

Scope Effects in Contingent Valuation: Does the Assumed Statistical Distribution of WTP Matter?



Nicolas Borzykowski^{a,b,*}, Andrea Baranzini^a, David Maradan^a

^a Haute École de Gestion Genève (HEG-Ge), University of Applied Sciences Western Switzerland (HES-SO), 17 rue de la Tambourine, Carouge-Geneva CH-1227, Switzerland

^b Institute of Economic Research, University of Neuchâtel, rue A.-L. Breguet 2, Neuchâtel CH-2000, Switzerland

ARTICLE INFO

JEL classification:

Q23

Q57

Keywords:

Contingent valuation

Scope effects

Forests

Scope elasticities

Plausibility

ABSTRACT

Economic theory assumes that willingness to pay (WTP) increases with the quantity of the consumed good. This implies that there should be a scope effect in contingent valuation studies. However, in previous issues of *Ecological Economics*, several authors criticized the contingent valuation (CV) method for the absence of such effect or its inadequacy. In this paper, we contribute to this ongoing debate by proposing to systematically apply several WTP statistical distribution assumptions to test for scope effects and check its plausibility, following Whitehead's (2016) recent recommendations. We perform this approach using data from a Swiss case study assessing the WTP for an increased surface of forest reserves. We find that both mean WTP and scope effects are sensitive to the statistical distribution assumption. Regarding plausibility, scope elasticities provide mixed result and also depend on the assumed statistical distribution of WTP. For small sample size CV studies, a non-parametric analysis, a spike model or an open-ended format can thus be better suited to reveal scope effects than the classical parametric dichotomous choice analysis. We thus recommend to systematically apply several statistical distribution assumptions of WTP to test for scope effects and their plausibility.

1. Introduction

The supposed insensitivity of the willingness to pay (WTP) to the scope of the assessed good in stated preferences methods, and all the more in contingent valuation (CV) studies, is a hot controversy in the literature. Kahneman and Knetsch (1992) were the first to find no significant impact of scope in their CV study, thus casting fundamental doubts on this valuation method. Indeed, if Kahneman and Knetsch (1992, p.1) were right in assuming that “contingent valuation responses reflect the WTP for the moral satisfaction of contributing to public goods, not the economic value of these goods”, then the CV method would be fundamentally flawed and could not be used in valuation studies. However, Carson and Mitchell (1993a) reviewed Kahneman and Knetsch's (1992) study, as well as other CV studies, and observed that the median is always higher for the WTP related to the bigger scope. Also, several studies (e.g. Smith and Osborne, 1996; Carson, 1997; Bandara and Tisdell, 2005) have indicated that a scope effect is detected when correcting WTP for the difference in sub-samples

characteristics and applying the appropriate significance test, even with inexpensive survey methods such as face-to-face interviews (Whitehead et al., 1998). Meta-analyses have also shown that the scope of the assessed good has an impact on WTP (Richardson and Loomis, 2009; Ojea and Loureiro, 2011; Hjerpe et al., 2015). Recently, Bishop et al. (2017a) included two scopes of avoided injuries and found evidence of responsiveness to scope in WTP studies to avoid oil spills like the 2010 Deepwater Horizon in the Gulf of Mexico. However, the WTP difference is quite small relative to the hypothetical change in scope and their results are discussed as well (Baron, 2017; Bishop et al., 2017b).

Since the NOAA panel guidelines (Arrow et al., 1993), testing scope effects should be part of the standard validity tests for a contingent valuation survey. However, the recent debate between Chapman et al. (2016) and Desvousges et al. (2016), the contribution of Whitehead (2016) in *Ecological Economics* and Hausman (2012), clearly demonstrate that scholars have not yet reached a consensus on the issue of “adequate” scope responsiveness. Moreover, the general suspicion that the CV method simply does not “pass” the classical scope test still casts

* The authors would like to thank the staff and students of the Laboratory of Market Studies (LEM) at the Haute Ecole de Gestion de Genève (HEG-Genève) for their valuable support for the survey, members of the Geneva green spaces service for their comments, Milad Zarin-Nejadan and Bengt Kriström for their precious suggestions and four anonymous reviewers, whose suggestions considerably helped in improving the paper. We bear the sole responsibility of all remaining errors. Financial support from the Swiss National Research Program 66 “Resource wood” is acknowledged.

* Corresponding author.

E-mail address: nicolas.borzykowski@hesge.ch (N. Borzykowski).

doubt on the very CV method (Hausman, 2012). Indeed, studies that do not robustly detect scope effects are not rare (see e.g. Boyle et al. (1994), McFadden (1994), Veisten et al. (2004), Desvousges et al. (2012) and Frontuto et al. (2017)) and the lack of systematic scope tests may also induce distrust for stated preferences methods in general (see Heberlein et al., 2005).

Although, as recommended in the literature the tests of adequacy or plausibility are legitimate, in this paper we maintain that one should firstly address the basic issue of statistically significant scope responsiveness. We thus deal with the following question: Does the assumed statistical distribution of WTP matter to detect scope effects? We argue that, while sensitivity of mean WTP estimates to distributional assumption is acknowledged (Bengochea-Morancho et al., 2005), split-sample comparisons, identification of scope effects and the result of plausibility tests can also be affected. Because of the lower number of degrees of freedom and the higher variability, this is even more pronounced in small samples and valuing complex environmental amenities (Veisten et al., 2004).¹

We apply a split-sample CV survey to measure the scope effects on the WTP and its plausibility, for a program aiming at increasing the surface of forests reserves in Switzerland to protect fauna and flora. Since the program implies some access restriction to recreationists, we fundamentally assess non-use values of these reserves. We analyze this issue using different econometric approaches and ad-hoc external scope tests: (i) we begin by calculating mean WTP estimates, accounting for sub-samples differences in characteristics with several parametric models and hence different WTP statistical distribution assumptions; (ii) we then add information regarding “real zeros” and carry out a parametric spike model (Kriström, 1997) with an asymmetric distribution; (iii) afterwards, we compute a non-parametric estimation, free of WTP distribution assumption; and finally (iv) we analyze the answers of an open-ended follow-up question. To test the plausibility of the responsiveness to scope, we follow Whitehead’s (2016) recent recommendation and thus calculate the scope arc elasticities.

Our results show that individuals are, on average, willing to pay more for larger size (Swiss) forests than for smaller (Geneva) forests. However, the significance level of this difference is largely affected by the assumed statistical distribution regarding WTP. Indeed, parametric estimations from the dichotomous choice elicitation format, such as log-logistic and log-normal, fail to detect statistically significant scope effects; results are mixed with the logistic distribution and depend on the inclusion of covariates; and finally non-parametric models, spike models and open-ended estimates robustly reveal significant scope effects. The test of plausibility proposed by Whitehead (2016) does not rule out nor confirm the appropriateness of our scope responsiveness, but the outcome of the test is also largely affected by the assumed statistical distribution of WTP. We thus suggest that CV studies have to test for scope effects and their plausibility using several distribution assumptions of WTP, especially in small samples studies.

We introduce the CV method and our questionnaire in Section 2. The empirical approach and descriptive statistics are provided in Section 3. Section 4.1 presents the results from different parametric distributions for the dichotomous choice format and different tests for scope effects. Section 4.2 analyzes the results from the Turnbull non-parametric estimation, while Section 4.3 uses the open-ended follow-up question to provide WTP estimates. We estimate scope elasticities in Section 4.4. Section 5 discusses and concludes.

2. Survey Design

The CV questionnaire is composed of three parts. The first aims at

¹ Small samples are not rare in the CV literature. In the meta-analysis of Meshrey et al. (2014), 39 out of the 105 estimates elicited with the CV method and dichotomous choice are drawn from less than 300 observations. For these studies, scope effects would be even harder to detect.

understanding preferences on, knowledge of and behavior in forests. The second part is the contingent valuation of the proposed increase in forest reserves. The third part gathers individuals’ socioeconomic characteristics.² It is worth noting that, in general, Swiss forests are in good health and growing. However, in some regions human activities are in conflict with biodiversity.

We base the scenario for the contingent valuation on an actual Federal program, part of the Swiss Forest Policy 2020 (FOEN, 2013), as recommended in Arrow et al. (1993). This program aims principally at fostering endangered biodiversity by increasing the surface of protected forests, from 5% to 10% of the total forest surface³. Transforming productive forests in forest reserves involves opportunity costs for the forest industry, as well as some access restrictions to recreationists.

To analyze the existence of a scope effect, we use the split-sample approach as in Berrens et al. (2000): To a sub-sample (the “Swiss forest sub-sample”, CH) composed of 228 individuals randomly picked in the whole population, we ask if the respondent’s household would be willing to pay an amount of CHF X to increase the surface of protected forests in Switzerland.

We then administer to another sub-sample from the same population (the “Geneva forest sub-sample”, GE) exactly the same question, but referring to Geneva forests only. To avoid part-whole bias (Whitehead et al., 1998), an issue emerging when individuals believe that the program will apply to a larger scale, we remind the Geneva forest sub-sample that the program will apply to Geneva forests only. It is worth noting that Geneva forests account for about 0.2% of total forest surface in Switzerland, so that the difference in scope between the two versions of the program is substantial. Geneva being part of Switzerland, we consider the two program versions to be perfectly embedded, as recommended in Kahneman and Knetsch (1992). The Geneva program is indeed geographically nested in the Swiss program.⁴ However, this assumption can be discussed. Indeed, our programs would not be perfectly embedded if individuals do not believe that their money will be used the same way in Geneva and in Switzerland and if they have a different perception of Cantonal vs. Federal taxes. However, a Greek study (Remoundou et al., 2012) has found a non-significant impact of the providing institution on the WTP, even with low and significantly different trust levels. In addition, in Switzerland, trust in the institutions is relatively high at the Cantonal and Federal level⁵, and thus this issue should not be very relevant.

We build the question as an advisory referendum, specifying that the results would be used to implement the policy. To reduce the hypothetical bias (Hausman, 2012), we follow an ex ante approach (Loomis, 2011) and thus insist on consequentiality of the survey (Carson and Groves, 2007), with a reminder that the respondent might concretely contribute to the program. We also follow Mitchell and Carson (2013) and Kotchen and Reiling (1999) and add an income constraint reminder to make the respondent aware of the opportunity costs she faces.

WTP is thus elicited through the Single-Bounded-Dichotomous-Choice (SBDC) approach, as is recommended by the NOAA panel (Arrow et al., 1993) and by most recent studies (Carson and Groves, 2007; Mueller, 2014; Bishop et al., 2017a), because of its incentive compatibility, its ability to avoid non-response and outliers and its lower cognitive burden (Bateman et al., 2002). Indeed, by asking if the

² A travel cost analysis and other information about the survey results are available in Baranzini et al. (2015). The questionnaire is available upon request.

³ The use of absolute instead of relative changes could have improved the understanding of the contingent question, which might have led to a more robust scope effect identification (Ojea and Loureiro, 2011). However, the use of relative changes is not rare in CV and can be better suited to value complex environmental amenities. The Federal Policy is also expressed in relative terms.

⁴ It is worth noting that, given the small size of Switzerland, each inhabitant lives relatively close to a forest and that all Swiss regions contain wide forest areas.

⁵ According to the World Value Survey, 65.1% of the Swiss inhabitants trust the government in 2009 (WVS, 2017).

respondent is willing to purchase the service at a given price, this format allows mimicking a market-type situation and is thus relatively simple to understand. Furthermore, as Swiss people are often consulted for referenda, this type of question seems particularly appropriate in our context.

We use a Federal or Cantonal lump-sum tax as payment vehicle, because of forests public good characteristics: as the benefits of forest reserves are non-rival, the appropriate payment vehicle must request contribution from everyone. The off-site survey also requires a payment vehicle that includes forests non-users. In a CV survey on tropical forests, Baranzini et al. (2010) indeed show that a tax provides a higher WTP than a voluntary payment in a forest fund, the latter being subject to free-riding.

With the SBDC approach, a selection of tax amounts (bids) is randomly assigned between respondents. As Haab and McConnell (2002) mention, the selection of bids is of particular importance. A carefully selected bid vector can considerably improve the efficiency of WTP estimates. However, the optimal bid vector can only be designed if the true WTP distribution is known. But, obviously, if true mean WTP is known, there is no reason to derive an optimal bid vector (Haab and McConnell, 2002, p. 129). An exhaustive literature review and a meta-analysis of 47 CV studies on forest in developed countries published between 1993 and 2014 (Meshreky et al., 2014) revealed that the mean WTP for forest programs ranges from USD (2011, PPP) 0 to 650 per household, per year, with a mean at 100 and a median at 68. We therefore select our 6 bids (CHF 10, 60, 100, 250, 500 and 1000)⁶ accordingly and confirm them by a preliminary open-ended qualitative questionnaire discussed in focus groups, as recommended in Kanninen (1993) and Haab and McConnell (2002).

After the referendum question, we administer an open-ended follow-up question, which consists of asking maximum WTP for the program, as Garcia et al. (2007). We use this approach to compute an open-ended estimate of WTP. This estimate, however, suffers from the incentive incompatibility issue (Carson and Groves, 2007). Indeed, because of the public good nature of the program, respondents usually tend to strategically understate their true WTP with open-ended questions. The bid proposed in the referendum question also creates anchoring, which pushes the maximum WTP towards the proposed bid.

We design a second follow-up question to distinguish protest bids (Jorgensen and Syme, 2000) from “real zeros”. If the answer to the previous question is zero, then the respondent has to state the reason in a closed-ended question. We identify protesters if the reason for not contributing is unrelated to the value of forest⁷. Other reasons are considered “real zeros”.

3. Empirical Approach and Descriptive Statistics

Based on the Random Utility Model (McFadden, 1974), first applied by Hanemann (1984) for CV, the individual's probability to answer “Yes” to the bid can be modeled by the bid itself, a vector of explanatory variables and an error term, and estimated with a binary model. However, before estimating the model, the statistical distribution of WTP still has to be assumed. This assumption can be made using different plain parametric models (Bishop and Heberlein, 1979), such as probit, logit, log-logit, log-normal or mixture models such as the spike model (Kriström, 1997). This choice is not necessary when using non-parametric models, since the latter do not require to assume any particular WTP distribution (Kriström, 1990). The logit model is the most frequently used model, because the calculation of mean WTP is

easier (Bateman et al., 2002) and it is often confirmed with a non-parametric approach. For example, Bishop et al. (2017a), use a probit model and confirm the estimations by a non-parametric model.

We test different distribution assumptions, by using several parametric and non-parametric models, either including covariates or not. Adding independent variables to the models should not change mean WTP as the latter is evaluated at covariates mean. However, it allows to control for heterogeneous characteristics of split-samples, which may, in our case, have a different effect on mean WTP. The inclusion of covariates in the model thus allows to better extrapolate the WTP from one sample to another, which enables estimating what respondents from one sub-sample would be willing to pay for the program proposed to the other sub-sample. Covariates may also help in decreasing unobserved heterogeneity, thus decreasing the variance. The final model choice is thus based on differences in split-samples characteristics and intuition regarding the effect of variables on bid acceptance.⁸ The sample is composed of 419 independent observations from which the sub-sample GE has 191 and CH 228 observations. The final matrix of explanatory variables is described in subsection A, in the appendix.

As Table 1 shows, each bid has been proposed to 16 persons minimum. The acceptance rates, unsurprisingly, decrease with the bid amount in both sub-samples and, while the acceptance rate of the Geneva program falls to 6% for a cost of CHF 1000, it reaches 18% for the Swiss program for the same cost. This may indicate that the range of the bid vector could have been wider for the Swiss program. We observe a non-monotonous decrease in the acceptance rates in both sub-samples, which is not a particularly rare result with small sample sizes (Kriström, 1990; Bateman et al., 2002).

Protest rates are stable across bids⁹ and are not significantly different across sub-samples.¹⁰ It is interesting, but not surprising, to observe that the mean maximum WTP elicited with the open-ended follow-up question usually increases with the bid proposed for the Swiss program, revealing an anchoring effect.

4. Results

4.1. Parametric Estimations

The use of parametric modeling techniques implies to impose a statistical distribution of WTP. Normal and logistic distributions are often used because of their relative ease to handle. However, these distributions suffer from an important drawback: since they are defined symmetrically over $]-\infty; +\infty[$, it includes the possibility of negative or infinite willingness to pay. As Bateman et al. (2002) mention, if an individual does not value the improvement in the provision of the good, we expect a zero WTP. A negative WTP is acceptable only if the program can be considered as a deterioration (see Boman and Bostedt

⁸ Because the low response rate for the income question drastically reduces the number of observations, we do not include any income variable in the final model. For information, we provide the descriptive statistics of income in both sub-samples in Table A.1 in the appendix. While income and its distribution could be drivers of WTP (Baumgärtner et al., 2017), there is no reason to think that it may impact scope. Indeed, mean income and income distribution in both sub-samples are not statistically different.

⁹ With the exception of the relatively low protest rate of the 500 bid in CH and the 10 bid in GE.

¹⁰ Protest bids should not be removed unless the sub-sample of protest bidders reveals the same characteristics as other respondents (Halstead et al., 1992). In the CH sub-sample, protesters are on average significantly older, they are less member of environmental associations, live in a more urban environment and go less often to forests than non-protesters. In the GE sub-sample, protesters are more often men and live in a more urban environment. Protesters characteristics are apparently different from non-protesters and dropping those observations may hence bias our estimates because of self-selection. However, considering them as real zeros would create a downward bias. There is no universally acknowledged simple method to deal with protest bids (see Strazzera et al. (2003)). In the following analysis we hence exclude protesters but, in terms of scope effects, results are similar when protesters are included as real zeros. Results including protesters are available upon request.

⁶ On December 31st 2014, CHF 1 = EUR 1.20 = USD 0.98.

⁷ “I would like more information on this topic before accepting paying this amount”, “I do not trust the state. I have no guarantee that the money will indeed be used to finance the program.”, “I already pay enough taxes”, “Forest is a public good, so it is not reasonable to ask me to pay for it.” and “The state, the forests owners or forest users should pay, not me.”

Table 1
Answers structure to bids.

| | | 10 | 60 | 100 | 250 | 500 | 1000 | Total |
|----|---------------------------------|---------|---------|----------|----------|----------|----------|----------|
| CH | Yes | 26 | 22 | 18 | 12 | 6 | 6 | 90 |
| | No | 16 | 12 | 24 | 29 | 29 | 28 | 138 |
| | (incl. protester) | (9) | (7) | (10) | (12) | (2) | (5) | (45) |
| | (incl. real zeros) | (7) | (4) | (8) | (9) | (11) | (10) | (49) |
| | N | 42 | 34 | 42 | 41 | 35 | 34 | 228 |
| | Acceptance rate | 0.62 | 0.65 | 0.43 | 0.29 | 0.17 | 0.18 | 0.39 |
| | Protest rate | 0.21 | 0.21 | 0.24 | 0.29 | 0.06 | 0.15 | 0.20 |
| | Follow-up max. WTP ^a | 23.44 | 54.07 | 87.81 | 124.48 | 171.67 | 303.10 | 126.84 |
| | (Std. dev.) | (34.60) | (27.63) | (110.85) | (115.22) | (226.11) | (386.98) | (212.07) |
| GE | Yes | 34 | 25 | 19 | 3 | 3 | 1 | 85 |
| | No | 6 | 15 | 20 | 37 | 13 | 15 | 106 |
| | (incl. protester) | (3) | (7) | (6) | (12) | (6) | (5) | (39) |
| | (incl. real zeros) | (3) | (7) | (9) | (8) | (2) | (1) | (30) |
| | N | 40 | 40 | 39 | 40 | 16 | 16 | 191 |
| | Acceptance rate | 0.85 | 0.63 | 0.49 | 0.08 | 0.19 | 0.06 | 0.45 |
| | Protest rate | 0.08 | 0.18 | 0.15 | 0.30 | 0.38 | 0.31 | 0.29 |
| | Follow-up max. WTP ^a | 46.24 | 40.50 | 60.40 | 84.68 | 203.00 | 141.82 | 79.30 |
| | (Std. dev.) | (54.87) | (34.10) | (54.88) | (83.02) | (216.75) | (136.59) | (101.90) |

^a Mean of the follow-up open-ended question about maximum WTP, excluding protest answers.

(1999) and the wolves example). In addition, an individual's WTP shall not be higher than her income and WTP should thus lie in the interval $[0 : y]$ (with y the income) in most cases. Asymmetric distributions such as the log-normal or log-logistic can take care of the negative WTP issue but cannot rule out infinite WTP. Mixture models such as spike models or truncated models could also be used and are recommended in Bateman et al. (2002).¹¹

Since there is no consensus on the statistical distribution and given that we are more interested in testing the scope effects and its robustness across distributions, rather than in the value of WTP, we decide to run different parametric models (logistic, log-normal, log-logistic) on our split-samples and apply different tests for scope effects such as the Z-test, the non-overlapping confidence intervals test (Park et al., 1991), and the Complete Combinatorial (CC) test (Poe et al., 2005). Since logistic and normal distributions usually give similar results we rule out the normal distribution thanks to the prediction accuracy. The logistic distribution is modeled by a logit, the log-normal and log-logistic distributions are, as usual, computed with probit and logit models respectively and applying the logarithm on the Bid variable.

Full estimation results are presented in Table B.1 in the appendix. We observe that the coefficients associated with the Bid variables are always significantly negative, which was expected. All remaining the same, the probability of accepting the bid thus significantly decreases with its amount in both sub-samples. Effects of covariates are commented in subsection B, in the appendix.

Before calculating WTP estimates, we check whether both sub-samples respond differently to the bid proposed and if coefficients are similar across sub-samples. We therefore run pooled models on top of models on split-samples and test for poolability using the Likelihood Ratio test (LR) as in Berrens et al. (2000) or Veisten et al. (2004). We find that sub-samples should not be pooled, which confirms our split-sample methodology¹².

As Poe et al. (1994) stress, if mean WTP does not reveal any scope effects, one should anyway check if statistical distributions are different. An analysis of mean, median and WTP distribution is therefore necessary to analyze scope effects. The calculation of WTP central

tendency (mean and median) depends on the distributional assumption (see Aizaki et al., 2014).

Mean and median WTP resulting from parametric estimations are presented in Table 2.

As expected, estimates are very sensitive to the distributional assumption and range from CHF 277 to 2064 for the Swiss program and CHF 183 to 335 for the Geneva program. WTP arising from log-distributions are higher, because of the asymmetry and the right skewness of these distributions. This is a standard result, as acknowledged in Bengochea-Morancho et al. (2005).

According to these parametric approaches, we find that respondents from our sample are willing to pay on average more for new protected areas in Swiss forests than for the same program, but applied to Geneva forests only. However, using Z-tests, the difference is statistically significant for the logit model with covariates only. The difference in median WTP values for log-distributions do not reach the significance level either.

Similarly, the non-overlapping confidence interval method (Park et al., 1991) using Krinsky and Robb (1986) confidence intervals at 95% does not reveal any significant scope effect, since the major part of the intervals overlap for each WTP distribution.

A complete combinatorial approach, as proposed in Poe et al. (2005) aims at testing the difference between two distributions. This methodology requires Krinsky-Robb simulation technique. We simulate 1000 replications of WTP for both sub-samples and subtract each possible combination of these WTP. The proportion of positive differences can be interpreted as the p-value for $H_0: WTP_{CH} = WTP_{GE}$. This test rejects H_0 with 90% confidence for the log-normal models with and without covariates. In all other models H_0 is not rejected, and the test concludes that no scope effects are observed.

To check if differences in samples characteristics play a role in the determination of scope effects, we follow Carson and Mitchell (1993b)'s procedure and evaluate WTP for the Geneva program at CH sub-sample covariates mean (\overline{CH}), to get an estimate of what respondents from the CH sub-sample would be willing to pay for the Geneva program. The results are shown in rows GE at \overline{CH} . When differences in sub-samples characteristics are taken into account, estimates display similar results as if this difference was not corrected for. Thus, no significant scope effects can be observed in this case either, except for the logistic distribution. We can therefore conclude that the correction for sub-samples differences in characteristics does not help in revealing scope effect and that these differences do not play any significant role in its determination in this case.

¹¹ In this paper, we do not artificially truncate the parametric distributions. However, results regarding scope effects are similar with a truncation point above CHF 1000.

¹² The LR test is written as $LR = -2[\ln L_{Pooled} - (\ln L_{CH} + \ln L_{GE})] \sim \chi^2(10)$, where $\ln L_{Pooled}$ is the log-likelihood from the pooled model, $\ln L_{CH}$ the log-likelihood from the CH model, $\ln L_{GE}$ the log-likelihood from the GE model. The test statistic follows a χ^2 with 10 degrees of freedom, the number of equality restrictions.

Table 2
WTP estimates from the plain parametric distributions.

| | Scope | Mean WTP | Std. Err. | CI ^a | ΔCH-GE | Med. WTP | Std. Err. | CI ^a | ΔCH-GE | Covariates | N |
|--------------|----------|-------------------|-----------|-----------------|---------|-------------------|-----------|-----------------|--------|------------|-----|
| Logistic | CH | 276.69 | (53.97) | [175; 421] | 93.57 | n.a. ^b | - | - | - | No | 183 |
| | GE | 183.11 | (42.99) | [125; 540] | | - | - | - | - | No | 152 |
| | CH | 303.02 | (51.45) | [179; 442] | 120.68* | n.a. ^b | - | - | - | Yes | 183 |
| | GE | 182.34 | (38.22) | [120; 420] | | - | - | - | - | Yes | 150 |
| | GE at CH | 188.17 | (45.67) | [120; 420] | 114.85* | - | - | - | - | Yes | 150 |
| Log-normal | CH | 2064.38 | (1990.39) | [582; 64939] | 1729.73 | 132.26 | (31.23) | [79; 216] | 21.32 | No | 183 |
| | GE | 334.65 | (142.99) | [186; 1417] | | 110.94 | (19.47) | [78; 164] | | No | 152 |
| | CH | 1578.97 | (1344.96) | [542; 53227] | 1298.42 | 148.04 | (33.95) | [90; 249] | 40.22 | Yes | 183 |
| | GE | 280.55 | (107.90) | [167; 1271] | | 107.82 | (17.94) | [75; 160] | | Yes | 150 |
| | GE at CH | 289.05 | (124.17) | [167; 1271] | 1289.92 | 111.09 | (24.11) | [75; 160] | 36.95 | Yes | 150 |
| Log-logistic | CH | n.a. ^c | - | - | - | 136.08 | (31.87) | [81; 220] | 24.08 | No | 183 |
| | GE | - | - | - | - | 112.00 | (18.48) | [80; 163] | | No | 152 |
| | CH | n.a. ^c | - | - | - | 151.25 | (35.64) | [89; 257] | 40.49 | Yes | 183 |
| | GE | - | - | - | - | 110.76 | (17.62) | [77; 165] | | Yes | 150 |
| | GE at CH | - | - | - | - | 111.48 | (23.64) | [77; 165] | 39.77 | Yes | 150 |

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.
 Std. err. computed with the Delta Method in parenthesis.
^a Krinsky and Robb CI at 95%, computed with 1000 replications.
^b Logistic distributions are symmetrical and hence mean WTP equals the median.
^c Mean WTP is undefined for $\sigma > 1$ for the log-logistic.

Table 3
WTP estimates from the spike log-normal distribution.

| | CH | GE | CH | GE |
|------------------|------------|-----------|------------|-----------|
| Cond. mean WTP | 832.56 | 294.88 | 746.96 | 234.75 |
| (Std. err.) | (342.37) | (82.57) | (259.32) | (27.69) |
| CI ^a | [462;3025] | [206;714] | [432;2742] | [209;407] |
| ΔCH-GE | 538.12 | | 512.21** | |
| Uncond. mean WTP | 609.63 | 236.68 | 546.95 | 188.42 |
| (Std. err.) | (250.70) | (66.28) | (171.87) | (22.23) |
| ΔCH-GE | 372.95 | | 358.53** | |
| Cond. med. WTP | 336.66 | 198.49 | 388.64 | 202.85 |
| (Std. err.) | (65.57) | (32.18) | (77.46) | (21.23) |
| ΔCH-GE | 138.17* | | 185.79** | |
| Covariates | No | No | Yes | Yes |
| Observations | 134 | 122 | 134 | 120 |

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.
 Std. err. computed with the Delta Method.
^a Krinsky and Robb CI at 95%, computed with 1000 replications.

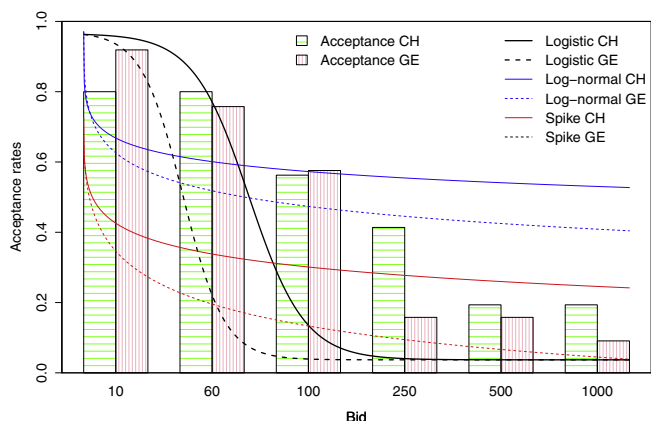


Fig. 1. Acceptance rate and assumed WTP distributions (estimated with covariates).

Krström (1997) proposes a spike model that split real-zeros and positive WTP in two groups. This distribution thus has two parts: a spike at zeros to account for excess real-zeros and an asymmetric distribution for strictly positive bidders. The advantage of this model is that, by allowing a distributional break at 0, it may better fit the real distribution and therefore better reveal scope effect, assuming there is

one. Coefficients from the log-normal model applied on positive bidders are presented in Table B.2 in the appendix¹³. They will allow to calculate a mean WTP conditioned on positive WTP.

We obtain the unconditional mean WTP by multiplying the conditional mean by the proportion of positive bidders. WTP resulting from the spike log-normal model are presented in Table 3 and show an important difference between the mean WTP for the programs. When adding covariates, this difference is significant for conditional and unconditional WTP estimates. Furthermore, the CC approach rejects the hypothesis of same distribution with and without covariates. By giving more information, the spike model with log-normal distribution on positive bidders thus reveals scope effects, even with a lower number of observations, contrary to the “plain” log-normal model. This again proves that one should test more sophisticated parametric models and try to better fit the real WTP distribution, particularly with small samples.

Fig. 1 represents the acceptance rate for each bid (excluding testers) and the estimated survival function of the assumed WTP distributions. We observe that none of the assumed statistical distribution perfectly fits the data and that the choice of one statistical distribution over the others is impossible. This fact further highlights the importance of testing different statistical distributions of WTP in this context.

4.2. Non-parametric Estimation

We follow the Turnbull non-parametric approach for binary data, also known as the Ayer estimator (Ayer et al., 1955), which has been developed in Krström (1990) for the CV method. The advantage of this approach is that a specific distribution assumption for WTP is not necessary. The only assumptions are that the probability of accepting a bid at CHF 0 is 100%, which rules out negative WTP and that the probability of accepting some high bid, the truncation point, is 0. The truncation point should be chosen according to respondents' preferences or income constraints, information that is usually missing. The relative strength or weakness of these assumptions depend on the program being valued.

We build the survival function of bid acceptance for discrete choice WTP data as in Bateman et al. (2002). The estimated points of the survival function are thus calculated as:

¹³ Spike models also reject the null hypothesis for poolability given by the LR test.

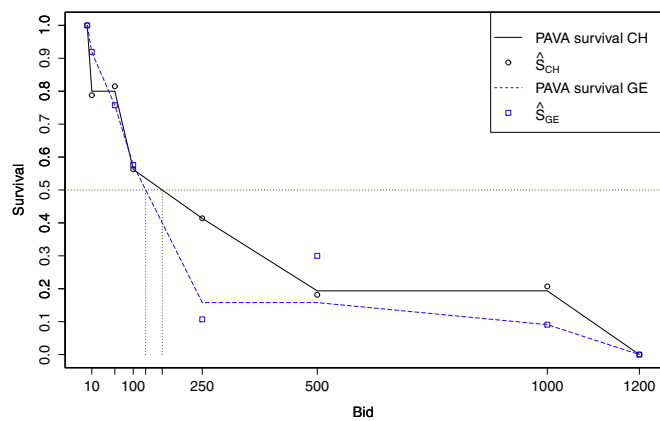


Fig. 2. PAVA survival function of WTP.

Table 4
Non-parametric WTP estimates.

| | CH | GE |
|----------------|-----------|-----------|
| Mean WTP | 341.52*** | 243.96*** |
| Std. err. | (27.32) | (26.54) |
| Δ CH-GE | | 97.56*** |
| Median WTP | 163.04*** | 127.20*** |
| Std. err. | (28.21) | (23.40) |
| Δ CH-GE | | 35.85 |
| Observations | 183 | 152 |

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

$$\hat{S}(Bid_j) = \frac{n_j}{N_j} \tag{1}$$

where Bid_j is the bid level ($j = 1 \dots 6$), N_j is the number of persons whom the bid has been proposed to, n_j the number of persons who said “Yes” to the given bid and \hat{S} the estimated survival function.

A valid survival function, from well-behaved preferences, has to be monotonously decreasing. As this is not the case for some bid levels, we correct for this issue using the Pooled Adjacent Violators Algorithm (PAVA) method, proposed by Robertson et al. (1988) and also called Turnbull Self Consistency Algorithm.¹⁴ This method, presented in Haab and McConnell (2002), pools the Bid_j with Bid_{j-1} if the acceptance rate for Bid_j is higher than for Bid_{j-1} .

As for the parametric estimation, we do not include protesters. Following Kriström (1990), we interpolate linearly between bids, but a step function, as proposed in Bateman et al. (2002) is also applicable. We arbitrarily truncate our survival function at 1200¹⁵, which is likely to underestimate the true WTP, because the last bid and the truncation point are close. However, this applies to both sub-samples estimations and should therefore not affect scope effects. The resulting survival functions are illustrated in Fig. 2, which shows that the survival function for the Swiss forest sub-sample (plain line) is usually higher than the Geneva forest sub-sample’s survival function (dashed line). It is interesting to see that the survival functions are close at low bids and seem to diverge only after a certain threshold.

To compare the survival functions in Fig. 2 and test for differences between these two functions, we use the non-parametric Kolmogorov-

¹⁴ For the sake of replication ease, the data used to estimate the non-parametric model are available in Table B.3, in the appendix.

¹⁵ The maximum WTP, as stated in the open-ended follow-up question, indicates that no respondent is willing to pay more than 1000. Since the acceptance rate of the 1000 bid is still high (18% and 6% for CH and GE respectively) and given the incentive incompatibility of the open-ended format, it would be unrealistic to truncate at 1000. We thus choose 1200 to stay conservative. The estimations with a truncation at 1000 display the same results in terms of mean WTP scope responsiveness.

Table 5
Weighted average maximum WTP from the open-ended follow-up.

| | CH | GE |
|----------------|----------|---------|
| Mean WTP | 126.84 | 97.36 |
| Std. dev. | (212.07) | (48.83) |
| Δ CH-GE | | 29.47* |
| Observations | 182 | 119 |

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Smirnov test (KS). This test does not reveal any significant difference between CH and GE distributions as a whole. However, a KS test concludes that the survival functions between the bids 250 and 1000 are significantly different at the 5% confidence level. In our bid design, since the bids are not equidistant, there may be troubles with the KS test as all points of the estimated survival function have the same weight, while the highest bids have the strongest impact on WTP estimates. To correct for that issue, we interpolate \hat{S} with 6 hypothetical equidistant bids and test again for significant difference in survival function. This manipulation does not distort respondents’ preferences. Indeed, if the individual accepts to pay 250, she should also accept to pay 166.66. This procedure allows to reject the hypothesis of same WTP distributions for both sub-sample at the 95% confidence level, revealing that WTP distribution is, on average, statistically higher for the Swiss program than for the Geneva program.

Using the Turnbull approach, the median WTP for the Swiss forest sub-sample on graph 2 corresponds to the point where the function hits 0.5 on the Y axis. Mean WTP can be calculated as the area under the survival function.

Table 4 presents the WTP central tendency estimates of the non-parametric approach. CH mean WTP is again larger than GE. Furthermore, a Z-test reveals that the CH mean WTP is significantly higher than GE WTP at the 99% confidence level. Hence, where parametric estimates fail to reveal scope effects by lack of efficiency, the Turnbull estimator and the associated tests manage to distinguish the WTP difference in both sub-samples.

4.3. Using the Anchored Open-ended Follow-up

As mentioned earlier, our survey has a follow-up open-ended question asking the maximum WTP for the program. While the first aim of this question was to identify protest bids, the answers also give information about (stated) maximum WTP in an anchored context¹⁶. Protest bids were again excluded from the sample to avoid the protest bias. As highlighted in Desvousges et al. (1987), although the preceding dichotomous choice question should moderate this issue, open-ended questions may also suffer from outliers. This seems not to be the case in our sample, since the stated maximum WTP never overcomes CHF 1000, which corresponds to the highest bid proposed and to a little fraction of respondents mean stated income.

As our bids are the same in both sub-samples, we assume that each anchor (each bid) has the same effect on the maximum stated WTP in both sub-samples. There are no clear evidence to support this assumption in the literature. However, there is no theoretical reason to believe that the anchoring effect would be different either. Therefore, we run a simple analysis of weighted means to test for scope effects. The weights are computed to keep the exact same proportion of each bids in both sub-samples, to ensure the same anchoring effect. As shown in Table 5, we again find a larger mean WTP for the Swiss forest sub-

¹⁶ In addition to the anchoring bias, these answers suffer from incentive incompatibility (Carson, 2012) and may thus be affected by a strategic bias. Indeed the open-ended format does not give the incentive to respondents to truthfully reveal their preferences. Yet, we assume that this bias is similar in both sub-samples. Therefore, the strategic biases should cancel out when looking at WTP differences.

Table 6
Arc scope elasticities.

| Statistical distribution | WTP central measure | Covariates | ϵ_Q | $IC_{95\%}$ |
|--------------------------|---------------------|------------|--------------|---------------|
| Logistic | Mean | No | 0.20 | [−0.08;0.49] |
| Logistic | Mean | Yes | 0.25** | [0.0010;0.50] |
| Log-normal | Mean | No | 0.72*** | [0.23;1.22] |
| Log-normal | Median | No | 0.088 | [−0.20;0.38] |
| Log-normal | Mean | Yes | 0.70*** | [0.23;1.17] |
| Log-normal | Median | Yes | 0.16 | [−0.11;0.43] |
| Log-logistic | Median | No | 0.098 | [−0.18;0.38] |
| Log-logistic | Median | Yes | 0.16 | [−0.12;0.43] |
| Spike | Conditional mean | No | 0.48** | [0.10;0.86] |
| Spike | Unconditional mean | No | 0.44** | [0.049;0.84] |
| Spike | Conditional median | No | 0.26** | [0.027;0.49] |
| Spike | Conditional mean | Yes | 0.52*** | [0.28;0.76] |
| Spike | Unconditional mean | Yes | 0.49*** | [0.24;0.74] |
| Spike | Conditional median | Yes | 0.32*** | [0.12;0.52] |
| Non-parametric | Mean | No | 0.12 | |
| Non-parametric | Median | No | 0.12 | |
| Open-ended | Weighted mean | No | 0.13 | |

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.
Confidence intervals computed with the Delta Method.

sample. Applying a Welch test, the difference is significant at the 90% confidence level. More efficient estimates produced by the open-ended format are thus better able to reveal scope effects in our case.

4.4. Plausibility of the Scope Responsiveness

As noted by the NOAA panel, findings from CV must show “adequate” responsiveness to scope to be considered reliable (Desvousges et al., 2012), an issue that needs to be tested with an adding-up test (Diamond and Hausman, 1994). Unfortunately, as the Geneva and Swiss programs are not defined incrementally¹⁷, we cannot discuss the adequacy of the scope responsiveness with this method. Very recently, Whitehead (2016) proposed to calculate the scope elasticity of WTP as an indicator of economic significance or “plausibility”, rather than only statistical significance. Scope elasticities (ϵ_Q) are defined for an infinitesimal change in quantity (the scope: Q) such that $\epsilon_Q = \frac{dWTP}{dQ} \frac{Q}{WTP}$. Hence a continuous WTP-scope function is necessary. Our split sample methodology does not allow to derive a continuous WTP-scope function since only two WTP points were estimated. We therefore calculate the variation of WTP in terms of “arc elasticity” such that: $\epsilon_Q = \frac{\Delta WTP}{\Delta Q} \frac{\bar{Q}}{\bar{WTP}}$ (Whitehead, 2016, p. 20), with \bar{Q} and \bar{WTP} the mean values of Q and WTP respectively. Arc elasticities thus represent an average elasticity between two distant points. For our case:

$$\epsilon_Q = \frac{(WTP_{CH} - WTP_{GE})}{(Q_{CH} - Q_{GE})} \frac{(Q_{CH} + Q_{GE})}{(WTP_{CH} + WTP_{GE})} \tag{2}$$

With WTP_{CH} the mean WTP for the Swiss program, WTP_{GE} the mean WTP for the Geneva program, Q_{CH} the scope of the Swiss program and Q_{GE} the scope of the Geneva program (in hectares). Q_{CH} is equal to 10% of the Swiss forest surface and Q_{GE} equals 10% of Geneva forest surface (i.e. 126'040 hectares and 301 hectares respectively in 2014 (FSO, 2016)). Elasticities whose confidence intervals fall into the [0;1] range are considered plausible, according to Whitehead (2016), since they would respect the positive but decreasing marginal utility theory.

The resulting arc scope elasticities and their confidence intervals at

95% are presented in Table 6. Confidence intervals of the logistic with covariates, log-normal and spike model are strictly positive. However, elasticities from the log-normal model may exceed 1, which would contradict the decreasing marginal utility theory. These elasticities, according to Whitehead (2016), tend to show that WTP elicited through plain models do not robustly show plausible responsiveness to scope, with the exception of the logit model with covariates. However, elasticities from the spike distributions fall in the plausible range. Hence, by adding more information on the WTP distribution and thus reducing the variance, the spike model results in more plausible scope responsiveness in our case.

The arc scope elasticities confirm that the responsiveness to scope is sensitive to the distribution assumption of WTP. Moreover, not only the statistical significance, but also the plausibility, is affected by this assumption. Similarly to the statistical significance, the logit model with covariates and the spike models are better able to reveal plausible responsiveness to scope.

5. Conclusion

We test for scope effects and evaluate its plausibility applying several WTP statistical distribution assumptions: parametric estimations, a non-parametric estimation and an estimation based on the open-ended format. We apply this approach on data from a CV survey assessing the WTP for a program aiming at increasing the surface of protected Swiss and Geneva forests. While the sensitivity of mean WTP to the statistical distribution assumption is acknowledged, we note that it also has an impact on split-samples comparison and on the ability to detect scope effects. Non-parametric models such as the Turnbull model, which assume no a priori WTP statistical distribution, are better able to reveal scope effects than plain parametric models (logit, log-logit, log-normal).

More sophisticated models such as the spike model, by giving more information about individuals' WTP distribution, are also more powerful in revealing statistically significant and plausible scope effects. Open-ended formats, despite biases, could also reasonably be used for this purpose. For small sample sizes, a non-parametric analysis, a spike model or an open-ended format can

¹⁷ The adding-up test requires to value 3 scopes such that $A = B + C$ (Whitehead, 2016). We only have information about one part (B) and the whole (A) but lack C to implement the adding-up test.

therefore constitute better options than the classical parametric dichotomous choice analysis for comparing two WTP estimations, in particular for complex non-market goods. Since the results will depend on the real WTP distribution, we suggest that CV studies have to systematically apply various statistical distribution of WTP, and different ad-hoc statistical tests, paying particular attention to both differences in point estimates such as the mean and the median, but also to differences in statistical distributions.

Since the NOAA panel guidelines (Arrow et al., 1993), testing scope effects should be part of the standard validity tests for a contingent valuation survey. However, some studies do not successfully detect

scope effects and argue that the very CV method may be unreliable. The debate has recently gone further in *Ecological Economics* with Whitehead (2016), Chapman et al. (2016) and Desvousges et al. (2016) about adequacy, rather than the very existence of the scope responsiveness. We argue that the debate on the existence of scope effects is not closed yet and that studies with mixed conclusions will continue to appear also with Whitehead's (2016) new plausibility test. We therefore recommend to pay particular attention to the assumed statistical distribution of WTP, since it has a major incidence on the detection of scope effects.

Appendix A. Descriptive Statistics

Table A.1

Summary statistics of covariates for Swiss (CH) and Geneva (GE) sub-samples.

| | Variable | Mean | Std. dev. | Min. | Max. | N |
|----|-----------------------|--------|-----------|--------|---------|-----|
| CH | Age | 39.070 | 14.448 | 18 | 81 | 228 |
| | Green member | 0.158 | 0.365 | 0 | 1 | 228 |
| | Members in household | 2.996 | 1.349 | 1 | 6 | 227 |
| | Urban | 0.732 | 0.444 | 0 | 1 | 228 |
| | Visit frequency | 2.412 | 1.492 | 0 | 4 | 228 |
| | Distance | 6.722 | 17.201 | 0 | 150 | 228 |
| | Forest concerned | 1.232 | 1.166 | 0 | 3 | 228 |
| | Swiss wood | 0.118 | 0.324 | 0 | 1 | 228 |
| | (Income) ^a | 76,108 | 40,599 | 17,500 | 180,000 | 88 |
| GE | Age | 42.089 | 13.231 | 19 | 81 | 191 |
| | Green member | 0.164 | 0.371 | 0 | 1 | 189 |
| | Members in household | 2.555 | 1.208 | 1 | 6 | 191 |
| | Urban | 0.801 | 0.400 | 0 | 1 | 191 |
| | Visit frequency | 2.251 | 1.629 | 0 | 4 | 191 |
| | Distance | 7.191 | 14.028 | 0 | 110 | 191 |
| | Forest concerned | 1.419 | 1.193 | 0 | 3 | 191 |
| | Swiss wood | 0.0524 | 0.223 | 0 | 1 | 191 |
| | (Income) ^a | 76,250 | 34,326 | 17,500 | 140,000 | 74 |

^a We provide the descriptive statistics of the income variable but do not include it in the final model.

Age represents the age of the respondent. We included this variable because it is statistically different across sub-samples. Also, as this variable is usually correlated with income, it may affect WTP. *Green member* is a binary variable taking the value 1 if the individual is member of or donates to an environment friendly association. The proportion of members is similar in both sub-sample. We expect this variable to be positively correlated with WTP. *Members in household* is the number of persons that composes the household. Households in the CH sub-sample are composed of a significantly higher number of individuals, which is why we included this variable in the model. *Urban* is a binary variable taking the value 1 if the individual lives in an urban area within the Canton of Geneva; *Visit frequency* is a categorical variable representing the annual frequency of visits in a forest; *Distance* is the distance in kilometers from respondent's home to the most visited forest; *Forest concerned* is an index representing the perception of the state of Swiss forests¹⁸. We added these variables because of their potential effect on WTP for our program. Finally, *Swiss wood* is a dummy indicating whether the respondent favors more expensive Swiss wood rather than cheaper imported wood. This variable is significantly higher in the Swiss forest sub-sample. *Income* corresponds to the middle point of the income class proposed.

Appendix B. Estimations

Ceteris paribus, respondents accept more often the bid if they are older, member of environment friendly associations and concerned about the state of Swiss forests. The number of household's member is negatively correlated with the bid acceptance for the Swiss program and positively for the Geneva program. Since larger households usually recreate more often in forests and do it in the closest forest, they may put a higher value on the closer Geneva forests and consider their income constraints when asked for Swiss forests. This point is also reflected in the coefficients associated with the *Visit frequency* variable. Indeed a frequent forest visitor is more likely to accept the bid for the Geneva program but less likely for the Swiss program, which tend to highlight that the higher the usage, the higher the value attributed to Geneva protected forests. Preferences for Swiss wood rather than imported wood is positively correlated with the bid acceptance for the Swiss program: an unsurprising result since domestic wood is considered as environmentally more friendly than imported wood (only 30% of respondents think that the exploitation of wood may threaten biodiversity).

¹⁸ This index is created from answers to the following questions: "According to you, in the last 20 years, the general health of Swiss forests has: a) improved, b) stayed the same, c) degraded", "According to you, in the last 20 years, the surface of forests in Switzerland has: a) increased, b) stayed the same, c) decreased", "According to you, in the last 20 years, biodiversity in Swiss forests has: a) improved, b) stayed the same, c) degraded". This index ranges from 0 to 3, the higher the index, the more concerned the respondent is about the state of Swiss forests.

Table B.1
Results from the plain parametric estimations.

| | Logistic | | | Log-normal | | | Log-logistic | | | | | |
|-----------------------------|---------------------------|-------------------------|---------------------------|-------------------------|-----------------------|-----------------------|-----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| | CH | GE | CH | CH | GE | CH | CH | GE | CH | GE | CH | GE |
| <i>Bid</i> | -0.00311*** (0.000749) | -0.00717** (0.00314) | -0.00364*** (0.000832) | -0.00801** (0.00328) | | | | | | | | |
| <i>ln(Bid)</i> | | | | | -0.427*** (0.0754) | -0.673*** (0.1115) | -0.460*** (0.0828) | -0.723*** (0.135) | -0.721*** (0.140) | -1.203*** (0.240) | -0.777*** (0.159) | -1.326*** (0.285) |
| <i>Age</i> | | | 0.0186 (0.0166) | 0.0273* (0.0155) | | 0.00793 (0.00843) | 0.0149 (0.0105) | 0.0152 (0.0149) | 0.0152 (0.0149) | | 0.0152 (0.0149) | 0.0286 (0.0189) |
| <i>Green member</i> | | | 0.902* (0.514) | 0.242 (0.527) | | 0.363 (0.296) | 0.238 (0.330) | 0.318 (0.496) | 0.318 (0.496) | | 0.318 (0.496) | 0.318 (0.575) |
| <i>Members in household</i> | | | -0.243 (0.153) | 0.280 (0.174) | | -0.191** (0.0844) | 0.211** (0.107) | -0.307** (0.149) | -0.307** (0.149) | | -0.307** (0.149) | 0.369** (0.187) |
| <i>Urban</i> | | | -1.259*** (0.407) | 0.177 (0.526) | | -0.562** (0.241) | 0.148 (0.329) | -0.936** (0.410) | -0.936** (0.410) | | -0.936** (0.410) | 0.367 (0.605) |
| <i>Distance</i> | | | 0.0155 (0.00953) | -0.00619 (0.0126) | | 0.00971 (0.00662) | -0.00276 (0.00771) | 0.0161 (0.0105) | 0.0161 (0.0105) | | 0.0161 (0.0105) | -0.00456 (0.0126) |
| <i>Visit frequency</i> | | | -0.0954 (0.132) | 0.238* (0.125) | | -0.0708 (0.0788) | 0.141* (0.0785) | 0.229* (0.140) | 0.229* (0.140) | | 0.229* (0.140) | 0.229* (0.138) |
| <i>Forest concerned</i> | | | 0.119 (0.176) | 0.0405 (0.177) | | 0.106 (0.100) | 0.0246 (0.102) | 0.0815 (0.175) | 0.177 (0.175) | | 0.177 (0.175) | 0.0815 (0.182) |
| <i>Swiss wood</i> | | | 1.405** (0.652) | -0.182 (1.254) | | 0.768** (0.352) | -0.404 (0.561) | 1.284** (0.604) | 1.284** (0.604) | | 1.284** (0.604) | -0.633 (1.020) |
| <i>Constant</i> | 0.859*** (0.223) | 1.313*** (0.364) | 1.656* (0.977) | -1.126 (0.920) | 2.084*** (0.396) | 3.169*** (0.546) | 2.787*** (0.622) | 1.742*** (0.663) | 3.541*** (0.738) | 5.677*** (1.142) | 4.598*** (1.107) | 3.177*** (1.258) |
| <i>N</i> | 183 | 152 | 183 | 150 | 183 | 152 | 183 | 150 | 183 | 152 | 183 | 150 |
| <i>Pseudo-R²</i> | 0.142 | 0.204 | 0.271 | 0.247 | 0.165 | 0.283 | 0.277 | 0.331 | 0.167 | 0.287 | 0.274 | 0.338 |
| <i>AIC</i> | 221.6 | 170.1 | 204.9 | 175.4 | 215.8 | 153.6 | 203.4 | 158.0 | 215.3 | 152.6 | 204.0 | 156.5 |

Heteroskedasticity-robust standard errors in parentheses.
p < 0.1, ** p < 0.05, *** p < 0.01.

Table B.2
Results from the spike model.

| | CH | GE | CH | GE |
|-------------------------------|----------------------|----------------------|-----------------------|----------------------|
| <i>ln(Bid)</i> | −0.743*** (0.120) | −1.124*** (0.223) | −0.875*** (0.145) | −1.850*** (0.358) |
| <i>Age</i> | | | 0.00812 (0.0119) | 0.0348* (0.0179) |
| <i>Green member</i> | | | 0.172 (0.385) | −1.220*** (0.468) |
| <i>Members in household</i> | | | −0.167 (0.112) | −0.00859 (0.166) |
| <i>Urban</i> | | | −0.657** (0.306) | 0.0647 (0.519) |
| <i>Distance</i> | | | 0.0441 (0.0340) | −0.0140 (0.0186) |
| <i>Visit frequency</i> | | | −0.219* (0.123) | 0.232* (0.129) |
| <i>Forest concerned</i> | | | −0.0000585 (0.135) | −0.564** (0.258) |
| <i>Swiss wood</i> | | | 0.482 (0.480) | 0.878 (1.056) |
| Constant | 4.324*** (0.654) | 5.946*** (1.091) | 5.958*** (0.980) | 9.060*** (2.190) |
| <i>N</i> | 134 | 122 | 134 | 120 |
| Pseudo- <i>R</i> ² | 0.317 | 0.484 | 0.410 | 0.613 |
| <i>AIC</i> | 119.9 | 81.28 | 120.1 | 77.31 |

Heteroskedasticity-robust standard errors in parentheses.
 $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B.3
Data used in the non-parametric model.

| | | 10 | 60 | 100 | 250 | 500 | 1000 | Total |
|----|---------------|------|------|------|------|------|------|-------|
| CH | Yes | 26 | 22 | 18 | 12 | 6 | 6 | 90 |
| | % of Yes | 0.79 | 0.81 | 0.56 | 0.41 | 0.18 | 0.21 | |
| | PAVA % of Yes | | 0.80 | 0.56 | 0.41 | | 0.19 | |
| | No | 7 | 5 | 14 | 17 | 27 | 23 | 93 |
| | <i>N</i> | 33 | 27 | 32 | 29 | 33 | 29 | 183 |
| GE | Yes | 34 | 25 | 19 | 3 | 3 | 1 | 85 |
| | % of Yes | 0.92 | 0.76 | 0.58 | 0.11 | 0.30 | 0.09 | |
| | PAVA % of Yes | 0.92 | 0.76 | 0.58 | | 0.16 | 0.09 | |
| | No | 3 | 8 | 14 | 25 | 7 | 10 | 67 |
| | <i>N</i> | 37 | 33 | 33 | 28 | 10 | 11 | 152 |

This data can be derived from Table 1 when excluding protesters.

References

Aizaki, H., Nakatani, T., Sato, K., 2014. Stated Preference Methods using R. CRC Press.

Arrow, K., Solow, R., Portney, P.R., Leamer, E.E., Radner, R., Schuman, H., 1993. Report of the NOAA panel on contingent valuation. Fed. Regist. 58 (10), 4601–4614.

Ayer, M., Brunk, H.D., Ewing, G.M., Reid, W.T., Silverman, E., 1955. An empirical distribution function for sampling with incomplete information. Ann. Math. Stat. 26 (4), 641–647.

Bandara, R., Tisdell, C., 2005. Changing abundance of elephants and willingness to pay for their conservation. J. Environ. Manag. 76 (1), 47–59.

Baranzini, A., Borzykowski, N., Maradan, D., 2015. La forêt vue par les Genevois : Perceptions et valeurs économiques de la forêt. Schweiz. Z. Forstwes. 166, 306–313.

Baranzini, A., Faust, A.-K., Huberman, D., 2010. Tropical forest conservation: attitudes and preferences. Forest Policy Econ. 12 (5), 370–376.

Baron, J., 2017. Contingent valuation: flawed logic? Science 357 (6349) 363–363.

Bateman, I.J., Carson, R.T., Day, B., Hanemann, M., Hanley, N., Hett, T., Jones-Lee, M., Loomes, G., Mourato, S., Özdemiroglu, E., Pearce, D.W., 2002. Economic Valuation with Stated Preference Techniques: A Manual. Edward Elgar Publishing Ltd.

Baumgärtner, S., Drupp, M.A., Meya, J.N., Munz, J.M., Quaas, M.F., 2017. Income inequality and willingness to pay for environmental public goods. J. Environ. Econ. Manag. 85, 35–61.

Bengochea-Morancho, A., Fuertes-Eugenio, A.M., del Saz-Salazar, S., 2005. A comparison of empirical models used to infer the willingness to pay in contingent valuation. Empir. Econ. 30 (1), 235–244.

Berrens, R.P., Bohara, A.K., Silva, C.L., Brookshire, D., McKee, M., 2000. Contingent values for New Mexico instream flows: with tests of scope, group-size reminder and temporal reliability. J. Environ. Manag. 58 (1), 73–90.

Bishop, R.C., Boyle, K.J., Carson, R.T., Chapman, D., Hanemann, W.M., Kanninen, B., Kopp, R.J., Krosnick, J.A., List, J., Meade, N., Paterson, R., Presser, S., Kerry Smith, V., Tourangeau, R., Welsh, M., Wooldridge, J.M., DeBell, M., Donovan, C., Konopka, M., Scherer, N., 2017a. Putting a value on injuries to natural assets: the BP oil spill. Science 356 (6335), 253–254.

Bishop, R.C., Boyle, K.J., Carson, R.T., Chapman, D., Hanemann, W.M., Kanninen, B., Kopp, R.J., Krosnick, J., List, J., Meade, N., Paterson, R., Presser, S., Smith, V.K., Tourangeau, R., Welsh, M., Wooldridge, J.M., De Bell, M., Donovan, C., Konopka, M., Scherer, N., 2017b. Contingent valuation: flawed logic?—response. Science 357 (6349), 363–364.

- Bishop, R.C., Heberlein, T.A., 1979. Measuring values of extramarket goods: are indirect measures biased? *Am. J. Agric. Econ.* 61 (5), 926–930.
- Boman, M., Bostedt, G., 1999. Valuing the wolf in Sweden: are benefits contingent on the supply? In: Boman, M., Brännlund, R., Kriström, B. (Eds.), *Topics in Environmental Economics*. Springer Netherlands, Dordrecht, pp. 157–174.
- Boyle, K.J., Desvousges, W.H., Johnson, F.R., Dunford, R.W., Hudson, S.P., 1994. An investigation of part-whole biases in contingent-valuation studies. *J. Environ. Econ. Manag.* 27 (1), 64–83.
- Carson, R.T., 1997. Contingent valuation surveys and tests of insensitivity to scope. In: *Determining the Value of Non-marketed Goods*. Springer, pp. 127–163.
- Carson, R.T., 2012. Contingent valuation: a practical alternative when prices aren't available. *J. Econ. Perspect.* 26 (4), 27–42.
- Carson, R.T., Groves, T., 2007. Incentive and informational properties of preference questions. *Environ. Resour. Econ.* 37 (1), 181–210.
- Carson, R.T., Mitchell, R.C., 1993a. The issue of scope in contingent valuation studies. *Am. J. Agric. Econ.* 75 (5), 1263–1267.
- Carson, R.T., Mitchell, R.C., 1993b. The value of clean water: the public's willingness to pay for boatable, fishable, and swimmable quality water. *Water Resour. Res.* 29 (7), 2445–2454.
- Chapman, D.J., Bishop, R.C., Hanemann, W.M., Kanninen, B.J., Krosnick, J.A., Morey, E.R., Tourangeau, R., 2016. On the adequacy of scope test results: comments on Desvousges, Mathews, and Train. *Ecol. Econ.* 130, 356–360.
- Desvousges, W., Mathews, K., Train, K., 2012. Adequate responsiveness to scope in contingent valuation. *Ecol. Econ.* 84, 121–128.
- Desvousges, W., Mathews, K., Train, K., 2016. Reply to On the adequacy of scope test results: comments on Desvousges, Mathews, and Train. *Ecol. Econ.* 130, 361–362.
- Desvousges, W.H., Smith, V.K., Fisher, A., 1987. Option price estimates for water quality improvements: a contingent valuation study for the Monongahela River. *J. Environ. Econ. Manag.* 14 (3), 248–267.
- Diamond, P.A., Hausman, J.A., 1994. Contingent valuation: is some number better than no number? *J. Econ. Perspect.* 8 (4), 45–64.
- FOEN, 2013. *Politique forestière 2020*. <http://www.bafu.admin.ch/wald/01152/11490/index.html?lang=fr>. Last view: 07.07.2015.
- Frontuto, V., Dalmazzone, S., Vallino, E., Giaccaria, S., 2017. Earmarking conservation: further inquiry on scope effects in stated preference methods applied to nature-based tourism. *Tour. Manag.* 60, 130–139.
- FSO, 2016. *Indicateur d'environnement- surfaces forestières*. https://www.bfs.admin.ch/bfs/fr/home/statistiques/espace_environnement/ressources/systeme_indicateurs_environnement/etat_environnement/surfaces_forestieres.html. Last view: 07.12.2016.
- Garcia, S., Harou, P., Montagne, C., Stenger, A., 2007. Valuing forest biodiversity from a national survey in France: a dichotomous choice contingent valuation. In: *Technical report Laboratoire d'économie Forestière*. AgroParisTech-INRA.
- Haab, T.C., McConnell, K.E., 2002. *Valuing Environmental and Natural Resources: The Econometrics of Non-market Valuation*. Edward Elgar Publishing.
- Halstead, J.M., Luloff, A., Stevens, T.H., 1992. Protest bidders in contingent valuation. *Northeastern J. Agric. Resour. Econ.* 21 (2), 160–169.
- Hanemann, W.M., 1984. Welfare evaluations in contingent valuation experiments with discrete responses. *Am. J. Agric. Econ.* 66 (3), 332–341.
- Hausman, J., 2012. Contingent valuation: from dubious to hopeless. *J. Econ. Perspect.* 26 (4), 43–56.
- Heberlein, T.A., Wilson, M.A., Bishop, R.C., Schaeffer, N.C., 2005. Rethinking the scope test as a criterion for validity in contingent valuation. *J. Environ. Econ. Manag.* 50 (1), 1–22.
- Hjerpe, E., Hussain, A., Phillips, S., 2015. Valuing type and scope of ecosystem conservation: a meta-analysis. *J. For. Econ.* 21 (1), 32–50.
- Jorgensen, B.S., Syme, G.J., 2000. Protest responses and willingness to pay: attitude toward paying for stormwater pollution abatement. *Ecol. Econ.* 33 (2), 251–265.
- Kahneman, D., Knetsch, J.L., 1992. Valuing public goods: the purchase of moral satisfaction. *J. Environ. Econ. Manag.* 22 (1), 57–70.
- Kanninen, B.J., 1993. Optimal experimental design for double-bounded dichotomous choice contingent valuation. *Land Econ.* 69 (2), 138–146.
- Kotchen, M.J., Reiling, S.D., 1999. Do reminders of substitutes and budget constraints influence contingent valuation estimates? Another comment. *Land Econ.* 75 (3), 478–482.
- Krinsky, I., Robb, A.L., 1986. On approximating the statistical properties of elasticities. *Rev. Econ. Stat.* 68 (4), 715–719.
- Kriström, B., 1990. A non-parametric approach to the estimation of welfare measures in discrete response valuation studies. *Land Econ.* 66 (2), 135–139.
- Kriström, B., 1997. Spike models in contingent valuation. *Am. J. Agric. Econ.* 79 (3), 1013–1023.
- Loomis, J., 2011. What's to know about hypothetical bias in stated preference valuation studies? *J. Econ. Surv.* 25 (2), 363–370.
- McFadden, D., 1974. Conditional logit analysis of qualitative choice behavior. *Front. Econ.*
- McFadden, D., 1994. Contingent valuation and social choice. *Am. J. Agric. Econ.* 76 (4), 689–708.
- Meshreky, A., Baranzini, A., Maradan, D., 2014. Contingent valuation of forests: a meta-analysis. In: HEG Working Paper.
- Mitchell, R.C., Carson, R.T., 2013. *Using Surveys to Value Public Goods: The Contingent Valuation Method*. Routledge.
- Mueller, J.M., 2014. Estimating willingness to pay for watershed restoration in Flagstaff, Arizona using dichotomous-choice contingent valuation. *Forestry* 87 (2), 327–333.
- Ojea, E., Loureiro, M.L., 2011. Identifying the scope effect on a meta-analysis of biodiversity valuation studies. *Resour. Energy Econ.* 33 (3), 706–724.
- Park, T., Loomis, J.B., Creel, M., 1991. Confidence intervals for evaluating benefits estimates from dichotomous choice contingent valuation studies. *Land Econ.* 67 (1), 64–73.
- Poe, G.L., Giraud, K.L., Loomis, J.B., 2005. Computational methods for measuring the difference of empirical distributions. *Am. J. Agric. Econ.* 87 (2), 353–365.
- Poe, G.L., Severance-Lossin, E.K., Welsh, M.P., 1994. Measuring the difference (X-Y) of simulated distributions: a convolutions approach. *Am. J. Agric. Econ.* 76 (4), 904–915.
- Remoundou, K., Kountouris, Y., Koundouri, P., 2012. Is the value of an environmental public good sensitive to the providing institution? *Resour. Energy Econ.* 34 (3), 381–395.
- Richardson, L., Loomis, J., 2009. The total economic value of threatened, endangered and rare species: an updated meta-analysis. *Ecol. Econ.* 68 (5), 1535–1548.
- Robertson, T., Wright, F., Dykstra, R.L., Robertson, T., 1988. *Order Restricted Statistical Inference*. Wiley New York.
- Smith, V.K., Osborne, L.L., 1996. Do contingent valuation estimates pass a “scope” test? A meta-analysis. *J. Environ. Econ. Manag.* 31 (3), 287–301.
- Strazzera, E., Genius, M., Scarpa, R., Hutchinson, G., 2003. The effect of protest votes on the estimates of WTP for use values of recreational sites. *Environ. Resour. Econ.* 25 (4), 461–476.
- Veisten, K., Hoen, H.F., Navrud, S., Strand, J., 2004. Scope insensitivity in contingent valuation of complex environmental amenities. *J. Environ. Manag.* 73 (4), 317–331.
- Whitehead, J.C., 2016. Plausible responsiveness to scope in contingent valuation. *Ecol. Econ.* 128, 17–22.
- Whitehead, J.C., Haab, T.C., Huang, J.-C., 1998. Part-whole bias in contingent valuation: will scope effects be detected with inexpensive survey methods? *South. Econ. J.* 65 (1), 160–168.
- WVS, 2017. *World Values Survey wave 5: 2005-2009*. <http://www.worldvaluessurvey.org/wvs.jsp>. Last view: 26.06.2017.