspectives, Vol. 8, (1994) pp. 45-64.

Diamond, P., Hausman, J., Leonard, G. and Denning, M. 'Does contingent valuation measure preferences? Experimental evidence', In: Hausman J (ed) Contingent valuation: a critical assessment (Amsterdam: North Holland Publishers, 1993).

Dziegielewska, D.A., Mendelsohn, R. 'Does "no" mean "no"? A protest methodology', Environmental and Resources Economics, Vol. 38, (2007) pp. 71-87.

Gardner, R. 'L. V. Kantorovich: The Price Implications of Optimal Planning', Journal of Economic Literature, American Economic Association, Vol. 28, (1990) pp. 638-48.

Gómez-Baggethun, E., de Groot, R., Lomas, P.L. and Montes, C. 'The history of ecosystem services in economic theory and practice: from early notions to markets and payment schemes', Ecological Economics, Vol. 69, (2010) pp.1209-1218.

Hanemann, W. H. 'The economic conception of water, In: P.P. Rogers, P.P., Llamas, M. R., Martinez-Cortina, L., (eds) Water Crisis: myth or reality? (London: Taylor & Francis plc., 2006).

Halvorsen, B. 'Ordering Effects in Contingent Valuation Surveys. Willingness to Pay for Reduced Health Damage from Air Pollution', Environmental and Resources Economics, Vol. 8, (1996) pp. 485-499.

Hammitt, J.K., and Graham, J. D. 'Willingness to pay for health protection: inadequate sensitivity to probability' Journal of Risk and Uncertainty, Vol. 8, (1999) pp. 33-62.

Harrison, G.W. 'Valuing public goods with the contingent valuation method: a critique of Kahneman and Knetsch', Journal of Environmental Economics and Management, Vol. 23, (1992) pp. 248–257.

Hart, A. 'Mann-Whitney test is not just a test of medians: differences in spread can be important'. Bmj Vol323 (7309) (2001): 391-393

Hatton Donald, D., Morrison, M. D. and Barnes, M. B. 'Willingness to pay and willingness to accept compensation for changes in urban water customer standards', Water Resources Management, Vol. 24, (2010) pp. 3145–3158

Hausman, J. A. 'Contingent valuation: from dubious to hopeless', Journal of Economic Perspectives, Vol. 26, (2012) pp. 43–56.

Hensher, D., Shore, N. and Train, K. 'Households' willingness to pay for water service attributes', Environmental Resources Economics, Vol. 32, (2005) pp. 509–531.

Heberlein, T. A., Wilson, M. A., Bishop, R. C. and Schaeffer, N. C. 'Rethinking the scope test as a criterion for validity in contingent valuation', Journal of Environmental Economics and Management, Vol. 50, (2005) pp. 1–22.

Jones-Lee, M. W., Hammerton, M. and Phillips, P.R. 'The value of safety: results of a national sample survey', The economic journal, Vol. 95, (1985) pp. 49–72

Kahneman, D. and Knetsch, J. L. 'Valuing public goods: the purchase of moral satisfaction', Journal of Environmental Economics and Management, Vol. 22, (1992) pp. 57-70.

Kahneman, D., Ritov, I., and Schkade, D. 'Economic Preferences or Attitude Expressions? An Analysis of Dollar Responses to Public Issues', Journal of Risk and Uncertainty, Vol. 19, (1999) pp. 203–235.

Kahneman, D., Slovic, P., Tversky, A. Judgement Under Uncertainty: Heuristics and Biases. (New York: Cambridge University Press, 1982).

Kallas, Z., Gómez-Limón, J. A., Barreiro-Hurle, J. 'Descomposing the value of agricultural multifuncionality: combining contingent valuation and the analytical hierarchy process', Agricultural Economics, Vol. 58, (2007) pp. 218-241

Longo, A., Hoyos, D. and Markandya, A. 'Testing for ordering effects in a single bounded dichotomous choice contingent valuation survey', 17th Annual Conference European Association of Environmental and Resource Economists, (2009) Amsterdam, NL.

Martin-Ortega, J., Giannoccaro, G. and Berbel, J. 'Environmental and Resource Costs Under Water Scarcity Conditions: An Estimation in the Context of the European Water Framework Directive', Water Resources Management, Vol. 25, (2011) pp. 1615–1633

Mesa-Jurado, M. A., Martin-Ortega, J, Ruto, E. and Berbel, J. 'The economic value of guaranteed water supply for irrigation under scarcity conditions' Agricultural Water Management, Vol. 113, (2012) pp. 10-18.

Mitchell, R.C. and Carson, R.T. Using surveys to value public goods: the contingent valuation method (Washington, DC: Resource for the Future, 1989).

Nielsen, J. S. and Kjær, T. 'Does question Order Influence Sensitivity to Scope? Empirical Findings From a Web-Based Contingent Valuation Study', Journal of Environmental Planning and Management, Vol. 54, (2011) pp. 369–381.

Ostrom, E. and Gardner, R. 'Coping with Asymmetries in the Commons: Self-Governing Irrigation Systems Can Work', Journal of Economic Perspectives, Vol. 7, (1993) pp. 93-112.

Payne, J. W., Schkade, D. A., Desvousges, W. H. and Aultman, C. 'Valuation of multiple environmental programs', Journal of Risk and Uncertainty, Vol. 21 (1), (2000) pp. 95–115.

Powe, N. A. and Bateman, I. 'Ordering Effects in Nested 'Top-Down' and 'Bottom-Up' Contingent Valuation Designs', Ecological Economics, Vol. 45, (2003) pp. 255–270.

Ready, R. C., Navrud, S., Dubourg, W. R., 'How do respondents with uncertain willingness to pay answer contingent valuation questions?' Land Economics, Vol. 77, (2001) pp. 315-326.

Schläpfer, F. 'Contingent valuation: a new perspective' Ecological economics, Vol. 64, (2008) pp. 729-740.

Slovic, P. 'The construction of preference', American Psychologist, Vol. 50, (1995) pp. 364-371.

Spangenbergh, J. H. and Settele, J. 'Precisely incorrect? Monetising the value of ecosystem services', Ecological Complexity, Vol.7, (2010) pp. 327-337.

Veisten, K., Hoen, H. F. and Strand, J. ‘Sequencing and the adding-up property in contingent valuation of endangered species: Are contingent non-use values economic values?’ Environmental and Resources Economics, Vol. 29, (2004) pp. 419-433.

Venkatachalam, L. ‘The Contingent Valuation Method: A Review’, Environmental Impact Assessment Review, Vol. 24, (2004) pp. 89-124.

ANNEX 1: Literature on order effects in contingent valuation using exclusive lists of goods

Study	Type of CV	Sample size and elicitation type	Environmental goods	Disclosure method: AD=Advance disclosure SD=Stepwise disclosure	Payment vehicle	Order of WTP questions	Results of tests for order effects	Analysis of influence of order effects on sensitivity to scope (yes/no)
Boyle et al., 1993	Single bounded dichotomous choice	506 mail survey	Recreational value (white-water boating) of the Colorado River (US) under different flow levels	Not clear (mail survey)	An increase in trip expenditure	8 CV questions (1 for their actual trip and 7 involving hypothetical scenarios). Split sample: reversed order of the 7 hypothetical scenarios	No evidence of order effects for experienced boaters; potential order effects for non-experienced boaters.	No
Powe and Bateman, 2003	Double bounded dichotomous choice	1,592 in-person interviews	Riverine saline flood alleviation scheme in the Norfolk Broads (UK) Preservation of wetlands: 5 small areas (a set of P goods) included in a large area (the W good)	SD	An increase in general taxation	Split sample: bottom-up vs. top-down	Do find order effects (except in D area): values for the smaller goods are highest when that good is presented first in a list, while values for the larger good are larger when it is presented second in a list	They test sensitivity to scope but do not include further discussion on the issue
Bateman et al., 2004	Open-ended	Lab trial: 149 students in-person interviews	Improvements to an open access lake located within the grounds of the University of East Anglia through 3 nested schemes: F (filter), P (plants), D (dredging) 1. WTP (F) 2. WTP (F+P) 3. WTP (F+P+D)	Split sample: AD and SD	Increases in rental charges to campus shops which would result in higher prices for students	2x2 split sample: bottom-up vs. top-down and AD vs. SD	SD: WTP for the small good is consistently higher when asked first; WTP for the large good is higher if asked first in 1 out of 3 tests AD: procedural invariance cannot be rejected	Yes. The degree of scope sensitivity changes with the order in the stepwise sub-sample, but not in advance disclosure sub-sample.
	Not specified in paper. Assumed it is open ended	Field trial: 139 in-person interviews	1: improve River Tame water quality in Birmingham, UK: 3 nested schemes: small, medium and large improvement 2: prevent saline flooding within the Norfolk Broads, UK (wetland): 2 nested schemes: PA (protected a part of overall Broads area) and WA (protected the whole area)	AD (1) and SD (2)	An annual increase in council tax (1) and an increase in general taxation (2)	Split sample: bottom-up vs. top-down	1:No order effects 2: WTP for the small good is higher if asked first WTP for the large good does not depend on order	Yes. 1: The degree of scope is not affected by the order 2: Equivocal results: order affects degree of sensitivity in median WTP, but not mean WTP

Study	Type of CV	Sample size and elicitation type	Environmental goods	Disclosure method: AD=Advance disclosure SD=Stepwise disclosure	Payment vehicle	Order of WTP questions	Results of tests for order effects	Analysis of influence of order effects on sensitivity to scope (yes/no)
Veisten et al., 2004[a]	Half sample open-ended. Half sample payment card	1,019 in-person interviews	Preservation of all and a sub-set of endangered species in Norwegian forests: (A) Package of environmental preservation; (b) Protection of endangered cryptogams; (c) Protection of endangered bird species; (P) Multi-package of 6 environmental projects	No information provided in the paper, but appears SD	A tax on wood products	4 sub-samples. Sub-sample I valued A directly. Sub-sample II valued (b) before A. Sub-sample III first valued (c), then sub-sample III valued (b), and then finally the valued A. Sub- sample IV had a different extra introduction about a multi-package of six environmental tasks (P).	The smaller goods (sub-set) obtained higher values when asked first	No
Bateman et al., 2006	Open-ended	675 in-person interviews	Improving River Tame water quality in Birmingham, UK: 3 nested schemes: small, medium and large improvement	AD	An annual increase in council tax	Split sample: bottom-up vs. top-down	No order effects	No
Clark and Friesen, 2008	Single bounded dichotomous choice	235 in-person interviews	Lab experiments where participants are asked to value sequences of nested good for purchase using an incentive-compatible mechanism a) Induced value goods b) Actual private goods c) Identical private goods that are to be donated to charities Goods are: orange juice; a pack of pens and corrector tape; and a camera.	SD	Part of the endowment given in the experiment	Split sample: bottom-up vs. top-down (for each treatment a, b and c)	Order effects for private goods when valued for own use, but not when valued for donation. No order effects for induced-value goods	Yes. The order effects found for private goods in the valuation are not sufficient to affect the degree of sensitivity to scope

Study	Type of CV	Sample size and elicitation type	Environmental goods	Disclosure method: AD=Advance disclosure SD=Stepwise disclosure	Payment vehicle	Order of WTP questions	Results of tests for order effects	Analysis of influence of order effects on sensitivity to scope (yes/no)
Longo et al., 2009[b]	Single bounded dichotomous choice	1,000 in-person interviews	WTP for three programs aimed at decreasing the greenhouse gases of the in a region of North of Spain. WTP RE (promotion of Renewable Energy production) WTP EE (Energy Efficient measures) WTP BPCCC (combination of EE+RE)	AD	New taxes	Split sample: bottom-up vs. top-down	Order effects are present, but disappear when respondents are given the possibility to revise their WTP at the end	Results pass the scope test but no further analysis/interpretation is provided
Nielsen and Kjær, 2011	Open-ended WTP question following a card-sorting	1,559 web-based survey (internet panel)	WTP for 2 independent gains in life expectancy caused by a change in air pollution (three months and six months)	AD	An increase in the cost of living for the rest of their lives	Split sample: bottom-up vs. top-down	WTP for the small good is higher if asked first WTP for the large good does not depend on order	Yes. Their results demonstrate an order effect which generates an observed difference in the degree of scope sensitivity
Andersson and Svensson, 2013	Open-ended	200 undergraduate students	WTP to reduce mortality risk of being involved in a fatal bus accident (two levels of risk mortality)	AD	An increase in bus fares	2x2 split sample: bottom-up vs. top-down and two different initial bus fares to test for anchoring effects	Absence of order effects cannot be rejected	Yes They conclude that the bottom-up treatment is related to answers closer to near-proportionality between the large and small mortality risk reduction

Source: Own elaboration, adapted and expanded from Longo et al. (2009). Longo et al. only look at the literature on ordering effects and not specifically on how do they affect sensitivity to scope. Like the rest of this manuscript, the table is focused on exclusive lists of goods. Desvousges et al. (2012) can be consulted for a review of CV studies conducing scope tests.

[a]Verstein et al. (2004) refer to 'sequence effects' but our understanding from reading the paper is that they are valuing goods in an exclusive list.

[b]In Longo et al. (2009) a mix of inclusive and exclusive lists of goods is analysed.

ANNEX 2: Payment card