Visible Choice Sets and Scope Sensitivity in Contingent Values

Bateman, I.J.[1,2], Cole, M.[3], Cooper, P.[2], Georgiou, S.[2], Hadley, D.[4] and POE, G.L.[1,2,5]

Paper to be Presented at the
American Agricultural Economics Association's Annual Meetings
Chicago, August 2001.

1. Capital letters indicate designated presenter.
2. Centre for Social and Economic Research on the Global Environment (CSERGE), University of East Anglia and University College London. Contact details: I.J. Bateman: email i.bateman@uea.ac.uk, tel: 044 01603 593125, fax: 044 01603 507719. P. Cooper: P.cooper@uea.ac.uk, S. Georgiou: email: s.georgiou@uea.ac.uk.
3. Department of Economics, University of Birmingham.
4. Department of Geography, University of Birmingham.
5. Department of Applied Economics and Management, Cornell University, New York, and Visiting Fellow at the Jackson Environmental Institute, University of East Anglia and CSERGE. Email: glp2@cornell.edu. Tel: 607 255-4707.

Abstract

Through a combination of experimental and field analyses, we demonstrate that varying the visible choice set (i.e., set of goods which, at any given point in a valuation exercise, the respondent perceives as being the full extent of purchase options which will be made available in the course of that exercise) strongly affects observed scope. These results challenge the conclusion that contingent values are plagued by embedding and ordering effects, suggesting that such effects can be eliminated when the respondent is completely informed about the entirety of valuation tasks prior to answering the first elicitation question.

Acknowledgements

The authors have benefited greatly from discussions with Professors Robin Cubitt, Graham Loomes and Robert Sugden (University of East Anglia). The usual disclaimer applies. Funding for this research was provided by the Commission of the European Community (CEC) through the EMERGE project, Framework V Ref. No. EVK1-1999-00159, also by The University of Birmingham, The University of East Anglia and USDA regional project W-133.

Visible Choice Sets and Scope Sensitivity in Contingent Values

Introduction

Contingent valuation research has long been burdened by the related issues that reported values are invariant to the magnitude or scope of the good being valued, i.e. "all contingent valuation studies provided estimates which approximate some fixed amount, say $30" (Boyle et al., 1994 p. 65), and that the order in which these goods "are presented to respondents influences the values ascribed to each, with the goods valued first receiving higher values than later-mentioned goods – other things being equal" (Mitchell and Carson, 1989, p. 44). While such concerns have been longstanding (Tolley et al., 1983; Kahneman, 1986), they gained prominence as a fundamental challenge to the use of contingent values in benefit-cost analyses with the publication of an influential paper by Kahneman and Knetsch (1992). In that paper the authors reported further evidence of insensitivity to scope and contended that estimates of willingness to pay "for the same particular good differ …[substantially]… depending on the scope of the initial question"(p. 62), an effect that they term embedding but we shall later classify as a 'list' effect. In considering these findings, Kahneman and Knetsch conclude that such patterns of response primarily "reflect the willingness to pay for the moral satisfaction of contributing to public goods, not the economic value of these goods" (p.57).

Concern over scope sensitivity has fuelled a burgeoning empirical debate[1], primarily centred on the relative importance of moral satisfaction/warm glow versus more conventional economic-commodity value motivations. In this paper, we argue that this body of research is fundamentally incomplete in that the effect of study design upon observed scope sensitivity has largely been ignored. In particular we highlight the frequently overlooked fact that in many common study designs the choice set offered to respondents (which, following Cubitt and Sugden (forthcoming), we term the 'visible choice set') changes as they progress through a valuation exercise. The issue of changes in the visible choice set is raised by Smith (1992) who identifies the failure to completely inform respondents "about what is to come *before* they are asked to value the first improvement" (p.74) as a primary challenge to the Kahneman and Knetsch results. Indeed, in a footnote defending their treatment of zero responses, Kahneman and Knetsch touch upon this issue by acknowledging that they could have "informed respondents in advance that they will have an opportunity to allocate each contribution to an inclusive good among its separate constituents" (p.61). However, they

[1] Concern over scope was institutionalised within the subsequent NOAA-panel report which maintained that a scope test should be required to assess whether a contingent valuation study is consistent with neoclassical economic-theoretical predictions (Arrow, et al., 1993). In subsequent years a heated empirical debate over scope has permeated the environmental economics literature: while some studies have demonstrated scope sensitivity (e.g., Carson and Mitchell, 1993; Smith and Osborne, 1996; Carson, 1997; Smith, Zhang, and Palmquist, 1999), others have not (e.g., Boyle et al., 1994; Schkade and Payne, 1994; Diamond et al. 1993; Beattie et al., 1998; Hammitt and Graham, 1999), and still others show that it is possible to observe scope and scope insensitivity within the same study (Loomis, Lockwood, and Delacey, 1993; Rollins and Lyke, 1998; Schulze et al., 1998; Giraud, Loomis and Johnson, 1999). A separate set of papers has sought to reinterpret of the Kahneman and Knetsch (1992) data by using alternative statistical analyses, among other critiques on survey design and implementation. See Smith (1992), Harrison (1992) and Nickerson (1995). Other research has been directed towards developing an economic theoretic framework to explain the dependency of values on the sequence in which they are asked (Hoehn and Randall, 1989; Carson and Mitchell, 1995; Carson, Flores, and Haneman. 1998) and supporting these theoretical models with empirical evidence for public (Hoehn and Loomis, 1993) and private goods (Randall and Hoehn, 1996; Bateman et al., 1997). In addressing a critique of the Kahneman and Knetsch study, Brown et al. (1995) demonstrate that providing more informative and detailed descriptions of the commodities to be valued does not change the conclusions drawn in the Kahneman and Knetsch article.

argue that it seems "highly implausible that this minor procedural change would significantly alter results". Similarly, in defending their demonstration of embedding effects, Brown et al. note that "we might have better alerted respondents to what was to come before they were asked the three WTP questions. Specifically, respondents might have been told that they would be asked to provide three estimates of WTP, one for improved environmental services, then one for the subset of improved natural resource services, and then one for the subset of natural areas protection" (p.8). Echoing Kahneman and Knetsch, they express confidence that such a change would not alter the findings reported in their study.

This assertion that changes in the visible choice set will not impact upon the degree of scope sensitivity reported in contingent valuation studies is an open empirical question which to date has, perhaps somewhat surprisingly, not been tested. In this paper we present a formal test of the impacts of varying the visible choice set showing that, counter to the expectations of Kahneman and Knetsch and Brown et al., this has a highly significant impact upon willingness-to-pay values and scope sensitivity. As a result of these findings we argue that these previous studies provide only a partial test of the correspondence between contingent values and economic theory.

In the following section we develop terms that are necessary to characterise those dimensions of study design which are pertinent to tests of scope sensitivity and which allow us to form expectations that arise from the economic-theoretic construct. Building upon this lexicon, we then describe an experiment that isolates and tests hypotheses concerning the role of the visible choice set while controlling for other dimensions. We then explore selected conjectures about influences upon scope sensitivity which arise when the visible choice set is varied. The implications of this research are provided in the final section.

Definitions

One of the key impediments to progress in the debate over scope sensitivity is the absence of a complete and agreed upon lexicon of study design dimensions relevant to scope sensitivity tests. Various commentators have suggested a litany of terms including perfect embedding, part-whole bias, nesting, etc. In this paper we build upon the clarifying work of Carson and Mitchell (1995) which facilitated a fundamental advance in common understanding of the issues underpinning scope sensitivity. In particular, we restate, almost verbatim, their categorisation of types of nesting and somewhat extend their discussion of sequencing and ordering. Our work is distinguished from the Carson and Mitchell paper in that we identify two further dimension of study design which pertain to the issue of scope sensitivity; the visible choice set and how 'lists' of goods are constructed, both of which terms we define below.

(1) Quantitative vs. Categorical nesting.

Carson and Mitchell motivate their discussion through consideration of two goods, stating that if one is a proper subset of the other then together they constitute a set of "nested" goods. They extend this discussion to consider two types of nesting, which we define as follows. *Quantitative nesting* occurs when the goods in a list, say A, B and C, are distinguished only by the magnitude of one argument in a multivariate utility function (e.g., in the valuation of improved visibility A is 30 days of improved visibility, B is 15 of those days, and C is 5 of those days). As such, the primary economic concern in scope tests involving quantitatively

nested goods is the degree of satiation, a point that was explicitly addressed in a recent paper by Rollins and Lyke (1998). *Categorical nesting* occurs when such goods are distinguished by changes in more than one argument in a multivariate utility function (e.g., good A might be improved visibility, wildlife viability, and lower noise levels, B only contains two of those attributes, and C only one). Here therefore the primary economic concern in scope tests involving categorically nested goods is the issue of substitution and complementarity relationships (Carson, Flores and Hanemann, 1998). Although the distinction between types of nesting becomes pertinent in the following discussion of types of lists, we should reiterate the Carson and Mitchell correctly caution that the distinction between quantitative and categorical nesting is a relative one that is, in many ways, in the eye of the beholder; "At one level of detail, everything becomes categorical. Indeed, quantitative nesting can always be treated as categorical nesting by treating increments of the good of interest as separate goods. At another level of detail, everything might be treated as numerical" (ibid., p.157).

(2) Inclusive vs. Exclusive lists.

Carson and Mitchell fail to distinguish between the type of list, within which goods are presented to a respondent, and the effect which the form of that list may have upon our expectations regarding stated values. While this might be regarded as semantics, we identify two fundamentally distinct types of lists, each of which has unique economic-theoretic properties and expectations.

In an *inclusive list* goods are presented as additions to (or subtractions from) any good(s) presented previously in that list. In this manner, adopting the nomenclature of Kahneman and Knetsch (1992) as extended by Carson and Mitchell (1995), the broader good, A, is thought of as being composed of good B plus its complement B*. Similarly B is composed of C and its complement C*. Since the value stated by a respondent for any given good is dependent upon their current endowment of goods, it is readily seen that, for example, the value for good B as the first good presented to an individual will be different from the value stated when the good appears after good C. In the latter case the value stated for good B actually only refers to C*. This *sequencing effect* is an expected prediction of economic theory (Carson and Mitchell, 1995; Randall and Hoehn, 1996) and is one of the earliest findings of empirical CV research (Randall, Hoehn and Tolley, 1981; Hoehn and Randall, 1982; Hoehn, 1983; Tolley et al. 1983). Indeed it is this sequencing effect, which Kahneman and Knetsch term an embedding effect, that is purportedly most damaging to the practical application of CV, "because willingness to pay for the same good can vary over a wide range depending on whether the good is assessed on its own or embedded as part of a more inclusive package". (ibid. p. 57). However, a number of authors (e.g., Smith, 1992; Harrison, 1992; Carson and Mitchell, 1995; Carson, Flores and Hanemann, 1998) have sought to show, with varying degrees of mathematical sophistication, that such context dependence is to be expected in inclusive lists (although the plausibility of the degree to which such effects occur in contingent valuation research remains a concern).

In an *exclusive list*, which is the kind of list that choice theory typically addresses, goods are presented as alternatives to any other goods given in that list. Here the expressed value for a good valued at any position in such a list always refers to the same unit of that good irrespective of its position in that list. For example, a respondent may initially be asked to value good C, after which the endowment which serves as a reference point for the utility function is returned to its original position (i.e. good C is no longer held) and, say, good B is valued. Here, according to standard theory, the expressed value for good B is independent of

its position in the list (i.e., in this example, the stated value for B does not refer to the value of C* as per the inclusive list case given above) as the reference income, prices, level of public goods and utility level across valuation questions remains constant. Any residual variation associated with presentation is therefore a bias (most probably of cognitive psychological origin) and can be termed an *ordering effect.* Empirical evidence for the presence of such effects in CV studies is mixed (Boyle, Welsh, and Bishop, 1993).

(3) List Direction

Carson and Mitchell (1995) further make a directional distinction between the presentation of lists. A *bottom-up* list presents individuals with a succession of nested goods from the 'smallest' to the 'largest'. So, using the example given above, good C would be presented before good B after which good A is presented. In an exclusive list at each instance the reference level of utility is identical, being the initial endowment. In an inclusive list respondents, as before, value good C, then B then A. However, while the initial endowment defines the reference level of utility for the first valuation task, the endowment then changes for the second valuation task as discussed above and changes again for the last valuation.

A *top-down* list reverses the presentation of goods from that detailed above (i.e., the presentation now becomes good A then B and finally C). Such a presentation is logical for exclusive lists but, as Carson and Mitchell note, it becomes poorly defined for inclusive lists and can only partially accomplished by using a value partitioning approach such as that adopted by Kahneman and Knetsch (1992) and Brown et al. (1995).

To the extent that an exclusive list is provided, then economic theory suggests that, in contrast to inclusive lists, values should be invariant to the direction of the list. We return to this issue after we define the visible choice set.

(4) The Visible Choice Set.

We depart from the nomenclature of Carson and Mitchell by defining a new dimension through which study design may influence scope sensitivity; *the visible choice set*. Mirroring recent theoretical developments by Cubitt and Sugden (forthcoming), we define the visible choice set as that set of goods which, at any given point in a valuation exercise, the respondent perceives as being the full extent of purchase options which will be made available in the course of that exercise. The important point to note here is that in some study designs the extent of the visible choice set is varied throughout the course of the experiment. This may occur in bottom-up, top-down, inclusive and/or exclusive list study designs. For example, in a bottom-up, exclusive list respondents might be told that they are going to presented with three goods, C, B and A prior to any values being elicited; an approach which we will term an *advance disclosure* visible choice set. A similar framework could be adopted for top-down designs. Conversely, under what is in all other respects the same design, respondents may be presented initially with only good C and a value elicited on the basis of that visible choice set alone; then they are told about good B (i.e., the visible choice set changes relative to that held at the initial valuation) and a further valuation elicited; finally they are presented with good A (i.e., the visible choice set is further changed) and a value elicited. Similarly, but perhaps less dramatically, a top-down approach could be perceived as altering the realm of the possible choice set by unfolding new opportunities for valuing subsets of the more inclusive goods. We shall characterise such approaches as exhibiting a *stepwise disclosure* visible choice set, with valuation tasks being interspersed *between* each

expansion of the choice set. This contrasts with such valuation tasks being undertaken *after* full revelation of the full choice set as per the advance disclosure approach. Note that in the stepwise approach each valuation task is undertaken in ignorance of the subsequent expansion of the choice set. As such the additional choices can be seen as an unanticipated 'surprise' to the respondent[2].

As indicated in Cubitt and Sugden, "conventional decision theory has nothing to say about surprise choices" (p.19 draft manuscript)[3]. According to standard choice theory, preferences between options are independent of the choice set; an individual has a fixed preference ordering which applies to all choice sets. Therefore choice theory would lead us to expect no difference in stated values elicited from either a stepwise or advance disclosure presentations in an exclusive list[4]. Similarly, theory does not lead us to expect a difference between stepwise or advance disclosure treatments for inclusive lists. However, as previously discussed, we would still expect values obtained from exclusive and inclusive lists to differ.

---

[2] At issue is the degree to which changes in the visible choice set are unanticipated. We posit that in most scope and list tests conducted to date, individuals have substantially uninformed priors concerning potential subsequent expansions in the visible choice set prior to any given valuation task. We suggest that this lack of prescience stems from at least two sources. First, as raised by Bishop and Welsh (1992) in the context of the valuation of endangered, but little known species, information gathering is costly, and one way to ration scarce information gathering resources is to ignore information that is not relevant to the current choice set. Since contingent valuation choice exercises presented to a respondent typically addresses a novel issue to respondents it is naïve to expect that respondent to anticipate the realm of possible alternative scenarios that could potentially be offered. Second, the choices and the lists to be presented are controlled exclusively by the interviewer, who is, by the nature of the exchange, endowed with the "ability to construct arbitrary sequences of trading opportunities, to condition these opportunities on events and to construct arbitrary sequences on their potential victims" (Cubitt and Sugden, draft manuscript p.5). These factors lead us to believe that it is not unreasonable to assume that in most stepwise progressions respondents do not anticipate subsequent valuation questions once the initial valuation task has been completed. Unfortunately, the existing literature does not provide enough information to allow us to test this formally although some evidence to support this contention is presented in the present paper.

[3] As such, the notion of "surprise choices as entirely unanticipated presupposes a concept of radical uncertainty that is foreign to conventional Bayesian decision theory. In a Bayesian analysis, the closest analogue to a surprise choice is a choice that is offered only in an event which the agent initially regards as highly improbably" (Cubitt and Sugden, p. 14 draft manuscript).

[4] Intuitively one might expect that unanticipated expansions of the choice set could create the possibility of substitute and/or complementary relationships with previously unconsidered public goods and hence alter the valuation or 'virtual price' attached to some public good contained within the initial (unexpanded) choice set. If this were true then, following propositions (1) to (5) in Carson, Flores and Hanemann (1998), we would expect that a good valued early in a list would have a different value than the same good valued later in a list. However, the Carson, Flores and Hanemann framework only applies to inclusive lists in which individuals can be seen as 'purchasing' a public good and hence changing the endowment of public goods held when subsequent valuation tasks are undertaken. As such, their context dependence arguments only pertain to situations in which the endowment of goods has changed, not the choice set. For exclusive lists the only way in which adding an unanticipated public good to a choice set might affect the value placed on another public good is if that new good were to be made available at a price less than the respondent's willingness to pay for that good. To illustrate this point, suppose that an individual's willingness to pay for public goods B and C was $100 and $120, respectively. That is, the individual is just indifferent to either remaining in the current situation (without B or C), 'purchasing' good B for $100, or 'purchasing' good C for $120. Further suppose that these two goods are substitutes in the sense that if the level of good C increases then the virtual price for good B declines (see Madden, 1991). Under these conditions if good C were made available at $110 then we would expect the willingness-to-pay value reported for good B to be somewhat less than $100. However, this conflicts with what the researcher is asking a respondent to do in a contingent valuation exercise using an exclusive list format wherein the respondent is asked to provide their maximum willingness to pay for each exclusive good. Following this reasoning, the substitute and complementarity issues raised by Carson, Flores and Hanemann are not relevant to changes in the choice set within an exclusive list format and are not to be expected in the various experiments reported subsequently in this paper.

Developing an Experiment to Test for Study Design Effects

Given these definitions we can now characterise the experiment reported by Kahneman and Knetsch, and by extension outline an alternate set of empirical tests of the correspondence between contingent values and the economic-theoretic construct. Using our lexicon, the Kahneman and Knetsch experiment can best be characterized as consisting of a qualitatively nested set of goods, presented in an inclusive, top-down list and employing a step-wise disclosure format. As presented in their Table 1, the most embracing variant first elicited a value for the most inclusive good A which they call "Environmental services". They then asked respondents to indicate how much of their value for A was attributed to a subset good B called "Improve disaster preparedness". Finally respondents were asked to state how much of the value of good B was attributed to a further subset good C called "Improve rescue equipment, personnel". We refer to this process as value partitioning. In a separate top-down sample they first elicited a value for good B and then asked respondents what part of that value was attributed to good C. A further sample was asked to provide values for good C only. Denoting WTP(X) as the willingness to pay for good X then, using tests of medians and means Kahneman and Knetsch could not reject the hypothesis that WTP(A) = WTP(B) = WTP (C) when these goods are presented first in a list. Conversely, they did find that the value for good C was substantially and significantly lower when presented third in a list than when it was presented either second or first. While these results appear on the surface troubling, it is important to realise that the issue of whether scope sensitivity is found is essentially an empirical matter associated with a particular list of goods rather than one of theoretical necessity. However, while the direction of the effects reported by Kahneman and Knetsch is in accordance with proposition (5) of Carson, Flores and Hanemann, for public goods which are substitutes for each other, we concur with the former that the magnitude of effect reported in their study seems implausibly large.

Despite the above, we do take issue with the conclusion drawn by Kahneman and Knetsch that their results appear "to invalidate a basic assumption of CVM; that standard value theory applies to the measures obtained by this method" (p.68). As evident from the definitions set out above, their empirical design corresponds with only a relatively narrow subset of those dimensions of study design which may impact upon observed scope sensitivity. To attempt to generalise from their findings one must assume that study design dimensions such as inclusive or exclusive lists, or stepwise versus advance disclosure are inconsequential in terms of their impact upon scope. Further, given the inclusive list nature of the Kahneman and Knetsch experiment, there are no clear *a priori* expectations unless one has information regarding the substitution and complementarity relations which prevail between the goods concerned.

Given these caveats, our underlying objective in this research was to provide experimental tests for which clear economic expectations exist. Following from the above, it should be evident that exclusive lists provide much cleaner *a priori* expectations than do inclusive lists. Specifically the value placed upon a given good should be invariant to its position within an exclusive list; a property which does not necessarily hold for inclusive lists. Consequently we adopt an exclusive list format for the experiments reported in this paper.

Variation in the visible choice set provides another pertinent dimension of study design for which clear theoretical expectations exist. As discussed above, standard choice theory indicates that stated values should be invariant to whether a stepwise or advance disclosure approach is adopted. However, our discussion of Smith (1992), Kahneman and Knetsch

(1992) and Brown et al., (1995) indicates that this remains an open empirical issue which we duly investigate in this paper. Given the exclusive nature of the choices under investigation, clear theoretical expectations also exist that changes in the directional presentation of lists (top-down or bottom-up) will also not impact upon stated values, which itself provides an additional testable hypothesis. Additionally, the conjunction between directional and visible choice set dimensions provides a further focus of empirical interest for which there is a clear theoretical expectation that there should be no interaction effect in terms of stated values for exclusive lists.

Whilst it would be desirable from the perspective of our definitional categorization to explore the role, if any, that nesting type has upon reported values, we are more modest in our experimental design, narrowing our focus to the role of visible choice set and direction while holding nesting type constant. This decision is motivated in part by the notion that in most cases distinctions between categorical and quantitative nesting are difficult to discern objectively and may vary from respondent to respondent (Carson and Mitchell, 1995).

To summarise thus far, the theoretical tests identified in the above discussion compel us towards an empirical design which adopts an exclusive list format to evaluate changes in visible choice set interacted with directional differences, holding nesting type constant throughout. To explore these relationships within a relatively controlled situation, we developed a contingent valuation questionnaire concerning students' WTP for improvements in an open access lake located within the grounds of the University of East Anglia (UEA). These surveys were administered to a total sample of 150 students allocated across a split sample design as detailed below using an open-ended elicitation format. A coercive payment vehicle was employed wherein improvements would be undertaken by the University authorities and costs recouped via increases in rental charges to campus shops which would in turn be permitted to pass on charges in the form of higher prices to students. Accounting measures were employed to prevent over-charging and subjects were asked to state maximum WTP per annum via this payment vehicle. A novel feature of the design was that respondents were encouraged to express, in their own words, the factors influencing their responses. Individual analysis of these qualitative data suggested that scenario rejection was not a problem in this study.

Respondents were provided with a structured presentation regarding three nested schemes for improving the lake. The schemes are summarised here as follows, with greater detail provided in Appendix A.:

Scheme F = Filter runoff water from the UEA campus into the lake.

Scheme P = Scheme F plus the planting of reedbeds around the lake.

Scheme D = Scheme P plus the dredging of sediment from the lake.

As we move from the current situation to F to P to D, the biological effects, in terms of water quality and macrofauna diversity, increase in a desirable direction.

To examine the effect of varying the extent of the visible choice set upon observed scope sensitivity. To facilitate this we initially divide our sample in two, with one half presented with the full choice set prior to answering any valuation questions. The other half is initially presented with just one of these schemes, asked to value it, and is then presented with the second scheme and so on. This in turn flags up a further design choice concerning the

direction in which the three schemes are presented. While we expect that direction in which questions are asked should not have a significant effect in advance disclosure visible choice sets (i.e. whether, prior to any valuation question, the top-down presentation of schemes D then P then F is used or a bottom-up presentation of schemes F then P then D), the stepwise effects described previously may influence expressed values in incomplete visible set designs with different directional presentations (i.e. whether in this case the initial choice set consists of scheme D or scheme F. Following from our previous discussions, the implementation of both top-down and bottom-up formats necessitates the application of an exclusive list format in order to maintain economic-theoretic comparability in the values elicited for each scheme. The joint consideration of varying the visible choice set and the possibility of direction related ordering effects dictate a 2x2 split-sample design details of which are presented together with sub-sample abbreviations and corresponding sample sizes in Table 1.

Table 1: Experimental Design and Sub-Sample Labels

		Disclosure	
		Advance	Stepwise
List direction	Bottom-up	ABU (*F*, *P*, *D*) n = 36	SBU (*F*, P, D) n = 35
	Top-down	ATD (*D*, *P*, *F*) n = 43	STD (*D*, P, F) n = 36

Note: All lists are exclusive. Parentheses indicate the presentation of schemes with *italic* type indicating the initial visible choice set.

In order to assess the impact of 'surprise' associated with an unexpected expansion of the visible choice set, we need to permit respondents to revise valuation responses while still preserving those initial stated values expressed prior to expansion of the choice set. To facilitate this, while experimental procedures prevented respondents changing their initial responses, once these were elicited subjects were then allowed to state revised values for any or all of the schemes. Note that respondents were not informed of the possibility of revision until all initial values were elicited.

Experimental Results and Hypothesis Tests of the Economic Theoretic Construct

Given that conventional economic theory does not recognise choice set as having a pertinent role to play in determining valuations within an exclusive list format, the appropriate test is that all of the valuation curves generated by the four treatments used in our experiment should be consistent with each other.

The raw mean and median WTP of initial (unrevised) valuation responses are reported in Table 2 and presented in Figures 1 and 2. Cursory inspection of these results suggests that the economic-theoretic expectation that changes in the visible choice set and list direction should not affect reported values is not supported. Notably, there appear to be substantial

directional effects in the stepwise variants but not in the advance disclosure treatments. Further, the degree of scope sensitivity appears to vary between the two stepwise treatments but is consistent across the advance disclosure samples. Finally, these results and to some extent the implications of this research, is dependent upon whether measures of the mean or the median are used. We investigate these results more systematically by formulating a series of hypotheses and testing these below.

Table 2: Mean and Median WTP for Three Lake Improvement Schemes

Sample	Mean WTP (£) For three schemes			Median WTP (£) for three schemes		
	F	P	D	F	P	D
ABU	16.12 (3.26)	28.18 (4.68)	35.67 (5.68)	8.00	15.00	20.50
ATD	12.94 (3.45)	23.15 (3.95)	36.63 (5.58)	8.40	15.00	25.00
SBU	33.75 (4.20)	40.30 (4.63)	48.79 (5.88)	30.00	32.75	40.00
STD	19.54 (5.27)	40.07 (10.04)	66.56 (18.38)	5.00	15.00	27.50

Note: numbers in parentheses are standard errors

Figure 1: Mean WTP: Scope Across Treatments, Prior to Opportunity for Revision

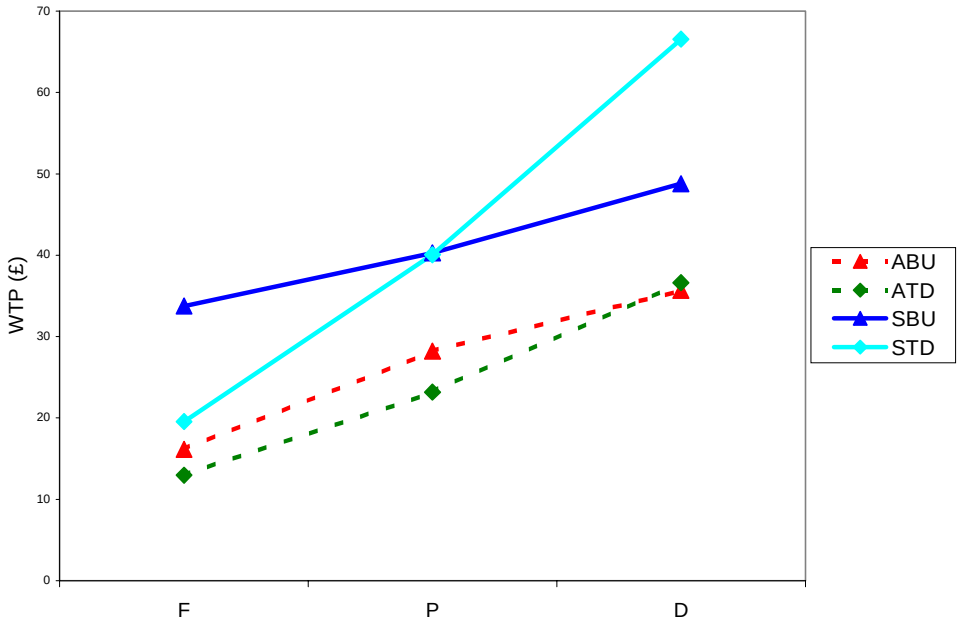
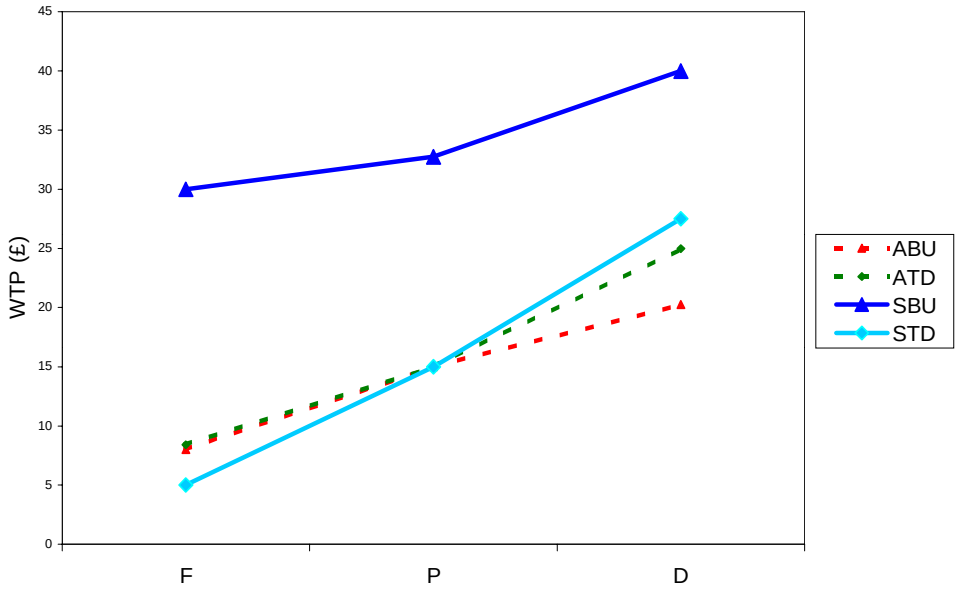


Figure 2: Median WTP: Scope Across Treatments, Prior to Opportunity for Revision



Considering first a formal test of equality between valuation curves then, letting X denote the three levels of protection to be valued (i.e., F, P, and D) and j, k denote the four possible treatments (SBU, STD, ABU, ATD), the economic-theoretic hypothesis of procedural invariance can be expressed as,

$$H^o_1: \quad WTP_j(X) = WTP_k(X) \; \forall j \neq k, X=X$$

A multivariate "within subjects" ANOVA (using the SPSS Manova feature)procedure that controlled for list direction and scheme was used to test this hypothesis of equality of means across experimental treatments. The estimated test statistic indicates that H^o_1 was rejected at the 10 percent level of significance (F=6.11, p=0.015). Therefore, taken as a whole, these results are not in line with standard expectations. Further pairwise testing of individual values for each scheme and treatment combination is reported in Appendix A.

Although the fundamental test of theoretical expectations, the test of H^o_1 provides little insight into the patterns that may be observed in the data. Rejection of this hypothesis requires only one inequality out of the 18 possible pairwise comparisons between stated values of schemes elicited through our various treatments. To refine the focus of our investigation we use the graphical results illustrated in Figures 1 and 2 as a rough guide for identifying a critical set of hypotheses to facilitate improved understanding of the factors influencing responses. Specifically we report results from a series of tests which contrast effects observed within the stepwise treatments with those seen in the advance disclosure treatments as consideration of Figures 1 and 2 suggests that the visible choice set distinction appears to promise certain insight into our empirical results.

Internal and External Tests of Scope Sensitivity:

A scope test looks at whether respondents are willing to pay more for a good that is somehow larger, either in a quantitative or qualitative sense. Tests of scope sensitivity can take the form of either internal or external hypothesis tests. In an internal, or within subjects, test, the same respondents are asked to value a series of goods that vary in inclusivity. An external, or across subjects, test compares values across samples for the same series of goods.

As detailed in Appendix B, each of the treatments exhibit significant internal scope sensitivity ($p < 0.01$). However, it is widely recognised that passage of internal tests is relatively facile and possibly related to the observation that respondents may simply be trying to be "internally consistent" in their reported values (Carson, Flores and Mitchell forthcoming). In light of this, commentators on both sides of the contingent valuation debate have argued for a comparison of the first valuation response of samples facing with differing treatments as the strongest test of scope sensitivity (Kahneman and Knetsch, 1992; Carson and Mitchell, 1995). Holding the type of disclosure constant (i.e. stepwise or advance), two such tests can be identified as follows:

$$H^o_2 : WTP_{ABU}(F) = WTP_{ATD}(D)$$

$$H^o_3 : WTP_{SBU}(F) = WTP_{STD}(D)$$

For each hypothesis difference mean and median WTP responses were tested via independent samples difference of means tests and the Mann-Whitney test, respectively. Reflecting our experimental design and our prior conjectures, both tests used one-tail levels of significance. The results presented in Table 3 confirm highly significant scope sensitivity within advance disclosure treatments. However, results are more equivocal for the stepwise sample with a significant difference between means but not between medians.

Table 3: First Response External Tests of Scope: Holding Disclosure Type Constant and Varying Presentation Order [a,b]

H^o_2 : $WTP_{ABU}(F) = WTP_{ATD}(D)$	H^o_3 : $WTP_{SBU}(F) = WTP_{STD}(D)$
0.001 **0.000**	**0.045** 0.474

a. The top entry in each cell is the p value associated with the difference of means test. The bottom entry in each cell is the p value associated with a difference of medians test. Both tests use a one-tailed level of significance.

b. **Bold** indicates that the probability value is significant at the 10% level.

With the exception of the medians test associated with H^o_3, the results from this experiment demonstrate both internal and external scope. In itself, these results are not that interesting, nor should they be regarded as tests of conformity with economic-theoretic predictions, as it is hard to envision a theory of choice, economic or otherwise, that does not imply a sensitivity to scope. Either finding or failing to find scope sensitivity should not of itself be regarded as proving or disproving any such theory. As highlighted by Rollins and Lyke (1998) a given change in provision of any good may or may not have a discernable impact upon an individuals utility. In effect therefore, scope sensitivity is an empirical rather than a theoretical question, dependent upon the characteristics of each case. However, the finding of significant scope sensitivity in the advance treatments compared to the more equivocal results for the stepwise samples is interesting as the stepwise disclosure format tested here most closely mirrors the external scope test results reported by Kahneman and Knetsch (1992). As discussed previously, they failed to find scope sensitivity in both mean and median comparisons "for the public good mentioned in the first question posed to respondents" (p.61) faced with such a stepwise format.

Ordering Effects:

Economic-theoretic arguments suggest that, for exclusive lists, there should not be ordering effects regardless of whether a stepwise or advance disclosure format is used. However, Figures 1 and 2 suggest that visible choice set plays a dominant role in whether ordering effects are observed or not.

Focusing only on the extreme schemes (i.e., F and D) the hypothesis of procedural invariance is formalised as follows for the advance disclosure treatments;

$$H^o_4 : WTP_{ABU}(F) = WTP_{ATD}(F)$$

H^o_5 : $WTP_{ABU}(D) = WTP_{ATD}(D)$

and equivalently for the stepwise versions,

H^o_6 : $WTP_{SBU}(F) = WTP_{STD}(F)$

H^o_7 : $WTP_{SBU}(D) = WTP_{STD}(D)$

Two sided tests of the medians and means are used because economic-theoretic arguments do not provide a directional expectation. Results are presented in Table 4.

Table 4: Ordering effects by disclosure treatment

H^o_4 : $WTP_{ABU}(F)=WTP_{ATD}(F)$	H^o_5 : $WTP_{ABU}(D)=WTP_{ATD}(D)$	H^o_6 : $WTP_{SBU}(F)= WTP_{STD}(F)$	H^o_7 : $WTP_{SBU}(D)=WTP_{STD}(D)$
0.510	0.905	**0.039**	0.372
0.976	0.693	**0.000**	0.162

a. The top entry in each cell is the p value associated with the difference of means test. The bottom entry in each cell is the p value associated with a difference of medians test. Both tests use a two-tailed level of significance.
b. **Bold** indicates that the probability value is significant at the 10% level.

As depicted, procedural invariance cannot be rejected in the advance disclosure formulations, and hence economic-theoretic predictions are not violated. However, findings are more equivocal for the stepwise treatments with equivalence of both means and medians being rejected for the 'smallest' scope Scheme F.

The Degree of Scope Sensitivity: Holding Disclosure Constant

Economic theoretic arguments indicate that the definition of visible choice set should not impinge upon the observed degree of scope sensitivity. However, this is an empirical question and inspection of Figure 1 and 2 suggests that such differences may be relevant within our data. As suggested in our introduction, there is clearly a division of opinions on this matter, with defenders of empirical evidence (e.g. Kahneman and Knetsch, 1992; Brown et al., 1992) arguing that such procedural variance is unlikely, and critiques of research (e.g., Smith 1992) arguing that such procedural variance may be expected to play an influential role.

Considering this analysis for the advance disclosure treatment generates hypotheses H^o_8 to H^o_{10}:

H^o_8 : $\{WTP_{ABU}(P) - WTP_{ABU}(F)\} = \{WTP_{ATD}(P) - WTP_{ATD}(F)\}$

H^o_9 : $\{WTP_{ABU}(D) - WTP_{ABU}(P)\} = \{WTP_{ATD}(D) - WTP_{ATD}(P)\}$

H^o_{10} : $\{WTP_{ABU}(D) - WTP_{ABU}(F)\} = \{WTP_{ATD}(D) - WTP_{ATD}(F)\}$

while repeating the analysis for the stepwise treatment generates hypotheses H^o_{11} to H^o_{13}:

$$H^o_{11} : \{WTP_{SBU}(P) - WTP_{SBU}(F)\} = \{WTP_{STD}(P) - WTP_{STD}(F)\}$$

$$H^o_{12} : \{WTP_{SBU}(D) - WTP_{SBU}(P)\} = \{WTP_{STD}(D) - WTP_{STD}(P)\}$$

$$H^o_{13} : \{WTP_{SBU}(D) - WTP_{SBU}(F)\} = \{WTP_{STD}(D) - WTP_{STD}(F)\}$$

An independent samples difference of means test was used to test H^o_8 through H^o_{13} while a Mann-Whitney test was used to compare medians. Since there is no economic theoretic basis for assuming directional effects across disclosure formats, a two-tailed level of significance is used. Results from hypothesis tests for the advance disclosure treatments (comparing ABU with ATD) are presented in the top portion of Table 5. Results for the stepwise disclosure treatments (comparing SBU with STD) are reported for the lower cells in Table 5.

Table 5: Scope Consistency across disclosure treatments [a,b]

H^o_8 : $\{WTP_{ABU}(P) - WTP_{ABU}(F)\} = \{WTP_{ATD}(P) - WTP_{ATD}(F)\}$	H^o_9 : $\{WTP_{ABU}(D) - WTP_{ABU}(P)\} = \{WTP_{ATD}(D) - WTP_{ATD}(P)\}$	H^o_{10} : $\{WTP_{ABU}(D) - WTP_{ABU}(F)\} = \{WTP_{ATD}(D) - WTP_{ATD}(F)\}$
0.549 0.851	0.153 0.441	0.518 0.972
H^o_{11} : $\{WTP_{SBU}(P) - WTP_{SBU}(F)\} = \{WTP_{STD}(P) - WTP_{STD}(F)\}$	H^o_{12} : $\{WTP_{SBU}(D) - WTP_{SBU}(P)\} = \{WTP_{STD}(D) - WTP_{STD}(P)\}$	H^o_{13} : $\{WTP_{SBU}(D) - WTP_{SBU}(F)\} = \{WTP_{STD}(D) - WTP_{STD}(F)\}$
0.042 **0.014**	**0.049** **0.001**	**0.035** **0.004**

a. The top entry in each cell is the p value associated with the difference of means test. The bottom entry in each cell is the p value associated with a difference of medians test. Both tests use a two-tailed level of significance.

b. **Bold** indicates that the probability value is significant at the 10% level.

Inspection of the results detailed in Tables 5 reveals a clear and highly consistent pattern. While the degree of scope sensitivity exhibited by the two advance disclosure treatments is statistically identical throughout the full extent of the bid curve, exactly the opposite is true of the two stepwise treatments which exhibit significantly different degrees of scope sensitivity throughout all of the stated values elicited.

Testing the impact of varying the visible choice set in field trials.

While the experimental evidence presented above is, we feel, persuasive of there being a substantial impact from varying the visible choice set, contingent valuation is typically a field-based methodology. In order to test whether our laboratory results carry over into the field, two contingent valuation surveys were undertaken, the first adopting an advance

disclosure format while the second employed a stepwise approach[5]. In both cases both top-down and bottom-up list directions were used.

The advance disclosure treatments (ABU and ATD) were applied to a study examining inner city river water quality improvements for the River Tame in Birmingham, which is currently classified as being of very poor quality (Environment Agency, 1998). Survey data collection was carried out by in-person, at-home interview of residents in the local Birmingham area. A total sample size of 675 respondents was collected, of which 329 (49%) faced the ABU treatment while 346 (51%) were presented with the ATD format. As in our experimental analyses, three nested water quality improvement schemes, centring on ecological and recreational changes, were presented to respondents within a mutually exclusive list (i.e. all valuations were conducted with respect to the current status quo endowment). The three schemes were classified into a Small Improvement (scheme S), a Medium Improvement (scheme M), and Large Improvement (scheme L). To ensure the advance disclosure nature of this design, respondents were provided with full details of all three schemes prior to being asked to state WTP for each in turn. Full details of this study are provided in Georgiou *et al* (2000).

The stepwise disclosure treatments (SBU and STD) were applied to a study examining schemes to prevent saline flooding within the Norfolk Broads wetland. Survey data collection was carried out by in-person, on-site interview of visitors to the wetland area. A sample size of 139 respondents was collected, of which 66 (47%) faced the SBU treatment while 73 (53%) were presented with the STD format. Two nested flood prevention schemes, centring on ecological and recreational changes, were presented to respondents within a mutually exclusive list (i.e. again all valuations were conducted with respect to the current status quo endowment). The three schemes were classified into a Partial-Area Flood Prevention Scheme (scheme PA) and a Whole-Area Flood Prevention Scheme (scheme WA). The stepwise nature of this design was ensured by only informing respondents of a given scheme immediately before asking them to value it such that respondents were unaware that they were to be asked about a second scheme when they provided WTP responses to the first scheme presented to them. The survey instrument used was a simple extension of that employed by Bateman et al., (1995) who also provide further details regarding scheme WA while Powe (1999) provides details of scheme PA.

To expedite an efficient discussion of results from these various surveys we follow the order of discussion given for our experimental findings, comparing for our field and lab trials at each stage. An initial flavour of our results are given in Figures 3 and 4 which illustrate mean WTP bid curves from each scheme in the River Tame and Norfolk Broads surveys respectively (tests of both mean and median WTP are discussed subsequently).

[5] Resource constraints prevented these being undertaken within the same survey.

Figure 3: Mean WTP (£) for all schemes considered in the River Tame study

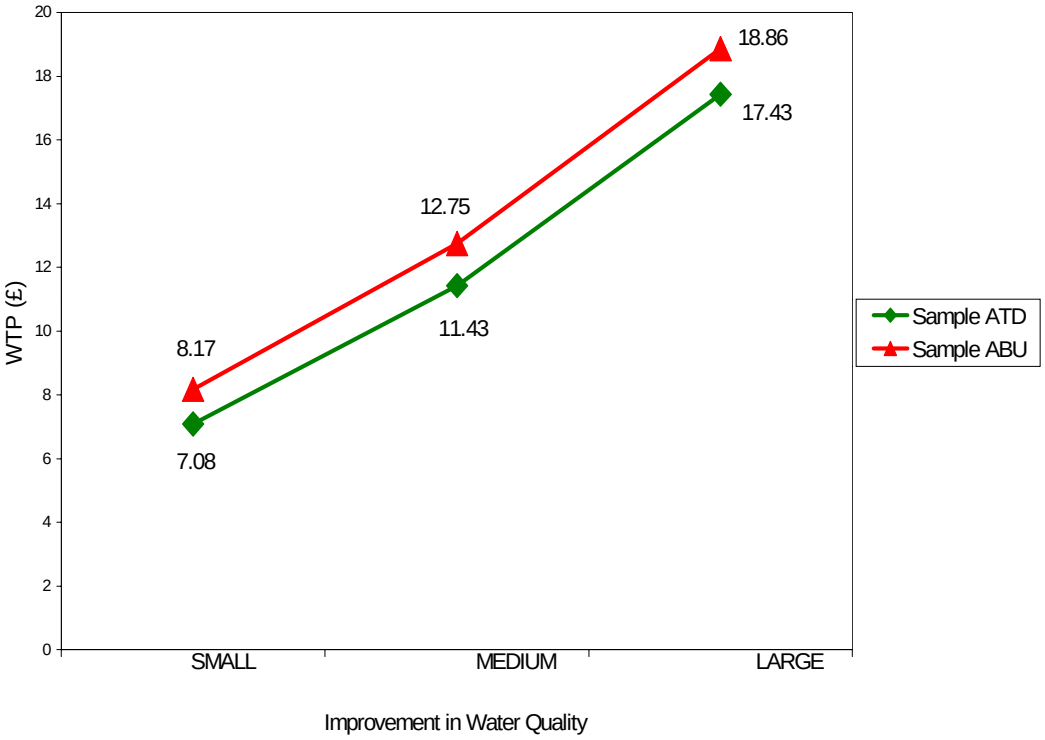
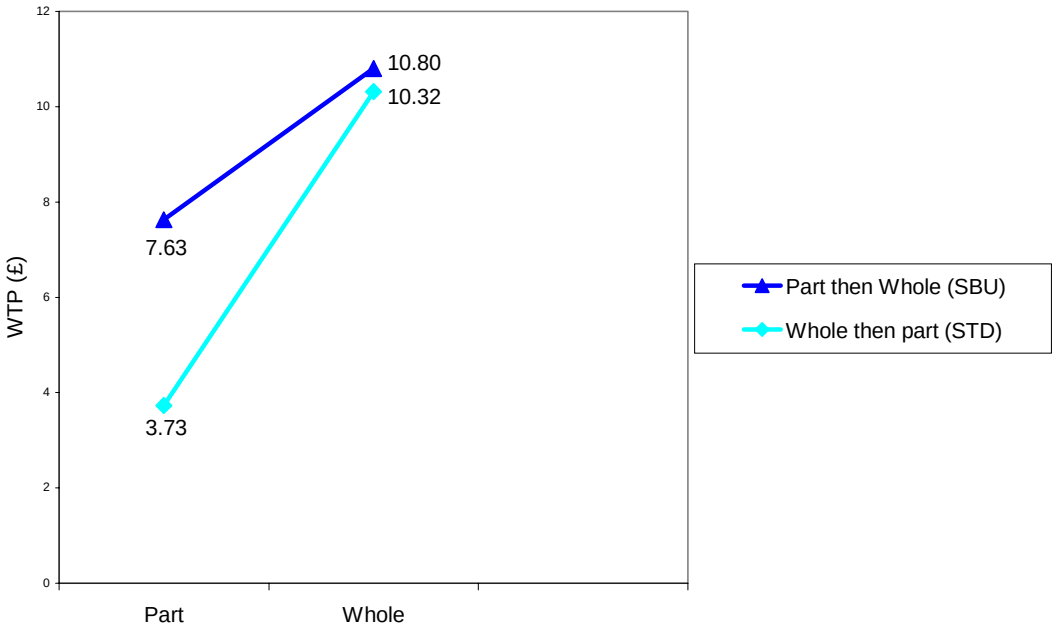


Figure 4: Mean WTP (£) for all schemes considered in the Norfolk Broads study



A cursory inspection of Figures 3 and 4 suggests that the patterns observed in our experimental tests appear to have been replicated in our field trails. While values for given schemes seem invariant to list direction in the advance disclosure format of the River Tame survey this does not appear to be the case in the stepwise disclosure Norfolk Broads study. We now consider more formal tests of these speculations.

All four of the treatments employed in our experimental analyses reveal significant internal scope. This result is replicated between <u>all</u> schemes considered in the advance disclosure study of improvements to the River Tame ($p<0.001$ for both mean and median WTP in all cases including comparisons involving Scheme M, the intermediate scheme). However, this is not the case for our stepwise disclosure study of flood prevention in the Norfolk Broads. Although internal scope sensitivity is found for the STD treatment ($p = 0.045$ for the difference between mean WTP for schemes PA and WA while $p = 0.051$ for a comparison of medians) this is very definitely not the case for the SBU treatment ($p = 0.32$ for means and $p = 0.97$ for medians)[6].

An external scope comparison of first responses (retesting hypotheses H^o_2 and H^o_3 for our field data) reveals a relatively similar pattern to that found in our experiments. Comparing $WTP_{ABU}(S)$ with $WTP_{ATD}(L)$, a one-sided test of the difference of means yields $p = 0.000$ while a difference of medians test gives $p = 0.038$. Our experimental stepwise treatments only detected significant external scope when means were compared and not in the comparison of medians. However, such external tests of scope are substantially less significant in our field stepwise survey where comparisons of $WTP_{SBU}(PA)$ with $WTP_{STD}(WA)$ $p = 0.4044$ and $p = 0.0976$ for the mean and median, respectively.

Tests for ordering effects (retesting hypotheses H^o_4 to H^o_7) also revealed strong similarities between our experimental and field results. As per our experimental findings, the advance disclosure field study of the River Tame revealed no significant difference between the mean or median WTP value of any of the three improvement schemes elicited from either the ABU or ATD treatment ($p<0.1$). Our experimental analysis found that the both mean and median WTP for the smallest scope scheme (F) elicited from the SBU treatment was significantly larger than that for the same scheme obtained from the STD sample (H^o_6 gave $p = 0.039$ for means and 0.000 for medians). A similar significant difference is found for our field stepwise treatments for the smallest scope (PA) scheme ($p = 0.081$ for means and 0.000 for medians). Interestingly both experimental and field analyses do not find significant differences between stepwise values for the largest scope schemes considered (H^o_7, testing scheme D in our experiment and scheme WA in the field trial[7]).

Our final tests concerned the degree of scope sensitivity revealed by each treatment and again we find field support for our laboratory findings. Our experimental analyses found no significant difference between the degree of scope exhibited by our advance treatment measures but did find significant differences for the stepwise treatments. These results are repeated in our field trials. The River Tame study testing our two variants of the advance disclosure treatment yield consistent measures of scope sensitivity when tested using either mean or median measures (with p varying between 0.81 and 0.93 for means and between 0.40

[6] Note that this result replicates the findings of Poe, Giraud and Loomis (forthcoming) who, in reanalysing the data reported in Giraud, Loomis and Johnson (1999), find significant scope sensitivity in what we define as a STD format, but not in an SBU treatment study of WTP to maintain habitat for endangered species.

[7] Tests for a difference between $WTP_{SBU}(W)$ and $WTP_{STD}(W)$ give $p = 0.8992$ for means and $p = 0.4554$ for medians.

and 0.47 for medians for field retests of H^o_8 to H^o_{10}). Conversely our Norfolk Broads stepwise treatments (retesting of H^o_{11} to H^o_{13}) exhibit significantly different degrees of scope when tested using a comparison of medians ($p = 0.08$) and cannot rejected using a test of means ($p = 0.39$).

In summary, we find many similarities between our field and experimental results which together help us to clarify the observed trends. In each of the above tests our advance disclosure field results are identical with those obtained in the laboratory. Both sets of findings exhibit both internal and external scope sensitivity, values for given schemes are invariant to ordering effects and consequently the degree of scope sensitivity does not varying with list presentation. Our laboratory and field stepwise treatments exhibit very similar (if not perfectly identical) characteristics to each other, although now these are not the theoretically expected qualities listed previously[8]; indeed the stepwise format appears to fare even worse in the field than it did in the laboratory. While our experimental tests reveal internal scope within both stepwise responses this only holds for our field STD treatment, with its SBU counterpart failing to exhibit significant scope. In a similar vein, while our lab results indicate somewhat equivocal external scope sensitivity (rejected using median but not mean measures) this translates to a clear cut rejection of scope in our field trials compounded by clear ordering effects found in the value elicited for the smallest scope good (scheme PA). Finally we find evidence that the general degree of scope responsiveness again differs significantly between the two field stepwise treatments, again echoing our experimental findings.

Explorations

The Kahneman and Knetsch top-down stepwise experiment limited their speculations regarding the factors influencing the observed lack of scope sensitivity. Our consideration, within the same experiment, of the interaction of both bottom-up and top-down formats with visible choice set allows us to propose and explore a richer set of potential explanations for the response patterns we observe. However, before moving to these speculations, it is important to note that many critics have ascribed the results such as those exhibited in Kahneman and Knetsch to be a consequence of the hypothetical nature of contingent valuation questions. While it may be that the values stated in hypothetical markets may differ in an absolute sense from those obtained in real market situations, it is not clear how the scenarios presented within the stepwise frameworks are any less credible than those given in the advance disclosure treatments. Furthermore, it has been our experience that several of the anomalous findings reported in the contingent valuation literature (e.g. part-whole effects, reference point impacts, warm glow, etc.) can readily be replicated within real trading situations (Bateman et al., 1997a,b; Ferarro, Rondeau and Poe, 2001).

A common theme running throughout our experimental and field results is that, contrary to economic expectations, scope sensitivity is indeed related to the visible choice set. We observe that within stepwise formats the degree of scope sensitivity varies dramatically and significantly depending upon the directional order in which values are elicited. Such procedural variance does not occur within advance disclosure frameworks. Here we return to our laboratory findings to explore factors which may engender such procedural variance

[8] It should be re-emphasised however that standard theory says nothing regarding the effect of varying the visible choice set and so the formal test is that all values from stepwise and advance disclosure treatments should be consistent.

within the stepwise treatment. We then extend this discussion to eventually speculate upon why the advance disclosure approach appears to be immune to this issue.

Kahneman and Knetsch (1992) report no significant difference in stated values for goods of differing scope where those values were obtained from first responses in a stepwise disclosure format even when the goods being valued vary substantially in terms of their inclusivity. Both they and others attribute such deviations from the economic theoretic construct to the hypothesis that people are willing to "dump their good cause account" (p.64) into the good being valued (see also Harrison, 1992; Cummings, 1989). Hence when a single good is being valued they contend that the reported value is largely comprised of the "moral satisfaction" associated with giving. If this is true then we would expect in our situation that the first stepwise response would be elevated compared to its advance disclosure counterpart value for that scheme. Our results do lend some support to this conjecture. A one-sided t-test rejects the hypothesis that $WTP_{ABU}(F) = WTP_{SBU}(F)$ for both means and medians with $p<0.01$ in each case. However, we obtain a more equivocal result when testing the hypothesis that $WTP_{ATD}(D) = WTP_{STD}(D)$. A one-tail[9] difference of means test indicates that they are significantly different ($p= 0.063$) but a similar test of medians cannot reject the null hypothesis of equality ($p=0.383$).

Such inferences are consistent with findings in the experimental economics literature wherein warm glow (Andreoni, 1990) and other regarding behaviour (Goeree et al., 1999) motives have been shown to represent a significant proportion of actual contributions in single-shot, public good contributions games using real trading situations (Ferraro, Rondeau and Poe, 2001). Further, in moving from single-shot to repeated round public good contribution games, the frequently observed decline in contributions across rounds, which has often been attributed to a learning effect (e.g., Palfrey and Prisbey, 1997) may instead reflect a decline in other-regarding behaviour over rounds (Andreoni, 1995; Ferraro, Rondeau, and Poe, 2001). The extent to which such effects carry over to hypothetical situations could impact the degree of observed scope sensitivity.

Stated alternatively, using the terminology adopted in Sugden (1999), reported values for public goods derive from both instrumental values associated with choices among goods as a means of satisfying preferences and expressions of the consumers sense of identity. The latter 'expressive values' contain many of the elements identified as other regarding behaviour. It seems plausible that expressive values do not have to be maintained over subsequent questions in the sense that the respondent has already shown that they care. To the extent that such conjectured effects exist, there will be an impact upon the observed degree of scope sensitivity in that a top-down stepwise disclosure design will appear to be more scope sensitive than its bottom-up counterpart[10], i.e. the first value in a stepwise list will tend to overstate the economic-commodity value relative to subsequent values. We offer no conjectures as to the relationship of these values in advance disclosure lists.

The differences in the slopes of the stepwise valuation curves are also reminiscent of the gains/losses asymmetry which has been the subject of much contemporary debate and

[9] We elected to adopt a one-tail test of significance here because we have directional expectations formed by the moral satisfaction argument.

[10] Another conjecture is that warm glow may be somehow allocated across the extent of the initial visible choice set. Here the entire good causes account is allocated to the initial valuation response in stepwise disclosure designs, but is spread across the wider set of goods available from the outset in advanced disclosure formats. Such an effect would not preclude the simultaneous action of the other effects outlined in the text.

frequently observed in empirical studies using both contingent and real trading markets (Thaler, 1980; Knetsch and Sinden, 1984; Samuelson and Zeckhauser, 1988; Mitchell and Carson, 1989; Hanemann, 1999; Sugden, 1999). Such asymmetry is purported to arise if an individual's valuation curve is 'kinked' at a status quo *reference level* such that 'losses…loom larger than corresponding gains (Tversky and Kahneman, 1991, p.1047) from that reference level. In our experiment, STD subjects faced with reductions in the scope of the good under consideration exhibit reductions in WTP which are significantly larger than the change in values stated by the SBU subjects for similar gains in the scope of that good. However, this is a subtly different form of 'loss aversion' than observed in previous studies which have typically focussed upon the contrast between WTP and willingness to accept measures. Here, all our values are WTP measures, i.e. strictly speaking they all relate to gains from the present actual holding of the good. However, whereas the SBU measures are elicited in the context of successively 'larger' goods than that initially offered to the respondent, the STD measures are framed in a context where respondents are being presented with successively 'smaller' goods than that initially offered. In stepwise cases then the psychological reference level for second and third valuations is the initially valued scheme (rather than present actual holdings[11]) and 'gains' and 'losses' are determined relative to that reference level. Hence, the concept of loss aversion and reference levels in such situations appears to be primarily psychological, associated with the framing of how goods are presented (Sugden, 1992), with the initially offered gain in the good being seen as the reference level which conditions subsequent responses.

In order to sustain such a line of argument we need to also explain the apparent lack of gains/losses asymmetry observed within the advance disclosure treatments. This appears relatively straightforward as, using the same argument as above, in the advance treatment the reference level will be the set of three schemes available to the individual for purchase. Here then there is no movement (neither 'gain' nor 'loss') outside that reference level and hence we observe symmetry between the valuation curves produced by our two advance disclosure treatments. Following such an argument we find it more convenient to think not of reference levels but rather *reference sets*. For respondents in the stepwise treatment, the reference set consists of just the initially encountered single scheme. For those facing the advance treatment, the reference set consists of all three schemes, as they are presented prior to any valuation task being undertaken. We return to this reference sets argument subsequently.

The difference in slopes may also be in part a product of surprise. Bateman et al., (2001) show via both quantitative and qualitative analyses that contingent valuation respondents react negatively to the introduction of an unanticipated second and third WTP question resulting in a lowering of stated values. Analogously, the broadening of the visible choice set inherent in stepwise designs may plausibly 'surprise' respondents such that they have a greater propensity to want to revise earlier valuations based upon what retrospectively appears to be incomplete choice sets. We can speculate that the degree of surprise and consequent value revision may be stronger in the SBU treatment, where the change in the visible choice reveals an expansion in the scope of the goods available to the respondent, than in the STD sample were the change in the choice set reveals that a good of smaller scope can also be purchased. This is because, in the latter case, the respondent does in effect know about the 'wider picture' in terms of the possibilities of provision but is initially unaware that a smaller subset of those goods may be available. However, in the former case the respondent

[11] Note that the reference level for the initial valuation response may well be the actual present holding of the good.

initially knows relatively little of this wider picture and the consequent feeling of surprise when the visible choice set is revealed may well be tinged with resentment at the apparently deliberate decision to keep this from they until after the initial valuation is elicited. In a manner that reinforces the effects of warm glow wearing off across repeated valuation questioning, this may also diminish scope sensitivity in bottom-up stepwise question formats and exaggerate scope sensitivity in top-down stepwise designs as in both cases second and subsequent valuations are depressed relative to first responses.

Our findings strongly bear out these expectations with the *proportion* of respondents revising at least one value being 8% in the ABU sample and 9% for ATD rising sharply to 22% for the STD treatment and then nearly doubling to 41% amongst SBU respondents with the exception of the two advance disclosure versions these proportions are significantly different ($p<0.10$). Further testing regarding the *direction and magnitude* of revision indicated that the revealed that the only significant changes in occurred in the measures of $WTP_{SBU}(F)$ and $WTP_{SBU}(P)$ where revised values were on average £16.93 and £7.86 lower than respective initial values ($p<0.001$ in both cases). Further details of these analyses are given in Appendix C.

All of the above explorations bear out the apparently important role which variations in the visible choice set have for stated values and it is upon the wider implications of this finding that we focus upon in our concluding remarks.

Conclusions

The experimental and field studies reported above constitute a richer series of investigations into the relation between study design and scope sensitivity than previously undertaken. We conclude that changes in study design and in variation in the visible choice set may have highly significant impacts upon observed sensitivity to scope and may underpin many of the results (both positive and negative) reported in the literature to date. Specifically we have noted that tests such as that reported by Kahneman and Knetsch are incomplete in that they fail to investigate the role that direction and visible choice set have upon scope sensitivity. Hence assertions that their findings show that CV results do not correspond with standard theory are based upon insufficient evidence.

The paper opened by developing a lexicon of definitions, setting out those dimensions of study design which are liable to be pertinent to observed scope sensitivity. Building upon the work of Carson and Mitchell (1995) we offer definitions of types of nesting (quantitative and categorical) and the directions (bottom-up and top-down) in which those nested goods can be described to respondents. We then extended this lexicon to present definitions of inclusive and exclusive lists of nested goods and of the visible choice set; that set of goods which, at any given point in a valuation exercise, the respondent perceives as being the full extent of purchase options which will be made available in the course of that exercise. Our experimental design presented respondents with a categorically nested set of goods within an exclusive list framework. This was used to examine the impact upon stated values of either revealing the full extent of the final choice set of goods from the outset (the advance disclosure treatment), or starting with a subset of those goods and progressively extending the visible choice set (the stepwise treatment). The experiment also allowed inspection of an interaction with the direction of presentation.

Findings from the experiment indicate that the advance disclosure approach yields results which exhibit scope sensitivity both within and across presentation orderings. Furthermore, expressed values are also invariant to the direction of presentation and consequently provide statistically similar degrees of scope sensitivity. By contrast, while the stepwise treatment yields values which are sensitive to scope within a given direction, the treatment exhibits none of these other theoretically consistent characteristics. Stepwise values are not invariant across presentation orderings and neither, unsurprisingly, is the degree of scope sensitivity. External scope tests are at best equivocal and at worst clearly failed.

Our field trials of both advance and stepwise disclosure formats provide fairly consistent corroboration of our experimental findings. Advance formats remain invariant to presentation orderings and reveal consistent degrees of scope sensitivity. Neither of these attributes apply to the stepwise field study results.

In considering possible explorations of our observed results we appeal to a number of speculations derived from the literature concerning anomalies and departures from the predictions of standard theory. These include gains/loss asymmetry, warm glow effects, the impact of surprise and reference point effects. Many of these explorations focus upon the procedural variance observed within responses to stepwise disclosure treatments. However, in focussing our attention upon the stepwise format we do not mean to imply that the values obtained from advance disclosure designs are unbiased measures that somehow reflect 'true WTP'. As we have emphasised throughout this paper, the most comprehensive test of theoretical expectations is that all of the value functions should be invariant to the study design dimensions of presentation ordering and changes in the visible choice set tested in our analyses. Although response patterns in the advance formats appear to correspond with predictions derived from the economic-theoretic construct, it may be simply that respondents are constructing stated values in a way which appears consistent with received economic theory. Clearly further investigations of these effects are warranted, perhaps using induced value, real money experiments.

References.

Andreoni, J., (1990) Impure altruism and donations to public goods, *Economic Journal*, 100: 464-77.

Andreoni, J., (1995) Cooperation in public-goods experiments: Kindness or confusion, *American Economic Review*, 85(4):891-904.

Arrow, K., R. Solow, E. Leamer, R. Radner, and H. Schuman (1993). Report of the NOAA Panel on contingent valuation, *Federal Register,* 58(10): 4601-14.

Bateman, I.J., Langford, I.H., Jones, A.P. and Kerr, G.N., (2001) Bound and path effects in multiple-bound dichotomous choice contingent valuation, *Resource and Energy Economics*, in press.

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. (1997a) Does part-whole bias exist? An experimental investigation, *Economic Journal*, 107(441): 322-332

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

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

Bishop, R. C. and M. P. Welsh (1992). "Existence Values in Benefit-Cost Analysis and Damage Assessments, *Land Economics*, 68:405-417.

Boyle, K. J., M. P. Welsh, and R. C. Bishop, (1993) The Role of Question Order and Respondent Experience in Contingent-Valuation Studies, *Journal of Environmental Economics and Management*, 25 S-80-S-99.

Boyle, K. J., R. Johnson, D. W. McCollum, W. H. Desvousges, R. Dunford, and S. Hudson, 1994. "An Investigation of Part-Whole Biases in Contingent Valuation Studies", *Journal of Environmental Economics and Management,* 27: 64-83.

Brown, Thomas G., Susan C. Barro, Michael J. Manfredo, and George L. Peterson, 1995. *Journal of Environmental Management*, 441-10.

Carson, R.T. (1997) Contingent valuation surveys and tests of insensitivity to scope, in Kopp, R.J., Pommerehne, W.W. and Schwarz, N. (eds.) *Determining the Value of Non-Marketed Goods: Economic, Psychological, and Policy Relevant Aspects of Contingent Valuation Methods*, Kluwer Academic Publishers, Boston.

Carson, R.T., Flores, N.E. and Hanemann, W.M. (1998) Sequencing and valuing public goods, *Journal of Environmental Economics and Management*, 36; 314-323.

Carson, R.C., N. E. Flores, and N. F. Meade, (forthcoming), "Contingent Valuation: Controversies and Evidence", *Environmental and Resource Economics*.

Carson, R. T., and R. C. Mitchell (forthcoming) Public preferences toward environmental risks: The case of trihalomethanes, in Alberini, A., Bjornstad, D. and Kahn, J., *Handbook of Contingent Valuation*, Edward Elgar, Brookfield, VT.

Carson, R.T. and R.C. Mitchell, 1995. "Sequencing and Nesting in Contingent Valuation Surveys." *Journal of Environmental Economics and Management* 28(2): 155-173

Carson, R. T., and R. C. Mitchell. 1993. "The Issue of Scope in Contingent Valuation Surveys." *American Journal of Agricultural Economics* 75:5, pp. 1263-1267.

Cubitt, R., and R. Sugden (forthcoming). "On Money Pumps", *Games and Economics Behavior*.

Diamond, P.A., J.A. Hausman, G.K. Leonard and M.A. Denning, 1993. Does Contingent Valuation Measure preferences? Experimental Evidence, in *Contingent Valuation: A Critical Assessment*, edited by J.A., Hausman, North-Holland, Amsterdam.

Ferraro, P. J., D. Rondeau, and G. L. Poe, forthcoming. "Detecting Other-Regarding Behavior With Virtual Players*", Journal of Economic Behavior and Organization.*

Georgiou, S., Bateman, I.J., Cole, M., Hadley, D. (2000) Contingent ranking and valuation of river water quality improvements: Testing for scope sensitivity, ordering and distance decay effects, *CSERGE Global Environmental Change Working Paper GEC2000-18*, Centre for Social and Economic Research on the Global Environment, University of East Anglia and University College London, October 2000.

Giraud, K. L., J. B. Loomis, and R. L. Johnson (1999). "Internal and External Scope in Willingness-To-Pay Estimates for Threatened and Endangered Wildlife", *Journal of Environmental Management*, 56:221-229.

Goeree, J.K., C. A. Holt, and S. K. Laury, "Private Costs and Public Benefits: Unraveling the Effects of Altruism and Noisy Behavior", *Unpublished Manuscript*, Dept. of Economics, University of Virginia, Charlottesville.

Hammitt, J. K., and J. D. Graham, 1999. "Willingness to Pay for Health Protection: Inadequate Sensitivity to Probability?", *Journal of Risk and Uncertainty*, 8:33-62.

Hanemann, W.M. (1999) Neo-classical economic theory and contingent valuation, in Bateman, I.J. and Willis, K.G. (eds.) *Valuing Environmental Preferences: Theory and*

Practice of the Contingent Valuation Method in the US, EU and Developing Countries, Oxford University Press, Oxford, pp 42-96.

Harrison, G. W. (1992). "Valuing Public Goods with the Contingent Valuation Method: A Critique of Kahneman and Knetsch" *Journal of Environmental Economics and Management* 23:248-257.

Hoehn, J. P., 1983. The Benefits-Costs Evaluation of Multi-Part Public Policy: A Theoretical Framework and Critique of Estimation Methods, *Ph.D. Dissertation*, University of Kentucky.

Hoehn, J. P. , and J. B. Loomis, 1993. "Substitution Effects in the Valuation of Multiple Environmental Programs", *Journal of Environmental Economics and Management*, 25:56-75.

Hoehn J. P., and A. Randall, 1982. "Aggregation and Disaggregation of Program Benefits in a Complex Policy Environment: A Theoretical Framework and Critique of Estimation Methods", paper presented at the annual meetings of the American Agricultural Economics Association, Logan, Utah.

Hoehn, J.P. and A. Randall, 1989. "Too Many Proposals Pass the Benefit Cost Test." *American Economic Review* 79(3): 544-51.

Kahneman, D. 1986. "Comments", in R. C. Cummings, D. S. Brookshire, and W. D. Schulze, eds., *Valuing Environmental Goods*, Totawa, N. J., Roman and Allenheld.

Kahneman, D. and J.L. Knetsch (1992). "Valuing Public Goods: the Purchase of Moral Satisfaction." *Journal of Environmental Economics and Management* 22(1): 57-70.

Knetsch, J.L. and Sinden, J.A. (1984) Willingness to pay and compensation demanded: Experimental evidence of an unexpected disparity in measures of value, *Quarterly Journal of Economics*, 99: 507-521.

Loomis, J.B., M. Lockwood and T. DeLacy. 1993. "Some Empirical Evidence on Embedding Effects in Contingent Valuation of Forest Protection," *Journal of Environmental Economics and Management* 25: 45-55.

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

Nickerson C.A.E., 1993. "Valuing Public-Goods - A Comment On Harrison Critique Of Kahneman And Knetsch", *Journal of Environmental Economics and Management*, 25 (2): 93-102.

Palfrey, T. P., and J. E. Prisbey, 1997. "Anomalous Behavior in Public Goods Experiments: How Much and Why? *American Economic Review* 87:829-846.

Payne, John W., James R. Bettman and Eric J. Johnson (1992). 'Behavioral Decision Research: A Constructive Processing Perspective', *Annual Review of Psychology* 42, 87-131.

Poe, G.L., Giraud, K.L., Loomis, J.B. and Johnson R.L. (forthcoming) Simple algorithms for measuring differences in simulated distributions: Internal and external scope tests in dichotomous choice contingent values, working paper, Cornell University.

Powe, N.A. (1999) Using contingent valuation to value nested goods: a case of the Broadland flood alleviation scheme, *PhD Thesis*, School of Environmental Sciences, University of East Anglia.

Randall, A., and J. P. Hoehn, 1996. "Embedding in Market Demand Systems", *Journal of Environmental Economics and Management,* 30:369-380.

Randall, A., J. P. Hoehn, and G. S. Tolley, 1981. "The Structure of Contingent Markets: Some Empirical Results", Paper Presented at the Annual Meeting of the American Economic Association, Washington D. C.

Rollins, K.S. and A. J. Lyke. 1998. "The Case for Diminishing Marginal Existence Values," *Journal of Environmental Economics and Management* 36: 324-344.

Samuelson, W. and Zeckhauser, R. (1988) Status quo bias in decision making, *Journal of Risk and Uncertainty*, 1: 7-59.

Schulze, W.D., G.H. McClelland, J.K. Lazo, and R.D. Rowe. "Embedding and Calibration in Measuring Non-Use Values," *Resource and Energy Economics* 20(2): 163-178, June 1998.

Smith, V.K., 1992. Arbitrary Values, Good Causes, and Premature Verdicts, *Journal of Environmental Economics and Management*, 22(1): 71-89.

Smith V.K. and L. Osborne. 1996. Do Contingent Valuation Estimates Pass a Scope Test? A Meta-Analysis. *Journal of Environmental Economics and Management,* 31(3):287-301.

Smith, V. K., X. Zhang, and R. B. Palmquist, 1999. Marine Debris, Beach Quality, and Non-Market Values, *Environmental and Resource Economics*, 10:223-247.

Sugden, R. (1992) How people choose, in Hargreaves-Heap, S., Hollis, M., Lyons, B., Sugden, R. and Weale, A. *The Theory of Choice: A Critical Guide*, Blackwell, Oxford, pp36-50.

Sugden, R. (1999) Public goods and contingent valuation, in Bateman, I.J. and Willis, K.G. (eds.) *Valuing Environmental Preferences: Theory and Practice of the Contingent Valuation*

Method in the US, EU and Developing Countries, Oxford University Press, Oxford, pp 131-151.

Thaler, R. (1980) Toward a positive theory of consumer choice, *Journal of Economic Behavior and Organization*, 1: 39-60.

Tolley, G. S., A. Randall, G. Blomquist, R. Fabian, G. Fishelson, A. Frankel, J. Hoehn, R. Krumm, and E. Mensah, 1983. "Establishing and Valuing the Effects of Improved Visibility in the Eastern United States," interim report to the U.S. Environmental Protection Agency.

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

Appendix A

Additional insights into the rejection of equivalence between values for schemes elicited from differing treatments can be achieved by exploring the "within scheme" pairwise comparisons of means and medians, which are indicated in the shaded cells of Table A1 and the "across scheme" pairwise comparisons indicated in the unshaded cells. In an ideal situation in which scope sensitivity is present and procedural invariance holds, the pairwise comparisons in the shaded cells should be significantly different (indicating scope sensitivity) while those comparisons in unshaded cells should not be significantly different. As discussed, values from the two advance disclosure treatments are consistent with theoretical expectations while those from the stepwise treatments are not.

Table A1: Pair-Wise Comparisons of Means and Medians by Scheme and Treatment

			TREATMENT								
			SBU			ATD			STD		
			F	P	D	F	P	D	F	P	D
TREATMENT	ABU	F	**0.001** **0.000**	**0.000** **0.000**	**0.000** **0.000**	0.510 0.976	**0.092** **0.013**	**0.001** **0.000**	0.582 0.756	**0.014** **0.035**	**0.005** **0.000**
		P		**0.070** **0.019**	**0.004** **0.002**		0.414 0.839	0.130 **0.075**		0.288 0.883	**0.025** **0.067**
		D			0.113 **0.039**			0.905 0.693			0.116 0.619
	SBU	F				**0.000** **0.000**	**0.036** **0.005**	0.347 0.334	**0.040** **0.000**	0.285 **0.046**	**0.045** 0.474
		P					**0.006** **0.001**	0.313 **0.078**		0.984 **0.022**	**0.085** 0.228
		D						0.141 **0.023**			0.362 0.164
	ATD	F							0.299 0.526	**0.007** **0.013**	**0.004** **0.000**
		P								0.124 0.894	**0.013** **0.022**
		D									0.127 0.767

Notes: The upper number in each cell is the significance level for a t-test for equality of means. The lower number corresponds to the significance level of difference of medians using Mann-Whitney statistics. Shaded cells indicate tests where the scheme is the same for both treatments and we therefore have no expectations of a difference and consequently employ two tailed significance tests. Unshaded cells indicate tests where the scheme differs across the two treatments and we therefore have no expectations of a difference and employ one-tailed significance levels. Numbers in **bold** indicate that the test statistic is significant at the 10 percent level.

Appendix B

Our experimental design explicitly facilitated both internal and external scope tests. For internal tests values for scenarios F, P, and D were compared within each of the four treatments that constituted the experiment. External scope tests were conducted by holding the approach to visible choice set disclosure constant (i.e. within advance or stepwise treatments) and vary the presentation of goods.

Letting j represent treatment (i.e., ABU, ATD, SBU, STD), then, if non-satiation and scope sensitivity were present, we would expect the following relationship across values: $WTP_j(F) < WTP_j(P) < WTP_j(D)$ where letters in parentheses refer to the three lake improvement schemes discussed previously. This implies three testable internal scope tests hypotheses as follows:

$$H^o_{B1} : WTP_j(F) = WTP_j(P)$$
$$H^o_{B2} : WTP_j(P) = WTP_j(D)$$
$$H^o_{B3} : WTP_j(F) = WTP_j(D)$$

One-sided tests are adopted to reflect our experimental design and results for each treatment are provided in Table B1. Here the upper value in each cell reports the significance level for differences between sample mean WTP (calculated using paired samples difference of means t-tests) while the lower value reports the significance of differences between median WTP (calculated using the Wilcoxon Signed Ranks test).

Table B1: Internal Scope Sensitivity Tests[a,b]

Sample	H^o_{B1} : $WTP_j(F) = WTP_j(P)$	H^o_{B2} : $WTP_j(P) = WTP_j(D)$	H^o_{B3} : $WTP_j(F) = WTP_j(D)$
ABU	**0.000** **0.000**	**0.010** **0.000**	**0.000** **0.000**
ATD	**0.000** **0.000**	**0.000** **0.000**	**0.000** **0.000**
SBU	**0.000** **0.000**	**0.001** **0.000**	**0.000** **0.000**
STD	**0.003** **0.000**	**0.004** **0.000**	**0.003** **0.000**

a. The top entry in each cell is the p value associated with the difference of means test. The bottom entry in each cell is the p value associated with a difference of medians test.
b. **Bold** indicates that the probability value is significant at the 10% level.

Inspection of Table B1 shows that significant internal scope sensitivity is present within all treatments, with the extremely strong statistical results being attributed, at least in part, to the fact that such tests control for across subject variation.

Appendix C: Reactions to Extension of the Visible Choice Set: Surprise and Value Revision

As outlined in the text, we have clear expectations regarding the impact of extending the visible choice set from that initially presented to respondents as implicit in the stepwise treatment. We expect that the surprise caused by such extension will result in higher rates of value rejection amongst respondents faced with a stepwise treatment than those facing the advance disclosure approach were subjects see the final choice set from the outset. Furthermore, as discussed we expect greater resistance to such extensions within the SBU treatment than in the STD sample as the former respondents are initially unaware of the wider picture of potential provision change. Our findings strongly bear out these expectations with the proportion of respondents revising at least one value being 8% in the ABU sample and 9% for ATD rising sharply to 22% for the STD treatment and then nearly doubling to 41% amongst SBU respondents.

Letting $P(r)_j$ denote the proportion of respondents in treatment j who revise at least one of their values these hypotheses can be formalised as follows:

$$H^o_{C1}: P(r)_{SBU} = P(r)_{ABU}$$

$$H^o_{C2}: P(r)_{STD} = P(r)_{ATD}$$

$$H^o_{C3}: P(r)_{SBU} = P(r)_{STD}$$

each of which has an alternate one-sided hypothesis. That is, based on the above, our alternative hypotheses are that $P(d)_{SBU} > P(d)_{ABU}$, $P(d)_{STD} > P(d)_{ATD}$, and $P(d)_{SBU} > P(d)_{STD}$. For completeness, we also investigate the following hypothesis:

$$H^o_{C4}: P(r)_{ABU} = P(r)_{ATD}$$

Because we have no prior theoretical expectations on the direction that differences in proportions might take for the advance versions, this hypothesis test is assessed using a two tail test of significance. In evaluating each of hypotheses C1-C4 a standard difference of proportions t-test is used, with the results of these tests presented in Table C1.

Table C1: Difference of Proportions t-test of Respondents Changing One or More Values

H^o_{C1}: $P(r)_{SBU} = P(r)_{ABU}$	H^o_{C2}: $P(r)_{STD} = P(r)_{ATD}$	H^o_{C3}: $P(r)_{SBU} = P(r)_{STD}$	H^o_{C4}: $P(r)_{ABU} = P(r)_{ATD}$
0.001	**0.063**	**0.046**	0.882

b. **Bold** indicates that the probability value is significant at the 10% level.

These results demonstrate that the stepwise format leads to greater proportion of revision than their advance counterparts and that the bottom up version of the stepwise disclosure format indicates a greater proportion of revision than the top down version. As such, the notion that the stepwise bottom up induces the greatest surprise and desire for revisions is supported by the data.

In addition to the simple proportion of revisions, the direction and the magnitude of revision are of interest. At issue is whether revisions systematically occur in one direction so as to

always low or raise a value. Formally, letting Δ denote change, the following hypothesis is posited:

$$H^{o}_{C5}: \ \Delta WTPj(k)| \ r = 0,$$

where k refers to the scheme. Although we might expect greater downward revision of first responses in the SBU format, such a directional expectation is primarily conjecture. Hence, we use a two-tailed alternative hypothesis for each j,k combination. With exception of SBU(F) and SBU(P), H^{o}_{18} could not be rejected at the 10 percent level of significance. Mean values of revision (standard errors) and p values for these two versions are –16.93 (4.20) p=0.001 and –7.86 (1.87) p =0.001, respectively.

Alternatively, the combined effect of the proportion of revisions and the conditional revision might be examined by conducting within subject comparisons of original (WTP^{o}) and revised WTP (WTP^{r}) for each scheme and each experimental treatment. Formally,

$$H^{o}_{C6} \ : \ WTP^{o}_{j} = WTP^{r}_{j}$$

The central measures of tendency for the restated values are reported in Table 8 and hypothesis tests H^{o}_{1} to H^{o}_{13} were rerun. While p values in some of these tests changed slightly in some of tests of revised values, none of the acceptance/rejection decisions changed using a 10 percent level of significance. And hence, these test statistics are not reported here.

Support for these expectations can initially be obtained by comparing the summary WTP statistics described in Table 2 with those based upon respondents stated WTP once they are allowed to revise their initial bids as detailed in Table C2. These show that while most values are relatively stable, those from the SBU treatment are, as expected, the most changed. The first response value ($WTP_{SBU}(F)$) is the most altered with mean WTP being reduced by just over 20%.

Table 7: Mean and Median WTP for Three Lake Improvement Schemes Once Respondents are Permitted to Revise their Bids

Sample	Mean WTP (£) for three schemes			Median WTP (£) for three schemes		
	F	P	D	F	P	D
ABU	16.45 (3.27)	27.35 (4.23)	34.96 (5.34)	6.50	15.00	24.50
ATD	13.23 (3.45)	22.80 (4.00)	36.60 (5.64)	10.00	15.00	20.00
SBU	26.77 (3.42)	37.06 (4.47)	49.23 (6.05)	25.92	32.75	40.00
STD	20.65 (5.29)	41.19 (10.05)	70.28 (18.43)	5.00	15.00	27.50

Note: numbers in parentheses are standard errors