The Australian Journal of Agricultural and Resource Economics, 51, pp. 91–107

Ambivalent statements in contingent valuation studies: inclusive response formats and giving respondents time to think

Henrik Svedsater[†]

A recent concern in the valuation literature is the uncertainty respondents feel when posed with willingness-to-pay questions for environmental amenities in hypothetical market scenarios. Using a multiple-bounded discrete-choice format, the results indicate that respondents become less ambivalent when allowed considerable time to think about the valuation task before a response is elicited. In particular they tend to reduce the reported willingness to pay associated with low certainty of paying, hence resulting in more conservative welfare estimates. Implications for the application of environmental valuation techniques are discussed.

Key words: contingent valuation, environmental values, imprecise preferences, multiple bounded choice formats, time to think.

1. Introduction

Past research has revealed ambivalence in respondents' willingness-to-pay (WTP) statements in contingent-valuation (CV) studies (Dubourg *et al*. 1994; Gregory *et al*. 1995; Ready *et al*. 1995). In an attempt to more closely examine this phenomenon, Welsh and Poe (1998) compared a traditional dichotomous-choice (DC) question with a multiple-bounded discrete-choice (MBDC) format where respondents express their payment certainty for a wide range of WTP amounts. The latter format resulted in a wide range between amounts the respondents were sure of paying and amounts they were sure not to pay, suggesting that people experience a large uncertainty when asked to provide economic values for environmental goods and services. Some alternative payment mechanisms have since been developed that aim to capture and measure this type of ambivalence (Li and Mattsson 1995; Champ *et al*. 1997; Alberini *et al*. 2003).

However, there is nothing in these studies that suggests how people's uncertainty may be reduced by alternative valuation procedures that differ in their layout and contextual framing. We may have reason to question people's ability to provide precise monetary estimates to comprehensive issues that otherwise are difficult to reduce to a single measure. Hence, a large uncertainty

[†]Henrik Svedsater (email: henrik.svedsater@psy.gu.se), Department of Psychology, Gothenburg University, Box 500, 40530 Gothenburg, Sweden.

doi: 10.1111/j.1467-8489.2007.00346.x

may arise when respondents are asked to resolve and express attitudes toward a wide array of aspects incorporated by the public good, including their attitude toward the suggested payment scheme, their sense of responsibility, and ability to pay. Furthermore, given that 'rapid' responses to complex issues are unreliable (Schkade and Payne 1994; Fischhoff *et al*. 1999), and rather should be interpreted as context-dependent constructions, an important question is: are WTP statements assessed on the spot valid representations of underlying preferences? A CV question is also rather novel to most people as it proposes a way of making environmental priorities that is unfamiliar, and which does not correspond to how public issues are normally decided.

This paper takes notice of these potential caveats or limitations. Willingness to pay is first assessed using an MBDC format that allows the respondents to express their uncertainty about stated WTP. This format is then compared with a more inclusive response format in which respondents express attitudes toward multiple dimensions of the public good, rather than solely stating their support through a single estimate. Furthermore, in one setting respondents are given considerable time to think about the valuation task before a response is elicited. The main hypothesis is that people's opinions toward environmental resources are at best only vaguely represented in monetary terms (Blamey 1998; Kahneman *et al*. 1999). The paper provides specific tests in order to understand how the above manipulations reduce any ambivalence experienced by the respondents. The results confirm that responses become more precise when respondents are given substantial time to think about the valuation issue. This format also tends to lower the WTP for the environmental good being valued. No effect, however, is demonstrated from using a more inclusive response format.

1.1 Inclusive response formats and avoidance of 'yea-saying'

When faced with a DC question, people may be inclined to answer affirmatively (Ready *et al*. 1999). Blamey *et al*. (1999) denote this tendency as 'yea-saying', arising from people's inclination to agree with statements regardless of their knowledge or consent. Although the majority of CV proponents argue that this is primarily a statistical nuisance, others view this type of response acquiescence as 'inherent' in human decision-making (Bachman and O'Malley 1984). When measuring people's attitudes toward novel issues in a manner that does not necessarily have a bearing on everyday decision-making, this effect is likely to be even more pronounced (Schwarz 1994). Support for this assumption is found in Svedsäter (2003), who demonstrated that some respondents tend to provide WTP amounts unassociated with their underlying preferences or despite their disapproval of a monetary valuation of the public good.

Another reason for 'yea-saying' in CV contexts, and equally the uncertainty of stated WTP, is that respondents incorporate aspects and dimensions other than just instrumental value in their answers. The subordination of outcome-based considerations in favour of expressive motivations, particularly

when conflicting with each other, may lead respondents to experience uncertainty in their open-ended value statements. Similarly, when the latter motive adds to the former, overstatements of WTP are likely to occur. Hence, rather than confining ambivalence to indifference as suggested by Ready *et al*. (1995), imprecise responses may be much more complex phenomena that arise when the individual tries to resolve and express multiple and often conflicting opinions (Zaller and Feldman 1992).

Blamey *et al*. (1999) propose a response format allowing respondents to reflect and respond separately to various distinct attributes of the public good that are anticipated to reflect multiple dimensions of something that we call total value. Apart from making the respondents aware of important factors underlying the valuation task, they are permitted to state which particular aspects of the policy proposal they favour or oppose. For example, people may support an environmental project due to its benefits, but be reluctant to provide any monetary value if they believe it does not lie within their responsibility to do so. Similarly, they may like the idea of the project, but be sceptical of whether the proposed intervention will achieve what is intended to achieve.

1.2 Giving respondents time to think in valuation studies

Apart from in mail surveys, CV respondents are usually asked to provide rapid answers to the valuation question. They are hence not allowed the opportunity to thoroughly reflect on their opinions outside the interview context. One way to investigate the reliability of CV estimates is to perform so-called test–retests of WTP, whereby the same respondents are asked to provide answers to a particular valuation question at two distinct points of time. CV respondents seem fairly consistent in these assessments (Loomis 1989; Teisl *et al*. 1995). However, a problem arises if respondents recall what they said on the first occasion and anchor their estimates accordingly during the second interview. Despite the fact that the elapsed time between interviews was up to 9 months in some cases, it is still likely that the respondents recall their previous responses. The use of control groups does not solve this problem, as they only indicate that different individuals provide similar estimates, not how individuals are affected by the elapsed time between the presentation of the environmental problem and assessment of value.

To our knowledge, there is only one study that has explicitly addressed this issue. Whittington *et al*. (1992) investigated whether WTP from respondents who were given time to evaluate a proposed water system in three Nigerian villages, differed from those who were not. Their findings indicate that respondents in the former group were willing to pay significantly less than the latter. Other attitudinal research has shown that opinions tend to become polarised as a result of thinking through an issue, particularly among individuals who initially were less involved with the issue, and when the information and arguments presented are novel (Tesser and Conlee 1975; Harton and Latané 1997). There are also examples of studies that have investigated

attitudes at two different points of time in a manner that minimises the carry-over effects of previous value elicitations. Such experiments show that responses can change considerably if questions are repeated after some time has elapsed (Jagodzinski and Kuhnel 1988).

An individual's opinion may change simply as a result of thinking through an issue. Also, discussion with friends and relatives that is facilitated during this process is likely to have an impact on responses. (Latané 1981; Kruglanski and Mayseless 1987). Social interaction is envisioned to have at least two consequences for value assessment. First, when little or no information exists about a particular issue, and when people are unfamiliar with the valuation task, the consultation with other people may be useful in constructing a judgement (Kaplowitz and Hoehn 2001; Svedsäter 2003). Second, other people's judgements are important in order to develop and validate opinions about issues that the individual is already familiar with. The views of others reinforce the subjective validity of our own beliefs (Festinger 1950; Zimbardo and Leippe 1991), for instance, by providing supportive or non-supportive reasons for a particular attitude.

2. Design of experiment and model

Data were collected among undergraduate and graduate students at the London School of Economics (LSE). One group of respondents was asked during class teaching if they wanted to participate in the study. If they were, they were asked to sign up on an attendance list and were later notified through email. Another group of respondents was randomly selected from the school's register and were told about the study through email. From the list of self-selected subjects, a second email was sent that suggested various times for them to attend. Altogether, ten sessions with between 8 and 17 respondents in each were run. Each session lasted for approximately 30 min, and the respondents were reimbursed with £GBP5 for their participation. The participants came from a variety of fields in the social sciences and different course categories were evenly represented across all subsamples.

2.1 Experimental design

The questionnaire consisted of between 13 and 15 questions depending on the particular response format. It included instructions for the valuation task, some questions related to socio-economic characteristics, knowledge questions about the amenity being valued, attitude questions addressing the appropriateness of an economic valuation of natural resources, and finally, a description of the proposed environmental project. The valuation task was specified as a voluntary contribution for saving the African elephant, a campaign currently run by the World Wildlife Fund (WWF).

In order to compare the influence of various response formats, three variations were presented to independent groups. In all subsamples, the respondents

Table 1 Multiple-bounded response format

	£2	£5	£7	£10	£15	£20	£30	£50	£100	£200	£400
I am **definitely sure** that I will pay											
I am **almost certain** (90 per cent sure) that I will pay											
I am **rather certain** (75 per cent sure) that I will pay											
It is **equally likely** (50 per cent sure) that I will pay											
I am **rather certain** (75 per cent sure) that I will **not** pay											
I am **almost certain** (90 per cent sure) that I will **not** pay											
I am **definitely sure** that I will **not** pay											

were presented with an MBDC valuation format that presented an ordered sequence of WTP amounts (Welsh and Poe 1998). The difference compared with a standard DC question is that in the format used here they were told to indicate how certain they were of paying each of the amounts suggested in the valuation question. This allows respondents to express their degree of uncertainty associated with each bid threshold, from one bid that they are definitely sure of paying (lower bound), to one that they are definitely sure that they will not pay (upper bound). What is also different from previous studies is that each certainly level, apart from the 'definitely yes' and the 'definitely no' response, is associated with a numerical probability, specified as 90, 75, 50, 25, and 10 per cent sure of paying. We argue that this avoids potential problems of subjective probabilities associated with phrases such as 'probably yes', 'almost certain', etc. It furthermore enables a direct estimation of weighted models that summarise all response categories into one function (Evans *et al*. 2004). The instructions preceding the task of indicating payment certainties of various WTP amounts are presented below, and Table 1 presents the specific certainty thresholds and bid amounts used.

> In the table below you are presented with 11 different amounts. We want you to state how sure you are of paying each of these as a contribution to the WWF campaign for saving the African elephant. Please tick the appropriate box for each suggested amount. The willingness to pay is an annual payment. Take your time and try to consider the following before answering:
>
> - Your income and/or grants
> - Your current expenses
> - Your possible future use of your income

2.1.1 Baseline scenario

The respondents were told that the amount should represent their maximum WTP associated with each threshold, or alternatively, an amount beyond

which they are not willing to pay. Thus, rather than providing a single point estimate, the format discloses a range of possible WTP amounts that the individual is more or less sure of paying. In addition to facilitating direct estimations of mean and median WTP, the design reveals whether different approaches to estimating WTP make respondents more or less ambivalent about their answers. The particular range of amounts was chosen on the basis of the results of a CV study using an open-ended (OE) elicitation format, assessed prior to the main study. The median WTP of this format was £GBP25, with a mean WTP of £GBP41.80 (n = 35).

2.1.2 Inclusive response format

Exactly the same design, range of bids, and certainty thresholds as presented in Table 1 were used in all three experimental conditions. In one of these, the questionnaire was formulated as above. In the scenario that allows respondents to reflect several dimensions of the environmental project, rather than merely stating their WTP, the following questions preceded the valuation questions:

To save the African elephant is worth something to me.	☐	Agree
	☐	Disagree
To save the African elephant is generally an important issue.	☐	Agree
	☐	Disagree
I cannot afford to pay too much for this issue.	☐	Agree
	☐	Disagree
I do not believe the particular campaign suggested will be efficient in saving the African elephant.	☐	Agree
	☐	Disagree
I do not think this lies within my responsibility. Poachers and other responsible parties should pay.	☐	Agree
	☐	Disagree
There are other environmental issues that are more important and to which I rather contribute.	☐	Agree
	☐	Disagree
Although being worth a lot to me, I do not think it is appropriate to base policies on the public's WTP.	☐	Agree
	☐	Disagree

The format thus highlights various potentially important aspects of the valuation task and permits respondents to consider each of them explicitly. It is hypothesised that this results in different and possibly less ambivalent responses to the subsequent WTP question. The format is based on Blamey *et al*. (1999), arguing that a more comprehensive elicitation format, capturing a wide array of attributes and dimensions of the public good and its provision, would facilitate more conservative WTP responses as the individual, for example, may express that they find the environmental issue important, but (for other reasons) are reluctant to pay for it, at least in this manner. Although the format used here does not allow the respondents to respond with a monetary value to each specific aspect of the environmental good, nor facilitate a decoupling of economic commitments from a more general support, we are still inclined to argue that simply reminding the respondents that there are many

reasons why one may support (or not support) the campaign will affect WTP. Specifically, the tendency of 'yea-saying' is assumed to reduce within this format.

2.1.3 Time-to-think response format

In the final subsample, the respondents were given considerable time to think about the environmental project and the valuation task before a WTP was elicited. The following information preceded the valuation question for this group of respondents:

> We want you to consider the question below. However, we do not want you to answer it now. Instead, you will be given a week or so to think about a monetary contribution to the WWF campaign. During this time we encourage you to discuss the environmental problem, as well as an economic valuation of this, with friends, spouse, relatives, etc. We also want you to think of your opinion when similar (environmental or public) issues are presented in the media. Although receiving valuable comments from other sources, keep in mind that it is your own opinion that we are interested in.

The questionnaire was separated into two parts and the experiment was conducted in two sessions. In the first, the respondents were informed about the WWF campaign and answered some related knowledge questions. They were also presented with the valuation question and the range of WTP amounts illustrated in Table 1, but were told to only use this as a guide for their subsequent responses, being urged not to answer the question as we did not want them to commit to any response at this stage. After completing this part of the questionnaire, they were asked to bring with them the first part of the questionnaire to the next occasion we met, at which time they provided their WTP. Respondents were given between 7 and 10 days to think about the issue. The subjects were paid £GBP5 for each occasion they turned up.

To summarise the design of the study, three different valuation scenarios or formats were applied: (i) an MBDC question with certainty thresholds similar to the one used by Welsh and Poe (1998) (in the following denoted *MBDC baseline*); (ii) an MBDC question that presented several questions relating to various dimensions of the environmental good (*MBDC inclusive*); and finally, (iii) an MBDC question where the respondents were given time to think before they provided their WTP (*MBDC time to think*).

2.2 Econometric models

The data were analysed using a multiple-bounded generalisation of double-bounded models (Hanemann *et al*. 1991). These models assume that an individual's WTP is situated somewhere in the interval between a 'no' and a 'yes' response. For instance, if the individual replies 'yes' to a first bid of $10 and 'no' to a follow-up bid of $20, WTP lies somewhere between $10 and $20. Translated to the MBDC technique utilised here, for a given certainty level,

WTP lies somewhere in the switching interval between a 'yes' and a 'no' response. Hence, the lower bound of WTP, X_{iL}, is the maximum amount the individual is willing to pay, whereas the upper bound, $X_{iU,}$ is the next higher bid. By letting $F(X_i; \beta)$ represent a statistical distribution function of WTP_i with parameter vector β, the probability that individual i would vote no to a specific amount, X, is simply $F(X_i; \beta)$, whereas the corresponding probability for a yes-vote is 1-$F(X_i; \beta)$. The probability that WTP_i falls between two bid amounts is $F(X_{iU}; \beta) - F(X_{iL}; \beta)$, which gives the following log-likelihood function:

$$\ln(L) = \sum_{i=1}^{n} \ln[F(X_{iU}; \beta) - F(X_{iL}; \beta)] \quad (1)$$

If the individual responds 'yes' to all amounts provided, $X_{iU} = \infty$, and if she responds 'no' to all amounts, $X_{iL} = -\infty$. In order to facilitate analysis of such cases, the same approach is used as in the analysis of discrete choice models and interval data from payment cards. When the upper bound of WTP is unobserved, such as when a yes response is obtained for all bids, the probability associated with this observation is $1 - F(X_{iL}; \beta)$. When on the other hand the lower bound of WTP is unobserved (the respondent answers no to the lowest bid), the associated probability is $(X_{iU}; \beta)$. Incorporating the latter cases in Equation (1) results in the following log-likelihood function:

$$\ln(L) = \sum_{i=1}^{n} \ln[1 - F(X_{iL}; \beta)] + \sum_{i=1}^{n} \ln[F(X_{iU}; \beta) - F(X_{iL}; \beta)] + \sum_{i=1}^{n} \ln[F(X_{iU}; \beta)] \quad (2)$$

On the basis of this function, two different estimation techniques were employed. First, we used the same approach as Welsh and Poe (1998) where a separate analysis is made of each certainty level. For example, in the 'definitely sure' model, the lower bound is set at the highest amount that the respondent accepts to pay and the upper bound at the next higher bid.

Additionally, we used an alternative approach suggested by Evans *et al.* (2004) that weights responses of each certainty level. Instead of analysing each certainty level separately, this technique results in a single estimate of WTP. In this analysis, the log-likelihood function of the '100 per cent sure' model is multiplied by 1, the '90 per cent sure' model by 0.9, the '75 per cent sure' model by 0.75, etc. (see Equation 3 below).

3. Results

Altogether 111 students participated in the main study, with 37 students randomly allocated to each of the three response formats. Logit and probit models were run separately for each certainty level. These models included intercept variables capturing the impacts of the bid amounts and response formats tested, and two variables representing the influence of gender and

income (presented in Appendix). Students with higher incomes and women were more likely to support the WWF campaign. The size and direction of these effects are similar across all models and response formats, and are therefore excluded from the following analyses.

Table 2 presents the parameter estimates of regressions associated with our response formats and certainty levels. Responses were analysed using Equations (2) and (3). For ease of interpretation, only estimates associated with the 100, 75, 50, and 25 per cent-sure model are presented here. The weighted model correspondingly includes these four certainty levels.[1] Hence, Equation (2) is in the analysis modified to be:

$$\ln(L) = \left(\sum_{i=1}^{n} \ln[1 - F(X_{iL};\beta)] + \sum_{i=1}^{n} \ln[F(X_{iU};\beta) - F(X_{iL};\beta)] + \sum_{i=1}^{n} \ln[F(X_{iU};\beta)] \right)$$
$$+ 0.75^{*} \left(\sum_{i=1}^{n} \ln[1 - F(X_{iL};\beta)] + \sum_{i=1}^{n} \ln[F(X_{iU};\beta) - F(X_{iL};\beta)] + \sum_{i=1}^{n} \ln[F(X_{iU};\beta)] \right)$$
$$+ 0.5^{*} \left(\sum_{i=1}^{n} \ln[1 - F(X_{iL};\beta)] + \sum_{i=1}^{n} \ln[F(X_{iU};\beta) - F(X_{iL};\beta)] + \sum_{i=1}^{n} \ln[F(X_{iU};\beta)] \right)$$
$$+ 0.25^{*} \left(\sum_{i=1}^{n} \ln[1 - F(X_{iL};\beta)] + \sum_{i=1}^{n} \ln[F(X_{iU};\beta) - F(X_{iL};\beta)] + \sum_{i=1}^{n} \ln[F(X_{iU};\beta)] \right) \quad (3)$$

Table 2 Estimated econometric models (standard errors of parameter estimates are presented in brackets)

	MBDC benchmark		MBDC inclusive		MBDC time to think	
	α	β	α	β	α	β
Definitely sure	1.061**	–0.187**	1.325**	–0.219**	0.916*	–0.194**
	(0.390)	(0.027)	(0.442)	(0.021)	(0.385)	(0.029)
75 per cent certain	1.576**	–0.107**	1.301*	–0.100**	1.397**	–0.121**
	(0.450)	(0.014)	(0.546)	(0.009)	(0.420)	(0.016)
50 per cent certain	1.547**	–0.064**	1.163*	–0.061**	1.480**	–0.092**
	(0.541)	(0.009)	(0.545)	(0.006)	(0.492)	(0.016)
25 per cent certain	1.727**	–0.035**	1.223*	–0.024**	1.611**	–0.060**
	(0.603)	(0.006)	(0.599)	(0.003)	(0.567)	(0.01)
Weighted model	1.034**	–0.076**	0.944**	–0.073**	1.054**	–0.105**
	(0.230)	(0.006)	(0.266)	(0.004)	(0.217)	(0.009)
n	37		37		37	

Note: * and ** denote significance at the 0.05 and 0.01 level, respectively.

[1] The remaining two certainty levels (90 and 10 per cent sure) were also analysed but added nothing to the analysis.

Results in Table 2 show that the estimated intercept (α) and slope (β) of each model are significant at the 0.05 or 0.01 level. As expected, for each response format, the slopes of the functions become less steep as the level of certainty declines. In other words, the less certainty that is specified, the more likely the respondent accepts higher bids. Furthermore, parameter estimates of the weighted model show that both the intercepts and slope coefficients are somewhat lower than in the models representing specific certainty levels.

Our main hypothesis that the parameter estimates of various response formats come from different distributions is evaluated through likelihood ratio (LR) tests. Hence, to what extent the coefficients of two different models are statistically different is defined by:

$$\mathrm{LR} = -2^{*}[\ln L_{pooled} - (\ln L_{restrictedA} + \ln L_{restrictedB})] \sim \chi^2(r) \tag{4}$$

where $\ln L_{restrictedA}$ and $\ln L_{restrictedB}$ are the log-likelihood functions associated with each of the specific models, and $\ln L_{pooled}$ is the log-likelihood function associated with the pooled model of A and B. Table 3 presents the results of LR tests.

Table 3 Likelihood-ratio tests

	MBDC benchmark vs MBDC inclusive	MBDC benchmark vs MBDC time to think	MBDC inclusive vs MBDC time to think
Definitely sure	0.45	0.22	0.70
75 per cent certain	0.33	1.11	0.79
50 per cent certain	0.78	4.35	3.41
25 per cent certain	2.78	10.37**	18.38**
Weighted model	0.12	6.16*	6.84*

Note: * and ** denote significance at the 0.05 and 0.01 level, respectively.

At higher levels of certainty (i.e., '100% sure' and '75% sure'), the null hypothesis of equality between coefficients cannot be rejected. However, for certainty levels '50 per cent sure' and '25 per cent sure', the coefficients of the MBDC time-to-think format differ from the MBDC benchmark format. Furthermore, the parameter estimates of the weighted model are significantly different between the MBDC time-to-think and the MBDC benchmark formats, and between the MBDC time-to-think and MBDC inclusive formats ($P < 0.05$). The MBDC time-to-think format thus produces different estimates than the other two response formats, at least at lower levels of certainty. We may also conclude that the MBDC inclusive format does not yield significantly different parameter estimates than the MBDC benchmark model. Thus, a more inclusive response format, as utilised here, does not seem to alter the results of a traditional CV format.

Given these results, we are interested in estimating mean and median values of each response format and certainty level. These are estimated via the non-negative mean (Mean WTP = $[-\ln(1 + \exp\alpha)/\beta]$), and the analytical

Table 4 Mean and median values in GBP(£) associated with different response formats

	MBDC benchmark		MBDC inclusive		MBDC time to think	
	Mean	Median	Mean	Median	Mean	Median
Definitely sure	7.25	5.66	7.12	6.04	6.46	4.73
	(1.20)	(1.68)	(1.25)	(1.69)	(1.13)	(1.65)
75 per cent certain	16.54	14.78	15.35	12.95	13.36	11.53
	(2.38)	(3.05)	(3.33)	(4.60)	(1.99)	(2.63)
50 per cent certain	27.24	24.22	23.62	19.15	18.39	16.16
	(4.32)	(5.89)	(5.38)	(7.72)	(2.60)	(3.54)
25 per cent certain	53.89	49.23	60.49	49.95	29.66	26.65
	(7.77)	(10.46)	(13.11)	(19.43)	(4.43)	(6.06)
Weighted model	17.64	13.63	17.36	12.87	12.83	9.99
	(1.57)	(2.38)	(2.07)	(3.19)	(1.00)	(1.53)

Note: standard errors of estimates in brackets.

median (Median WTP = $[-\alpha/\beta]$) and are presented in Table 4. Standard errors are calculated using the Delta method (Cameron 1991).

The MBDC time-to-think format yields lower mean and median values. Consistent with patterns of previous statistical analyses, the difference is greater the lower the certainty. For instance, whereas median values are only marginally lower in the '100 per cent sure' and '75 per cent sure' model, for the '50 per cent sure' and '25 per cent sure' model these amount to roughly 67 and 54 per cent of the MBDC benchmark format, respectively. The values of the weighted model also indicate that mean (but not median) WTP are significantly lower in the MBDC time-to-think model than for the other two formats ($P < 0.05$). Thus, the MBDC time-to-think format results in a significant decrease of mean WTP when assessed simultaneously for all certainty levels.

The differences of mean and median values between the MBDC inclusive and MBDC benchmark formats are less consistent, where the latter produces higher values for the '25 per cent sure' level, but lower values for the remaining three certainty levels. However, neither of these results is statistically significant, as indicated in Table 3. The similarity of mean and median values between these response formats is also demonstrated in the weighted model.

Another important insight is that the weighted MBDC time-to-think model produces mean and median values that are lower than those of the '75 per cent sure' model, whereas the weighted models of the other two response formats result in WTP estimates that are higher than the '75 per cent sure' model. This implies that the weighted MBDC time-to-think model results in more conservative estimates of WTP. This response format also generates a narrower range of mean and median values between the '100 per cent sure' and the '25 per cent sure' models.

Overall, the mean and median values of the weighted model are lower than those by Evans *et al*. (2004). However, given that Evans *et al*. (2004) did not

include exactly the same certainty levels as used here, and because the characteristics of the public goods being valued and the respondents of each study may be (very) different, the results are not directly comparable.

The final analysis is concerned more specifically with the ambivalence bounds associated with various certainty levels. From the estimated logit regressions, ambivalence regions were constructed for each response format. The lower ambivalence bound is defined as the amount to which 50 per cent of the respondents would respond with 'I am definitely sure that I will pay' (i.e., 100% sure), whereas the upper bound is defined as an amount to which 50 per cent would respond with 'I am definitely sure that I will not pay'. The bounds are calculated with other variables in the model are set equal to their mean values in the sample. In order to take account of the possible sensitivity of the results of these categorisations to what is considered a 'definitely yes' response and what is considered as a 'definitely no' response, two alternative definitions of lower and upper bounds were used, defined as the amounts to which 60 and 70 per cent of the respondents would respond with 'I am definitely sure that I will pay' and 'I am definitely sure that I will not pay', respectively.

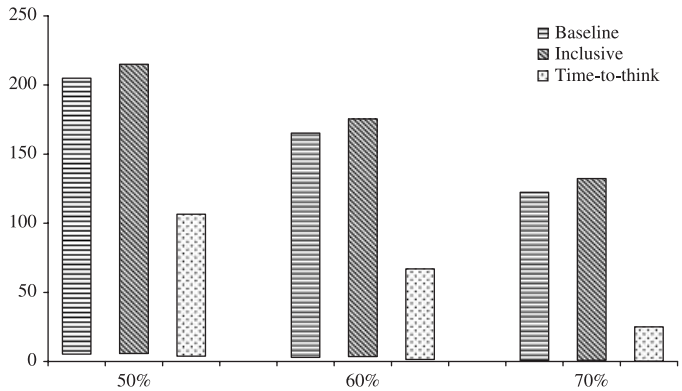

Figure 1 Ambivalence bounds of WTP, measured in GBP(£).

The difference between lower bounds of WTP in Figure 1 is not as pronounced as between the upper bounds, a result that is consistent across all three definitions of ambivalence regions (i.e., for 50, 60, as well as 70%).[2] Whereas the MBDC time-to-think format results in much narrower ambivalence regions, these are approximately the same for the other two formats. Therefore, the region of ambivalence among the respondents who were given

[2] However, due to the scale of the value axis, the precise difference between the lower bounds across formats is concealed in the figure. The exact ambivalence regions were as follows, presented in the order MBDC benchmark, MBDC inclusive, and MBDC time to think; 50 per cent (5.4–199.5, 5.8–209.2, 3.9–102.7); 60 per cent (3–162.3, 3.5–172, 1.5–65.5); 70 per cent (0.5–121.9, 0.9–131.4, 0.1–24.9).

time to think about the valuation task is, according to the above definitions, half or less than half than that of the remaining two subsamples, leading us to conclude that the respondents in this subgroup provide less ambivalent estimates of WTP. More specifically, they tend to revise downwards their responses toward bids that are associated with lower certainty levels. The parameter estimates of the models presented in Table 2, as well as mean and median values presented in Table 4, support this interpretation.

4. Discussion

The fact that people provide answers in surveys and interviews, despite being uncertain about their opinions, has long been recognised in social research (Converse 1970; Nadeau and Niemi 1995). In a CV context, Svedsäter (2003) has shown that stated WTP amounts do not necessarily reflect whether people consent to the valuation procedure or not, and there seems to be a large degree of uncertainty involved regarding what would properly reflect individual economic value. This study examines this type of ambivalence, expressed as the difference between various thresholds of subjective certainty regarding the likelihood that the individual will actually pay the WTP bid presented. Three variants of an MBDC format were applied, one which posed questions aimed at capturing various multidimensional aspects of the proposed project and the valuation task, one that allowed the respondents considerable time to think about the issue before responding, and an MBDC format employed in previous research that relies on a more traditional CV design.

The most important finding is that people's responses turn out to be different when they are given time to think about the valuation task, particularly at lower levels of certainty where respondents adjust their WTP downwards. Whereas a respondent's immediate reaction to the valuation question seems to be 'I would probably do X, and may even do Y', this now becomes 'I would still probably do X, but now I don't think I would do Y'. As a consequence, the range of WTP estimates between lower and higher levels of certainty are narrower among this group of respondents. The ambivalence regions, measured as the difference between a bid that the individual is sure of paying and one that she is sure not to pay, is less than half as wide among the respondents given time to think about the issue, compared with the other two MBDC response formats. It is assumed that this process enables respondents to put the issue in a broader context, whereby competing public issues, personal responsibility, and budget constraints are realised.

Vossler *et al*. (2003) have shown that results of the MBDC format mimic actual purchasing decisions only when the 'definitely yes' and 'probably yes' levels are treated as 'yes' responses and other levels as 'no' responses. We may therefore question to what extent giving respondents time to think will in fact have any implications for value estimation. However, respondents allowed time to think provide lower WTP estimates also at higher levels of certainty.

This has important implications for weighted models of benefit estimation that aim to produce single WTP estimates (Evans *et al*. 2004). Additionally, the MBDC approach is not the sole technique of benefit estimation. When using standardised DC formats without certainty thresholds, we do not exactly know what level of uncertainty a response entails, and unless people are very conservative in their assessments, such formats are likely to bias estimates upward.

The results furthermore correspond closely with Whittington *et al*. (1992), thus indicating that the findings are not necessarily limited to the sample group used here, the good presented, or the context in which it is evaluated. Findings in other areas of social science indicate that responses to attitude surveys tend to become more polarised and conservative when respondents are allowed time to evaluate the attitude object, contributing to their stability over time (Kaplan and Miller 1977). In this process, participants are also likely to place more focus on instrumental outcomes associated with the good, whereby affective reactions and symbolic values are downplayed. Ultimately this leads to a lower WTP. Other studies show that responses can be considerably different if some time elapses between the presentation of questions and elicitation of answers (Jagodzinsky and Kuhnel 1988).

An implication of the findings presented here is that statements of WTP for a public good are only vaguely represented in people's minds. Therefore, these should not be treated as point estimates, but rather as measures that fall within a wider region of ambivalence that may (or may not) capture any underlying 'true' value. Furthermore, when individuals are unsure of how much they are actually willing to pay, they tend to bias their responses upwards. One way of dealing with these problems is to apply an elicitation format that reveals the width of this ambivalence upon which approximate upper and lower bounds of welfare estimates may be calculated. Such formats also enable an investigation into which particular environmental issues people seem to possess more crystallised attitudes and values. Apart from the possibility of distinguishing well-founded values from non-attitudes in this sense, this procedure is in itself also likely to remind the respondents of how certain they actually are of paying the amounts suggested.

The next step is to develop methodologies that aim to reduce the uncertainty that the respondents feel when answering a CV question. Encouragement and inducement to take more time in answering the valuation question seems like a fruitful approach in order to fulfil this aim. This procedure also has the capacity to reveal whether there exists a predisposition among individuals to respond in a consistent manner toward a given attitude object. In order to speak of meaningful attitudes, the same question ought to produce consistent responses over time. By comparing results from immediate responses with those that are assessed after people have been thinking through the valuation task, it is possible to distinguish stable values from context-dependent temporary constructions.

References

Alberini, A., Boyle, K. and Welsh, M. (2003). Analysis of contingent valuation data with multiple bids and response options allowing respondents to express uncertainty, *Journal of Environmental Economics and Management* 45, 40–62.

Bachman, J.G. and O'Malley, P.M. (1984). Yea-saying, nay-saying, and going to extremes: black–white differences in response styles, *Public Opinion Quarterly* 48, 491–509.

Blamey, R.K. (1998). Decisiveness, attitude expression and symbolic responses in contingent valuation surveys, *Journal of Economic Behaviour and Organization* 34, 577–601.

Blamey, R.K., Bennett, J.W. and Morrison, M.D. (1999). Yes-saying in contingent valuation surveys, *Land Economics* 75, 126–141.

Cameron, T.A. (1991). Interval estimates of non-market resource values from referendum contingent valuation surveys, *Land Economics* 67, 413–421.

Champ, P.A., Bishop, R.C., Brown, T.C., and McCollum, D.W. (1997). Using donation mechanisms to value nonuse benefits from public goods, *Journal of Environmental Economics and Management*, 33, 151–162.

Converse, P.E. (1970). Attitudes and non-attitudes: continuation of a dialogue, in Tufte, E.R. (ed.), *The Quantitative Analysis of Social Problems.* Addison–Wesley, Reading, MA, pp. 168–189.

Dubourg, W.R., Jones-Lee, M.W. and Loomes, G. (1994). Imprecise preferences and the WTP–WTA Disparity, *Journal of Risk and Uncertainty* 9, 115–133.

Evans, M.F., Flores, N. and Boyle, K.J. (2004). *A decision theoretic approach to modeling multiple-bounded uncertainty choice data*, Working Paper, Center for Economic Analysis, Department of Economics, University of Colorado, Boulder.

Festinger, L. (1950). Informal social communication, *Psychological Review* 57, 271–282.

Fischhoff, B., Welch, N. and Frederick, S. (1999). Construal processes in preference assessment, *Journal of Risk and Uncertainty* 19, 139–164.

Gregory, R., Lichtenstein, S., Brown, T.C., Peterson, G.L. and Slovic, P. (1995). How precise are monetary representations of environmental improvements? *Land Economics* 71, 462–473.

Hanemann, W.M., Loomis, J. and Kanninen, B. (1991). Statistical efficiency of double-bounded dichotomous choice contingent valuation, *American Journal of Agricultural Economics* 72, 1255–1263.

Harton, H.C. and Latané, B. (1997). Information and thought-induced polarization: the mediating role of involvement in making attitudes extreme, *Journal of Social Behavior and Personality* 12, 271–300.

Jagodzinski, W. and Kuhnel, S.M. (1988). Estimation of reliability and stability in single indicator multiple-wave models, *Sociological Methods and Research* 15, 219–258.

Kahneman, D., Ritov, I. and Schkade, D. (1999). Economic preferences or attitude expressions?: an analysis of dollar responses to public issues, *Journal of Risk and Uncertainty* 19, 203–235.

Kaplan, M.F. and Miller, C.E. (1977). Judgments of group discussion: effect of presentation and memory factors, *Sociometry* 40, 337–343.

Kaplowitz, M.D. and Hoehn, J.P. (2001). Do focus groups and individual interviews reveal the same information for natural resource valuation? *Ecological Economics* 36, 237–247.

Kruglanski, A.W. and Mayseless, O. (1987). Motivational effects in the social comparison of opinions, *Journal of Personality and Social Psychology* 53, 834–842.

Latané, B. (1981). The psychology of social impact, *American Psychologist* 36, 343–356.

Li, C.Z. and Mattsson, L. (1995). Discrete choice under preference uncertainty: an improved structural model for contingent valuation, *Journal of Environmental Economics and Management* 28, 256–269.

Loomis, J.B. (1989). Test–retest reliability of the contingent valuation method: a comparison of general population and visitor responses, *American Journal of Agricultural Economics* 71, 76–84.

Nadeau, R. and Niemi, R.G. (1995). Educated guesses: the process of answering factual knowledge questions in surveys, *Public Opinion Quarterly* 59, 323–346.

Ready, R.C., Navrud, S. and Dubourg, W.R. (1999). *How do respondents with uncertain willingness to pay answer contingent valuation questions?* Paper presented at the 9th EAERE Annual Conference, 25–27 June 2001, Oslo, Norway.

Ready, R.C., Whitehead, J.C. and Blomquist, G.C. (1995). Contingent valuation when respondents are ambivalent, *Journal of Environmental Economics and Management* 29, 181–196.

Schkade, D.A. and Payne, J.W. (1994). How people respond to contingent valuation questions: a verbal protocol analysis of willingness to pay for an environmental regulation, *Journal of Environmental Economics and Management* 26, 88–109.

Schwarz, N. (1994). Judgment in a social context: biases, shortcomings, and the logic of conversation, *Advances in Experimental Social Psychology* 26, 123–162.

Svedsäter, H. (2003). Economic valuation of the environment: how people make sense of contingent valuation questions, *Land Economics* 79, 122–135.

Teisl, M.F., Boyle, K.J., McCollum, D.W. and Reiling, S.D. (1995). Test–retest of contingent valuation with independent sample pretests and post-test control groups, *American Journal of Agricultural Economics* 77, 613–619.

Tesser, A. and Conlee, M.C. (1975). Some effects of time and thought on attitude polarization, *Journal of Personality and Social Psychology* 31, 262–270.

Vossler, C.A., Ethier, R.G., Poe, G.L. and Welsh, M.P. (2003). Payment certainty in discrete choice contingent valuation responses: results from a field validity test, *Southern Economic Journal* 69, 886–902.

Welsh, M.P. and Poe, G.L. (1998). Elicitation effects in contingent valuation: comparisons to a multiple bounded discrete choice approach, *Journal of Environmental Economics and Management* 36, 170–185.

Whittington, D., Smith, V.K., Okorafor, A., Okore, A., Long Liu, J. and McPhail, A. (1992). Giving respondents time to think in contingent valuation studies: a developing country application, *Journal of Environmental Economics and Management* 22, 205–225.

Zaller, J. and Feldman, S. (1992). A simple theory of the survey response: answering questions versus revealing preferences, *American Journal of Political Science* 36, 579–616.

Zimbardo, P.G. and Leippe, M.R. (1991).*The Psychology of Attitude Change and Social Influence.* McGraw–Hill Inc, New York.

Appendix

Unrestricted logit models

Variable	Definitely sure	90% certain	75% certain	50% certain	25% certain	10% certain
Intercept	0.831**	1.329**	1.406**	1.659**	1.789**	1.935**
	(0.256)	(0.244)	(0.228)	(0.218)	(0.206)	(0.206)
Bid amount	–0.175**	–0.153**	–0.111**	–0.073**	–0.036**	–0.021**
	(0.016)	(0.013)	(0.009)	(0.006)	(0.003)	(0.002)
MBDC inclusive format	0.081	0.094	0.036	–0.204	0.187	–0.071
	(0.219)	(0.207)	(0.196)	(0.191)	(0.194)	(0.194)
MBDC time-to-think format	–0.254	–0.244	–0.336†	–0.621**	–0.680**	0.843**
	(0.221)	(0.207)	(0.195)	(0.189)	(0.184)	–(0.185)
Gender (1 if male)	–0.485**	–0.374*	–0.337*	–0.273†	–0.337*	–0.278†
	(0.183)	(0.171)	(0.161)	(0.154)	(0.153)	(0.151)
Income	0.0005*	0.0004*	0.0005*	0.0004*	0.0003†	0.0004†
	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Log-likelihood function	–371.6217	–418.3056	–467.4314	–507.5095	–528.1344	–547.2296
n	111	111	111	111	111	111

Note: standard errors of estimates are presented in brackets; †, * and ** denote significance at the 0.1, 0.05, and 0.01 levels, respectively.