

Bateman, I.J. and Langford, I.H. (1997) Budget constraint, temporal and ordering effects in contingent valuation studies, *Environment and Planning A*, 29(7): 1215-1228.

**BUDGET CONSTRAINT, TEMPORAL AND QUESTION ORDERING EFFECTS
IN CONTINGENT VALUATION STUDIES**

by

Ian J. Bateman^{1,2} and Ian H. Langford^{1,3}

¹ Centre for Social and Economic Research on the Global Environment, University of East Anglia, Norwich and University College, London, United Kingdom.

² School of Environmental Sciences, University of East Anglia, Norwich, United Kingdom.

³ School of Health Policy and Practice, University of East Anglia, Norwich, United Kingdom.

Abstract

A split sample approach is employed to test three potential design options for contingent valuation studies. A survey of unpriced woodland recreation is undertaken to test: (i) the impact of introducing budget constraint questions; (ii) the effect of varying temporal extent from willingness to pay (WTP) per annum to WTP per visit; and (iii) the consequences of varying the order in which such WTP questions are presented to respondents. Some significant design effects are detected. Alternative explanations of such findings are considered and implications discussed.

1. Introduction

If contingent valuation (CV) assessments of public goods are to have credibility in the decision making process then it must be clearly established that respondents take into account their personal circumstances and constraints when giving valuations, and that these valuations are appropriately responsive to changes in the contingent market and not simply the product of alternate survey design. A number of papers have considered whether or not most CV studies conform to these criteria. In comments upon the first of these issues, Arrow et al., (1993) criticise the majority of CV studies for not incorporating specific questions regarding respondents budget constraints prior to asking valuation questions. Reporting on an empirical investigation into the effects of altering the payment scenario, Kahneman and Knetsch (1992) found that respondents' willingness to pay (WTP) answers were unresponsive to significant changes in the length of time to which valuation questions related. Furthermore, Brookshire *et al.* (1981) report results indicating that where more than one good is presented in a valuation sequence, then the ordering of that sequence could affect the value assigned to any particular good.

This paper attempts to test for evidence of these three issues of budget, temporal and ordering effects within a single split-sample experiment in which:

- i. some respondents were presented with budget constraint questions prior to eliciting a subsequent WTP valuation response;
- ii. each respondent was asked their WTP under two temporal payment scenarios;
- iii. the order in which these scenarios were presented to respondents was varied so that ordering effects might be assessed.

In addition, the study was designed to allow inspection of interaction effects between the above. To avoid the complexities raised by non-use values or information problems, it was decided to conduct the experiment upon visitors to an open-access woodland recreation site.

1.1. Budget Constraint Effects

Clearly, ability to pay should affect WTP. However, in many empirical studies income does not appear as a particularly strong factor within estimated bid functions¹. One potential reason for such a result may be that gross income is not the relevant determining budget constraint. Deaton and Muellbauer (1980) discuss the concept of two-stage budgeting wherein

individuals firstly divide their net household income into separate mental accounts (housing; food; recreation; etc.) and then subdivide each account into individual expenditures. Here, in the context of a CV evaluation of a specific recreation resource, the relevant budget constraint is given by the mental account for all recreation goods and services rather than the overall household income.

Mitchell and Carson (1989) suggest that encouraging respondents to assess their relevant available income for the good in question should reduce what they term "budget constraint bias". However, the impact of such consideration upon stated WTP is, *a priori*, uncertain and depends upon whether the assessed mental account is higher, lower or the same as that implicitly used prior to such explicit consideration.

Only a relatively small number of CV studies have asked respondents to consider, prior to stating their WTP, the implications which their responses may have upon consumption of other goods. Exceptions include Burness *et al.*, (1983), Schulze *et al.*, (1983), Willis and Garrod (1993) and Loomis *et al.*, (1994). In this latter study, which was prompted in direct response to the concerns raised by Arrow *et al.* (1993), roughly half the sample was reminded, prior to the WTP question, that any money spent on the good under evaluation (wilderness fire prevention) "would reduce the amount of money your household will have available to spend on the other environmental problems (mentioned) as well as on the everyday products you buy" (Loomis *et al.*, 1994, 502). Comparison with respondents who did not face this prior statement indicated no significant difference in WTP between the two groups. While this is an interesting result we have two concerns regarding this approach. First, the overall household budget used may not conform to the response-relevant mental account discussed by Deaton and Meullbauer (1980). Secondly, the general nature of this instruction may not motivate respondents sufficiently to calculate the extent of their mental account. Accordingly in our study, prior to any WTP question, subsamples of respondents were invited to calculate their annual expenditure upon all recreation goods i.e. their relevant mental account.

1.2 Temporal Effects

The possibility of a 'temporal embedding' problem is considered in Kahneman and Knetsch's (1992) study of a proposed toxic waste treatment facility. Here no significant difference was found between median WTP elicited as a once-and-for-all lump sum payment,

and that given in response to an annual payment question with a stipulated five year commitment period. Kahneman and Knetsch accordingly conclude that respondents failed to discriminate between the differing temporal inclusiveness of these two payment schedules. However, this particular study has been criticised. Uncertainties, income constraints and the discounting question may mitigate against lump-sum approaches such that comparison with annual WTP bids is not straightforward (Smith, 1992). Furthermore, studies by Carson and Mitchell (1993) and Carson *et al.* (1995) both report an absence of temporal effects over three and two year periods respectively. Nevertheless the possibility of such effects is interesting and deserves further consideration.

An obvious temporal division was suggested by the valuation literature regarding non-market woodland recreation. Here, studies have frequently used either per annum or per visit measures, sometimes within the same study (e.g. Bishop, 1992). Such a division seemed to have considerable advantages with respect to scenario comprehension and credibility. Furthermore, if WTP bids did fail to differ across these designs this would give strong evidence of a serious temporal embedding problem for CV². However, we were aware that the switch in payment vehicles necessitated by such an experiment might have a confounding effect upon subsequent interpretation (a criticism which can also be levelled at the Kahneman and Knetsch study). We conclude that the advantages of scenario clarity and credibility afforded by such a comparison (and the virtual impossibility of altering payment periods without in effect altering perceived payment scenarios) meant that such an experiment would be of interest, although findings would have to be interpreted with some caution.

1.3. Ordering Effects

A basic assumption of rational choice theory is the principle of 'extensionality' (Arrow, 1982) or 'invariance' (Kahneman and Tversky, 1984; Tversky et al., 1988) under which preferences should not depend upon (amongst other things) the order in which options are presented to the individual. The problem of "sequence" or "question-order" bias (Cummings, et al., 1986; Mitchell and Carson, 1989, respectively) arises when this principle is contravened.

Again, relatively little empirical work has been carried out in this area. In an important early study, Brookshire et al., (1981) found significant question ordering effects regarding WTP for three levels of air quality. In contrast Boyle et al., (1990) report no significant ordering

impacts in WTP for hunting several different species of wildlife. However, in a more recent study, Boyle et al., (1993) suggest a possible inverse link between ordering effects and respondents experience of the good under investigation. Here, in a study of WTP for water flows in the Colorado River, inexperienced white-water rafters were found to exhibit significant ordering effects, a result which was not repeated for experienced rafters.

A thorough appraisal of theoretical expectations regarding ordering effects is presented by Carson and Mitchell (1995) who place such effects within the context of the expanding literature on embedding (e.g. Kahneman and Knetsch, 1992; Desvousges et al., 1993; Diamond et al., 1993). In particular they focus on nested goods, i.e. those which are subsets of more inclusive goods. Carson and Mitchell demonstrate that, where good B is a quantitative subset (or "part") of good A (the "whole") with values for both being elicited from top-down (i.e. whole then part; sample I) *and* bottom-up (i.e. part then whole; sample II) sequences of WTP questions, then economic theory expects:

$$A \geq B \quad (1a)$$

and

$$B_I \leq A_I \quad (1b)$$

where the superscripts indicate whether the good was the first or second in the valuation sequence and the subscripts identify samples.

However, drawing upon earlier work regarding differences that respondents perceive in the available choice sets implied by top-down and bottom-up sequences (Carson et al., 1992; Randall and Hoehn, 1992¹), Carson and Mitchell show that theory also implies that:

$$A \geq A_I \quad (2a)$$

and

$$B_I \geq B \quad (2b)$$

This therefore rejects the argument, implicit in the equality tests of (2) conducted by critics of CV (e.g. Kahneman and Knetsch, 1992), that valuations for given goods should be

¹A version of this paper is forthcoming in the Journal of Environmental Economics and Management.

largely context free. Rather such inequalities are quite feasible and not inconsistent with economic theory. As an alternative, Carson and Mitchell suggest the following 'component sensitivity' test:

$$A \geq B_i \quad (3)$$

However, Carson and Mitchell recognise that while (3) is necessary it is not sufficient to demonstrate the validity of CV studies and argue for the inclusion of further tests with specific emphasis being placed upon theoretical validity investigations of bid functions examining whether they conform to expectations.

In our study we wished to assess each of (1) to (3) above (accepting the argument of Carson and Mitchell that only (3) can have a strong theoretical expectation of non-equality). These ordering effect tests were combined with our investigation into temporal effects by asking half the sample to state their WTP per annum (i.e. good A above) and then their WTP per visit (i.e. good B above), and then reversing this question order for the remainder of the sample. By this device we obtain both unpreceded and preceded estimates of both WTP measures.

2. Study Design and Survey

2.1. Study Design

In addition to the above we wished to examine interactions between effects. To facilitate this a split-sample design was used in which respondents were divided into two groups, each of which was further divided into two subgroups as follows:

Group B: Prior to any WTP question, respondents were asked to calculate and state their annual recreational budget.

Group NB: No budget question asked prior to any WTP response.

Subgroup 1: WTP per annum (tax payment vehicle) asked prior to WTP per visit (entrance fee payment vehicle) question.

Subgroup 2: WTP per visit (fee) asked prior to WTP per annum (tax) question.

In order to ease the sample size requirements of such an experiment, valuations were elicited using open-ended WTP questions³. Questionnaire testing had been achieved through an earlier study of an open-access wetland (Bateman et al., 1995a) which, in a large pilot exercise (433 respondents), tested all the budget and payment questions employed in the present analysis but had not employed a split sample approach, thus making controlled comparison tests infeasible. A small pilot analysis was undertaken (32 respondents) which confirmed the transferability of such questions to an open-access woodland setting.

The above design gave us four subsamples each of which provided both per annum (tax) and per visit (fee) WTP responses for which we defined a series of testable hypotheses concerning the effects under investigation. Testing was achieved via various approaches. Simple comparisons of means and standard (normal) confidence intervals were undertaken. However, while of interest, such statistics are potentially biased by necessary distributional assumptions. To combat this non-parametric confidence intervals for mean WTP were calculated via the BCa percentile bootstrap method (Efron and Tibshirani, 1993) using 999 simulations. This method is based on a refined normal approximation which corrects for bias and skewness in the distribution of mean WTP and is hence an improvement upon the basic non-parametric bootstrap. This is of importance with the samples of WTP values considered which are both skewed and truncated at zero.

2.2. Survey Details

The study was conducted between 26th March and 25th April 1993 at Lynford Stag, a typical Forestry Commission (State owned) open-access woodland recreation site within East Anglia's Thetford Forest. The site has a single entrance and sufficient experienced interviewers were employed such that all visitors could be asked to participate in the survey. Of the 475 parties of visitors approached, 351 agreed to be interviewed giving an initial participation rate of 73.9%. At the start of each interview respondents (unbeknown to themselves) were randomly allocated to one of the four subsamples described above such that roughly one-quarter of the total sample was in each subsample.

Prior to the budget and WTP questions, respondents were asked whether or not they were willing to pay anything at all. This 'payment principle' question was included because we felt that interviewees who did not wish to pay might feel inhibited from stating such a response if they were directly asked the WTP question⁴. Some 96 respondents (27.6% of all respondents) stated that they were not prepared to pay at least some amount for the recreational facilities provided at Thetford Forest⁵, leaving 255 respondents to answer the budget and WTP questions. To prevent overstating sample WTP (and avoid problems caused by somewhat uneven numbers in each subsample accepting the payment principle), in later calculations these refusals were allocated evenly between the four subsamples and treated as zero's⁶.

In addition to the valuation responses the survey also elicited information regarding relevant visit, socioeconomic and interview condition variables necessary for subsequent validity analyses.

3. Results

Here we present results from the analysis of, first, per annum and, secondly, per visit WTP responses. In discussing each we concentrate upon testing for budget and ordering effects while comparison of the two allows us to consider temporal/payment vehicle effects.

3.1. *WTP Per Annum (Tax) Responses*

Table 1 details mean WTP per annum (via taxes) for each of our four subsamples. For notational simplicity we can refer to the subsamples described in the upper row as NB1_a and NB2_a (left and right hand cells respectively; subscript *a* indicating per annum (tax) response) and similarly those on the lower row as B1_a and B2_a. Below each mean (in curved brackets) we report 95% confidence intervals calculated via the BCa percentile bootstrap method, while below these (in square brackets and small print) we report standard normal 95% confidence intervals for comparison.

Table 1 indicates that the inclusion or exclusion of the recreational budget question, and/or changes in the ordering of payment vehicle presentation, results in apparently consistent and major impacts upon stated WTP. For ordering scenario 1 (tax then fee), the inclusion of the

budget question (i.e. moving from cell NB1_a to cell B1_a) raised mean annual WTP (tax) by a factor of 2.60, while for ordering scenario 2 (fee then tax) inclusion of the budget question (i.e. moving from cell NB2_a to cell B2_a) raised mean annual WTP (tax) by a factor of 2.15. However, examination of BCa confidence intervals shows that only the first of these differences is clearly significant (i.e. 95% BCa confidence intervals do not overlap).

Considering the impact of changing the order of payment questions upon per annum responses, in those subsamples not given the prior budget question, asking for per visit WTP before the per annum question (cell NB2_a) lowered the latter to just 60.7% of stated annual WTP when not preceded by a per visit question (cell NB1_a). For those subsamples who were given a prior budget question, this disparity increased so that annual WTP preceded by a per visit question (cell B2_a) is just 50.2% of annual WTP not so preceded (cell B1_a). Again, examining BCa confidence intervals indicates that only one of these differences (the latter) is significant, suggesting in this case that the prior per visit question has substantially reduced the subsequent stated per annum WTP.

Consideration of the diagonals of table 1 shows that where the apparently negative effect of including a prior per visit WTP question is combined with the positive impact of a prior budget question (cell B2_a), then the resultant per annum WTP statement is quite clearly not significantly different from that elicited in the absence of both priors (cell NB1_a). However, comparison of stated per annum WTP when preceded solely by the apparently negative effect of a prior per visit question (cell NB2_a) with annual WTP preceded solely by the positive impact of a prior budget question (cell B1_a) shows a highly significant difference in WTP responses.

Comparison of the BCa and standard (normal) confidence intervals is also interesting. The distributional assumption underlying the latter does not prevent negative WTP values and the presence of significant numbers of zero's (payment principle refusals), alongside a distribution of non-zero responses containing some relatively high values, results in unreliable confidence intervals. These problems are corrected for by the BCa approach by using empirically derived estimates of bias and skewness which are calculated for each subsample. Upper and lower percentile points are then calculated accordingly. Here we can see that reliance upon conventional (normal) confidence intervals would overemphasise the significance of differences between subsamples.

3.2. *WTP Per Visit (Fee) Responses*

Table 2 details mean WTP via per visit fees for each subsample and 95% confidence intervals (as previously described). In subsequent discussions subsample notation is generally as before with the subscript v indicating per visit (fee) responses.

Considering table 2 we can see that the design effects detected in the per annum experiments have been repeated in the per visit studies. Again the inclusion of a prior question regarding recreation budgets seems to lead to increases in subsequent per visit WTP responses which are significant in both cases⁷. Table 2 also shows that the prefixing of per visit WTP questions by per annum questions apparently increases per visit WTP bids. However, examination of BCa confidence intervals indicates that only one of these two differences is statistically significant.

Consideration of the diagonals of table 2 again tells a consistent story regarding the interplay of budget and ordering effects. Where these move respondents in the same direction (i.e. comparing both negative influences in cell NB2 $_v$ with both positive influences in cell B1 $_v$) confidence intervals indicate highly significant differences, but where they move in opposition (comparing cell NB1 $_v$ with cell B2 $_v$) equality cannot be rejected.

Finally as before, reliance upon normal confidence intervals would generally lead us to overestimate the significance of these results.

3.3. *Validation*

Validation of survey results was carried out in accordance with the criteria set by Mitchell and Carson (1989). A central notion here is the concept of construct validity which is in turn composed of two analyses: convergent and theoretical validation. In practice convergent validity testing has generally been achieved by comparing benefits with those of other studies while theoretical validity has been analysed through the estimation of bid functions and analysis of their consistency with theoretical expectations.

Two types of convergent validity test were undertaken. In the first, results from the NB subgroups of this study were compared with 48 other estimates of UK woodland recreation value (there were no studies comparable with the B format subgroups). Numerous statistical tests including a meta-analysis across estimates showed that the results obtained in the present study strongly conformed to expectations from prior research (details in Bateman *et al.*, 1995b). A

second test compared the unprecedented per annum WTP result with a selection of similar format studies. It was found that results across these studies were logically related to both substitute availability and the change in provision presented in the contingent market and that the findings of the present study were consistent with the model produced from these expectations (details in Bateman *et al.*, 1995c).

Theoretical validation of our results was carried out via statistical investigation of the bid functions underlying WTP responses. A semi-log (dependent) functional form provided the best fit for the per annum data as reported in equation (4):

$$\begin{aligned} \ln WTP_{tax} = & 1.20 + 1.50 \text{ BUDGET} - 0.633 \text{ ORDER} + 0.390 \text{ GREEN} \\ & (10.6) \quad (11.17) \quad (4.76) \quad (1.66) \\ & + 1.08 \text{ NONCAR} + 0.574 \text{ SUPERB} \end{aligned} \quad (4)$$

(3.35) (2.88)

$R^2(\text{adj}) = 33.7\%$ Figures in brackets are modulus t-values

where:

- $\ln WTP_{tax}$ = natural logarithm of WTP per annum (tax vehicle)
- BUDGET = 1 if respondent had been asked to state annual recreational budget prior to WTP questions; = 0 otherwise
- ORDER = 1 if respondent faced a prior per visit WTP question (ordering scenario 2); = 0 otherwise
- GREEN = 1 if respondent was a member of at least one of various countryside/wildlife organisations; = 0 otherwise
- NONCAR = 1 if the respondent did not travel to the site by car; = 0 otherwise
- SUPERB = 1 if the respondent rated scenery at the site on the top of a four point scale; = 0 otherwise

Mitchell and Carson (1989) recommend that a CV bid function should only be accepted as proof of theoretical validity provided that it explains at least 15% of the variation in WTP responses, a criteria which is easily satisfied here. Equation (4) also bears out the significant influence of the budget constraint and question ordering effects upon per annum responses, a

result which is repeated in the per visit bid function shown in equation (5), which again readily satisfies the Mitchell and Carson criteria regarding the degree of explanation. Here a linear form fitted the data best, reflecting the clumping of bids around two round-figure amounts (50 pence and £1).

$$\begin{aligned} \text{WTPfee} = & 0.618 + 0.167 \text{ BUDGET} - 0.167 \text{ ORDER} - 0.299 \text{ GREEN} \\ & (8.12) \quad (2.48) \quad (1.94) \quad (3.05) \\ & + 0.397 \text{ CAMP} \quad (5) \\ & (3.16) \end{aligned}$$

$R^2(\text{adj}) = 26.4\%$ Figures in brackets are modulus t-values

where:

WTPfee = WTP per visit (entrance fee vehicle)
 CAMP = 1 if respondent often camps in the area;
 = 0 otherwise
 Other variables as defined above.

In general the above findings are unremarkable with one exception: the dramatic change in the influence of the explanatory variable GREEN which is positively related to per annum bids (although only significant at the 10% level), but negatively associated with per visit bids (significant at the 1% level). We discuss this and other findings below.

4. Discussion

4.1. Budget Constraint Effects

In both our per visit and per annum responses the inclusion of a prior budget constraint question results in a very substantial increase in subsequent stated WTP. Three of the four comparisons which make up this analysis indicate that this difference is statistically significant; a result of some importance for CV research. Comparison with the results of Loomis *et al.*, (1994) suggests that our use of response-relevant mental accounts rather than general household budgets and the process of getting respondents to calculate the extent of these mental accounts may well explain the difference in findings.

The direction of impact is also interesting. Most commentators (Mitchell and Carson, 1989; Willis and Garrod, 1993) discuss cases in which, *a priori*, we would expect that respondents' consideration of annual expenditure upon recreation and consequent budget constraints would lead to a reduction in stated WTP compared to statements made without such consideration. However, here we observe a very strong opposite effect whereby respondents who are asked to calculate their present annual expenditure state significantly higher WTP sums than those not asked the prior budget question.

Why has this effect occurred? It seems to us that two interpretations are possible, one generally supportive of CV and the other critical. The former argues that respondents forced to overtly consider their annual recreational budget find that, on average, this accounts for a significant portion of their total annual expenditure, perhaps more than they realised without such consideration. Certainly stated annual recreational budgets were not insignificant. The mean budget (£227.30) was considerably affected by the skewed nature of this distribution. Nevertheless, the median value of £120 shows that most respondents had considerable annual recreation budgets. Following this argument then, after considering the apparent importance of recreation in their preference sets, such respondents gave higher WTP sums than would otherwise have been stated. If we accept such a line of reasoning then a supplementary question arises as to which WTP measure (with, or without, the prior budget question) is correct. The argument would seem to suggest that answers formulated following the consideration of available budgets will be less susceptible to mental accounting problems and therefore preferable⁸.

A more critical interpretation of our findings, however, argues that the calculation of the annual budget (which is relatively high compared to WTP) acts as an anchor for subsequent WTP statements. Kahneman *et al.*, (1982) suggest that such an effect is most likely to occur where individuals are inexperienced and face considerable uncertainty in forming their response. Here, then, our use of an open-ended WTP elicitation approach may have exacerbated such an effect as individuals do not have as much experience of setting prices as opposed to reacting to them⁹.

Clearly such findings gives us pause for thought regarding the degree to which WTP responses may be manipulated by small and apparently defensible changes in questionnaire

design. The responsiveness of stated WTP to the inclusion of the budget question is remarkable and a matter of significant concern for future CV studies.

4.2. Temporal and Ordering Effects

Given our choice of goods (per visit and per annum recreation) our temporal effect test becomes a subset of our wider investigation into ordering effects. The relevant tests for temporal effects in our study design are specified in equations (1) and (3). Comparison of the per annum and per visit WTP sums stated within each sample readily passes the inequality tests outlined in (1a) and (1b). However, such tests are weak in that a prior answer may enforce temporal scaling and apparent rationality upon a subsequent bid made by the same respondent. By using only those WTP responses which have not followed previous WTP questions (i.e. comparing annual WTP from subsamples NB1_a, B1_a with per visit WTP from subsamples NB1_v and B1_v) then, for any given budget question scenario, per annum responses clearly differ very significantly from per visit responses. In other words Carson and Mitchell's (1995) 'component sensitivity' test (3) is well satisfied by these results. Proponents of the temporal embedding idea may argue, quite defensibly, that such a result may say as much about the change in payment vehicle between the per annum and per visit experiments (of which more below) than about their temporal extent. However, any change in such extent implies a consequent change in payment scenario. We accept that it is difficult to disentangle the two in this experiment (although note that such separation is not clearcut elsewhere, e.g. Kahneman and Knetsch, 1992) and consequently we cannot claim that our result shows that temporal embedding, per se, is not an issue in CV. Rather all we can say is that there is no evidence for such a problem in this case.

Considering ordering effects within responses to per annum questions (i.e. testing equation (2a)) we can see that, irrespective of whether or not a budget constraint question was asked, the asking of a prior per visit WTP question reduced the magnitude of subsequent WTP per annum responses compared to its level when presented as the first valuation question, although only one of these differences is statistically significant. A first point to note regarding such ordering effects is that, as indicated in our discussion of both equation (2a) and (2b), these results are not inconsistent with theory, indeed the work of Carson et al. (1992) and Randall and Hoehn (1992) would lead to such an expectation. However, there are further (although not necessarily contradictory) explanations of these results. A somewhat simplistic interpretation of

such findings might be that such respondents were taking prior per-visit payments and extrapolating them to produce a per-annum sum¹⁰. However, this would imply that per annum responses made prior to per-visit bids were in error.

An alternative explanation of the apparent ordering effect is suggested by our observation that membership of "green" groups was positively correlated with WTP per annum but negatively related to WTP per visit. We suggest that this apparent disparity arises from a change, induced by the switch in payment vehicle, in the perceived nature of the good under evaluation. When presented with the unprecedented WTP per annum question (ordering scenario 1), respondents recognise a typical payment mechanism for funding public goods in the UK. Individuals understand the redistributive nature of most UK taxes and that such a payment would preserve the common property, public good, nature of recreation within Forestry Commission woodlands. Here then, payments ensure provision for both the payee and other members of society, both of which are likely to be valued by the respondent. However, respondents facing ordering scenario 2 are initially presented with a WTP per visit (entrance fee) question. Such payments only ensure access to the payee and imply the exclusion of non-payers. Here then the payment vehicle describes a private rather than public good. This perception is liable to be retained when, immediately after this, respondents are presented with the per annum WTP question. We can therefore view the apparent ordering effect in per annum responses as arising out of a payment vehicle induced, category shift in perceptions regarding the nature of the good under evaluation. The observed relationship of responses conforms to the perceived loss of services between ordering scenario 1 (recreation seen as a public good) and 2 (recreation seen as a private good).

If the difference in WTP statements is derived purely from the additional value aspects which respondents feel they obtain from woodland as a public good (bequest, altruism, etc.) than as a private good, then, while complicating the matter, this may be viewed as simply reflecting preferences. However, a number of commentators have argued that the evaluation of the same asset as either a public or private good may alter the underlying motivations upon which individual preferences lie. Schkade and Payne (1994) note that evaluations of public goods appear in part to reflect norms regarding civic duty and fairness. Furthermore, Brennan and Buchanan (1984) argue that such valuations may also be influenced by a self-image or

expressive value, derived from contributions towards goods which benefit not just the individual but also others in society.

In support of such an argument it is important to emphasise that the study was conducted midway through a high-profile, year-long public debate concerning (and generally opposing) the UK Governments plans to privatise the Forestry Commission estate, a resource which provides the largest area of free access recreational land in the UK. Countryside groups and their members were vociferous in their opposition to privatisation, as evidenced in the remarkable swing from positive to negative correlations with WTP as payment vehicles switch from those of a public to a private good. If normative and expressive values do underpin these differences then, as Sugden (forthcoming) argues, CV estimates must be considered as being context specific rather than as absolute valuations of the assets concerned.

A contrary and more critical explanation of the observed ordering effect again follows Kahneman et al. (1982) in arguing that relatively small prior per-visit WTP responses have here downwardly anchored subsequent per annum bids. In the context of our particular experiment, with one WTP response directly preceding another, such an effect is similar to the widely observed phenomena of starting point bias (Boyle et al, 1985). However, the remarkable and highly significant reversal in WTP correlation signs for members of green groups described above, makes us feel that the former public/private goods argument cannot be ignored here. Nevertheless, this does not preclude the possibility that the observed ordering effect has been heightened by anchoring/starting point bias, with consequent questions being raised regarding the validity of such results.

Each of the theoretical expectation, public/private goods and anchoring arguments can also be applied as explanations of the observed ordering effects in per visit WTP responses (i.e. testing equation (2b)). Here the line of effect is reversed in that the introduction of a prior per annum WTP question raises per visit WTP (although again only one of these effects is statistically significant). This could be following theoretical expectations (i.e. satisfying the inequality allowed in equation (2b)), perhaps enhanced by the per annum approach inducing respondents to think of this as being a public as opposed to private good, or it could be that the relatively high prior per annum response upwardly anchors subsequent per visit responses, or a mixture of both.

5. Conclusions

This analysis has applied a split-sample approach to the investigation of budget constraint, temporal and ordering effects in CV studies. In three out of four tests significant budget constraint effects were detected. Interpretation of such results is not straightforward as it may either be viewed as the expected consequence of respondents revising bids in the light of further reflection, or as evidence of an anchoring bias. While both explanations may have some validity their implications for future studies are in direct conflict. If budget constraint questions induce respondents to consider more fully their personal circumstances then, following the recommendations of Arrow et al., (1993), some variant of these questions should be included prior to WTP questions. Conversely if the responses to budget constraint questions anchor subsequent WTP bids then this suggests that they should be avoided.

Turning to consider temporal effects our study has highlighted difficulties in distinguishing these from the implicit payment vehicle effects which changes in temporal extent often imply. However, accepting this potential coincidence, we find no evidence of temporal embedding in this experiment and note that results satisfy Carson and Mitchell's (1995) component sensitivity test.

Two of the four tests of ordering effects indicate that significant differences were observed. Again at least two explanations of these results can be posited. Following Carson *et al.* (1992), economic theory allows for divergence between measures of the same good elicited at different point in a valuation sequence. Such differences are likely to be exacerbated if the sequence itself induced differing subsamples to view the resource under evaluation as either a public or a private good. Following such an explanation, the divergence in valuations can either be seen as reflecting the differing attributes of such goods, or as arising from a consequent change in the motivations underlying the preferences expressed. However, as with our budget constraint experiment, these divergencies can also be interpreted as evidence of prior responses anchoring subsequent bids.

For a view on the implications of these findings for the wider use of CV within cost-benefit and other decision-support frameworks, we presented our findings to two eminent Professors in this field. Reactions differed markedly; while one felt that the observed variation in

valuations across subsamples invalidated the CV method, the other argued that this gave useful limits upon the extent of such variation. Our own view is that the CV method provides many useful insights into the complex motivations underpinning individuals preferences for environmental goods and that the results presented here are not demonstrably inconsistent with theory. However, in order for the method to provide reliable valuations of such goods we must use the information elicited to develop models of economic behaviour which capture and reflect that complexity of motivation and context. While there have been theoretical advances in this area, empirical models have lagged behind. In particular, we need to go beyond simple models of the rational (selfish), economic individual which typifies many CV applications, to recognise the diversity of motivations underpinning valuations, the context of the contingent market and the psychological pressures of the interview situation. Only studies based upon such well-informed and sophisticated models of behaviour will be capable of deconstructing responses and revealing true valuations.

References

- Arrow, K.J. (1982) Risk perception in psychology and economics, *Economic Inquiry*, 20:1-9.
- Arrow, K.J., Solow, R., Portney, P.R., Leamer, E.E., Radner, R. and Schuman, E.H. (1993) Report of the NOAA Panel on Contingent Valuation, *Federal Register*, 58(10):4602-14
- Bateman, I.J. and Langford, I.H. (1995a) Design effects in contingent valuation studies, *CSERGE Global Environmental Change Working Paper 95-02*, Centre for Social and Economic Research on the Global Environment, University of East Anglia, Norwich and University College London.
- Bateman, I.J., Brainard, J.S. and Lovett, A.A. (1995b) Modelling woodland recreation demand using geographical information systems: A benefit transfer study, *CSERGE Global Environmental Change Working Paper 95-06*, Centre for Social and Economic Research on the Global Environment, University of East Anglia and University College London.
- Bateman, I.J., Langford, I.H. and Graham, A.W.K. (1995c) A survey of non-users willingness to pay to prevent saline flooding in the Norfolk Broads, *CSERGE Global Environmental Change Working Paper 95-11*, Centre for Social and Economic Research on the Global Environment, University of East Anglia and University College London.
- 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
- Bishop, K.D. (1992) Assessing the benefits of community forests: an evaluation of the recreational use benefits of two urban fringe woodlands, *Journal of Environmental Planning and Management*, 35(1): 63-76.
- Boyle, K.J., Bishop, R.C. and Welsh, M.P. (1985) Starting point bias in contingent valuation bidding games, *Land Economics*, 61(2): 188-194
- Boyle, K.J., Reiling, S.D. and Phillips, M.L. (1990) Species substitution and question sequencing in contingent-valuation surveys evaluating the hunting of several species of wildlife, *Leisure Science*, 12:103-118
- Boyle, K.J., Welsh, M.P. and Bishop, R.C. (1993) The role of question order and respondent experience in contingent-valuation studies, *Journal of Environmental Economics and Management*, 25:S-80-99.
- Brennan, G. and Buchanan, J.M. (1984) Voter choice, *American Behavioural Scientist*, 28:185-201
- Brookshire, D.S., d'Arge, R.C. Schulze, W.D. and Thayer, M.A. (1981) Experiments in valuing public goods, in Smith, V.K. (ed.) *Advances in Applied Economics: Volume 1*, JAI Press, Greenwich, CT.

- Burness, H.S., Cummings, R.G., Mehr, A.F. and Walbert, M.S. (1983) Valuing policies which reduce environmental risk, *Natural Resources Journal*, 23(3): 675-682.
- Carson, R.T. and Mitchell, R.C. (1995) Sequencing and nesting in contingent valuation surveys, *Journal of Environmental Economics and Management*, 28:155-173.
- Carson, R.T., Hanemann, W.M., Kopp, R.J., Krosnick, J.A., Mitchell, R.C., Presser, S., Ruud, P.A. and Smith, V.K. (1995) Temporal reliability of estimates from contingent valuation, *Discussion Paper 95-37*, Resources for the Future, Washington, D.C.
- Carson, R.T. and Mitchell, R.C. (1993) The value of clean water: The Public's willingness to pay for boatable, fishable and swimmable quality water, *Water Resources Research*, 29:2445-2454.
- Carson, R.T., Flores, N.E. and Hanemann, W.M. (1992) On the creation and destruction of public goods: The matter of sequencing, paper presented at the *Third Annual Conference of the European Association of Environmental and Resource Economists*, Cracow, Poland.
- Cummings, R.G., Brookshire, D.S. and Schulze, W.D. (1986) *Valuing Environmental Goods: An Assessment of the Contingent-Valuation Method*, Rowman and Littlefield, Savage, MD.
- Deaton, A. and Muellbauer, J. (1980) *Economics and Consumer Behaviour*, Cambridge University Press, Cambridge.
- Desvousges, W.H., Johnson, F.R., Dunford, R.W., Boyle, K.J., Hudson, S.P. and Wilson, K.N. (1993) Measuring natural resource damages with contingent valuation: Tests of validity and reliability, in Hausman, J.A. (ed.) *Contingent Valuation: A Critical Assessment*, North-Holland, Amsterdam.
- Diamond, P.A., Hausman, J.A., Leonard, G.K. and Denning, M.A. (1993) Does contingent valuation measure preferences?: Experimental evidence, in Hausman, J.A. (ed.) *Contingent Valuation: A Critical Assessment*, North-Holland, Amsterdam.
- Efron, B. and Tibshirani, R.J. (1993) *An Introduction to the Bootstrap*, Chapman and Hall, New York.
- Kahneman, D. and Knetsch, J.L. (1992) Valuing public goods: the purchase of moral satisfaction, *Journal of Environmental Economics and Management*, 22:57-70
- Kahneman, D., Slovic, P. and Tversky, A. (eds.) (1982) *Judgement Under Uncertainty: Heuristics and Biases*, Cambridge University Press, New York.
- Kahneman, D. and Tversky, A. (1984) Choices, values and frames, *American Psychologist*, 39:341-350
- Langford, I.H. (1994) Using a generalized linear mixed model to analyze dichotomous choice contingent valuation data, *Land Economics*, 70(4):507-514.

Loomis, J.B., Gonzalez-Caban, A. and Gregory, R. (1994) Do reminders of substitutes and budget constraints influence contingent valuation estimates?, *Land Economics*, 70(4): 499-506

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

Randall, A. and Hoehn, J.P. (1992) Embedding effects in contingent valuation: Implications for natural resource damage assessment, *Staff Paper 92-14*, Department of Agricultural Economics, Michigan State University.

Schkade, D.A. and Payne, P.W. (1994) How people respond to contingent valuation questions, *Journal of Environmental Economics and Management*, 26:88-109

Schulze, W.D., Cummings, R.G., Brookshire, D.S., Thayer, M.H., Whitworth, R.L. and Rahmatian, M. (1983) Experimental approaches to valuing environmental commodities: Vol II, draft final report for *Methods Development in Measuring Benefits of Environmental Improvements*, USEPA Grant # CR808-893-01.

Smith, V.K. (1992) Arbitrary values, good causes and premature verdicts, *Journal of Environmental Economics and Management*, 22:71-89

Sugden, R. (forthcoming) Public goods and contingent valuation, in Bateman, I.J. and Willis, K.G. (eds.) *Contingent Valuation of Environmental Preferences: Assessing Theory and Practice in the US, EU and Developing Countries*, Oxford University Press.

Tversky, A., Sattath, S. and Slovic, P. (1988) Contingent weighting in judgement and choice, *Psychological Review*, 95:371-384

Willis, K.G. and Garrod, G.D. (1993) Valuing landscape: a contingent valuation approach, *Journal of Environmental Management*, 37:1-22.

Endnotes

1. For a recent and relevant example (to which we return subsequently) see Boyle et al., (1993).
2. This would not be true for those who visited only once per annum. However, mean frequency exceeded 14 visits annually.
3. Other elicitation methods are considered in Langford (1994) and Bateman et al., (1995a). While open-ended approaches seem more prone to respondent uncertainty, this will be constant across the present experiment.
4. This addresses the 'conservative design' principles advocated by Arrow et al., (1993). As before this type of question was pretested in Bateman et al., (1995a).
5. Reasons for refusing to pay were mainly related to economic factors and are analysed further in Bateman and Langford (1995). At most only 2% of the sample gave refusal reasons which can be interpreted as in some way protesting against the valuation process.
6. The inclusion of such zero's reinforces the need to conduct non-parametric testing.
7. BCa (although not normal distribution) confidence intervals just fail to overlap, although we have to go to three decimal points (i.e. tenths of a penny) to confirm this result and, of course, statistical significance is to an extent a product of sample size as well as measured effect.
8. If such a pattern were to hold generally this would imply that the bulk of the CV literature, which has not incorporated explicit budget constraint questions, have produced conservative estimates of WTP.
9. See previous endnote regarding elicitation effects.
10. Factors such as discounting, uncertainty and risk mean that we would not expect a simple relationship between per visit and per annum WTP.

Acknowledgements

The Centre for Social and Economic Research on the Global Environment (CSERGE) is a designated research centre of the UK Economic and Social Research Council (ESRC).

This research was funded by ESRC Grant Number L320223014, CSERGE, The School of Environmental Sciences and the University of East Anglia (UEA) Research Promotion Fund.

This paper has benefitted greatly from comments by Kevin Boyle (University of Maine), John Loomis (Colorado State University), Tim O'Riordan (UEA), Robert Sugden (UEA) and two anonymous referees for which the authors are grateful. The usual disclaimer applies.

Table 1: Mean WTP (tax) per annum (£) and 95% confidence intervals for each subsample (including payment principle refusals as zeros)

| Budget question | Payment ordering scenario | |
|-------------------|---|---|
| | 1 (tax then fee) | 2 (fee then tax) |
| NB (not asked) | 12.55 (8.94 - 18.47) [8.11 - 16.99] | 7.62 (4.36 - 15.77) [2.87 - 12.37] |
| B (asked) | 32.60 (23.18 - 45.89) [21.76 - 43.43] | 16.37 (11.78 - 22.12) [11.19 - 21.55] |

Table 2: Mean per visit (fee) WTP (£) and 95% confidence intervals for each subsample (including payment principle refusals as zeros)

| Budget question | Payment ordering scenario | |
|-------------------|--|--|
| | 1 (tax then fee) | 2 (fee then tax) |
| NB (not asked) | 0.45 (0.35 - 0.57) [0.35 - 0.55] | 0.20 (0.12 - 0.32) [0.11 - 0.29] |
| B (asked) | 0.78 (0.57 - 1.09) [0.53 - 1.03] | 0.46 (0.33 - 0.66) [0.30 - 0.62] |