

Working Paper Series

No. 12

Hypothetical, real, and predicted real willingness to pay in open-ended surveys: experimental results

Anabela Botelho
Lígia Costa Pinto

September 2001

Published in *Applied Economics Letters*, 2002 (9), 993-996

**Núcleo de Investigação em Microeconomia Aplicada
Universidade do Minho**



Hypothetical, Real, and *Predicted* Real Willingness to Pay in Open-Ended Surveys: Experimental Results

Anabela Botelho
Universidade do Minho and NIMA

Lígia Costa Pinto*
Universidade do Minho and NIMA

July 2001

Abstract

This study reports the results of experiments designed to elicit, within a controlled laboratory environment, hypothetical and real willingness to pay for an environmental educational program using the open-ended question format. By maintaining both the good and the question format constant across the treatments, our experiments overcome the shortcomings of recently reported experimental results, providing a clean test for hypothetical bias in open-ended valuations. Having found a statistically significant difference between the hypothetical and real values, we turn into the question of whether hypothetical valuations may nonetheless provide useful statistical information concerning individuals' real valuations. This question, which is perhaps the key question in the current state of the debate surrounding the contingent valuation method, is answered affirmatively in this study.

Keywords: experimental methods, contingent valuation methods, calibration methods
JEL classification: C91, Q26

* Corresponding author: Escola de Economia e Gestão, Universidade do Minho, Campus de Gualtar, 4710-057 Braga, Portugal. E-mail: pintol@eeg.uminho.pt

I. INTRODUCTION

Any consideration of the ability of the contingent valuation (CV) method to measure real willingness to pay (WTP) for environmental goods must address two issues. First, do individuals' responses to hypothetical surveys differ from real surveys? Second, if responses do differ, can any useful information about individuals' real WTP be obtained using contingent valuation? Most of the empirical literature has focused on the first issue, usually (but not always) finding evidence of hypothetical bias in controlled experimental settings.¹ Recently, Johannesson *et al.* (1997) failed to find evidence of hypothetical bias in an experiment intended to replicate the study of Neil *et al.* (1994) who found that hypothetical values significantly exceed real values. Neil (1999), however, argued that no meaningful comparison of the results obtained in the two studies could be made given the differences in methods and procedures used in the studies. More recently, Frykblom (2000) argued that Johannesson *et al.* (1997) did not provide a clean test of hypothetical bias since they used questions formats that differed across the hypothetical and real treatments.²

In this paper we compare hypothetical and real WTP for an environmental educational program using the open-ended question format. By maintaining both the good and the question format constant across the hypothetical and real treatments, we may attribute any differences in the elicited values to a hypothetical bias. Having found a statistically significant difference between the hypothetical and real values, we test whether the hypothetical values are informative as to the real values, which is perhaps

¹ Hypothetical bias is the term used in the CV literature to categorize the difference between responses to real and hypothetical valuation questions. See, for example, Cummings and Harrison (1994) for a review of empirical studies bearing on the existence of hypothetical and other potential sources of bias in contingent valuation.

² Johannesson *et al.* (1997) used the open-ended question format to elicit hypothetical values, and the second-price auction to elicit real values. Frykblom (2000) reports the results of experiments designed to test the null hypothesis that the hypothetical open-ended format and the hypothetical second-price auction format generate equal valuations. The null hypothesis is rejected, thus suggesting that the question format is not neutral with respect to differences (or similarities) between the payment modes (hypothetical or real).

the key issue in the current state of the debate surrounding the CV method. In the following we describe the experimental design, and subsequently we report our results. Concluding remarks are contained in the final section.

II. DESIGN OF THE EXPERIMENT

We conduct between-subjects experiments using the open-ended question format, where one set of subjects faced a real valuation question, and a second set of subjects faced a hypothetical valuation question. All subjects were drawn from the undergraduate student population at the School of Economics and Management, University of Minho in Braga, Portugal.

The good used in the experiments was an informative leaflet on the otter in Neiva River³ produced by an environmental organization.⁴ The information in the leaflet consists of a description of the otter and the human actions that interfere with the otter. The leaflet was distributed to all subjects and they were informed that the approximate cost of producing each leaflet was 60 PTE.⁵

The survey consisted of three sections. The first section asked subjects questions concerning their economic and demographic characteristics. The second section presented information on the otter and on the leaflet. The third section contained the valuation question. In the hypothetical session subjects were asked *how much they would be willing to pay if their were asked to contribute to finance the production of the leaflet*. In the real session subjects were asked *how much they were actually willing to pay to finance the production of the leaflet*. The payment was done at the time of the

³ The Neiva River joins the sea approximately 30 kilometres from Braga.

⁴ The environmental organization is *Associação Rio Neiva* that was in the process of implementing a campaign to alert farmers and fisherman for some current misconceptions about the otter (the otter is currently one of the species protected by CITES – Convention on International Trade of Endangered Species).

⁵ The exchange rate with respect to the Euro is 1 Euro=200.482 PTE.

experiment and a 5-day interest free loan was available to subjects, but no subject used it.⁶

III. RESULTS

The results of the experiments as well as descriptive statistics on the subjects' characteristics collected in the experiments are summarized in Table 1. The sample sizes for the hypothetical and real treatments were 9 and 13, respectively.⁷ These sample sizes are quite small, so the *precise* quantitative results must be interpreted with caution. As can be seen from Table 1, both the mean and median WTP in the hypothetical treatment were substantially higher than those observed in the real treatment.

Table 1. *Experimental results and descriptive statistics of the subjects*

	Treatment	
	Hypothetical	Real
<i>A. Experimental Results</i>		
Sample size	9	13
Mean WTP	3611.11	313.85
Median WTP	1500.00	200.00
Standard deviation of WTP	3846.90	384.70
<i>B. Subject Characteristic (sample means)</i>		
AGE	20.67	22.08
SEX	0.44	0.31
INCOME	241022.00	191400.00
TV	0.56	0.69

Notes: AGE is in years; SEX is 1 for females; INCOME is reported monthly family income; TV is 1 if respondent usually watches environmental programs.

⁶ The experiments were conducted with the collaboration of *Associação Rio Neiva*, and the actual payments were to be used by the environmental organization for the purpose of producing the leaflet.

⁷ These samples sizes are comparable to those of Johannesson *et al.* (1997), which are 10 for the real treatment and 10 for the hypothetical treatment.

We first employed a two-tailed Kolmogorov-Smirnov test to assess the statistical significance of the apparent differences in the observed distributions of hypothetical and real WTP. Despite the small sample sizes, the test has a significance level of 0.1 percent, and we *reject* the null hypothesis that the two distributions come from the same population.

Table 2. *Tobit estimation results*

Variable	Pooled data	Treatment	
		Hypothetical	Real
REAL	-3763.071 (1036.007) [0.002]	—	—
AGE	67.313 (197.321) [0.737]	350.184 (614.940) [0.594]	45.710 (27.126) [0.126]
SEX	-1889.547 (1080.233) [0.098]	-6409.639 (2468.701) [0.048]	444.265 (199.916) [0.053]
INCOME	-0.002 (0.005) [0.735]	0.012 (0.011) [0.325]	-0.6×10 ⁻⁴ (0.8×10 ⁻³) [0.946]
TV	255.042 (1063.784) [0.813]	1219.912 (2083.753) [0.584]	-32.325 (197.571) [0.874]
Constant	3294.987 (4379.668) [0.462]	-4279.006 (13075.48) [0.757]	-798.558 (646.382) [0.248]
N	22	9	13
Log-likelihood	-200.248	-83.751	-91.032
χ ² statistic	11.87	5.57	8.55
P-value	0.037	0.234	0.073

Notes: Dependent variable is WTP. Coefficient estimates reported. Standard errors are in round brackets. P-values are in square brackets.

The same conclusion is reached by comparing the conditional distributions of WTP on the economic and demographic characteristics of the respondents. The conditional test for hypothetical bias proceeded by pooling the data from both treatments, and

regressing WTP on the subjects' characteristics including a dummy variable, REAL, in the specification.⁸ The results using the pooled data, and by treatment are reported in Table 2. The coefficient estimate for REAL is significant at the 0.2 percent level, implying that a bias due to treatment type exists.

Finding evidence for hypothetical bias, we investigate whether the hypothetical values are statistically informative as to the real values. While other approaches have been suggested in the literature to dealing with the hypothetical bias problem, here we follow the statistical calibration approach first suggested by Blackburn *et al.* (1994).⁹ They used within-subjects responses to dichotomous choice hypothetical and real valuation questions to estimate a bias function relating the differences in responses to subjects' characteristics, and applied the estimated function to calibrate the hypothetical responses of a different set of subjects valuing a different commodity. They found that the calibrated hypothetical responses successfully predicted the observed valuations in a paired real treatment.

Our test of the statistical "informativeness" of hypothetical surveys follows these lines in that we use the estimated coefficient vector from the real treatment (reported in Table 2) and the observed economic and demographic characteristics of the subjects in the hypothetical treatment to predict what they *would* have responded *if* they had been placed in the real treatment. The resulting mean and median predicted real WTP are 311.69 PTE and 243.56 PTE, respectively. These figures compare well with the mean and median real WTP: 313.85 PTE and 200.00 PTE, respectively.

The cumulative frequency distribution of this counterfactual predicted real WTP, as well as the cumulative frequency distributions of the observed real and hypothetical WTP are shown in Figure 1.

⁸ The variable REAL takes the value of 1 if the subject participated in the real treatment, 0 otherwise.

⁹ See, for example, Cummings and Taylor (1999) for an approach that focus on alternative designs for hypothetical surveys meant to induce truthful responses from the subjects to the valuation questions.

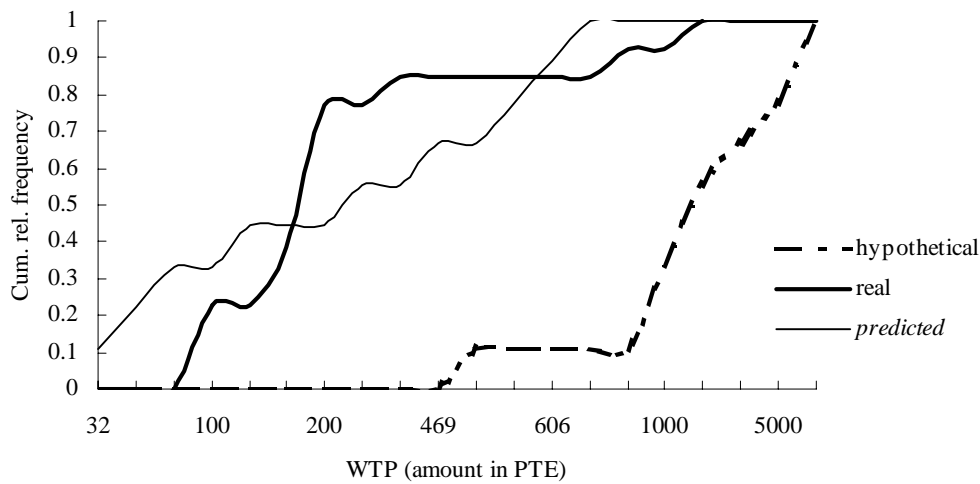


Figure 1. Hypothetical, Real and *Predicted Real* WTP

Qualitatively, the results not only confirm the finding of hypothetical bias, but also indicate that the distribution of predicted real WTP matches well the distribution of real WTP. Finally, we conducted a two-tailed Kolmogorov-Smirnov test of the null hypothesis that real WTP and *predicted* real WTP have the same distribution. The test yields a significance level of 59.6 percent, and we do not reject the null.¹⁰

IV. CONCLUDING REMARKS

In this paper we compare hypothetical and real WTP for an environmental educational program, both assessed with the open-ended question format. In contrast to the study by Johannesson *et al.* (1997), and in line with the results of Neil *et al.* (1994), we found that hypothetical WTP overstates real WTP by a substantial margin, and that the difference is statistically significant despite the small size of our samples. To a degree, this result adds weight to the argument of Frykblom (2000) that, due to

¹⁰ A two-tailed Kolmogorov-Smirnov test of the null hypothesis that hypothetical WTP and *predicted* real WTP have the same distribution yields a significance level of 0.2 percent, and we reject this null hypothesis.

confounding factors, Johannesson *et al.* (1997) cannot validly draw the conclusion of no difference between hypothetical and real WTP from their study.

In addition, our results suggest that, although biased, hypothetical valuations convey useful information about individuals' real WTP, and point to the need for further research on calibration techniques that might eliminate the biases in CV surveys.

REFERENCES

- Blackburn, M., Harrison, G.W. and Rutström, E.E. (1994) Statistical bias functions and informative hypothetical surveys, *American Journal of Agricultural Economics*, 76, 1084—88.
- Cummings, R.G. and Taylor, L.O. (1999) Unbiased value estimates for environmental goods: a cheap talk design for the contingent valuation method, *American Economic Review*, 89, 649—65.
- Frykblom, P. (2000) Willingness to pay and the choice of question format: experimental results, *Applied Economics Letters*, 7, 665—67.
- Johannesson, M., Liljas, B. and O’Conor, R.M. (1997) Hypothetical versus real willingness to pay: some experimental results, *Applied Economics Letters*, 4, 149—51.
- Neil, H.R. (1999) Hypothetical versus real willingness to pay: comment, *Applied Economics Letters*, 6, 267—69.
- Neil, H.R., Cummings, R.G., Ganderton, P., Harrison, G.W., and McGuckin, T. (1994) Hypothetical surveys and real economic commitments, *Land Economics*, 70, 145—54.