OOH la la: Testing the one-and-one-half bound dichotomous choice elicitation method for robustness to anomalies

Ian J. Bateman[a,*], Brett H. Day[a], Diane Dupont[b] , Stavros Georgiou[a] , N. Noceda-Matias and L. Subramanian

a. Programme in Environmental Decision Making (PEDM), Centre for Social and Economic Research on the Global Environment (CSERGE), School of Environmental Sciences, University of East Anglia, Norwich, NR4 7TJ, United Kingdom.

b. Department of Economics, Brock University St. Catharine's, Ontario, L2S 3A1, Canada.

* Contact details: Ian J. Bateman, e-mail i.bateman@uea.ac.uk, tel.: +44 (0)1603 593125, fax.: +44 (0)1603 507719.

Abstract

Although attractive in terms of its incentive compatibility, the standard single bound (SB) dichotomous choice technique for eliciting willingness to pay (WTP) responses in contingent valuation surveys has a major drawback in terms of its low statistical efficiency. While alternatives such as the double bound (DB) approach (which supplements an initial SB style question concerning a specified bid amount with a subsequent follow-up question concerning a different bid amount) offer improved statistical efficiency, they do so at the cost of compromised incentive compatibility and have also been shown to be vulnerable to a number of response anomalies. An innovative alternative, the one-and-one-half-bound (OOHB) dichotomous choice approach, has recently been proposed by Cooper, Hanemann and Signorello (2002). The OOHB differs from the DB in a number of important respects; the most important being that while each respondent is again exposed to two bid amounts, these are presented prior to any response as upper and lower limits on the cost of schemes. This preserves the incentive compatibility of responses concerning those two limits while generating most of the efficiency gains afforded by the DB method. However, Cooper, Hanemann and Signorello fail to test the method for robustness against response anomalies. Such a test is provided by the present paper. A number of theoretical consistency hypotheses are formulated by contrasting standard expectations with those derived from non-standard reference dependent utility theory. These are tested through the first application of the OOH method within its intended public goods context in a study concerning WTP for remediating impacts upon water quality associated with climate change. Data is collected through a face-to-face survey of over 1250 UK households. Results reject the theoretical consistency of elicited WTP responses showing that the OOHB is highly vulnerable to a number of anomalies. In particular acceptance rates for a given bid amount varied according to which other amount it was paired with and the order in which responses were elicited. We speculate upon the implications of these findings.

Keywords:
Contingent valuation, elicitation techniques, one-and-one-half bound, anomalies, willingness to pay, water quality.

JEL codes:
Q51; Q25; C25; D80

1. Introduction

More than fifty years ago, Ciriacy-Wantrup (1947, 1952) suggested that "appropriately constructed interviews" are capable of obtaining information about people's preferences for goods not ordinarily priced in the market. While it took some time for economists to embrace survey techniques (Davis, 1961; Randall et al., 1974), the contingent valuation (CV) method is now widely used to obtain willingness-to-pay (WTP) values for an assortment of environmental and other non-market goods (Carson, forthcoming). Indeed it is the valuation question itself which is the central feature of any CV study. However, given the hypothetical nature of the CV market, the method by which a valuation response is to be elicited is not immediately obvious. A wide variety of elicitation techniques have been proposed and tested ranging from early forays using simple open-ended questions (Brookshire et al., 1983) or bidding games (Davis, 1963; Randall et al., 1974) to payment ranges (Cameron and James, 1987), referenda (Carson et al., 1992) and more recently randomised card sorting (Carthy et al., 1999; Beattie et al., 1999). Comparisons across these elicitation techniques revealed significant differences between resultant benefit estimates (Bateman et al., 1995). However, rather than being random, these differences have been shown to be linked to a mixture of theoretical expectations (typically linked to differences in the incentive compatibility of differing approaches) and a range of theoretically unanticipated but empirically replicated anomalies (Carson, Groves and Machina, 2000; Bateman and Jones, 2003).

Given this diversity of techniques and consequent variation in estimated values, the identification of a clearly superior approach to the elicitation of WTP responses has been one of the most consistent themes in CV research and official guidance (Cummings et al., 1986; Mitchell and Carson, 1989; Arrow et al., 1993; Bateman et al., 2002). While different commentators have emphasised different attributes, in sum we can see three defining characteristics for an ideal elicitations method: (i) incentive compatibility, ruling out strategy space such that it is in the best interest of the respondent to answer truthfully; (ii) procedural invariance; in particular robustness to the range of preference anomalies identified in CV literature and; (iii) statistical efficiency, such that the technique can provide sufficient data in order to permit robust estimation of WTP without recourse to excessive sample size. Although all three criteria are important, it was the issue of incentive compatibility which dominated the landmark NOAA Blue-Ribbon Panel report on CV (Arrow et al., 1993). Here, building upon the work of Gibbard (1973) and Satterthwaite (1975) who establish the potential incentive compatibility of one-shot referenda[1], Arrow et al., recommend the use of the 'single bound' (SB) dichotomous choice elicitation technique[2] wherein each respondent is presented with a single question asking if they are willing to pay a specified sum, $X, for the good in question, to which they can only reply 'yes' or 'no'. By varying the amount X across the sample CV researchers can estimate decision compatible measures such as mean WTP.

The dichotomous nature of the SB approach is incentive compatible (Carson, Groves and Machina, 2000) and the technique is considered by many CV researchers to be robust against anomalies[3]. However, it is not statistically efficient, eliciting only whether a given respondent's WTP lies above or below the bid amount $X offered to them. In order to address this latter problem, Hanemann, Loomis, and Kanninen (1991) proposed a double-bounded (DB) format. Here, following an initial SB response, a second 'follow-up' or 'conditional' question is added to further probe the respondents' WTP. However, while this yields substantial gains in terms of statistical efficiency, subsequent empirical testing showed a number of response anomalies, in particular a lack of consistency between determinants of the first and second response and adverse reactions to the 'surprise' follow-up amount (Cameron and Quiggen, 1994; Herriges and Shogren, 1996; Alberini, Kanninen, and Carson, 1997; Bateman et al. 2001; DeShazo, 2002).

More fundamentally, theoretical analyses have identified that the DB approach fails incentive compatibility criteria (Mitchell and Carson, 1989; Carson, Groves and Machina, 2000). Specifically, the approach undermines the crucial face value interpretation of the bid amount as being the cost of providing the good in question. While this is credible in the initial question, it is no longer so in the follow-up. Clearly, only one of these amounts can be the real cost of the good. This loss of

[1] Within the CV context this argument is developed through Hoehn and Randall (1987) and Carson et al., (2000).
[2] First introduced by Bishop and Heberlein (1979)
[3] Note that the single response of the SB method is difficult to test for anomalies. Nevertheless, some commentators argue that it is not inherently robust to such problems (Green et al., 1998).

credibility compromises the incentive compatibility of the second question contributing to a characteristic reduction in acceptance rates relative to the first response.

In an innovative effort to retain much of the statistical efficiency gains of the DB approach, Cooper, Hanemann, and Signorello (2002; hereafter CHS) propose the one-and-one-half bound (OOHB) format. This makes a virtue of the difference between an initial and follow-up dichotomous choice amount by presenting both of these to the respondent as lower and upper (or upper and lower) bounds on the costs of providing the good in question. These bounds are presented prior to eliciting any valuation response thus ensuring incentive compatibility (providing these bounds are credible it remains in the best interests of the respondent to answer truthfully) and avoiding any elements of 'surprise' which may be engendered by the introduction of unexpected cost information. CHS show that, in comparison to the SB method, much of the statistical efficiency gains of the DB are retained by the OOHB approach. This combined with the incentive compatibility of the latter make it clearly superior to the DB and raise the promise of superiority over the standard SB method. However, CHS fail to test their method for procedural invariance against commonly observed preference anomalies. This paper undertakes such a test and in so doing provides the first application of the OOHB method within its intended public goods arena.

In the following section we provide a fuller account of the operation of the OOHB technique and provide a formal description of the regression models used to analyse our survey data. In Section 3 we derive a set of testable hypotheses concerning procedural invariance. In particular we draw upon the literature regarding prospect theory and reference dependent utilities (Kahneman and Tversky, 1979; Tversky and Kahneman, 1991; Bateman et al., 1997) to identify a set of anomalies relative to standard Hicksian theory which might arise from the process of providing survey respondents with prior information concerning the upper and lower bid amounts. Section 4 presents findings from our survey and reports results from our formal hypotheses tests. Finally, Section 5 discusses our findings and concludes.

2. The OOHB Elicitation Format

The OOHB approach, as proposed by CHS, is defined by two key features which distinguish it from either the SB or DB dichotomous choice techniques. The first of these features is that, prior to any valuation response being elicited, respondents are told that the price of the good in question is uncertain but lies between two stated values - a lower bound and an upper bound. A random process (unseen by the respondent) is used to determine whether the initial question posed will use the lower or upper of the pair of stated values. The chosen question is than presented to the respondent using the standard dichotomous format and a response recorded. The second distinguishing feature is that a follow-up question is only asked in two circumstances: when a "yes" response is given to a lower amount - which results in a second question using the upper amount; or when a "no" response is given to an upper amount - which results in a second question using the lower amount. When, in response to the initial question, an upper amount is accepted or a lower amount rejected then no follow-up question is asked. This process ensures that the respondent only faces amounts which are presented at the outset rather than being surprised by unexpected amounts which have dubious validity.

The CHS OOHB model begins by assuming that the an individual's WTP is a random variable with a cumulative distribution function (CDF) denoted $G(C_i;\theta)$. Define a range of bid values as [L,U] where \$L < \$U. An 'ascending' sequence presents the \$L bid first and presents the \$U bid only if a 'yes' response arises on the first question. We call this sequence LU. A 'descending' sequence presents the bid values in the opposite order and is denoted UL. There are three possible response probabilities for each of these two sequences, those for LU being detailed as Equation (1).

$$\begin{aligned} \pi_i^N &= \Pr\{C_i \le L\} = G(L;\theta) \\ \pi_i^{YN} &= \Pr\{L \le C_i \le U\} = G(U;\theta) - G(L;\theta) \\ \pi_i^{YY} &= \Pr\{C_i \ge U\} = 1 - G(U;\theta) \end{aligned} \tag{1}$$

The response probabilities for the UL sequence are given in Equation (2).

$$\begin{aligned} \pi_i^{Y} &= \Pr\{C_i \geq U\} &&= 1 - G(U;\theta) \\ \pi_i^{NY} &= \Pr\{L \leq C_i \leq U\} &&= G(U;\theta) - G(L;\theta) \\ \pi_i^{NN} &= \Pr\{C_i \leq L\} &&= G(L;\theta) \end{aligned} \tag{2}$$

The corresponding log-likelihood function is detailed in Equation (3).

$$\ln L^{OOHB}(\theta) = \sum_{i=1}^{N}\{d_i^{Y}\ln[1 - G(U;\theta) + d_i^{YN}\ln[G(U;\theta) - G(L;\theta) + d_i^{N}\ln[G(L;\theta)]\} \tag{3}$$

Given the symmetry of responses, it should be noted that $d_i^N = 1$ when either the starting bid is \$L and the response is (no) or the starting bid is \$U and the response is (no, no), and 0 otherwise. Similarly, $d_i^{YN} = 1$ when either the starting bid is \$L and the response is (yes, no) or the starting bid is \$U and the response is (no, yes), and 0 otherwise. Finally, $d_i^{YY} = 1$ when either the starting bid is \$L and the response is (yes, yes) or the staring bid is \$U and the response is (yes), and 0 otherwise. The log-likelihood function in Equation (3) is estimated using maximum likelihood techniques programmed in Gauss, described more fully in the results section of the paper.

3. Procedural invariance in the OOHB format: Formulating hypotheses

CHS show that the OOHB format exhibits efficiency gains over the SB format without indulging in the compromised incentive compatibility of the DB approach. However, they fail to test for robustness against a number of commonly observed response anomalies. Building upon the notions of prospect theory and loss aversion set out by Kahneman and Tversky (1979) and Tversky and Kahneman (1991), DeShazo (2002) offers a useful review of anomalies that affect the DB technique which in turn provides a basis for formulating the tests employed in our analysis of the OOHB approach. Most pertinently to the present analysis, when considering the DB format, DeShazo describes a model in which the respondent assumes that the good described in a valuation scenario will be provided with either certainty or with some subjective probability, say p. The respondent further assumes that she is being offered the good at the specified price. Prior to the asking of the first dichotomous choice question, the respondent's expected gain is the surplus value from the transaction times the probability that the good is provided. In Equation (4) WTP^T is the true underlying WTP, while EP is the expected price the respondent holds for the good.

$$EG = p[WTP^T - EP] \tag{4}$$

Now the respondent is asked the first dichotomous choice question. Suppose this involves a bid value of \$L. It is assumed that the respondent believes that she will receive the good if she says 'yes' to paying this amount. Under these circumstances she forms a new reference point equal to the updated expected consumer surplus shown in Equation (5).

$$EG = p[WTP^T - L] \tag{5}$$

The crucial point is that following a 'yes' response this new reference point becomes the relevant benchmark against which the gains or losses associated with any subsequent questions are assessed. On the other hand, suppose the respondent says 'no' to the first price offered. Here DeShazo argues that a new reference point is not formed since rejection of the good at the stated price reflects uncertainty on the part of the respondent both with respect to provision of the good and the price of such provision.

The existence of a reference point is relevant for the second question of a DB format but has different implications according to whether the questions are asked in an ascending sequence (the lower bid presented first then, iff a 'yes' responses is given the second, higher bid question is asked) or in a descending sequence (the higher bid presented first then, iff a 'no' responses is given the second, lower bid question is asked). In an ascending sequence, since the second bid is higher than the first (\$L<\$U), the person frames the follow-up bid as an expected loss (EL) defined by Equation (6):

$$EL \quad = \quad p[L \; - U] \tag{6}$$

As a result of this updating of the reference point, the overall expected utility change for the respondent in such an ascending sequence is given in Equation (7).

$$\Delta EU \;=\; U\,\{p\,[(WTP^{T} \quad - \quad L) + (L \quad - \quad U)\,] \tag{7}$$

DeShazo (2002) argues that respondents view such a follow-up question as negatively framed. Furthermore, there is a resulting piecewise construction of expected utility, which leads such respondents to believe that the expected utility in Equation (7) is less than the expected utility of an "unframed prospect" arising from a situation in which a respondent is asked to respond to \$U in the first instance. The expected consumer surplus from purchasing the good at the second price without prior information is given in Equation (8).

$$EG = \; p\,[WTP^{T} \quad - \quad U] \tag{8}$$

In contrast to the ascending sequence, DeShazo asserts that the descending sequence should not result in the respondent viewing a second, lower bid as being a positively framed question. This follows from the assumption that, after having said 'no' to the first (\$U) bid, the respondent does not form a new reference point.

The application of this loss aversion framework to explain response anomalies in a DB format leads to testable hypotheses regarding the relative proportions of respondents accepting a given bid level in a follow-up question dependent upon the relation of that bid amount to that posited in the initial WTP question. Specifically, it predicts that the probability of a respondent answering 'yes' to a follow-up question from an ascending sequence (negatively framed) is less than the probability of a respondent answering 'yes' to the same value presented in an initial valuation question. A further prediction is the presence of a downward bias in WTP in ascending sequences and an overall downward bias in sample mean WTP.

The reference dependent utility approach provides a useful framework for the analysis of responses to questions posed in the OOHB format. Recall that the OOHB approach tells respondents upfront that the costs to households could range between \$L and \$U. Suppose a respondent is presented with this range of bids and she believes that the lower bound value of these bids (\$L) will occur with probability q and thus, the upper bound value (\$U) is believed to occur with probability (1-q). We will assume that these limits are set so that $q \neq 0$ implying that each limit is to at least some degree credible. In this case, the expected gain *prior* to the first question being asked is the probability of the good being provided times a weighted average of the two possible consumer surpluses associated with the upper and lower bound bid values, as in Equation (9).

$$EG^{OOHB} \quad = \quad p\,\{E\,(\,q\,[WTP^{T} - \quad L] + (1-q)[WTP^{T} \; - U\,])\} \tag{9}$$

If the respondent believes at the outset that the two bid values are equally likely, then the EG^{OOHB} will be a simple average of the two possible consumer surpluses, as in Equation (10)[4].

$$EG^{OOHB} \quad = \quad p\,\{E\,(\,q\,[WTP^{T} - \quad \frac{(L-U)}{2}])\} \tag{10}$$

With the OOHB format the respondent then replaces the initial reference point of Equation (4) with that in Equation (9) *prior* to the first question being asked. This means that, unlike in a DB format, both questions in a sequence will be affected by framing. In order to investigate the implications of

[4] An examination of equation (7) reveals that people who believe the lower bound value has a higher weight (i.e. probability) will tend to have a larger EG or initial reference point since the weighted average of the two bids will be lower, the more weight is put on the lower bid. This means that they will feel the loss associated with receiving the upper bid in the second round even more keenly.

this situation we distinguish between two the different sequences of questioning arising under the OOHB format: that of the ascending (lower bound bid is asked first, followed by upper bound, only if appropriate) and descending sequence (upper bound bid is asked first, followed by lower bound, only if appropriate).

We begin with the first question in the *ascending sequence*. First, prior to any response being elicited, the respondent is informed of the lower and upper limits of the cost estimate. Recall that this generates the reference point EG^{OOHB} as in Equation (9). Now, the respondent faces the bid level of the first question in which she is told that she need pay only $L. This represents a positively framed outcome relative to the reference point.

Turning to the second question in the ascending sequence, we direct our attention to the subset of people who say 'yes' to the $L question[5]. After giving this positive response, the respondent updates her reference point to the gain arising from the difference between her true willingness to pay and the lower bound bid to which she has just said yes, as shown in Equation (11)[6].

$$EG \quad = \quad p[WTP^T - L] \tag{11}$$

In the ascending sequence of the OOHB format, the respondent is now asked to state whether they will or will not pay the upper amount $U. When compared to the updated reference point, this is a negatively framed question and seen as a loss.

We turn next to the *descending sequence*. Again, prior to any response being elicited, the respondent is informed of the lower and upper limits of the cost estimate. This will again result in respondents forming a first reference point based on some combination of these values. Insofar as respondents have the same utility function and homogeneous tastes, this reference point should be identical to the one formed by respondents who answered questions in an ascending sequence. It should be solely based on the values themselves and on the range between these values since, at the outset, the respondent does not know the order in which questions will be asked.

With the first question in the descending sequence the respondent is asked to state whether they will or will not pay the higher of the two values ($U). Even, though this is a first question, it is negatively framed as it represents a loss relative to the expected gain from the reference point based on some combination of the upper and lower bound values. Only those who refuse to pay this initial amount are asked the follow-up question concerning the lower amount ($L). Extending DeShazo's argument to the OOHB format, the expectation would be that, for such respondents, the reference level of utility remains linked to the initially perceived balance of the $U and $L amounts. Here then the lower amount ($L) encountered in the follow-up question constitutes a gain over this reference level. However, this gain will become larger if, in fact, the respondent alters their reference level to one based solely upon the larger $U amount encountered in the first question.

A final speculation concerns the influence of either the relative or absolute distance between the $U and $L amounts upon responses to either of these bounds. We postulate that increases in one or either of these measures may impinge directly upon the size of any reference framing effects. So we might expect that a given effect might be more pronounced where the difference between the $U and $L is 'large'. Whether the definition of a large effect should be in terms of relative or absolute amounts (and whether these interact or exhibit eventually diminishing sensitivity as might be implied from Tversky and Kahneman, 1992) is an open empirical question which we investigate subsequently.

Given this theoretical framework of positively and negatively framed questions we can formulate a series of testable hypotheses. We define Prob ($X=Yes | L-U) as the probability of a yes response to particular bid value, $X, when it is asked first in an ascending sequence, i.e. when $X is the lower bid value $L and is therefore presented within a positively framed question. Similarly, Prob

[5] Recall, those individuals who said 'no' to $L are not asked the second question, however, they would still be affected by the bid range in answering their first question.

[6] Note that this is equivalent to DeShazo's reference point after a first lower bound bid is asked in a DB format.

(\$X=yes | \$Z=yes | L-U) is the probability of a yes response to the same bid value (\$X), when it is asked second in an ascending sequence. In order for this question to be asked, it must be the case that the respondent has already said yes to a lower bound bid of \$Z. This means that \$X represents the upper bound value (\$U) for this particular sequence of questions and is therefore presented within a negatively framed question. Similarly, we can think about a particular \$X bid value as being asked as either the first question in a descending sequence (and so representing the upper bound value and subject to negative framing) or as the second question in a descending sequence (thereby representing the lower bound value and subject to positive framing). The associated probabilities of yes responses for these two cases are: Prob (\$X=yes | U-L) and Prob (\$X | \$Z=no | U-L). It should be remembered in this latter case that \$Z represents the upper bound value asked first which has been rejected by the respondent. Given this nomenclature we can now define four distinct hypotheses with which to test for preference anomalies and thereby examine procedural invariance within the OOHB format. In each case the null hypothesis H_o is defined in accordance with standard Hicksian theory which does not allow for reference dependent effects and therefore implies that the probability of accepting some given bid level \$X should be invariate to the frame within which it is presented and therefore should not vary according to whether that amount is presented within either and ascending or descending sequence or as the initial or follow-up question.

H_o^1: Prob (\$X=Yes | L-U) = Prob (\$X=yes | \$Z=yes | L-U). Here the null hypothesis of equality is contrasted with an alternative hypothesis derived from reference dependent utility theory which states that, for responses to ascending sequence questions the probability of agreeing to pay a given amount \$X will be higher when it is presented in the first (lower amount; positively framed) question than in the second (higher amount; negatively framed) question. Therefore we have H_a^1: Prob (\$X=Yes | L-U) > Prob (\$X=yes | \$Z=yes | L-U)

H_o^2: Prob (\$X | \$Z=no | U-L) = Prob (\$X=yes | U-L). Here the null hypothesis of equality is again contrasted with an alternative hypothesis derived from reference dependent utility theory which states that, for responses to descending sequence questions the probability of agreeing to pay a given amount \$X will be higher when it is presented in the second (lower amount; positively framed) question than in the first (higher amount; negatively framed) question. Therefore we have H_a^2: Prob (\$X | \$Z=no | U-L) > Prob (\$X=yes | U-L).

H_o^3: Prob (\$X | \$Z=no | U-L) = Prob (\$X=Yes | L-U). Here the null hypothesis of equality is contrasted with an alternative hypothesis which tests whether respondents update their frames between the first response and second response. Specifically we have H_a^3: Prob (\$X | \$Z=no | U-L) ≠ Prob (\$X=Yes | L-U) which argues that, for a given dollar value, \$X, we expect a differing probability of yes responses to the positive framed follow-up question within a descending sequence than to a similar positively framed response to an initial question in an ascending sequence.

H_o^4: Prob (\$X=yes | \$Z=yes | (L-U)*) = Prob (\$X=yes | \$Z=yes | (L-U)**) where (L-U)* < (L-U)**. Here the null hypothesis of equality is contrasted with an alternative hypothesis which tests whether the difference between \$U and \$L, whether measured in either absolute or relative terms, impacts upon response rates for any given amount \$X. Note that H_o^4 is written in terms of a response to a follow-up question from an ascending sequence. However, this test is also applicable to first response and descending sequence scenarios and is tested across all feasible combinations for both absolute and relative differences.

4. Data and Estimation Results

4.1. The survey instrument design and implementation process

The data used to examine the OOHB were collected using a face-to-face interview survey designed, in all other respects, in accordance best practice guidelines (Mitchell and Carson, 1989; Arrow et al., 1993; Bateman et al., 2002). Extensive use was made of focus groups to refine the description of the public good in question and formulate an appropriate contingent market which was conveyed using a combination of clear and concise text augmented by visual aids. The resultant survey instrument was tested through a pilot survey of some 100 households after which the final survey questionnaire was refined. Both focus group and pilot exercises were also used to define an appropriate vector of bid amounts across which a range of positive and negative responses might be expected. The final survey instrument was applied to a sample of randomly selected households in and around the city of Norwich, England. Surveying was conducted during a five week period in the summer of 2003, all interviews being undertaken in a face-to-face format by a team of trained interviewers.

The public good chosen to be valued was a typical focus for CV research, namely the remediation of eutrophication problems affecting nearby rivers and lakes. Survey respondents were presented with a proposal to address this problem by installing new technology at sewage works so as to remove phosphates from household sewage. Survey respondents were told that such treatment would increase their annual household water bill. This payment vehicle is attractive from a CV perspective as it is effectively universal and unavoidable thereby avoiding the strategic behaviour associated with discretionary payment vehicles (Bateman et al., 2002).

In accordance with the principles of the OOHB approach, survey respondents were informed in advance that the cost to their household of the phosphate removal scheme was between a specified lower and upper bound ($L and $U respectively). An unseen random process was used to allocate respondents to one of thirteen pairs of amounts. Of these seven described ascending sequences as follows: £10-£50; £25-£100; £50-£100; £75-£100; £100-£150; £100-£200; £48.50-£98.50. The remaining six pairs describe descending sequences as follows: £50-£10; £100-£25; £100-£50; £100-£75; £150-£100; and £200-£100. These pairs were chosen upon two criteria: first that they all fell within the distribution of bids implied by our focus group and pilot survey investigations; second, that they permitted ready and unambiguous testing of our various hypotheses. In particular the repetition of certain bid amounts, such as the £100 bid, across a variety of framing contexts permits simple non-parametric testing of hypotheses (although in the present paper we focus upon more embracing regression based tests to allow for the impact of other covariates upon acceptance rates). Note that the £48.50-£98.50 pair was only used in the ascending sequence to provide a side analysis comparison with the £50-£100 pair examining whether the implied greater accuracy of the former pair resulted in any significant impact upon acceptance rates. Although households were randomly allocated to bid pairs care was taken to ensure that roughly similar numbers were allocated to each pairing.

4.2 Results I: Acceptance rates across treatments.

In total 1254 households provided completed questionnaires[7]. Table 1 shows the resulting acceptance rates for each bid level from each of the thirteen bid pairings describing our various ascending and descending sequences. Comparison of identical bid amounts presented in either sequence as either the initial or follow-up amount reveals substantial differences across these various treatments. These differences provide our first commentary upon the various hypothesis tests described previously.

[7] In total a further 1067 households were approached but declined to take part in the survey giving a response rate of 54%. The most common reasons for refusing to take part were time constraints and a lack of interest in any survey (respondents were unaware of the subject matter of the study at the outset of the survey). Arguably this may mean that estimated values might overstate true WTP across the population. However, this was not the principle focus of the present research.

Table 1: Comparison of bid acceptance rates across treatments[1]
Ascending Sequence (LU)

Initial (lower) bid amount ($L)	Acceptance rate for $L (%)		Follow-up (Upper) bid amount ($U)	Acceptance rate for $U[2]		Absolute difference
10	90.1%		50	46.5%		40
50	51.6%		100	22.0%		50
100	43.5%		150	27.2%		50
100	47.4%		200	9.3%		100
25	82.1%		100	22.1%		75
75	42.4%		100	33.3%		25
48.50	52.6%		98.50	18.6%		50

Descending Sequence (UL)

Initial (upper) bid amount ($U)	Acceptance rate for $U (%)		Follow-up (Lower) bid amount ($L)	Acceptance rate for $L[2]		Absolute difference
50	60.0%		10	91.0%		40
100	37.4%		50	66.0%		50
150	41.6%		100	47.5%		50
200	28.7%		100	62.8%		100
100	30.5%		25	85.2%		75
100	30.0%		75	42.2%		25

1. Total sample size = 1254 households. Sample sizes within each treatment vary from a minimum of 90 to a maximum of 106 households.
2. Acceptance (rejection) rates for follow-up questions include as 'yes' ('no') responses those respondents who were not asked the second question because they had implicitly accepted (rejected) this amount as part of their response to the initial question.

A cursory inspection of the data presented in Table 1 suggests some substantial treatment effects arising from the various negative and positive frames provided by the OOHB process. Formal testing of our various hypotheses is given in the following section however an overview commentary is presented here.

With respect to the ascending sequences H_o^1 we note that acceptance rates for identical bid amounts are substantially lower when that amount is presented in the first question as the lower of the two amounts shown ($L) than when it is presented in the follow-up question as the higher of the two amounts ($U). For example, acceptance rates for the £100 bid amount vary between 43.5% and 47.4% when it is presented as the lower first amount but only between 22.0% and 33.3% when the same amount is presented in the follow-up up question as the higher of the two amounts. A similar although less extreme effect is noted with respect to he £50 bid level used in both the initial and follow-up questions. Taken overall, this suggests the rejection of the standard expectation given in H_o^1 in favour of the reference utility predictions underpinning H_a^1

Considering the descending sequence tested under H_o^2 we find acceptance rates for the £100 amount rising as high as 62.8% when this is presented in a follow-up question as the lower of two amounts, yet acceptance rates for the same bid amount fall as low as 30.0% when it is presented in the initial question as the higher of the two amounts seen by a given respondent. This again suggests a rejection of the standard expectation given by H_o^2 in favour of the reference utility based alternative hypothesis H_a^2. Again a similar but less extreme pattern is observed for the £50 bid level, a point to which we return subsequently.

Null hypothesis H_o^3 considers consistency across sequences. Here an interesting contrast can be seen between acceptance rates for initial questions, which are relatively similar across sequences, and those for follow-up questions which diverge remarkably across sequences. So, for example, considering responses to the £100 amount when presented in the initial question we see consistently higher acceptance rates for the ascending sequence (where this is presented in a positive frame as the lower of the two amounts resulting in rates between 43.5% and 47.4%) than in the descending sequence (where this is presented in a negative frame as the higher of the two amounts resulting in rates between 30.0% and 37.4%). However, these differences are dwarfed by those recorded for follow-up questions. Here, acceptance rates for the negatively framed ascending sequence range from 33.3% down to a low of just 22.0%. In contrast acceptance rates for the positively framed descending sequence range from a low of 47.8% to a high of 66.0%. Therefore acceptance rates can triple across ordering sequences suggesting that the OOHB approach clearly fails procedural invariance tests being highly susceptible to preference anomalies.

Hypothesis H_o^4 concerns the possible impact upon acceptance rates of either the absolute or relative difference between those amounts. Returning to the findings regarding H_o^1 and H_o^2 above recall that the treatment effects observed with respect to the £100 bid amounts are replicated but to a lesser extent within acceptance rates for the £50 bid level. However, whether this is due to the absolute size of these bid amounts or the absolute or relative difference between paired amounts is unclear. This uncertainty suggests that further investigation of H_o^4 is justified within our subsequent regression analysis.

Finally our comparison of the bid pairs (£48.50, £98.50) and (£50, £100) reveals no clear evidence that the former ‘more accurate’ pair resulted in any substantial impact upon acceptance rates.

4.3 Results II: Regression testing of hypotheses.

In order to examine these relationships further, we estimate a series of bid functions that allow us to permit the mean of the distribution to be shifted by the socioeconomic characteristics of respondents. We include variables for respondents’ sex, age, income and their annual frequency of visits to lakes and rivers in East Anglia. Our objective is to test the extent to which the distribution of WTP, once we have controlled for socioeconomic heterogeneity, might differ between four subsamples of valuation questions. These are as follows:

- Low amount asked as the First question in an ascending sequence (LOW1)
- Low amount asked as the Second question in a descending sequence (LOW2)
- High amount asked as the First question in a descending sequence (HIGH1)
- High amount asked as the Second question in an ascending sequence (HIGH2)

We specify the model to allow both the mean and variance of the WTP distributions to differ across the four treatments. The model must account for the fact that many, but not all, individuals provide two responses, one to a high amount and one to a low amount. Their responses were initially assumed to be drawn from a bivariate normal or lognormal distribution in which the correlation between responses is captured by the estimated parameter ρ. Respondents answering just the one valuation question are assumed to draw responses from a univariate normal or lognormal distribution and their answers do not contribute to the calculation of the correlation coefficient ρ. Initial analysis found that the underlying relationships were consistent across the normal and lognormal assumptions and so only the latter are reported here as this seems a theoretically more plausible assumption.

A further objective of the modelling exercise is to assess the extent to which the response to the second of the two valuation questions may be influenced by the difference between the two bid amounts. One testable hypothesis is that respondents will increasingly reject a high bid amount in the second question the greater the difference between this bid amount and the low amount offered in the first question (loss aversion arising from a negatively framed second question). Likewise, we might expect greater acceptance of a low amount in the second question the greater the decrease in this amount when compared to the high bid amount in the first question. We test two specifications, one that uses the absolute difference in bid amounts and another which uses the relative difference in bid amounts.[8] Such a finding would describe, for any given bid amount, a 'fanning out' of acceptance rates as respondents move through the various permutations of he OOHB bidding tree.

Tables 2 to 4 report a series of models in which different constraints are imposed on the WTP distributions of the four treatments; LOW1, LOW2, HIGH1, HIGH2. A number of preliminary comments will facilitate interpretation of the modelling results. First, each WTP distribution is determined by two parameters, a *location* parameter and a *scale* parameter. Roughly put, the location parameter determines where the WTP distribution is centred; higher values of the location parameter shift the distribution to higher WTP values, lower values shift the distribution to lower WTP values. The scale parameter determines the spread of the WTP distribution. Again, higher values of the scale parameter indicate the WTP distribution is more widely dispersed, lower values that it is more tightly concentrated. In terms of the parameters of the model, the location parameter is determined by the value of the CONSTANT. A constant can be entered for each treatment or constants for different treatments can be constrained to have equal values. To account for the effect of COVARIATES, we include variables reflecting socioeconomic characteristics of households. These act to shift around the location parameter. Again we can allow the effect of covariates to be different in the different treatments, or impose equality constraints across treatments. The scale parameter of the WTP distributions is parameterised as the coefficient on the BID variable. To be more exact, the negative of the coefficient on the BID variable. Once again, we can include separate BID variables for each treatment to determine the independent scale of the four WTP distributions or we can impose equality constraints between the different treatments. Finally, the degree of correlation that exists between the two responses from the same individuals is captured by the parameter RHO. If individuals are responding consistently to the two questions then we would expect the correlation in their answers to be nigh on perfect and rho would take a value of one. In this work we constrain rho to be identical across all pairs of treatments[9].

Table 2 describes results from three regression models, all of which impose certain constraints upon our coefficient estimates. Model 1 assumes that WTP follows a lognormal distribution and that all four WTP distributions (LOW1, LOW2, HIGH1 and HIGH2) are identical. The location parameter of this distribution, as determined by the CONSTANT, is positive and highly significant. Moreover, WTP is seen to increase significantly with income and the number of visits to lakes and rivers, decline significantly with age, but is not significantly determined by the sex of the respondent. The scale of the WTP distribution, as determined by the LnBID coefficient, is highly significant as expected. The correlation coefficient, RHO, is positive and significantly different from zero, implying that there is positive correlation between individuals' responses. More intriguingly, the correlation is far from perfect; RHO is highly significantly different from a value of 1 (t-stat 8.38).

[8] The code used to model this data was written in Gauss by the authors and is similar to that developed by Joe Cooper for the CSH (2002) paper. The differences between the two codes are the following. First, our code constrains the correlation coefficient to lie between –1 and 1. Second, our code replaces numerical with analytical calculation of the Hessian of the log-likelihood, providing improved speed of convergence and greater accuracy in the calculation of the variance-covariance matrix of the parameters. Finally and most importantly, our code allows a variety of parameter constraints that are required for the testing strategy we have adopted in this paper.

[9] Future work will investigate if the degree of correlation in responses is different in a LOW to HIGH pair of answers when compared to a HIGH to LOW pair, although it is considered unlikely that this will have much effect.

Table 2: Comparison of coefficients from constrained models

Variable	MODEL 1 (WTP distributions constrained to be identical: LOW1 = LOW2 = HIGH1 = HIGH2 for location and scale parameters)	MODEL 2 (WTP distributions allowed to differ for first and second bids only; (LOW1 = HIGH1) ≠ (LOW2 = HIGH2) for location and scale parameters)	MODEL 3 (WTP distributions allowed to differ for high and low bids; first and second bid distributions assumed to be the same; (LOW1 = LOW2) ≠ (HIGH1 = HIGH2) for location and scale parameters)
Ln BID	-0.7783 (0.0584) *	N/a	N/a
INCOME	0.0133 (0.0024) *	N/a	N/a
AGE	-0.0040 (0.0019) **	N/a	N/a
FEMALE	0.0652 (0.0680)	N/a	N/a
VISITOR	0.0009 (0.0004) **	N/a	N/a
CONSTANT	3.0917 (0.2777) *	N/a	N/a
Ln BID1 (first bid only)	N/a	-0.7433 (0.0623)*	N/a
Ln BIDLOW (low bids only)	N/a	N/a	-0.8058 (0.0556) *
INCOME	N/a	0.0137 (0.0029) *	0.0170 (0.0032) *
AGE	N/a	-0.0027 (0.0023)	-0.0027 (0.0025)
FEMALE	N/a	-0.0239 (0.0805)	-0.0763 (0.0895)
VISITOR	N/a	0.0007 (0.0005)	0.0004 (0.0005)
CONSTANT	N/a	2.9572 (0.3045) *	3.1773 (0.3115) *
Ln BID2 (second bid only)	N/a	-0.8883 (0.0951) *	N/a
Ln BIDHIGH (high bids only)	N/a	N/a	-0.5310 (0.1096) *
INCOME	N/a	0.0135 (0.0035) *	0.0103 (0.0030) *
AGE	N/a	-0.0052 (0.0029) **	-0.0052 (0.0025) **
FEMALE	N/a	0.2007 (0.1005) **	0.1998 (0.0884) **
VISITOR	N/a	0.0011 (0.0005) ***	0.0013 (0.0005) **
CONSTANT	N/a	3.5358 (0.4387) *	2.4243 (0.5119) *
RHO	0.4263 (0.1134) *	0.5025 (0.1155) *	0.4988 (0.1653) *
LLF	-1092.157	-1086.909	-1085.311

Figures in parentheses are standard errors. Confidence levels are: * 99 % ; ** 95 % ; *** 90 %

For Model 2 we allow the parameters of the WTP distributions for the first and second questions to be independently determined. A quick glance down the coefficient values reveals that they remain remarkably similar across the two distributions. The only difference of note is in the coefficient on FEMALE. In response to the first question males and females give the same answers. However, in response to the second question females are more likely to answer yes to any particular bid level than are males. However, overall Model 2 provides no significant improvement over the fully constrained model, Model 1. We cannot reject the hypothesis of equality of parameters between the first and second WTP distributions at the 95% or even 90% level of confidence (LR stat = 10.49, df = 6, p-value = .105).

For Model 3, we allow the parameters of the WTP distributions for the low and high questions to be independently determined. As Table 2 shows, this provides further evidence that the two distributions differ. The likelihood ratio test comparing model 1 with model 3 (LR stat = 13.69, df = 6, p-value = .03) rejects the hypothesis that the parameters of the two distributions are identical. From the coefficient on FEMALE in the second equation (high bid values only), it appears that women are more likely to say yes to higher bid values, regardless of whether these are asked in an ascending or descending sequence. Notice that RHO continues to take a value significantly different from one, indicating that given the constraints of the model, individual's responses fall far short of being perfectly correlated.

In both Models 2 and 3, it is also interesting to note that while the age of the respondent appears to be unimportant for answering the first question (or for low bid values), it is negative and significantly different for either the second question or for high bid values.

Table 3: Unconstrained model coefficients

Variable	MODEL 4 (WTP distributions unconstrained; allowed to differ for first and second bids and for high and low bids; LOW1 ≠ LOW2 ≠ HIGH1 ≠ HIGH2 for location and scale parameters)
Equation 1: Responses to First Question	
Respondents receiving low bids	
Ln BID	-0.7368 (0.0721) *
INCOME	0.0205 (0.0037) *
AGE	0.0008 (0.0029)
FEMALE	-0.1205 (0.1089)
VISITOR	0.0006 (0.0006)
CONSTANT	2.6446 (0.3247) *
Respondents receiving high bids	
Ln BID	0.1651 (0.0677)**
INCOME	-0.0130 (0.0055) **
AGE	-0.0052 (0.0043)
FEMALE	-.2635 (0.1598) ***
VISITOR	0.0000 (0.0010)
ABSDIFF	-0.0051 (0.0023) **
Equation 2: Responses to Second Question	
Respondents receiving low bids	
Ln BID	-0.7023 (0.0796) *
INCOME	0.0154 (0.0037) *
AGE	-0.0045 (0.0031)
FEMALE	0.1084 (0.1142)
VISITOR	0.0021 (0.0006) *
CONSTANT	3.2306 (0.3592) *
Respondents receiving high bids	
Ln BID	-0.0850 (0.0707)
INCOME	-0.0033 (0.0057)
AGE	-0.0026 (0.0046)
FEMALE	0.0488 (0.1691)
VISITOR	-0.0019 (0.0010) ***
ABSDIFF	-0.0850 (0.0707)
RHO	1.0000 (0.0000) *
LLF	-1057.234

Figures in parentheses are standard errors. Confidence levels are: * 99 % ; ** 95 % ; *** 90 %

In Model 4 (detailed in Table 3) we report a completely unconstrained model in which we allow each of the four treatments to have a different WTP distribution. Here, we have two equations corresponding to the first and second responses. Within these we take as our base case responses to the lower of the two amounts seen by each respondent ($L) and then calculate departures for responses to the higher of these two amounts ($U). This allows the locational parameters to vary. We also allow the scales of the distributions to vary by defining as separate variables the lnBID values faced by each treatment. The coefficients on the first and second Low bid WTP distributions are given by the LnBID values. The scale of the HIGH1 WTP distribution is found by adding the two LnBID coefficients under Equation 1, whilst the scale of the HIGH2 WTP distribution is found adding the two LnBID coefficients under Equation 2. Similarly, the coefficient on the INCOME variable in Equation 1 for the low bid shows the influence of this variable on the low bid WTP distribution when these low bid questions are answered first. To find the influence of income on the high bid WTP distribution we add this coefficient to the high bid income coefficient.

We experimented with a number of ways of specifying differences in the constants for the four distributions. The specification shown in Table 3 includes the absolute difference between the high and low bids encountered by a respondent as the variable ABSDIFF. This specification enforces the logical restriction that, if there is no difference between the two bids, then there would be no difference between the constants of the distributions. This specification results in the best fit for the data according to the log likelihood function values.[10]

Let us now examine the coefficients reported for this unconstrained model. First, observe the value of rho. Now that we have removed all constraints from the model, the correlation between an individual's responses is nigh on perfect at a value of 1. We can conclude that differences in the way individuals respond to the two questions have been captured by the specification of the model.

Now, observe the scale coefficients for the four different WTP distributions. Since the coefficient on LnBID for the high bid first responses is significantly different from zero (p-value .015), there is some evidence to suggest that there is somewhat less dispersion in WTP when respondents answer a high question first. However, closer examination of the scale coefficients for all four distributions shows that they are actually remarkably similar, a point to which we will return shortly in the next section.

Consider next the variables parameterising the constant of the location parameter of the WTP distributions. The coefficients on the CONSTANT variables in Equations 1 and 2 indicate the constants for the LOW1 and LOW2 WTP distributions. The coefficients on the ABSDIFF variables (indicating the absolute difference between high and low bids) capture the degree to which each £ difference between the LOW and HIGH bids shifts the HIGH1 and HIGH2 WTP distributions away from their respective LOW WTP distributions. Notice first that the CONSTANT in Equation 2 is somewhat larger than that in Equation 1. This suggests that WTP when estimated from the low bid asked as the first question tends to be smaller than WTP when estimated from the low bid asked as the second question. Moreover, observe that both of the coefficients on the ABSDIFF variables are negative and significantly different from zero. It seems that WTP when calculated from responses to High bids is significantly smaller than that calculated from Low bids and that this difference increases with the absolute difference between the low and high bids. This is a finding replicated in the iterative bid literature (DeShazo, 2002).

Finally consider the significance of the High-bid specific covariate coefficients. Only the coefficient on income in the first equation seems to differ between high and low distributions. There does not appear to be a great deal of evidence to support the contention that covariates influence responses to high bids differently than they do low bids. However, we return to this issue when we test the hypothesis that the covariates have the same influence on all the distributions in the next model.

[10] Two other approaches were tried. The first was to include a constant in both Equations specific to the HIGH responses so that the coefficients on these HIGH specific constants indicate the degree to which the HIGH distributions are shifted away from the LOW distributions. The limitation of this approach is that it assumes that the degree of shift in the distributions is constant. A second approach that allows the degree of shift to be determined by the difference in the bid amounts faced by any one individual was also tried. This was done by specifying a HIGH variable that measure the ratio of the high to the low bid in the first and second equations Comparing the maximised values of the log likelihood from these three specifications reveals that the specification we present in the Table best fits the data. The log likelihood for the constants model is -1065.55, for the ratios of bids model is -1061.45 and for the absolute differences in bids model is -1057.23.

A number of further speculations were explored before we refined our final model. First we test the hypothesis that the scales of the four different distributions are equal by constraining that the coefficient on lnBID is constant for all four treatment distributions. Carrying out a likelihood ratio test between the resultant model and that detailed in Table 3 model 5 (with unconstrained scales) indicated that we cannot reject the hypothesis of equal scales (LR stat = 5.82, df = 3, p-value = 0.12). A further model tested the hypothesis that the influence of the covariates on the locations of the four different distributions are equal. Again we cannot reject this hypothesis (LR stat = 17.20, df = 12, p-value = 0.14) and conclude that the covariates act upon the locations of the distributions in much the same way. Further tests showed that we cannot reject equality of the constants in the low and high WTP distributions arising solely from the first question responses and that the variable measuring the absolute difference between the high and low bid values is insignificant in explaining responses to the first question.

Given the above findings we now know that we can impose the constraints of equal scales (imposed through a single LnBID variable), equal covariate effects upon the location of the distributions and equality of the constants in the low and high WTP distributions arising solely from the first question responses and that the variable measuring the absolute difference between the high and low bid values is insignificant in explaining responses to the first question.. Our final model, reported in Table 4 imposes all of these constraints and findings and tests a final constraint of equality of the constant for the low and high WTP distributions in response to the second question. In contrast to our finding that the absolute difference variable is insignificant in equation one, the extreme significance of the coefficient in the second equation indicates that there is no such simplification possible with the high and low WTP distributions in response to the second question.

Table 4: Equal scale and covariates effects for all distributions; low and high intercepts for first equation constrained; intercepts for second question unconstrained

Variable	MODEL 5 (Covariate effects constrained to be equal across WTP distributions unconstrained; allowed to differ for first and second bids and for high and low bids; for locational parameters constrained to be equal for first but not second response ((LOW1 = HIGH1) ≠ LOW2 ≠ HIGH2); scale parameters fully constrained (LOW1 = LOW2 = HIGH1 = HIGH2))
Equation 1 (First question)	
CONSTANT (same for both low and high bids in first equation)	2.8975 (0.2509) *
Equation 2 (Second Question)	
ABSDIFF	-0.0109 (0.0015) *
CONSTANT	3.1106 (0.2371) *
Common coefficients constrained to be same for both equations	
Ln BID	-0.7312 (0.0504)*
INCOME	0.0142 (0.0025)*
AGE	-0.0034 (0.0020)***
FEMALE	0.0470 (0.0711)
VISITOR	0.0009 (0.0004) **
RHO	0.9991 (27.9580)
LLF	-1068.895

Figures in parentheses are standard errors. Confidence levels are: * 99 % ; ** 95 % ; *** 90 %

At first glance it may appear that the constants given for Equation 1 and 2 in Table 4 are not significantly different. However, the distribution of second question responses is fixed both by the locational parameter (the CONSTANT in Equation 2; which locates responses for the 'low' bids) and the ABSDIFF variable. This latter variable adjusts the location for responses to high bids; the degree of adjustment being given by multiplying its coefficient (-0.0109) by the absolute difference, in pounds, between the low and high bids. Note that, as the locational CONSTANT for the second responses (3.1106) is larger than that for the first (2.8975) this implies that, when ABSDIFF is very small, then for a given bid amount acceptance rates are higher in the second response than in the first. But as ABSIFF increases so its negative coefficient shifts the implied WTP distribution downwards, past that of the first response to acceptance levels which are substantially lower than those given for the same bid amount encountered in the first bound.

The significance of this difference between first and second responses was tested via a hypothesised further constraint of equality of constants between first and second responses. This was rejected due to a significant reduction in loglikelihood for such a model (-1077.24) and a likelihood ratio test reveals that we cannot accept the hypothesis of equality of parameters between the LOW2 and LOW1 and HIGH1 distributions (LR stat = 16.98, df = 1, p-value <.001). In contrast, the extreme significance of the ABSDIFF coefficient in Table 4 indicates that there is no such simplification possible with the high and low WTP distributions in response to the second question. We conclude that responses to the second question of a pair of questions under OOHB elicitation differ significantly according to whether this is the low or high bid in the pair.

The results given in Table 4 appears to support the “fanning out” hypothesis. That is, the WTP distributions in response to the first question are effectively the same regardless of whether respondents face the high or low bid of the bid pair. However, the implied distribution of WTP for those facing a low bid as the second question (LOW2) shifts up significantly, i.e. for a given bid amount, a positively framed second question elicits a higher probability of acceptance than is observed when that same amount is encountered within the initial question. Moreover, in contrast to the LOW2 WTP distribution, those facing a high bid as the second question (HIGH2) tend to express significantly lower WTP, and the degree of shift is determined by the difference between this bid and the low bid they received in the first question. This again supports the reference dependent utility framework of a negatively framed second question leading to a lower likelihood of accepting a given bid amount. Moreover, the greater the distance between the bid values, the lower the likelihood of a respondent accepting a negatively framed second bid.

A simple illustration of this result can be observed in Table 1. Here the £100 bid amount has a mean acceptance rate of between 44-47% when encountered in the first bound of the ascending sequence (i.e. within a positive frame as the $L amount). However, when encountered in the second bound of the same sequence (i.e. within a negative frame, being presented as the higher ($U) of the two bids seen by the respondent) the acceptance rate is consistently lower. Importantly this difference increases with the size of the absolute difference between low and high bids such that acceptance rate is 33% when preceded by $L = £75 (i.e. ABSDIFF = 25) but falls even further to 22% when preceded by $L = £50 (i.e. ABSDIFF = 50)[11]. Therefore, as per our model, as ABSDIFF increases so the difference between first and second bound WTP distributions increases.

5. Discussion and conclusions.

The OOHB approach is an innovative addition to the armoury of contingent valuation elicitation methods. It combines the response simplicity of a dichotomous choice approach yet offers substantial statistical efficiency gains over the more conventional single bound approach without incurring the incentive compatibility compromises of he double bounded method. Therefore the promise of the OOHB approach is extremely high. However, in order for this promise to be fulfilled the method has to be proved robust against common response anomalies such as those afflicting the double bound approach. This paper represents the first test of the theoretical consistency of the OOHB. It also represents the first application of the method to its intended target of environmental public goods. Data is collected through a very large sample survey using state of the art instrument procedures and high quality, resource intensive face-to-face interviewing techniques. Our findings reject the consistency of resulting expressed preferences with economic-theoretical expectations. On the contrary, evidence is found of a number of persistent response anomalies. These are extreme in nature. For example, mean acceptance rates for a given £100 bid amount vary from a

[11] Note that this acceptance rate then stays roughly constant for further increases in ABSDIFF. This may reflect a quadratic shape in the ABSDIFF effect upon the second WTP distribution. Alternatively, it may signify a residue of yea-sayers (respondents who say ‘yes’ irrespective of the amounts concerned) observed within dichotomous choice CV studies (Mitchell and Carson, 1989; Kanninen, 1995; Alberini and Carson, 2001).

low as 22% to as high as 66% across treatments which are equivalent from a standard theory perspective. Indeed response patterns are more consistent with non-standard reference dependent theories. Given these findings, further application of the OOHB approach cannot be justified for policy appraisal purposes until these anomalies are either rectified or explained in a theoretically consistent manner.

One useful line of future enquiry might be to attempt to disentangle the extent to which the relationship between the absolute difference between low and high bids is either triggering theoretically inconsistent framing effects, or is inducing increased uncertainty regarding whether the project is technically feasible or will indeed by undertaken. However, until such justification is forthcoming it may be preferable to treat the OOHB as a target for research effort rather than a tool for practical appraisal use. This is somewhat regrettable given the other highly desirable properties of this innovative approach and underlines the ongoing difficulties faced in constructing theoretically consistent, robust and statistically efficient elicitation methods for contingent valuation applications.

Acknowledgements
Funding for this research was provided by the Commission of the European Community (CEC) through the CLIME project, Framework V Ref. No. EVK1-CT-2002-00121, by the Economics for the Environment Consultancy (EFTEC) and by the PEDM which is funded by the UK Economic and Social Research Council (ESRC). Research assistance from Nuno Noceda-Matias and Loga Subramanian is gratefully acknowledged. An early version of this paper was presented at the Heartland Conference, University of Iowa, September 2003. We would like to thank the participants at that conference, particularly J.R. De Shazo and David Zilberman, for very helpful comments.

References
Alberini, A. and Carson, R.T., (2001) Yea-Sayers, Nay-Sayers or Just Plain Confused? Mixture of Populations in Contingent Valuation Survey Responses, paper presented at *The Eleventh Annual Conference of the European Association of Environmental and Resource Economists (EAERE),* University of Southampton, June 28th - June 30th 2001.

Alberini, A., Kanninen, B. and Carson, R.T., 1997, Modelling response incentive effects in dichotomous choice contingent valuation data, *Land Economics*, 73: 309-324.

Arrow, K., Solow, R., Portney, P.R., Leamer, E.E., Radner, R. and Schuman, H., 1993, Report of the NOAA panel on contingent valuation, *Federal Register,* 58: 4601-4614.

Bateman, I.J., Carson, R.T., Day, B., Hanemann, W.M., Hanley, N., Hett, T., Jones-Lee, M., Loomes, G., Mourato, S., Özdemiroğlu, E., Pearce, D.W., Sugden, R. and Swanson, J. (2002) *Economic Valuation with Stated Preference Techniques: A Manual*, Edward Elgar Publishing, Cheltenham

Bateman, I.J. and Jones, A.P., (2003) Contrasting conventional with multi-level modelling approaches to meta-analysis: An illustration using UK woodland recreation values, *Land Economics*, 79(2): 235-258.

Bateman, I.J., I.H. Langford, A.P. Jones, A.P. and G.N. Kerr, Bound and path effects in double and triple bounded dichotomous choice contingent valuation, *Resource and Energy Economics,* 23(3) (2001), 191-213.

Bateman, I.J., Langford, I.H., Turner, R.K., Willis, K.G. and Garrod, G.D. (1995) Elicitation and truncation effects in contingent valuation studies, *Ecological Economics*, 12(2):161-179.

Bateman, I.J., Munro, A., Rhodes, B., Starmer, C. and Sugden, R. (1997) A test of the theory of reference-dependent preferences, *Quarterly Journal of Economics,* 112(2): 479-505.

Beattie, J., Covey, J., Dolan, P., Hopkins, L., Jones-Lee, M., Loomes, G., Pidgeon, N., Robinson, A and Spencer, A. (1998) On the Contingent Valuation of Safety and the Safety of Contingent Valuation: Part 1 – *Caveat Investigator, Journal of Risk and Uncertainty,* 17, 5-25.

Bishop, R. and T. Heberlein (1979) Measuring values of extra market goods: are indirect measures biased?, *American Journal of Agricultural Economics*, 61(5), 926-930.

Cameron, T.A. and James, M.D. (1987) Efficient estimation methods for "closed-ended" contingent valuation surveys, *Review of Economics and Statistics*, 69(2), 269-276.

Cameron, T.A. and J. Quiggin, "Estimation Using Contingent Valuation Data from a Dichotomous Choice with Follow-Up Questionnaire," *Journal of Environmental Economics and Management* 27(3) (1994), 218-234.

Carson, Richard T. (forthcoming), *Contingent Valuation: A Comprehensive Bibliography and History*. Northampton, MA: Edward Elgar.

Carson, R., Groves, T. and Machina, M. (2000) "Incentive and Informational Properties of Preference Questions" Manuscript. University of California San Diego, February.

Carson, R.T., R.C. Mitchell, W.M. Hanemann, R.J. Kopp, S. Presser, and Paul A. Ruud (1992) *A Contingent Valuation Study of Lost Passive Use Values Resulting from the Exxon Valdez Oil Spill*, Report to the Attorney General of the State of Alaska.

Carthy T., Chilton S., Covey J., Hopkins L., Jones-Lee M., Loomes G., Pidgeon N. and Spencer A. (1999) On the contingent valuation of safety and the safety of contingent valuation: Part 2 - The CV/SG 'chained' approach. Journal of Risk and Uncertainty, 17(3):187-213.

Ciriacy-Wantrup, S.V. 1952. Resource Conservation Economics and Policies. Berkeley: University of California Press.

Ciriacy-Wantrup, S.V. 1947. "Capital Returns from Soil-Conservation Practices." *Journal of Farm Economics* 29:1181-96.

Cooper, J., W.M. Hanemann, and G. Signorello, "One-and-One-Half-Bound Dichotomous Choice Contingent Valuation," *Review of Economics and Statistics* 84(4) (2002), 742-750.

Cummings, R.G., Brookshire, D.S. and Schulze, W.D. (eds.) (1986) *Valuing Environmental Goods: A State of the Arts Assessment of the Contingent Method*, Rowman and Allenheld, Totowa, N.J.

Davis, R.K. (1963) Recreation planning as an economic problem, *Natural Resources Journal*, 3, 239-49.

DeShazo, J.R., "Designing Transactions without Framing Effects in Iterative Question Formats," *Journal of Environmental Economics and Management* 43 (2002), 360-385.

Gibbard A., 1973, Manipulation of Voting Schemes: A General Result, *Econometrica* 41, 587-601.

Green, D., Jacowitz, K.E., Kahneman, D. and McFadden, D. (1998) Referendum contingent valuation, anchoring, and willingness to pay for public goods, *Resource and Energy Economics*, 20 (2): 85-116.

Hanemann, W.M., J. Loomis, and B. Kanninen, "Statistical Efficiency of Double-Bounded Dichotomous Choice Contingent Valuation," *American Journal of Agricultural Economics* 73(4) (1991), 1255-1263.

Herriges, J. and J. Shogren, "Starting Point Bias in Dichotomous Choice Valuation with Follow-up Questioning," *Journal of Environmental Economics and Management* 30 (1996), 112-131.

Hoehn, J.P. and Randall, A., 1987, A satisfactory benefit cost indicator from contingent valuation. *Journal of Environmental Economics and Management* 14(3):226-247.

Kahneman, D. and A. Tversky, "Prospect Theory: An Analysis of Decisions Under Risk," *Econometrica* 47 (1979), 263-291.

Kanninen, B. J., (1995) Bias in discrete response contingent valuation, *Journal of Environmental Economics and Management*, 28, 114-125.

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

Randall, A., Ives, B.C. and Eastman, C., (1974) Bidding games for valuation of aesthetic environmental improvements, *Journal of Environmental Economics and Management*, 1:132-149.

Satterthwaite, M.A. (1975) Strategy-proofness and Arrow's conditions: Existence and correspondence theorems of voting procedures and social welfare functions, *Journal of Economic Theory*, 10: 187- 217.

Tversky, Amos and Daniel Kahneman (1991) Loss aversion in riskless choice: a reference-dependent model, *Quarterly Journal of Economics*, 106: 1039-1061.

Tversky, Amos and Daniel Kahneman (1992) Advances in prospect theory: Cumulative representation of uncertainty, *Journal of Risk and Uncertainty*, 5: 297-323.