UNIVERSITY OF COPENHAGEN

Mitigating hypothetical bias

a new approach

Ladenburg, Jacob; Olsen, Søren Bøye

Published in:
Symposium i Anvendt Statistik

Publication date:
2008

Document version
Publisher's PDF, also known as Version of record

Citation for published version (APA):
Ladenburg, J., & Olsen, S. B. (2008). Mitigating hypothetical bias: a new approach. In P. Linde (Ed.), *Symposium i Anvendt Statistik* (pp. 280-290). Copenhagen Business School Press.

Mitigating hypothetical bias – a new approach

Jacob Ladenburg,
Environmental Economics and Social Affairs, COWI, jalg@cowi.dk

and

Søren Bøye Olsen
Institute of Food and Resource Economics, University of Copenhagen,
sobo@life.ku.dk

Abstract

Hypothetical bias in terms of overstatement of Willingness-To-Pay is an essential problem which reduces the validity of the obtained welfare estimates for non-market goods in stated preference studies. "Cheap Talk" has previously been found to reduce hypothetical bias but empirical results are ambiguous. In the attempt to further mitigate hypothetical bias, we test the effect of improving Cheap Talk scripts by adding a new type of reminder. This addition is an objective short script presented prior to the choice sets, prompting the respondent to choose the opt-out alternative, if he/she finds the proposed policy generated alternatives in a choice set too expensive. In the survey, this "Opt-Out Reminder" is applied in conjunction with Cheap Talk. Results show that Willingness-To-Pay for re-establishment of a stream in an urban green area is significantly reduced when introducing the Opt-Out Reminder. Further, results show that respondent choices are influenced across the entire bid range applied. This suggests that an Opt-Out Reminder is an effective instrument for reducing hypothetical bias beyond the potential of Cheap Talk on its own.

Keywords: Cheap talk, Opt-Out Reminder, choice experiments, hypothetical bias, stream re-establishment

Introduction

Stated preferences methods such as the Contingent Valuation Method and Choice Experiments are generally known to suffer from hypothetical bias which drives a wedge between *true* and *hypothetical* Willingness-To-Pay (WTP) (Arrow et al. 1993; Carlsson, Frykholm and Lagerkvist 2005; Carlsson and Martinsson 2001; Diamond and Hausman 1994; Hanemann 1994; List, Sinha and Taylor 2006; Mitchell and Carson 1989). Across a broad spectrum of different goods, a substantial number of studies and reviews find that stated hypothetical maximum WTP is an overstatement of true maximum WTP, see e.g. Harrison and Rutstrom (Forthcoming), List and Gallet (2001) and Murphy et al. (2005). Furthermore, Lusk and Schroeder (2004) find that this overstatement of WTP might apply to the entire bid range provided in a choice

experiment study. These findings are supported by Murphy, Stevens and Weatherhead (2005) and Brown, Ajzen and Hrubes (2003). Regardless of the price of a particular beef steak, Lusk and Schroeder (2004) find that the hypothetical simulated market share is notably larger than the non-hypothetical simulated markets share. More noteworthy, their results emphasise that it is particularly at the lower end of the bid range that the largest differences in market shares, and thereby WTPs, appear.

In the attempt to mitigate hypothetical bias, Cummings and Taylor (1999) tested a reminder known as "Cheap Talk" (CT)[1] in three independent contingent valuation surveys and found CT to effectively eliminate the hypothetical bias. However, in subsequent studies the effectiveness of CT has been found to be sensitive to the bid range applied[2]. Brown, Ajzen and Hrubes (2003) and Murphy, Stevens and Weatherhead (2005) find that CT only has an effect on those respondents who are presented with bid levels in the higher end of the bid range in dichotomous choice and referendum surveys. As the two studies also find that the hypothetical bias is evident in the lower bid range, their results, together with the results of Lusk and Schroeder (2004), point towards that traditional CT might not be a *hypothetical bias panacea*.

To further remedy the hypothetical bias problem, this article contributes to the literature by suggesting the use of a so-called Opt-Out Reminder (OOR). The OOR explicitly reminds the respondent to choose the opt-out alternative[3] if he/she finds the proposed policy generated alternatives in the choice set too expensive.

Adopting the framework of Lusk and Schroeder (2004) we test the influence of an OOR on stated preferences and WTP in a choice experiment setup. Though, contrary to Lusk and Schroeder (2004), who compare stated hypothetical preferences with stated non-hypothetical preferences, our experimental setup is based on a comparison of two hypothetical treatments. We find that the OOR significantly reduces the propensity to choose a policy generated alternative at a given price. Relative to using a traditional CT, adding the OOR significantly reduces estimated WTP for the proposed change in the non-market environmental good. Moreover, our estimates show that the effect is significant over the entire price range which, relating to Lusk and Schroeder (2004), translates into reduced predicted demand regardless of the price level. This is an improvement compared to the CT price level sensitivity observed in Brown, Ajzen and Hrubes (2003) and Murphy, Stevens and Weatherhead (2005). Thus, compared to

---

[1] Cheap Talk simply explains the problem of hypothetical bias to respondents prior to the preference elicitation.

[2] Aadland and Caplan (2003), List (2001) and Lusk (2003) find that CT only seems to influence the preferences of specific sub-groups. Samnaliev, Stevens and More (2003) and Carlsson and Martinsson (2006) do not find CT to effectively reduce WTP, and Aadland and Caplan (2006) even find that the CT increases WTP.

[3] In the literature expressions such as the "opt-out" alternative, the "status quo" alternative, the "do nothing" alternative or the "no choice" alternative have been used more or less interchangeably. For simplicity we only use the term opt-out alternative in this article.

the generally recommended CT, our results suggest that adding an OOR more effectively reduces hypothetical bias.

Adding an opt-out reminder

The aim of this article is to test the applicability of a, to the authors' knowledge, new type of CT script in the form of adding an OOR to a traditional CT. The OOR is intended to reduce hypothetical bias by plainly directing the respondent's attention to the trade-off between attributes and cost with the status quo as an explicit benchmark. The OOR was shown to respondents prior to each choice set. The exact wording of the OOR was the following: "*If both prices are higher than what you think your household will pay, then you should choose the present situation* (the opt-out)*.*"

Inspired by Aadland and Capland (2003), Bulte et al. (2005), Carlsson, Frykholm and Lagerkvist (2005) and Lusk (2003), we apply two hypothetical treatments in otherwise identical environments to isolate the potential effect of the OOR. The only difference between treatments is that in one treatment respondents were provided with an OOR whereas they were not in the other[4].

The experimental setup of the survey is based on a questionnaire aimed at surveying local communities' preferences for streams in urban green areas. The survey aimed at examining preferences for re-establishing a stream, Lygte Å, which presently is running in a pipeline beneath an urban park known as Lersøparken in Copenhagen. Respondents were recruited from the population living in the three Copenhagen city districts Bispebjerg, Nørrebro and Østerbro, all located adjacent to Lersøparken[5]. From each city district, 2x200 respondents between the ages 18 and 70 were randomly drawn from the Danish Civil Registration System, summing up to a total of 1200 respondents.

The construction and validation of the questionnaire was carried out firstly by approaching people visiting Lersøparken in an informal manner, asking them about their perceptions and attitudes towards a potential re-establishment of the stream. Secondly, four focus groups were interviewed as part of developing the questionnaire and identifying the relevant attributes and attribute levels[6]. The final set of attributes used in the CE design as well as the associated attribute levels are displayed in table 1. A D-optimal fractional factorial design was generated entailing a total of 24 policy generated alternatives. The alternatives were paired into 12 choice sets which were randomly blocked in two. Consequently, each respondent evaluated six choice sets in total. Besides the two policy generated alternatives, each choice set contained a third

[4] In both samples, the scenario description included a short scripted CT.

[5] The associated benefits were expected to be strongly dependent on the use of the park.

[6] Approximately 20% of the population in the three city districts is non-natives who have Danish as their second language. Since this group might have difficulties in reading and understanding the questionnaire, one focus group consisted entirely of respondents from this part of the population.

alternative; the opt-out alternative, which entailed leaving the stream in the present pipeline at no extra cost.

Table 1. Attributes and Attribute Levels

Attribute	Levels	Coding
Course of the stream	Straight	0
	Meandering	1
Water level	One month dry-out per annum	0
	no dry-outs	1
Stream edges/banks	Covered with flagstones	0
	Covered with grass	1
Stream profile	Single	0
	Double	1
Price (tax)	9, 18, 36, 73, 127 and 200 US$/household/year	Continuous

Analysis

To test for equality of preferences across treatments we apply a random utility function. Let individual *i*'s utility of choosing alternative *j* be given by: $U_{ij} = V_{ij} + \varepsilon_{ij}$, where V_{ij} is the systematic part of the utility associated with the stream attributes and ε_{ij} is a stochastic element. Assuming V_{ij} is linear in parameters, the systematic utility of alternative *j* can be expressed as: $V_{ij} = \beta' X_{ij} + \varphi' A_{ij} + \eta' P_{ij}$. β's are the coefficients representing the utility associated with the attributes, X_{ij}, of the re-established stream, φ's are the coefficients of the two alternative specific constants, A_{ij}, representing the utility of the re-establishment alternatives relative to the opt-out alternative, and η represents the (dis-)utility of the price. The probability of an individual choosing alternative *j* from a choice set consisting of alternatives, *j*, *k* and *l* is given by: Prob($V_{ij} + \varepsilon_{ij} > V_{ik} + \varepsilon_{ik}$, $V_{ij} + \varepsilon_{ij} > V_{il} + \varepsilon_{il}$). Following Lusk and Shroeder (2004), Heteroscedastic Extreme Value (HEV), Multinomial Probit (MNP), and Random Parameter Logit (RPL) models are applied in the test for equal parameters across treatments. The specifications of the models in the present article are shortly presented in the following subsection. For more detailed presentations and discussions of models we refer to Lusk and Shroeder (2004).

The HEV model is fitted by estimating the error variances of alternative 0 and 2 relative to alternative 1, which is the minimum error variance alternative (Walker, Ben-Akiva and Bolduc *forthcoming*). The MNP model is fitted by assuming the error variance of alternative 1 and 2 to be uniform relative to the opt-out alternative (alternative 0). Compared to Lusk and Schroeder (2004), off-diagonal elements are not restricted to be zero. In the RPL model, the attribute parameters of the re-established stream are allowed to vary across respondents in that they are all assumed to be

independently and normally distributed in the population. The price coefficients and alternative specific constants are assumed to be fixed in the population. As opposed to the HEV and the MNP models, the RPL allows for panel specification which captures the repeated choice nature of the data set explicitly in the model. Hence, the parameters that enter utility are treated as being constant over the six choice observations for each respondent (Train 2003).

Results and Discussion

Response Rate

An initial response rate of 59% was obtained, but removing protest bids[7] and questionnaires that were not fully completed reduced the effective response rate to 30%. Further, 10% and 14% of the respondents in the effective without-OOR and with-OOR samples, respectively, were excluded from the analysis because they chose the opt-out alternative in all six choice sets (Alfnes et al. 2006). Thus, the parametric analysis is based on a total of 145 respondents in the sample without OOR and 175 respondents in the sample with OOR.

Participant Demographics

The datasets obtained from the without-OOR and the with-OOR samples are based on choices from two independent samples from the population. Differences in demographic background characteristics between the without-OOR and with-OOR samples might weaken the potential for inference with regard to the effect of the OOR, unless explicitly accounted for. Hence, before observed differences in preferences can be assigned to the effect of the OOR, it has to be ascertained whether the respondents constituting the two samples differ with regard to their demographic background characteristics. Pearson χ^2-tests applied within categories such as gender, age, education, income and city district cannot reject the null hypothesis of equality of distributions across samples in any of the cases. This suggests that the two respondent samples on average are homogeneous with regard to demographic characteristics[8].

[7] Reasons given for protesting included “I think the stream should be re-established, but I don’t want to pay more taxes”, “I cannot assess how much more I would be willing to pay in extra taxes”, “I didn’t consider the tax payment at all”, “I have not considered the tax payment, but I want to affect the policy decision”.

[8] Tests have been carried out for a wider range of demographic categories than that presented in table 2. These are available from the authors on request. All of these support the overall conclusion that the two samples do not differ significantly with regard to demographic background characteristics.

Thus, if a difference in preferences across the two samples is established in the following analyses, it can more likely be ascribed solely to the OOR.

Parametric Analysis

Hausman and McFadden (1984) tests clearly rejected the assumption of proportional substitution across alternatives, also known as the IIA property (Ben-Akiva and Lerman 1985). Hence, we do not present MNL models as these do not allow for violations of the restrictive IIA assumption. Inspired by Lusk and Schroeder (2004), the parametric modelling of choices is instead carried out using the less restrictive HEV, MNP and RPL models, as previously mentioned. These are displayed in table 3. The HEV and MNP models have been analysed using the MDC procedure in SAS (SAS 2005) and the RPL models have been analysed using Biogeme version 1.4 (Bierlaire 2003). As no closed form solutions exist for the MNP and RPL models, they are identified using simulated maximum likelihood estimation with 1000 pseudo-random replications.

Main Effect Models

A main effect model is specified and is then applied to the without-OOR and the with-OOR sample, respectively. Hence, for each overall type of model two different model specifications are considered in table 3. The parameter estimates associated with each of the stream attribute variables describe the shift in utility resulting from a shift in the underlying dummy variable from zero to one, see table 1. The parameter estimates reveal that respondents generally have significant preferences for improving the physical condition of the re-established stream from the low (zero) to the high (one) levels[9]. The respondents thus prefer a meandering stream opposed to a straight stream, a stream with constant water flow opposed to a stream which periodically dries out, a stream with grass banks opposed to flagstones, and a double profile stream opposed to a single profile stream. The ranking of stream attributes is similar across models. Respondents have strongest preferences for meandering and grass banks, followed by water level, whereas the stream profile is less important. The RPL model does however reveal that there is significant interpersonal preference heterogeneity. The estimate for the price parameter describes the shift in utility experienced from a one-unit increase in the price. Naturally, a negative sign would be expected for this parameter, which is also the case for all models in table 3. The alternative specific constants ASC1 and ASC2 express the utility associated with the two policy generated alternatives relative to the opt-out alternative. This utility is attributed to each alternative in itself and cannot be explained by other explanatory variables in the

[9] These results are in accordance with expectations based on focus group interviews.

model. Consequently, the parameter estimates for ASC1 and ASC2 describe the change in overall utility associated with moving from the present 'no stream' situation to a situation where the stream is re-established, all else being equal[10]. Not surprisingly, the signs of these parameter estimates generally turn out to be positive in the models.

Table 3. Comparison of Without-OOR and With-OOR Data: HEV, MNP, and RPL Models

	Heteroscedastic Extreme Value		Multinomial Probit		Random Parameters Logit	
	Main		Main		Main	
Parameter estimates	Without OOR	With OOR	Without OOR	With OOR	Without OOR	With OOR
Means						
Meandering	0.86	0.56	0.45	0.53	1.18	1.18
	(3.29)	(3.91)	(4.33)	(6.10)	(6.12)	(7.46)
Water level	0.76	*0.37*	0.39	0.33	0.91	0.69
	(2.79)	*(1.90)*	(5.18)	(3.86)	(3.60)	(3.23)
Grass banks	0.89	0.63	0.45	0.48	1.33	1.25
	(3.60)	(3.58)	(4.96)	(5.70)	(5.84)	(7.13)
Stream profile	0.40	*0.11*	0.21	*0.11*	0.46	*0.26*
	(2.23)	*(0.84)*	(2.65)	*(1.42)*	(3.10)	*(1.70)*
Price	-0.58	-0.49	-0.27	-0.29	-0.67	-0.73
	(5.61)	(10.55)	(12.44)	(12.23)	(7.43)	(9.59)
ASC1	0.70	*-0.10*	0.35	*0.09*	0.89	*0.43*
	(2.15)	*(0.73)*	(3.07)	*(0.83)*	(3.57)	*(1.64)*
ASC2	*0.12*	*-0.42*	*0.21*	*-0.22*	0.86	*0.18*
	(0.40)	*(1.44)*	*(1.37)*	*(1.60)*	(2.39)	*(0.53)*
Scale/standard deviation parameters						
Meandering					1.13	1.00
					(3.62)	(3.65)
Water level					1.57	1.22
					(5.58)	(6.74)
Grass banks					1.31	1.31
					(5.05)	(6.07)
Stream profile					0.95	1.15
					(2.93)	(4.30)
ASC0	1.15	*16.71*	0.96	0.83		
	(2.56)	*(0.45)*	(4.17)	(3.72)		
ASC2	0.52	0.77	1.00	1.00		
	(4.77)	(8.31)				
ASC1	1.00	1.00	1.00	1.00		
Covariance_0_21			0.64	0.67		
			(6.50)	(6.55)		
Log Likelihood	-653	-795	-659	-804	-623	-764
No. of observations	870	1050	870	1050	870	1050
LR-test for equality	52.3**		43.4**		35.6**	

[10] This interpretation entails the assumption that the model is well-specified and does not suffer from omitted variables.

With regard to possible differences in preferences between the two samples, the parameter estimates are not directly comparable across models due to potentially different scale parameters in the two samples (Louviere, Hensher and Swait 2000). Eyeballing estimates, it seems evident that the OOR does not affect preferences for the stream attributes markedly, though estimates in the HEV model appear to be more influenced by the OOR than is the case in the MNP and RPL models. It does however seem that the ASCs are affected in that they are more frequently insignificant in the with-OOR sample as opposed to the without-OOR sample. When not given an OOR, respondents generally prefer a stream to no stream, regardless of the attributes of the stream. Adding the OOR effectively removes this difference in ASC utility, and respondents are now indifferent between stream and no stream unless they have some saying on the physical attributes of the stream. This could be interpreted as an indication that the OOR does indeed make respondents pay more attention to the stream attributes and consider their tradeoffs relative to the opt-out benchmark more closely.

A more formal test of the hypothesis of identical preferences in the two samples is reported in the bottom line of table 3. This is the Likelihood Ratio (LR) test for pooling datasets with identical data generating processes. In the present case, the data generating process is ultimately respondents' preferences as expressed through choices. The null hypothesis of equal preferences across samples is tested by pooling the two samples after having controlled for scale differences (Swait and Louviere 1993; Louviere, Hensher and Swait 2000). As table 3 reports, overall preference equality across without-OOR and with-OOR samples is strongly rejected in all cases[11]. This is supported by comparing the WTP estimates displayed in table 4. Calculation of WTP entails cancelling out the scale parameters, thus enabling direct comparison between the two samples (Louviere, Hensher and Swait 2000).

Table 4. Comparison of Without-OOR and With-OOR Data: Willingness-To-Pay

	Main effect models								
	HEV			MNP			RPL		
Mean marginal WTP ($)[a]	Without OOR	With OOR	*p*-value[b]	Without OOR	With OOR	*p*-value	Without OOR	With OOR	*p*-value
Meandering	27.2	21.0	0.23	29.8	33.2	0.32	31.7	29.1	0.35
Water level	24.0	13.6	0.16	25.7	21.0	0.28	24.5	17.0	0.19
Grass banks	28.1	23.3	0.12	30.1	30.4	0.48	35.9	31.0	0.22
Stream profile	12.7	4.20	0.27	14.2	7.0	0.15	12.3	6.3	0.14
ASC1	22.2	-3.7	<0.01	23.2	5.5	0.02	23.9	10.6	0.06
ASC2	3.80	-15.5	0.10	14.3	-13.6	0.02	23.1	4.4	0.07

[a] For the readers convenience WTP estimates are reported in US$, entailing a conversion from DKK at exchange rate 1USD≈5.5DKK when deriving WTP estimates from the models in table 3.

[b] *p*-values report results of the one-sided t-test that $WTP_{Without\text{-}OOR} > WTPW_{With\text{-}OOR}$ for each corresponding parameter.

[11] Only minor insignificant differences with regard to the scale were identified in the LR test procedures.

Variances of the WTP point estimates were calculated using the using the Delta Method as described in Greene (2003) and Hanemann and Kanninen (1999).

The WTP estimates for the four stream attributes do not differ significantly for corresponding attributes across the two samples. However, all but two WTP estimates decrease numerically when adding the OOR. Furthermore, the ASC WTP estimates are strongly affected by adding the OOR. P-values reject WTP equality across samples for both of the ASCs at a 0.10 significance level or lower. Combined with the unidirectional tendency of decreasing WTP, this explains the significant LR test statistics in table 3.

In the without-OOR sample the grass parameter is significant, whilst insignificant in the with-OOR sample. Furthermore, in the RPL model, the same preference pattern is evident for the meandering attribute. Adding the OOR seems to make respondents more discriminative with regard to their trade-offs between attributes and cost of a re-established stream. More specifically, the respondents in the with-OOR sample only have significant preferences for a stream with a meandering course if it also has grass banks and vice versa. In the without-OOR sample, grass banks are significantly valued on their own. Meandering, even though not strictly significant at a 0.05 level in the HEV and MNP, also seems to be more valued on its own[12].

Conclusion

It is generally accepted and standard practice to include cheap talk in stated preferences studies to reduce hypothetical bias. However, previous studies have demonstrated that Cheap Talk might be less efficient in reducing hypothetical biases in the lower end of the bid range in dichotomous choice and referenda contingent valuation studies. Based on an empirical dataset from two hypothetical treatments, we find that introducing respondents to an enhanced CT script in the form of an explicit Opt-Out Reminder prior to each choice set reduces the WTP for re-establishing a stream in an urban green area and the marginal WTP for the stream characteristics. Assuming that the decrease in WTP and thereby drop in demand caused by the OOR is interpretable as stated preferences moving closer to the true preferences, the OOR, compared to CT on its own, seems to further reduce the hypothetical bias gap across the entire bid range. Thus, the OOR addresses the problem put forward by Brown, Ajzen and Hrubes (2003), Lusk and Schroeder (2004) and Murphy, Stevens and Weatherhead (2005). Parametric main effect models indicate that the OOR does not

[12] It is worth noting that, relative to the main effect models, adding the interaction parameter does not markedly change parameter estimates on attributes not included in the interaction term. The fact that the *meandering_grass* parameter is not confounded with other main effect attributes than *meandering* and *grass*, serves as a validation of the interaction effect models.

affect underlying preferences for stream attributes as such but only the overall WTP for re-establishing the stream.

Our results suggest that the underlying mechanism at work here is that the OOR makes respondents consider their preferences and tradeoffs more closely when choosing preferred alternatives from a choice set. Thus, the implication is that adding a relatively short-scripted OOR is an appropriate enhancement of CT that will reduce hypothetical bias more effectively than when using standard CT on its own. Thus, based on our findings it is recommendable to include an opt-out reminder in order to minimize hypothetical bias in future applications of choice experiments. However, it should be stressed that future research is warranted to further validate the effectiveness of the OOR, for instance where comparable market data exist.

References

Aadland, D., and A.J. Caplan. 2006. "Cheap Talk Reconsidered: New Evidence from CVM." *Journal of Economic Behavior & Organization* 60(4):562-578.

Aadland, D., and A.J. Caplan. 2003. "Willingness to Pay for Curbside Recycling with Detection and Mitigation of Hypothetical Bias." *American Journal of Agricultural Economics* 85(2):492-502.

Alfnes, F., A.G. Guttormsen, G. Steine, and K. Kolstad. 2006. "Consumers' Willingness to Pay for the Color of Salmon: A Choice Experiment with Real Economic Incentives." *American Journal of Agricultural Economics* 88(4):1050-1061.

Arrow, K., R. Solow, P.R. Portney, E.E. Leamer, R. Radner, and H. Schuman. 1993. *Report of the NOAA Panel on Contingent Valuation.* Washington DC: The National Ocean and Atmospheric Association's Damage Assessment and Restoration Program (DARP).

Ben-Akiva, M., and S.R. Lerman. 1985. *Discrete Choice Analysis. Theory and Application to Travel Demand.* Cambridge, MA: MIT Press.

Bierlaire, M. 2003. "BIOGEME: A Free Package for the Estimation of Discrete Choice Models." Proceedings of the 3rd Swiss Transportation Research Conference. Ascona, Switzerland.

Brown, T.C., I. Ajzen, and D. Hrubes. 2003. "Further Tests of Entreaties to avoid Hypothetical Bias in Referendum Contingent Valuation." *Journal of Environmental Economics and Management* 46(2):353-361.

Bulte, E., S. Gerking, J.A. List, and A. de Zeeuw. 2005. "The Effect of Varying the Causes of Environmental Problems on Stated WTP Values: Evidence from a Field Study." *Journal of Environmental Economics and Management* 49(2):330-342.

Carlsson, F., P. Frykblom, and C.J. Lagerkvist. 2005. "Using Cheap Talk as a Test of Validity in Choice Experiments." *Economics Letters* 89(2):147-152.

Carlsson, F., and P. Martinsson. 2001. "Do Hypothetical and Actual Marginal Willingness to Pay Differ in Choice Experiments." *Journal of Environmental Economics and Management* 41(2):179-192.

Carlsson, F., and P. Martinsson. 2006. "Do Experience and Cheap Talk Influence Willingness to Pay in an Open-Ended Contingent Valuation Survey." Working Papers in Economics no. 190, Dept. of Econ., School of Business, Economics and Law, Göteborg University.

Cummings, R.G., and L.O. Taylor. 1999. "Unbiased Value Estimates for Environmental Goods: A Cheap Talk Design for the Contingent Valuation Method." *American Economic Review* 89(3):649-665.

Diamond, P.A., and J.A. Hausman. 1994. "Contingent Valuation: Is Some Number better than No Number?" *The Journal of Economic Perspectives* 8(4):45-64.

Greene, W. H. 2003. *Econometric Analysis,* 5th. ed. New York: Prentice-Hall International Inc.

Hanemann, W.M. 1994. "Valuing the Environment through Contingent Valuation." *Journal of Economic Perspectives* 8(4):45-64.

Hanemann, W.M., and B. Kanninen. 1999. "The Statistical Analysis of Discrete-Response CV Data." In I.J. Bateman, and K.G. Willis, eds. *Valuing Environmental Preferences - Theory and Practice of the Contingent Valuation Method in the US, EU, and Developing Countries.* Oxford: Oxford University Press, pp. 302-441.

Harrison, G.W., and E.E. Rutström. Forthcoming. "Experimental Evidence on the Existence of Hypothetical Bias in Value Elicitation Methods." In C. Plott, and V.L. Smith, eds. *Handbook of Results in Experimental Economics*. New York: Elsevier Science.

Hausman, J.A., and D. McFadden. 1984. "Specification Tests for the Multinomial Logit Model." *Econometrica* 52(5):1219-1240.

List, J.A. 2001. "Do explicit Warnings Eliminate the Hypothetical Bias in Elicitation Procedures? Evidence from Field Auctions for Sportscards." *The American Economic Review* 91(3):1498-1507.

List, J.A., and C.A. Gallet. 2001. "What Experimental Protocol Influence Disparities Between Actual and Hypothetical Stated Values?" *Environmental and Resource Economics* 20(3):241-254.

List, J.A., P. Sinha, and M.H. Taylor. 2006. "Using Choice Experiments to Value Non-Market Goods and Services: Evidence from Field Experiments" *Advances in Economic Analysis and Policy* 6(2):1-37(article 2).

Louviere, J.J., D.A. Hensher, and J.D. Swait. 2000. *Stated Choice Methods. Analysis and Applications.* Cambridge: University Press.

Lusk, J.L. 2003. "Effects of Cheap Talk on Consumer Willingness-to-Pay for Golden Rice." *American Journal of Agricultural Economics* 85(4):840-856.

Lusk, J.L., and T.C. Schroeder. 2004. "Are Choice Experiments Incentive Compatible? A Test with Quality Differentiated Beef Steaks." *American Journal of Agricultural Economics* 86(2):467-482.

Mitchell, R.C., and R.T. Carson. 1989. *Using Surveys to Value Public Goods: The Contingent Valuation Method.* Washington D.C.: Resources for the Future.

Murphy, J.J., P.G. Allen, T. H. Stevens, and D. Weatherhead. 2005. “A Meta-analysis of Hypothetical Bias in Stated Preference Valuation.” *Environmental and Resource Economics* 30(3):313-325.

Murphy, J.J., T.H. Stevens, and D. Weatherhead. 2005. “Is Cheap Talk Effective at Eliminating Hypothetical Bias in a Provision Point Mechanism?” *Environmental and Resource Economics* 30(3):327-343.

Samnaliev, M., T.H. Stevens, and T. More. 2003. “A Comparison of Cheap Talk and Alternative Certainty Calibration Techniques in Contingent Valuation.” Working Paper No. 2003-11. Dept. of Res. Econ., University of Massachusetts, Amherst.

SAS. 2005. *The MDC procedure*. SAS Institute.

Swait, J., and J.J. Louviere. 1993. “The Role of the Scale Parameter in the Estimation and Comparison of Multinomial Logit Models.” *Journal of Marketing Research* 30(3):305-314.

Train, K. 2003. *Discrete Choice Methods with Simulation.* Cambridge: Cambridge University Press.

Walker, J.L, M. Ben-Akiva, and D. Bolduc. Forthcoming. “Identification of Parameters in Normal Error Component Logit-Mixture (NECLM) Models.” *Journal of Applied Econometrics,* in press.

Mitigating hypothetical bias – a new approach

Jacob Ladenburg,
Environmental Economics and Social Affairs, COWI, jalg@cowi.dk

and

Søren Bøye Olsen
Institute of Food and Resource Economics, University of Copenhagen,
sobo@life.ku.dk

Abstract

Hypothetical bias in terms of overstatement of Willingness-To-Pay is an essential problem which reduces the validity of the obtained welfare estimates for non-market goods in stated preference studies. "Cheap Talk" has previously been found to reduce hypothetical bias but empirical results are ambiguous. In the attempt to further mitigate hypothetical bias, we test the effect of improving Cheap Talk scripts by adding a new type of reminder. This addition is an objective short script presented prior to the choice sets, prompting the respondent to choose the opt-out alternative, if he/she finds the proposed policy generated alternatives in a choice set too expensive. In the survey, this "Opt-Out Reminder" is applied in conjunction with Cheap Talk. Results show that Willingness-To-Pay for re-establishment of a stream in an urban green area is significantly reduced when introducing the Opt-Out Reminder. Further, results show that respondent choices are influenced across the entire bid range applied. This suggests that an Opt-Out Reminder is an effective instrument for reducing hypothetical bias beyond the potential of Cheap Talk on its own.

Keywords: Cheap talk, Opt-Out Reminder, choice experiments, hypothetical bias, stream re-establishment

Introduction

Stated preferences methods such as the Contingent Valuation Method and Choice Experiments are generally known to suffer from hypothetical bias which drives a wedge between *true* and *hypothetical* Willingness-To-Pay (WTP) (Arrow et al. 1993; Carlsson, Frykholm and Lagerkvist 2005; Carlsson and Martinsson 2001; Diamond and Hausman 1994; Hanemann 1994; List, Sinha and Taylor 2006; Mitchell and Carson 1989). Across a broad spectrum of different goods, a substantial number of studies and reviews find that stated hypothetical maximum WTP is an overstatement of true maximum WTP, see e.g. Harrison and Rutstrom (Forthcoming), List and Gallet (2001) and Murphy et al. (2005). Furthermore, Lusk and Schroeder (2004) find that this overstatement of WTP might apply to the entire bid range provided in a choice

experiment study. These findings are supported by Murphy, Stevens and Weatherhead (2005) and Brown, Ajzen and Hrubes (2003). Regardless of the price of a particular beef steak, Lusk and Schroeder (2004) find that the hypothetical simulated market share is notably larger than the non-hypothetical simulated markets share. More noteworthy, their results emphasise that it is particularly at the lower end of the bid range that the largest differences in market shares, and thereby WTPs, appear.

In the attempt to mitigate hypothetical bias, Cummings and Taylor (1999) tested a reminder known as "Cheap Talk" (CT)[1] in three independent contingent valuation surveys and found CT to effectively eliminate the hypothetical bias. However, in subsequent studies the effectiveness of CT has been found to be sensitive to the bid range applied[2]. Brown, Ajzen and Hrubes (2003) and Murphy, Stevens and Weatherhead (2005) find that CT only has an effect on those respondents who are presented with bid levels in the higher end of the bid range in dichotomous choice and referendum surveys. As the two studies also find that the hypothetical bias is evident in the lower bid range, their results, together with the results of Lusk and Schroeder (2004), point towards that traditional CT might not be a *hypothetical bias panacea*.

To further remedy the hypothetical bias problem, this article contributes to the literature by suggesting the use of a so-called Opt-Out Reminder (OOR). The OOR explicitly reminds the respondent to choose the opt-out alternative[3] if he/she finds the proposed policy generated alternatives in the choice set too expensive.

Adopting the framework of Lusk and Schroeder (2004) we test the influence of an OOR on stated preferences and WTP in a choice experiment setup. Though, contrary to Lusk and Schroeder (2004), who compare stated hypothetical preferences with stated non-hypothetical preferences, our experimental setup is based on a comparison of two hypothetical treatments. We find that the OOR significantly reduces the propensity to choose a policy generated alternative at a given price. Relative to using a traditional CT, adding the OOR significantly reduces estimated WTP for the proposed change in the non-market environmental good. Moreover, our estimates show that the effect is significant over the entire price range which, relating to Lusk and Schroeder (2004), translates into reduced predicted demand regardless of the price level. This is an improvement compared to the CT price level sensitivity observed in Brown, Ajzen and Hrubes (2003) and Murphy, Stevens and Weatherhead (2005). Thus, compared to

---

[1] Cheap Talk simply explains the problem of hypothetical bias to respondents prior to the preference elicitation.

[2] Aadland and Caplan (2003), List (2001) and Lusk (2003) find that CT only seems to influence the preferences of specific sub-groups. Samnaliev, Stevens and More (2003) and Carlsson and Martinsson (2006) do not find CT to effectively reduce WTP, and Aadland and Caplan (2006) even find that the CT increases WTP.

[3] In the literature expressions such as the "opt-out" alternative, the "status quo" alternative, the "do nothing" alternative or the "no choice" alternative have been used more or less interchangeably. For simplicity we only use the term opt-out alternative in this article.

the generally recommended CT, our results suggest that adding an OOR more effectively reduces hypothetical bias.

Adding an opt-out reminder

The aim of this article is to test the applicability of a, to the authors' knowledge, new type of CT script in the form of adding an OOR to a traditional CT. The OOR is intended to reduce hypothetical bias by plainly directing the respondent's attention to the trade-off between attributes and cost with the status quo as an explicit benchmark. The OOR was shown to respondents prior to each choice set. The exact wording of the OOR was the following: "*If both prices are higher than what you think your household will pay, then you should choose the present situation* (the opt-out)*.*"

Inspired by Aadland and Capland (2003), Bulte et al. (2005), Carlsson, Frykholm and Lagerkvist (2005) and Lusk (2003), we apply two hypothetical treatments in otherwise identical environments to isolate the potential effect of the OOR. The only difference between treatments is that in one treatment respondents were provided with an OOR whereas they were not in the other[4].

The experimental setup of the survey is based on a questionnaire aimed at surveying local communities' preferences for streams in urban green areas. The survey aimed at examining preferences for re-establishing a stream, Lygte Å, which presently is running in a pipeline beneath an urban park known as Lersøparken in Copenhagen. Respondents were recruited from the population living in the three Copenhagen city districts Bispebjerg, Nørrebro and Østerbro, all located adjacent to Lersøparken[5]. From each city district, 2x200 respondents between the ages 18 and 70 were randomly drawn from the Danish Civil Registration System, summing up to a total of 1200 respondents.

The construction and validation of the questionnaire was carried out firstly by approaching people visiting Lersøparken in an informal manner, asking them about their perceptions and attitudes towards a potential re-establishment of the stream. Secondly, four focus groups were interviewed as part of developing the questionnaire and identifying the relevant attributes and attribute levels[6]. The final set of attributes used in the CE design as well as the associated attribute levels are displayed in table 1. A D-optimal fractional factorial design was generated entailing a total of 24 policy generated alternatives. The alternatives were paired into 12 choice sets which were randomly blocked in two. Consequently, each respondent evaluated six choice sets in total. Besides the two policy generated alternatives, each choice set contained a third

[4] In both samples, the scenario description included a short scripted CT.

[5] The associated benefits were expected to be strongly dependent on the use of the park.

[6] Approximately 20% of the population in the three city districts is non-natives who have Danish as their second language. Since this group might have difficulties in reading and understanding the questionnaire, one focus group consisted entirely of respondents from this part of the population.

alternative; the opt-out alternative, which entailed leaving the stream in the present pipeline at no extra cost.

Table 1. Attributes and Attribute Levels

Attribute	Levels	Coding
Course of the stream	Straight	0
	Meandering	1
Water level	One month dry-out per annum	0
	no dry-outs	1
Stream edges/banks	Covered with flagstones	0
	Covered with grass	1
Stream profile	Single	0
	Double	1
Price (tax)	9, 18, 36, 73, 127 and 200 US$/household/year	Continuous

Analysis

To test for equality of preferences across treatments we apply a random utility function. Let individual *i*'s utility of choosing alternative *j* be given by: $U_{ij} = V_{ij} + \varepsilon_{ij}$, where V_{ij} is the systematic part of the utility associated with the stream attributes and ε_{ij} is a stochastic element. Assuming V_{ij} is linear in parameters, the systematic utility of alternative *j* can be expressed as: $V_{ij} = \beta' X_{ij} + \varphi' A_{ij} + \eta' P_{ij}$. β's are the coefficients representing the utility associated with the attributes, X_{ij}, of the re-established stream, φ's are the coefficients of the two alternative specific constants, A_{ij}, representing the utility of the re-establishment alternatives relative to the opt-out alternative, and η represents the (dis-)utility of the price. The probability of an individual choosing alternative *j* from a choice set consisting of alternatives, *j*, *k* and *l* is given by: Prob($V_{ij} + \varepsilon_{ij} > V_{ik} + \varepsilon_{ik}$, $V_{ij} + \varepsilon_{ij} > V_{il} + \varepsilon_{il}$). Following Lusk and Shroeder (2004), Heteroscedastic Extreme Value (HEV), Multinomial Probit (MNP), and Random Parameter Logit (RPL) models are applied in the test for equal parameters across treatments. The specifications of the models in the present article are shortly presented in the following subsection. For more detailed presentations and discussions of models we refer to Lusk and Shroeder (2004).

The HEV model is fitted by estimating the error variances of alternative 0 and 2 relative to alternative 1, which is the minimum error variance alternative (Walker, Ben-Akiva and Bolduc *forthcoming*). The MNP model is fitted by assuming the error variance of alternative 1 and 2 to be uniform relative to the opt-out alternative (alternative 0). Compared to Lusk and Schroeder (2004), off-diagonal elements are not restricted to be zero. In the RPL model, the attribute parameters of the re-established stream are allowed to vary across respondents in that they are all assumed to be

independently and normally distributed in the population. The price coefficients and alternative specific constants are assumed to be fixed in the population. As opposed to the HEV and the MNP models, the RPL allows for panel specification which captures the repeated choice nature of the data set explicitly in the model. Hence, the parameters that enter utility are treated as being constant over the six choice observations for each respondent (Train 2003).

Results and Discussion

Response Rate

An initial response rate of 59% was obtained, but removing protest bids[7] and questionnaires that were not fully completed reduced the effective response rate to 30%. Further, 10% and 14% of the respondents in the effective without-OOR and with-OOR samples, respectively, were excluded from the analysis because they chose the opt-out alternative in all six choice sets (Alfnes et al. 2006). Thus, the parametric analysis is based on a total of 145 respondents in the sample without OOR and 175 respondents in the sample with OOR.

Participant Demographics

The datasets obtained from the without-OOR and the with-OOR samples are based on choices from two independent samples from the population. Differences in demographic background characteristics between the without-OOR and with-OOR samples might weaken the potential for inference with regard to the effect of the OOR, unless explicitly accounted for. Hence, before observed differences in preferences can be assigned to the effect of the OOR, it has to be ascertained whether the respondents constituting the two samples differ with regard to their demographic background characteristics. Pearson χ^2-tests applied within categories such as gender, age, education, income and city district cannot reject the null hypothesis of equality of distributions across samples in any of the cases. This suggests that the two respondent samples on average are homogeneous with regard to demographic characteristics[8].

---

[7] Reasons given for protesting included "I think the stream should be re-established, but I don't want to pay more taxes", "I cannot assess how much more I would be willing to pay in extra taxes", "I didn't consider the tax payment at all", "I have not considered the tax payment, but I want to affect the policy decision".

[8] Tests have been carried out for a wider range of demographic categories than that presented in table 2. These are available from the authors on request. All of these support the overall conclusion that the two samples do not differ significantly with regard to demographic background characteristics.

Thus, if a difference in preferences across the two samples is established in the following analyses, it can more likely be ascribed solely to the OOR.

Parametric Analysis

Hausman and McFadden (1984) tests clearly rejected the assumption of proportional substitution across alternatives, also known as the IIA property (Ben-Akiva and Lerman 1985). Hence, we do not present MNL models as these do not allow for violations of the restrictive IIA assumption. Inspired by Lusk and Schroeder (2004), the parametric modelling of choices is instead carried out using the less restrictive HEV, MNP and RPL models, as previously mentioned. These are displayed in table 3. The HEV and MNP models have been analysed using the MDC procedure in SAS (SAS 2005) and the RPL models have been analysed using Biogeme version 1.4 (Bierlaire 2003). As no closed form solutions exist for the MNP and RPL models, they are identified using simulated maximum likelihood estimation with 1000 pseudo-random replications.

Main Effect Models

A main effect model is specified and is then applied to the without-OOR and the with-OOR sample, respectively. Hence, for each overall type of model two different model specifications are considered in table 3. The parameter estimates associated with each of the stream attribute variables describe the shift in utility resulting from a shift in the underlying dummy variable from zero to one, see table 1. The parameter estimates reveal that respondents generally have significant preferences for improving the physical condition of the re-established stream from the low (zero) to the high (one) levels[9]. The respondents thus prefer a meandering stream opposed to a straight stream, a stream with constant water flow opposed to a stream which periodically dries out, a stream with grass banks opposed to flagstones, and a double profile stream opposed to a single profile stream. The ranking of stream attributes is similar across models. Respondents have strongest preferences for meandering and grass banks, followed by water level, whereas the stream profile is less important. The RPL model does however reveal that there is significant interpersonal preference heterogeneity. The estimate for the price parameter describes the shift in utility experienced from a one-unit increase in the price. Naturally, a negative sign would be expected for this parameter, which is also the case for all models in table 3. The alternative specific constants ASC1 and ASC2 express the utility associated with the two policy generated alternatives relative to the opt-out alternative. This utility is attributed to each alternative in itself and cannot be explained by other explanatory variables in the

[9] These results are in accordance with expectations based on focus group interviews.

model. Consequently, the parameter estimates for ASC1 and ASC2 describe the change in overall utility associated with moving from the present 'no stream' situation to a situation where the stream is re-established, all else being equal[10]. Not surprisingly, the signs of these parameter estimates generally turn out to be positive in the models.

Table 3. Comparison of Without-OOR and With-OOR Data: HEV, MNP, and RPL Models

	Heteroscedastic Extreme Value		Multinomial Probit		Random Parameters Logit	
	Main		Main		Main	
Parameter estimates	Without OOR	With OOR	Without OOR	With OOR	Without OOR	With OOR
Means						
Meandering	0.86	0.56	0.45	0.53	1.18	1.18
	(3.29)	(3.91)	(4.33)	(6.10)	(6.12)	(7.46)
Water level	0.76	*0.37*	0.39	0.33	0.91	0.69
	(2.79)	*(1.90)*	(5.18)	(3.86)	(3.60)	(3.23)
Grass banks	0.89	0.63	0.45	0.48	1.33	1.25
	(3.60)	(3.58)	(4.96)	(5.70)	(5.84)	(7.13)
Stream profile	0.40	*0.11*	0.21	*0.11*	0.46	*0.26*
	(2.23)	*(0.84)*	(2.65)	*(1.42)*	(3.10)	*(1.70)*
Price	-0.58	-0.49	-0.27	-0.29	-0.67	-0.73
	(5.61)	(10.55)	(12.44)	(12.23)	(7.43)	(9.59)
ASC1	0.70	*-0.10*	0.35	*0.09*	0.89	*0.43*
	(2.15)	*(0.73)*	(3.07)	*(0.83)*	(3.57)	*(1.64)*
ASC2	*0.12*	*-0.42*	*0.21*	*-0.22*	0.86	*0.18*
	(0.40)	*(1.44)*	*(1.37)*	*(1.60)*	(2.39)	*(0.53)*
Scale/standard deviation parameters						
Meandering					1.13	1.00
					(3.62)	(3.65)
Water level					1.57	1.22
					(5.58)	(6.74)
Grass banks					1.31	1.31
					(5.05)	(6.07)
Stream profile					0.95	1.15
					(2.93)	(4.30)
ASC0	1.15	*16.71*	0.96	0.83		
	(2.56)	*(0.45)*	(4.17)	(3.72)		
ASC2	0.52	0.77	1.00	1.00		
	(4.77)	(8.31)				
ASC1	1.00	1.00	1.00	1.00		
Covariance_0_21			0.64	0.67		
			(6.50)	(6.55)		
Log Likelihood	-653	-795	-659	-804	-623	-764
No. of observations	870	1050	870	1050	870	1050
LR-test for equality	52.3**		43.4**		35.6**	

[10] This interpretation entails the assumption that the model is well-specified and does not suffer from omitted variables.

With regard to possible differences in preferences between the two samples, the parameter estimates are not directly comparable across models due to potentially different scale parameters in the two samples (Louviere, Hensher and Swait 2000). Eyeballing estimates, it seems evident that the OOR does not affect preferences for the stream attributes markedly, though estimates in the HEV model appear to be more influenced by the OOR than is the case in the MNP and RPL models. It does however seem that the ASCs are affected in that they are more frequently insignificant in the with-OOR sample as opposed to the without-OOR sample. When not given an OOR, respondents generally prefer a stream to no stream, regardless of the attributes of the stream. Adding the OOR effectively removes this difference in ASC utility, and respondents are now indifferent between stream and no stream unless they have some saying on the physical attributes of the stream. This could be interpreted as an indication that the OOR does indeed make respondents pay more attention to the stream attributes and consider their tradeoffs relative to the opt-out benchmark more closely.

A more formal test of the hypothesis of identical preferences in the two samples is reported in the bottom line of table 3. This is the Likelihood Ratio (LR) test for pooling datasets with identical data generating processes. In the present case, the data generating process is ultimately respondents' preferences as expressed through choices. The null hypothesis of equal preferences across samples is tested by pooling the two samples after having controlled for scale differences (Swait and Louviere 1993; Louviere, Hensher and Swait 2000). As table 3 reports, overall preference equality across without-OOR and with-OOR samples is strongly rejected in all cases[11]. This is supported by comparing the WTP estimates displayed in table 4. Calculation of WTP entails cancelling out the scale parameters, thus enabling direct comparison between the two samples (Louviere, Hensher and Swait 2000).

Table 4. Comparison of Without-OOR and With-OOR Data: Willingness-To-Pay

	Main effect models								
	HEV			MNP			RPL		
Mean marginal WTP ($)[a]	Without OOR	With OOR	*p*-value[b]	Without OOR	With OOR	*p*-value	Without OOR	With OOR	*p*-value
Meandering	27.2	21.0	0.23	29.8	33.2	0.32	31.7	29.1	0.35
Water level	24.0	13.6	0.16	25.7	21.0	0.28	24.5	17.0	0.19
Grass banks	28.1	23.3	0.12	30.1	30.4	0.48	35.9	31.0	0.22
Stream profile	12.7	4.20	0.27	14.2	7.0	0.15	12.3	6.3	0.14
ASC1	22.2	-3.7	<0.01	23.2	5.5	0.02	23.9	10.6	0.06
ASC2	3.80	-15.5	0.10	14.3	-13.6	0.02	23.1	4.4	0.07

[a] For the readers convenience WTP estimates are reported in US$, entailing a conversion from DKK at exchange rate 1USD≈5.5DKK when deriving WTP estimates from the models in table 3.

[b] *p*-values report results of the one-sided t-test that $WTP_{Without\text{-}OOR} > WTPW_{With\text{-}OOR}$ for each corresponding parameter.

[11] Only minor insignificant differences with regard to the scale were identified in the LR test procedures.

Variances of the WTP point estimates were calculated using the using the Delta Method as described in Greene (2003) and Hanemann and Kanninen (1999).

The WTP estimates for the four stream attributes do not differ significantly for corresponding attributes across the two samples. However, all but two WTP estimates decrease numerically when adding the OOR. Furthermore, the ASC WTP estimates are strongly affected by adding the OOR. P-values reject WTP equality across samples for both of the ASCs at a 0.10 significance level or lower. Combined with the unidirectional tendency of decreasing WTP, this explains the significant LR test statistics in table 3.

In the without-OOR sample the grass parameter is significant, whilst insignificant in the with-OOR sample. Furthermore, in the RPL model, the same preference pattern is evident for the meandering attribute. Adding the OOR seems to make respondents more discriminative with regard to their trade-offs between attributes and cost of a re-established stream. More specifically, the respondents in the with-OOR sample only have significant preferences for a stream with a meandering course if it also has grass banks and vice versa. In the without-OOR sample, grass banks are significantly valued on their own. Meandering, even though not strictly significant at a 0.05 level in the HEV and MNP, also seems to be more valued on its own[12].

Conclusion

It is generally accepted and standard practice to include cheap talk in stated preferences studies to reduce hypothetical bias. However, previous studies have demonstrated that Cheap Talk might be less efficient in reducing hypothetical biases in the lower end of the bid range in dichotomous choice and referenda contingent valuation studies. Based on an empirical dataset from two hypothetical treatments, we find that introducing respondents to an enhanced CT script in the form of an explicit Opt-Out Reminder prior to each choice set reduces the WTP for re-establishing a stream in an urban green area and the marginal WTP for the stream characteristics. Assuming that the decrease in WTP and thereby drop in demand caused by the OOR is interpretable as stated preferences moving closer to the true preferences, the OOR, compared to CT on its own, seems to further reduce the hypothetical bias gap across the entire bid range. Thus, the OOR addresses the problem put forward by Brown, Ajzen and Hrubes (2003), Lusk and Schroeder (2004) and Murphy, Stevens and Weatherhead (2005). Parametric main effect models indicate that the OOR does not

[12] It is worth noting that, relative to the main effect models, adding the interaction parameter does not markedly change parameter estimates on attributes not included in the interaction term. The fact that the *meandering_grass* parameter is not confounded with other main effect attributes than *meandering* and *grass*, serves as a validation of the interaction effect models.

affect underlying preferences for stream attributes as such but only the overall WTP for re-establishing the stream.

Our results suggest that the underlying mechanism at work here is that the OOR makes respondents consider their preferences and tradeoffs more closely when choosing preferred alternatives from a choice set. Thus, the implication is that adding a relatively short-scripted OOR is an appropriate enhancement of CT that will reduce hypothetical bias more effectively than when using standard CT on its own. Thus, based on our findings it is recommendable to include an opt-out reminder in order to minimize hypothetical bias in future applications of choice experiments. However, it should be stressed that future research is warranted to further validate the effectiveness of the OOR, for instance where comparable market data exist.

References

Aadland, D., and A.J. Caplan. 2006. "Cheap Talk Reconsidered: New Evidence from CVM." *Journal of Economic Behavior & Organization* 60(4):562-578.

Aadland, D., and A.J. Caplan. 2003. "Willingness to Pay for Curbside Recycling with Detection and Mitigation of Hypothetical Bias." *American Journal of Agricultural Economics* 85(2):492-502.

Alfnes, F., A.G. Guttormsen, G. Steine, and K. Kolstad. 2006. "Consumers' Willingness to Pay for the Color of Salmon: A Choice Experiment with Real Economic Incentives." *American Journal of Agricultural Economics* 88(4):1050-1061.

Arrow, K., R. Solow, P.R. Portney, E.E. Leamer, R. Radner, and H. Schuman. 1993. *Report of the NOAA Panel on Contingent Valuation.* Washington DC: The National Ocean and Atmospheric Association's Damage Assessment and Restoration Program (DARP).

Ben-Akiva, M., and S.R. Lerman. 1985. *Discrete Choice Analysis. Theory and Application to Travel Demand.* Cambridge, MA: MIT Press.

Bierlaire, M. 2003. "BIOGEME: A Free Package for the Estimation of Discrete Choice Models." Proceedings of the 3rd Swiss Transportation Research Conference. Ascona, Switzerland.

Brown, T.C., I. Ajzen, and D. Hrubes. 2003. "Further Tests of Entreaties to avoid Hypothetical Bias in Referendum Contingent Valuation." *Journal of Environmental Economics and Management* 46(2):353-361.

Bulte, E., S. Gerking, J.A. List, and A. de Zeeuw. 2005. "The Effect of Varying the Causes of Environmental Problems on Stated WTP Values: Evidence from a Field Study." *Journal of Environmental Economics and Management* 49(2):330-342.

Carlsson, F., P. Frykblom, and C.J. Lagerkvist. 2005. "Using Cheap Talk as a Test of Validity in Choice Experiments." *Economics Letters* 89(2):147-152.

Carlsson, F., and P. Martinsson. 2001. "Do Hypothetical and Actual Marginal Willingness to Pay Differ in Choice Experiments." *Journal of Environmental Economics and Management* 41(2):179-192.

Carlsson, F., and P. Martinsson. 2006. "Do Experience and Cheap Talk Influence Willingness to Pay in an Open-Ended Contingent Valuation Survey." Working Papers in Economics no. 190, Dept. of Econ., School of Business, Economics and Law, Göteborg University.

Cummings, R.G., and L.O. Taylor. 1999. "Unbiased Value Estimates for Environmental Goods: A Cheap Talk Design for the Contingent Valuation Method." *American Economic Review* 89(3):649-665.

Diamond, P.A., and J.A. Hausman. 1994. "Contingent Valuation: Is Some Number better than No Number?" *The Journal of Economic Perspectives* 8(4):45-64.

Greene, W. H. 2003. *Econometric Analysis,* 5th. ed. New York: Prentice-Hall International Inc.

Hanemann, W.M. 1994. "Valuing the Environment through Contingent Valuation." *Journal of Economic Perspectives* 8(4):45-64.

Hanemann, W.M., and B. Kanninen. 1999. "The Statistical Analysis of Discrete-Response CV Data." In I.J. Bateman, and K.G. Willis, eds. *Valuing Environmental Preferences - Theory and Practice of the Contingent Valuation Method in the US, EU, and Developing Countries.* Oxford: Oxford University Press, pp. 302-441.

Harrison, G.W., and E.E. Rutström. Forthcoming. "Experimental Evidence on the Existence of Hypothetical Bias in Value Elicitation Methods." In C. Plott, and V.L. Smith, eds. *Handbook of Results in Experimental Economics*. New York: Elsevier Science.

Hausman, J.A., and D. McFadden. 1984. "Specification Tests for the Multinomial Logit Model." *Econometrica* 52(5):1219-1240.

List, J.A. 2001. "Do explicit Warnings Eliminate the Hypothetical Bias in Elicitation Procedures? Evidence from Field Auctions for Sportscards." *The American Economic Review* 91(3):1498-1507.

List, J.A., and C.A. Gallet. 2001. "What Experimental Protocol Influence Disparities Between Actual and Hypothetical Stated Values?" *Environmental and Resource Economics* 20(3):241-254.

List, J.A., P. Sinha, and M.H. Taylor. 2006. "Using Choice Experiments to Value Non-Market Goods and Services: Evidence from Field Experiments" *Advances in Economic Analysis and Policy* 6(2):1-37(article 2).

Louviere, J.J., D.A. Hensher, and J.D. Swait. 2000. *Stated Choice Methods. Analysis and Applications.* Cambridge: University Press.

Lusk, J.L. 2003. "Effects of Cheap Talk on Consumer Willingness-to-Pay for Golden Rice." *American Journal of Agricultural Economics* 85(4):840-856.

Lusk, J.L., and T.C. Schroeder. 2004. "Are Choice Experiments Incentive Compatible? A Test with Quality Differentiated Beef Steaks." *American Journal of Agricultural Economics* 86(2):467-482.

Mitchell, R.C., and R.T. Carson. 1989. *Using Surveys to Value Public Goods: The Contingent Valuation Method.* Washington D.C.: Resources for the Future.

Murphy, J.J., P.G. Allen, T. H. Stevens, and D. Weatherhead. 2005. "A Meta-analysis of Hypothetical Bias in Stated Preference Valuation." *Environmental and Resource Economics* 30(3):313-325.

Murphy, J.J., T.H. Stevens, and D. Weatherhead. 2005. "Is Cheap Talk Effective at Eliminating Hypothetical Bias in a Provision Point Mechanism?" *Environmental and Resource Economics* 30(3):327-343.

Samnaliev, M., T.H. Stevens, and T. More. 2003. "A Comparison of Cheap Talk and Alternative Certainty Calibration Techniques in Contingent Valuation." Working Paper No. 2003-11. Dept. of Res. Econ., University of Massachusetts, Amherst.

SAS. 2005. *The MDC procedure*. SAS Institute.

Swait, J., and J.J. Louviere. 1993. "The Role of the Scale Parameter in the Estimation and Comparison of Multinomial Logit Models." *Journal of Marketing Research* 30(3):305-314.

Train, K. 2003. *Discrete Choice Methods with Simulation.* Cambridge: Cambridge University Press.

Walker, J.L, M. Ben-Akiva, and D. Bolduc. Forthcoming. "Identification of Parameters in Normal Error Component Logit-Mixture (NECLM) Models." *Journal of Applied Econometrics,* in press.

Mitigating hypothetical bias – a new approach

Jacob Ladenburg,
Environmental Economics and Social Affairs, COWI, jalg@cowi.dk

and

Søren Bøye Olsen
Institute of Food and Resource Economics, University of Copenhagen,
sobo@life.ku.dk

Abstract

Hypothetical bias in terms of overstatement of Willingness-To-Pay is an essential problem which reduces the validity of the obtained welfare estimates for non-market goods in stated preference studies. "Cheap Talk" has previously been found to reduce hypothetical bias but empirical results are ambiguous. In the attempt to further mitigate hypothetical bias, we test the effect of improving Cheap Talk scripts by adding a new type of reminder. This addition is an objective short script presented prior to the choice sets, prompting the respondent to choose the opt-out alternative, if he/she finds the proposed policy generated alternatives in a choice set too expensive. In the survey, this "Opt-Out Reminder" is applied in conjunction with Cheap Talk. Results show that Willingness-To-Pay for re-establishment of a stream in an urban green area is significantly reduced when introducing the Opt-Out Reminder. Further, results show that respondent choices are influenced across the entire bid range applied. This suggests that an Opt-Out Reminder is an effective instrument for reducing hypothetical bias beyond the potential of Cheap Talk on its own.

Keywords: Cheap talk, Opt-Out Reminder, choice experiments, hypothetical bias, stream re-establishment

Introduction

Stated preferences methods such as the Contingent Valuation Method and Choice Experiments are generally known to suffer from hypothetical bias which drives a wedge between *true* and *hypothetical* Willingness-To-Pay (WTP) (Arrow et al. 1993; Carlsson, Frykholm and Lagerkvist 2005; Carlsson and Martinsson 2001; Diamond and Hausman 1994; Hanemann 1994; List, Sinha and Taylor 2006; Mitchell and Carson 1989). Across a broad spectrum of different goods, a substantial number of studies and reviews find that stated hypothetical maximum WTP is an overstatement of true maximum WTP, see e.g. Harrison and Rutstrom (Forthcoming), List and Gallet (2001) and Murphy et al. (2005). Furthermore, Lusk and Schroeder (2004) find that this overstatement of WTP might apply to the entire bid range provided in a choice

experiment study. These findings are supported by Murphy, Stevens and Weatherhead (2005) and Brown, Ajzen and Hrubes (2003). Regardless of the price of a particular beef steak, Lusk and Schroeder (2004) find that the hypothetical simulated market share is notably larger than the non-hypothetical simulated markets share. More noteworthy, their results emphasise that it is particularly at the lower end of the bid range that the largest differences in market shares, and thereby WTPs, appear.

In the attempt to mitigate hypothetical bias, Cummings and Taylor (1999) tested a reminder known as "Cheap Talk" (CT)[1] in three independent contingent valuation surveys and found CT to effectively eliminate the hypothetical bias. However, in subsequent studies the effectiveness of CT has been found to be sensitive to the bid range applied[2]. Brown, Ajzen and Hrubes (2003) and Murphy, Stevens and Weatherhead (2005) find that CT only has an effect on those respondents who are presented with bid levels in the higher end of the bid range in dichotomous choice and referendum surveys. As the two studies also find that the hypothetical bias is evident in the lower bid range, their results, together with the results of Lusk and Schroeder (2004), point towards that traditional CT might not be a *hypothetical bias panacea*.

To further remedy the hypothetical bias problem, this article contributes to the literature by suggesting the use of a so-called Opt-Out Reminder (OOR). The OOR explicitly reminds the respondent to choose the opt-out alternative[3] if he/she finds the proposed policy generated alternatives in the choice set too expensive.

Adopting the framework of Lusk and Schroeder (2004) we test the influence of an OOR on stated preferences and WTP in a choice experiment setup. Though, contrary to Lusk and Schroeder (2004), who compare stated hypothetical preferences with stated non-hypothetical preferences, our experimental setup is based on a comparison of two hypothetical treatments. We find that the OOR significantly reduces the propensity to choose a policy generated alternative at a given price. Relative to using a traditional CT, adding the OOR significantly reduces estimated WTP for the proposed change in the non-market environmental good. Moreover, our estimates show that the effect is significant over the entire price range which, relating to Lusk and Schroeder (2004), translates into reduced predicted demand regardless of the price level. This is an improvement compared to the CT price level sensitivity observed in Brown, Ajzen and Hrubes (2003) and Murphy, Stevens and Weatherhead (2005). Thus, compared to

[1] Cheap Talk simply explains the problem of hypothetical bias to respondents prior to the preference elicitation.

[2] Aadland and Caplan (2003), List (2001) and Lusk (2003) find that CT only seems to influence the preferences of specific sub-groups. Samnaliev, Stevens and More (2003) and Carlsson and Martinsson (2006) do not find CT to effectively reduce WTP, and Aadland and Caplan (2006) even find that the CT increases WTP.

[3] In the literature expressions such as the "opt-out" alternative, the "status quo" alternative, the "do nothing" alternative or the "no choice" alternative have been used more or less interchangeably. For simplicity we only use the term opt-out alternative in this article.

the generally recommended CT, our results suggest that adding an OOR more effectively reduces hypothetical bias.

Adding an opt-out reminder

The aim of this article is to test the applicability of a, to the authors' knowledge, new type of CT script in the form of adding an OOR to a traditional CT. The OOR is intended to reduce hypothetical bias by plainly directing the respondent's attention to the trade-off between attributes and cost with the status quo as an explicit benchmark. The OOR was shown to respondents prior to each choice set. The exact wording of the OOR was the following: "*If both prices are higher than what you think your household will pay, then you should choose the present situation* (the opt-out)*.*"

Inspired by Aadland and Capland (2003), Bulte et al. (2005), Carlsson, Frykholm and Lagerkvist (2005) and Lusk (2003), we apply two hypothetical treatments in otherwise identical environments to isolate the potential effect of the OOR. The only difference between treatments is that in one treatment respondents were provided with an OOR whereas they were not in the other[4].

The experimental setup of the survey is based on a questionnaire aimed at surveying local communities' preferences for streams in urban green areas. The survey aimed at examining preferences for re-establishing a stream, Lygte Å, which presently is running in a pipeline beneath an urban park known as Lersøparken in Copenhagen. Respondents were recruited from the population living in the three Copenhagen city districts Bispebjerg, Nørrebro and Østerbro, all located adjacent to Lersøparken[5]. From each city district, 2x200 respondents between the ages 18 and 70 were randomly drawn from the Danish Civil Registration System, summing up to a total of 1200 respondents.

The construction and validation of the questionnaire was carried out firstly by approaching people visiting Lersøparken in an informal manner, asking them about their perceptions and attitudes towards a potential re-establishment of the stream. Secondly, four focus groups were interviewed as part of developing the questionnaire and identifying the relevant attributes and attribute levels[6]. The final set of attributes used in the CE design as well as the associated attribute levels are displayed in table 1. A D-optimal fractional factorial design was generated entailing a total of 24 policy generated alternatives. The alternatives were paired into 12 choice sets which were randomly blocked in two. Consequently, each respondent evaluated six choice sets in total. Besides the two policy generated alternatives, each choice set contained a third

[4] In both samples, the scenario description included a short scripted CT.

[5] The associated benefits were expected to be strongly dependent on the use of the park.

[6] Approximately 20% of the population in the three city districts is non-natives who have Danish as their second language. Since this group might have difficulties in reading and understanding the questionnaire, one focus group consisted entirely of respondents from this part of the population.

alternative; the opt-out alternative, which entailed leaving the stream in the present pipeline at no extra cost.

Table 1. Attributes and Attribute Levels

Attribute	Levels	Coding
Course of the stream	Straight	0
	Meandering	1
Water level	One month dry-out per annum	0
	no dry-outs	1
Stream edges/banks	Covered with flagstones	0
	Covered with grass	1
Stream profile	Single	0
	Double	1
Price (tax)	9, 18, 36, 73, 127 and 200 US$/household/year	Continuous

Analysis

To test for equality of preferences across treatments we apply a random utility function. Let individual *i*'s utility of choosing alternative *j* be given by: $U_{ij} = V_{ij} + \varepsilon_{ij}$, *w*here V_{ij} is the systematic part of the utility associated with the stream attributes and ε_{ij} is a stochastic element. Assuming V_{ij} is linear in parameters, the systematic utility of alternative *j* can be expressed as: $V_{ij} = \beta' X_{ij} + \varphi' A_{ij} + \eta' P_{ij}$. β's are the coefficients representing the utility associated with the attributes, X_{ij}, of the re-established stream, φ's are the coefficients of the two alternative specific constants, A_{ij}, representing the utility of the re-establishment alternatives relative to the opt-out alternative, and η represents the (dis-)utility of the price. The probability of an individual choosing alternative *j* from a choice set consisting of alternatives, *j*, *k* and *l* is given by: Prob($V_{ij} + \varepsilon_{ij} > V_{ik} + \varepsilon_{ik}$, $V_{ij} + \varepsilon_{ij} > V_{il} + \varepsilon_{il}$). Following Lusk and Shroeder (2004), Heteroscedastic Extreme Value (HEV), Multinomial Probit (MNP), and Random Parameter Logit (RPL) models are applied in the test for equal parameters across treatments. The specifications of the models in the present article are shortly presented in the following subsection. For more detailed presentations and discussions of models we refer to Lusk and Shroeder (2004).

The HEV model is fitted by estimating the error variances of alternative 0 and 2 relative to alternative 1, which is the minimum error variance alternative (Walker, Ben-Akiva and Bolduc *forthcoming*). The MNP model is fitted by assuming the error variance of alternative 1 and 2 to be uniform relative to the opt-out alternative (alternative 0). Compared to Lusk and Schroeder (2004), off-diagonal elements are not restricted to be zero. In the RPL model, the attribute parameters of the re-established stream are allowed to vary across respondents in that they are all assumed to be

independently and normally distributed in the population. The price coefficients and alternative specific constants are assumed to be fixed in the population. As opposed to the HEV and the MNP models, the RPL allows for panel specification which captures the repeated choice nature of the data set explicitly in the model. Hence, the parameters that enter utility are treated as being constant over the six choice observations for each respondent (Train 2003).

Results and Discussion

Response Rate

An initial response rate of 59% was obtained, but removing protest bids[7] and questionnaires that were not fully completed reduced the effective response rate to 30%. Further, 10% and 14% of the respondents in the effective without-OOR and with-OOR samples, respectively, were excluded from the analysis because they chose the opt-out alternative in all six choice sets (Alfnes et al. 2006). Thus, the parametric analysis is based on a total of 145 respondents in the sample without OOR and 175 respondents in the sample with OOR.

Participant Demographics

The datasets obtained from the without-OOR and the with-OOR samples are based on choices from two independent samples from the population. Differences in demographic background characteristics between the without-OOR and with-OOR samples might weaken the potential for inference with regard to the effect of the OOR, unless explicitly accounted for. Hence, before observed differences in preferences can be assigned to the effect of the OOR, it has to be ascertained whether the respondents constituting the two samples differ with regard to their demographic background characteristics. Pearson χ^2-tests applied within categories such as gender, age, education, income and city district cannot reject the null hypothesis of equality of distributions across samples in any of the cases. This suggests that the two respondent samples on average are homogeneous with regard to demographic characteristics[8].

---

[7] Reasons given for protesting included “I think the stream should be re-established, but I don’t want to pay more taxes”, “I cannot assess how much more I would be willing to pay in extra taxes”, “I didn’t consider the tax payment at all”, “I have not considered the tax payment, but I want to affect the policy decision”.

[8] Tests have been carried out for a wider range of demographic categories than that presented in table 2. These are available from the authors on request. All of these support the overall conclusion that the two samples do not differ significantly with regard to demographic background characteristics.

Thus, if a difference in preferences across the two samples is established in the following analyses, it can more likely be ascribed solely to the OOR.

Parametric Analysis

Hausman and McFadden (1984) tests clearly rejected the assumption of proportional substitution across alternatives, also known as the IIA property (Ben-Akiva and Lerman 1985). Hence, we do not present MNL models as these do not allow for violations of the restrictive IIA assumption. Inspired by Lusk and Schroeder (2004), the parametric modelling of choices is instead carried out using the less restrictive HEV, MNP and RPL models, as previously mentioned. These are displayed in table 3. The HEV and MNP models have been analysed using the MDC procedure in SAS (SAS 2005) and the RPL models have been analysed using Biogeme version 1.4 (Bierlaire 2003). As no closed form solutions exist for the MNP and RPL models, they are identified using simulated maximum likelihood estimation with 1000 pseudo-random replications.

Main Effect Models

A main effect model is specified and is then applied to the without-OOR and the with-OOR sample, respectively. Hence, for each overall type of model two different model specifications are considered in table 3. The parameter estimates associated with each of the stream attribute variables describe the shift in utility resulting from a shift in the underlying dummy variable from zero to one, see table 1. The parameter estimates reveal that respondents generally have significant preferences for improving the physical condition of the re-established stream from the low (zero) to the high (one) levels[9]. The respondents thus prefer a meandering stream opposed to a straight stream, a stream with constant water flow opposed to a stream which periodically dries out, a stream with grass banks opposed to flagstones, and a double profile stream opposed to a single profile stream. The ranking of stream attributes is similar across models. Respondents have strongest preferences for meandering and grass banks, followed by water level, whereas the stream profile is less important. The RPL model does however reveal that there is significant interpersonal preference heterogeneity. The estimate for the price parameter describes the shift in utility experienced from a one-unit increase in the price. Naturally, a negative sign would be expected for this parameter, which is also the case for all models in table 3. The alternative specific constants ASC1 and ASC2 express the utility associated with the two policy generated alternatives relative to the opt-out alternative. This utility is attributed to each alternative in itself and cannot be explained by other explanatory variables in the

[9] These results are in accordance with expectations based on focus group interviews.

model. Consequently, the parameter estimates for ASC1 and ASC2 describe the change in overall utility associated with moving from the present 'no stream' situation to a situation where the stream is re-established, all else being equal[10]. Not surprisingly, the signs of these parameter estimates generally turn out to be positive in the models.

Table 3. Comparison of Without-OOR and With-OOR Data: HEV, MNP, and RPL Models

	Heteroscedastic Extreme Value		Multinomial Probit		Random Parameters Logit	
	Main		Main		Main	
Parameter estimates	Without OOR	With OOR	Without OOR	With OOR	Without OOR	With OOR
Means						
Meandering	0.86	0.56	0.45	0.53	1.18	1.18
	(3.29)	(3.91)	(4.33)	(6.10)	(6.12)	(7.46)
Water level	0.76	*0.37*	0.39	0.33	0.91	0.69
	(2.79)	*(1.90)*	(5.18)	(3.86)	(3.60)	(3.23)
Grass banks	0.89	0.63	0.45	0.48	1.33	1.25
	(3.60)	(3.58)	(4.96)	(5.70)	(5.84)	(7.13)
Stream profile	0.40	*0.11*	0.21	*0.11*	0.46	*0.26*
	(2.23)	*(0.84)*	(2.65)	*(1.42)*	(3.10)	*(1.70)*
Price	-0.58	-0.49	-0.27	-0.29	-0.67	-0.73
	(5.61)	(10.55)	(12.44)	(12.23)	(7.43)	(9.59)
ASC1	0.70	*-0.10*	0.35	*0.09*	0.89	*0.43*
	(2.15)	*(0.73)*	(3.07)	*(0.83)*	(3.57)	*(1.64)*
ASC2	*0.12*	*-0.42*	*0.21*	*-0.22*	0.86	*0.18*
	(0.40)	*(1.44)*	*(1.37)*	*(1.60)*	(2.39)	*(0.53)*
Scale/standard deviation parameters						
Meandering					1.13	1.00
					(3.62)	(3.65)
Water level					1.57	1.22
					(5.58)	(6.74)
Grass banks					1.31	1.31
					(5.05)	(6.07)
Stream profile					0.95	1.15
					(2.93)	(4.30)
ASC0	1.15	*16.71*	0.96	0.83		
	(2.56)	*(0.45)*	(4.17)	(3.72)		
ASC2	0.52	0.77	1.00	1.00		
	(4.77)	(8.31)				
ASC1	1.00	1.00	1.00	1.00		
Covariance_0_21			0.64	0.67		
			(6.50)	(6.55)		
Log Likelihood	-653	-795	-659	-804	-623	-764
No. of observations	870	1050	870	1050	870	1050
LR-test for equality	52.3**		43.4**		35.6**	

[10] This interpretation entails the assumption that the model is well-specified and does not suffer from omitted variables.

With regard to possible differences in preferences between the two samples, the parameter estimates are not directly comparable across models due to potentially different scale parameters in the two samples (Louviere, Hensher and Swait 2000). Eyeballing estimates, it seems evident that the OOR does not affect preferences for the stream attributes markedly, though estimates in the HEV model appear to be more influenced by the OOR than is the case in the MNP and RPL models. It does however seem that the ASCs are affected in that they are more frequently insignificant in the with-OOR sample as opposed to the without-OOR sample. When not given an OOR, respondents generally prefer a stream to no stream, regardless of the attributes of the stream. Adding the OOR effectively removes this difference in ASC utility, and respondents are now indifferent between stream and no stream unless they have some saying on the physical attributes of the stream. This could be interpreted as an indication that the OOR does indeed make respondents pay more attention to the stream attributes and consider their tradeoffs relative to the opt-out benchmark more closely.

A more formal test of the hypothesis of identical preferences in the two samples is reported in the bottom line of table 3. This is the Likelihood Ratio (LR) test for pooling datasets with identical data generating processes. In the present case, the data generating process is ultimately respondents' preferences as expressed through choices. The null hypothesis of equal preferences across samples is tested by pooling the two samples after having controlled for scale differences (Swait and Louviere 1993; Louviere, Hensher and Swait 2000). As table 3 reports, overall preference equality across without-OOR and with-OOR samples is strongly rejected in all cases[11]. This is supported by comparing the WTP estimates displayed in table 4. Calculation of WTP entails cancelling out the scale parameters, thus enabling direct comparison between the two samples (Louviere, Hensher and Swait 2000).

Table 4. Comparison of Without-OOR and With-OOR Data: Willingness-To-Pay

	Main effect models								
	HEV			MNP			RPL		
Mean marginal WTP ($)[a]	Without OOR	With OOR	*p*-value[b]	Without OOR	With OOR	*p*-value	Without OOR	With OOR	*p*-value
Meandering	27.2	21.0	0.23	29.8	33.2	0.32	31.7	29.1	0.35
Water level	24.0	13.6	0.16	25.7	21.0	0.28	24.5	17.0	0.19
Grass banks	28.1	23.3	0.12	30.1	30.4	0.48	35.9	31.0	0.22
Stream profile	12.7	4.20	0.27	14.2	7.0	0.15	12.3	6.3	0.14
ASC1	22.2	-3.7	<0.01	23.2	5.5	0.02	23.9	10.6	0.06
ASC2	3.80	-15.5	0.10	14.3	-13.6	0.02	23.1	4.4	0.07

[a] For the readers convenience WTP estimates are reported in US$, entailing a conversion from DKK at exchange rate 1USD≈5.5DKK when deriving WTP estimates from the models in table 3.

[b] *p*-values report results of the one-sided t-test that $WTP_{Without\text{-}OOR} > WTPW_{With\text{-}OOR}$ for each corresponding parameter.

[11] Only minor insignificant differences with regard to the scale were identified in the LR test procedures.

Variances of the WTP point estimates were calculated using the using the Delta Method as described in Greene (2003) and Hanemann and Kanninen (1999).

The WTP estimates for the four stream attributes do not differ significantly for corresponding attributes across the two samples. However, all but two WTP estimates decrease numerically when adding the OOR. Furthermore, the ASC WTP estimates are strongly affected by adding the OOR. P-values reject WTP equality across samples for both of the ASCs at a 0.10 significance level or lower. Combined with the unidirectional tendency of decreasing WTP, this explains the significant LR test statistics in table 3.

In the without-OOR sample the grass parameter is significant, whilst insignificant in the with-OOR sample. Furthermore, in the RPL model, the same preference pattern is evident for the meandering attribute. Adding the OOR seems to make respondents more discriminative with regard to their trade-offs between attributes and cost of a re-established stream. More specifically, the respondents in the with-OOR sample only have significant preferences for a stream with a meandering course if it also has grass banks and vice versa. In the without-OOR sample, grass banks are significantly valued on their own. Meandering, even though not strictly significant at a 0.05 level in the HEV and MNP, also seems to be more valued on its own[12].

Conclusion

It is generally accepted and standard practice to include cheap talk in stated preferences studies to reduce hypothetical bias. However, previous studies have demonstrated that Cheap Talk might be less efficient in reducing hypothetical biases in the lower end of the bid range in dichotomous choice and referenda contingent valuation studies. Based on an empirical dataset from two hypothetical treatments, we find that introducing respondents to an enhanced CT script in the form of an explicit Opt-Out Reminder prior to each choice set reduces the WTP for re-establishing a stream in an urban green area and the marginal WTP for the stream characteristics. Assuming that the decrease in WTP and thereby drop in demand caused by the OOR is interpretable as stated preferences moving closer to the true preferences, the OOR, compared to CT on its own, seems to further reduce the hypothetical bias gap across the entire bid range. Thus, the OOR addresses the problem put forward by Brown, Ajzen and Hrubes (2003), Lusk and Schroeder (2004) and Murphy, Stevens and Weatherhead (2005). Parametric main effect models indicate that the OOR does not

[12] It is worth noting that, relative to the main effect models, adding the interaction parameter does not markedly change parameter estimates on attributes not included in the interaction term. The fact that the *meandering_grass* parameter is not confounded with other main effect attributes than *meandering* and *grass*, serves as a validation of the interaction effect models.

affect underlying preferences for stream attributes as such but only the overall WTP for re-establishing the stream.

Our results suggest that the underlying mechanism at work here is that the OOR makes respondents consider their preferences and tradeoffs more closely when choosing preferred alternatives from a choice set. Thus, the implication is that adding a relatively short-scripted OOR is an appropriate enhancement of CT that will reduce hypothetical bias more effectively than when using standard CT on its own. Thus, based on our findings it is recommendable to include an opt-out reminder in order to minimize hypothetical bias in future applications of choice experiments. However, it should be stressed that future research is warranted to further validate the effectiveness of the OOR, for instance where comparable market data exist.

References

Aadland, D., and A.J. Caplan. 2006. "Cheap Talk Reconsidered: New Evidence from CVM." *Journal of Economic Behavior & Organization* 60(4):562-578.

Aadland, D., and A.J. Caplan. 2003. "Willingness to Pay for Curbside Recycling with Detection and Mitigation of Hypothetical Bias." *American Journal of Agricultural Economics* 85(2):492-502.

Alfnes, F., A.G. Guttormsen, G. Steine, and K. Kolstad. 2006. "Consumers' Willingness to Pay for the Color of Salmon: A Choice Experiment with Real Economic Incentives." *American Journal of Agricultural Economics* 88(4):1050-1061.

Arrow, K., R. Solow, P.R. Portney, E.E. Leamer, R. Radner, and H. Schuman. 1993. *Report of the NOAA Panel on Contingent Valuation.* Washington DC: The National Ocean and Atmospheric Association's Damage Assessment and Restoration Program (DARP).

Ben-Akiva, M., and S.R. Lerman. 1985. *Discrete Choice Analysis. Theory and Application to Travel Demand.* Cambridge, MA: MIT Press.

Bierlaire, M. 2003. "BIOGEME: A Free Package for the Estimation of Discrete Choice Models." Proceedings of the 3rd Swiss Transportation Research Conference. Ascona, Switzerland.

Brown, T.C., I. Ajzen, and D. Hrubes. 2003. "Further Tests of Entreaties to avoid Hypothetical Bias in Referendum Contingent Valuation." *Journal of Environmental Economics and Management* 46(2):353-361.

Bulte, E., S. Gerking, J.A. List, and A. de Zeeuw. 2005. "The Effect of Varying the Causes of Environmental Problems on Stated WTP Values: Evidence from a Field Study." *Journal of Environmental Economics and Management* 49(2):330-342.

Carlsson, F., P. Frykblom, and C.J. Lagerkvist. 2005. "Using Cheap Talk as a Test of Validity in Choice Experiments." *Economics Letters* 89(2):147-152.

Carlsson, F., and P. Martinsson. 2001. "Do Hypothetical and Actual Marginal Willingness to Pay Differ in Choice Experiments." *Journal of Environmental Economics and Management* 41(2):179-192.

Carlsson, F., and P. Martinsson. 2006. "Do Experience and Cheap Talk Influence Willingness to Pay in an Open-Ended Contingent Valuation Survey." Working Papers in Economics no. 190, Dept. of Econ., School of Business, Economics and Law, Göteborg University.

Cummings, R.G., and L.O. Taylor. 1999. "Unbiased Value Estimates for Environmental Goods: A Cheap Talk Design for the Contingent Valuation Method." *American Economic Review* 89(3):649-665.

Diamond, P.A., and J.A. Hausman. 1994. "Contingent Valuation: Is Some Number better than No Number?" *The Journal of Economic Perspectives* 8(4):45-64.

Greene, W. H. 2003. *Econometric Analysis,* 5th. ed. New York: Prentice-Hall International Inc.

Hanemann, W.M. 1994. "Valuing the Environment through Contingent Valuation." *Journal of Economic Perspectives* 8(4):45-64.

Hanemann, W.M., and B. Kanninen. 1999. "The Statistical Analysis of Discrete-Response CV Data." In I.J. Bateman, and K.G. Willis, eds. *Valuing Environmental Preferences - Theory and Practice of the Contingent Valuation Method in the US, EU, and Developing Countries.* Oxford: Oxford University Press, pp. 302-441.

Harrison, G.W., and E.E. Rutström. Forthcoming. "Experimental Evidence on the Existence of Hypothetical Bias in Value Elicitation Methods." In C. Plott, and V.L. Smith, eds. *Handbook of Results in Experimental Economics*. New York: Elsevier Science.

Hausman, J.A., and D. McFadden. 1984. "Specification Tests for the Multinomial Logit Model." *Econometrica* 52(5):1219-1240.

List, J.A. 2001. "Do explicit Warnings Eliminate the Hypothetical Bias in Elicitation Procedures? Evidence from Field Auctions for Sportscards." *The American Economic Review* 91(3):1498-1507.

List, J.A., and C.A. Gallet. 2001. "What Experimental Protocol Influence Disparities Between Actual and Hypothetical Stated Values?" *Environmental and Resource Economics* 20(3):241-254.

List, J.A., P. Sinha, and M.H. Taylor. 2006. "Using Choice Experiments to Value Non-Market Goods and Services: Evidence from Field Experiments" *Advances in Economic Analysis and Policy* 6(2):1-37(article 2).

Louviere, J.J., D.A. Hensher, and J.D. Swait. 2000. *Stated Choice Methods. Analysis and Applications.* Cambridge: University Press.

Lusk, J.L. 2003. "Effects of Cheap Talk on Consumer Willingness-to-Pay for Golden Rice." *American Journal of Agricultural Economics* 85(4):840-856.

Lusk, J.L., and T.C. Schroeder. 2004. "Are Choice Experiments Incentive Compatible? A Test with Quality Differentiated Beef Steaks." *American Journal of Agricultural Economics* 86(2):467-482.

Mitchell, R.C., and R.T. Carson. 1989. *Using Surveys to Value Public Goods: The Contingent Valuation Method.* Washington D.C.: Resources for the Future.

Murphy, J.J., P.G. Allen, T. H. Stevens, and D. Weatherhead. 2005. “A Meta-analysis of Hypothetical Bias in Stated Preference Valuation.” *Environmental and Resource Economics* 30(3):313-325.

Murphy, J.J., T.H. Stevens, and D. Weatherhead. 2005. “Is Cheap Talk Effective at Eliminating Hypothetical Bias in a Provision Point Mechanism?” *Environmental and Resource Economics* 30(3):327-343.

Samnaliev, M., T.H. Stevens, and T. More. 2003. “A Comparison of Cheap Talk and Alternative Certainty Calibration Techniques in Contingent Valuation.” Working Paper No. 2003-11. Dept. of Res. Econ., University of Massachusetts, Amherst.

SAS. 2005. *The MDC procedure*. SAS Institute.

Swait, J., and J.J. Louviere. 1993. “The Role of the Scale Parameter in the Estimation and Comparison of Multinomial Logit Models.” *Journal of Marketing Research* 30(3):305-314.

Train, K. 2003. *Discrete Choice Methods with Simulation.* Cambridge: Cambridge University Press.

Walker, J.L, M. Ben-Akiva, and D. Bolduc. Forthcoming. “Identification of Parameters in Normal Error Component Logit-Mixture (NECLM) Models.” *Journal of Applied Econometrics,* in press.