Australian Journal of Agricultural Economics, Vol. 32, Nos 2 and 3 (Aug/Dec 1988), pp. 98-112

EMPIRICAL TESTS OF HYPOTHETICAL BIAS IN CONSUMERS' SURPLUS SURVEYS

J. A. SINDEN*
University of New England, Armidale, NSW 2351

Questionnaire surveys are an established economic research method for eliciting consumers' surplus values. However, the necessarily hypothetical nature of surveys may not promote truthful responses. In 17 empirical tests, actual money donations were elicited and compared to hypothetical donations elicited previously. In all 17, there was no statistical difference between the actual and hypothetical willingness to pay. Hypothetical bias was therefore absent in all these tests.

Questionnaires are an established and important research method, despite their necessarily hypothetical nature. They have become particularly important in the assessment of preferences and estimation of benefit values for unpriced goods and services such as wildlife, clean air, and even the continued existence of soil resources. Nevertheless, many practical and theoretical issues surround their use.

Practical issues include question construction, question sequence, terminology, and of course the bias inherent in almost any wording. Statistical issues concern the choice of sample, sampling precision and sample size. Theoretical problems concern selection of a consumer's surplus concept, assessment for public goods, and the marginality and reversibility of changes. Progress has been made on most of these issues but much current interest concerns the validity of responses to hypothetical questions.

Following Cummings, Brookshire and Schulze (1986), hypothetical bias is defined as the difference between an hypothetical payment and a comparable actual payment. Differences can be caused by strategic behaviour, differences in perception of goods and circumstances, differences between hypothetical choices and actual choices, and differences in timing of the choices. Individuals are said to behave strategically when they respond untruthfully to hypothetical questions in an attempt to influence an outcome. Rowe and Chestnut (1983) argue that hypothetical bias is an important problem because surveys offer obvious opportunities for free-riding. Langford and Cocheba (1978) also argued that the hypothetical nature of questions is a major concern, and discussed ways to minimise misunderstanding and bias. Russell (1982) believed that this source of bias is a more important difficulty than theoretical differences in surplus concepts.

The objectives of this paper are to explore whether hypothetical willingness to pay from a questionnaire survey does equal actual willingness to pay, and whether hypothetical bias seems more prevalent in particular conditions. The vehicle for this exploration is a sequence of controlled

* The financial assistance of the National Soil Conservation Programme in funding the experiments is gratefully acknowledged, as is the cooperation of the Soil Conservation Service of New South Wales in helping with some of them. Professor Jack L. Knetsch, Dr J. W. Bennett, the Department of Economics Seminar at the University of Wollongong and participants at the 1987 Conference of the Australian Agricultural Economics Society all provided many helpful comments on earlier drafts. Two anonymous referees provided constructive suggestions on the present paper. The opinions and errors remain, of course, the author's responsibility.

experiments, and the service to be purchased is the knowledge that an increased quantity of soil is conserved.

Within this setting, three particular conditions are considered, namely:

(a) when specific 'advertising' information is supplied to the subjects;
(b) when the source of the actual money differs (from subject's own pocket, a free gift or government funds);
(c) when there are time lags between the hypothetical questionnaire and actual payment.

Ways to discourage hypothetical bias, and previous tests for it, are considered in the literature review which follows. The method is then defined and a sequence of four experiments is presented. The findings of the individual experiments are brought together in the results section. The final discussion examines possible reasons for the absence of hypothetical bias in the 17 tests contained in the experiments.

Literature Review

Before turning to previous empirical tests, the discouragement of hypothetical bias and ways to promote truthful responses are reviewed.

Discouraging bias

Despite careful wording, maximum realism and surveys of actual consumers, it is impossible to be sure that questionnaire responses provide conclusive estimates of true values (Knetsch and Davis 1966). Not surprisingly, therefore, much thought has gone into ways of promoting truthful responses and of checking accuracy. Bias should be removed from the final calculations when the type of person who acts strategically is identified and excluded (Bennett 1983). More truthful replies should follow when (a) the incentive to play games is removed (Bennett 1983; Schulze, d'Arge and Brookshire 1981), (b) iterative bids are used, (c) incentives for valid responses are provided, as for example in the Smith auction process (Smith 1979), and (d) the limits to relevant circumstances are observed. These limits are called the reference operating conditions (Cumming *et al.* 1986).

The accuracy of bids can be checked partially by comparing the hypothetical bids with (a) some standard (Brookshire, Ives and Schulze, 1976), (b) bids elicited by some other method, (c) payments of a related nature (Bennett 1982), and (d) stated preferences and attitudes (Bennett, 1984). Furthermore, an explanation for an unusual bid would lend credence to the whole set of bids.

Despite this range of procedures to discourage hypothetical bias and to check for its presence, the central problem remains, that is hypothetical choices are never truly comparable to actual choices. The existence and importance of hypothetical bias can only be assessed through actual tests which involve both kinds of choice.

Empirical tests for bias

Empirical tests for bias have compared hypothetical values from a questionnaire survey with either values from a second method or actual cash payments under the same conditions. The reported tests are now reviewed in these two categories.

The first kind of test approximates a true test of hypothetical bias where the second method is based on observed market transactions of some kind, and where the same persons are involved in the same context. This kind

of test assumes that the values from the second method are valid measures of actual willingness to pay, a very strong assumption. Cummings *et al.* (1986) report 15 such tests in their detailed assessment of the direct-question survey methods. Their second, or check, methods were the travel-cost technique and hedonic pricing. In 13 of the 15 tests, the survey and second-method values differed by 50 per cent or less. In the remaining two, the differences were 52 per cent and 60 per cent. The authors concluded that a 50 to 60 per cent divergence demonstrated 'a lack of significant difference' (p. 100) between actual and hypothetical payments, a conclusion that may not be shared by potential users of the value information.

Instead of comparing hypothetical bids with values from another method, bids could be tested against a relevant standard and Bennett (1987) used a normal distribution as his predetermined standard. Hypothetical bids with direct willingness to pay questions proved overstatements relative to this standard, whereas bids from a Smith auction process proved to be understatements.

Three studies have investigated hypothetical bias through the second category of test, by comparing hypothetical bids to actual cash payments under the same conditions. Bishop and Heberlein (1979) attempted to value the benefits of a 1978 early-season goose-hunting permit in Wisconsin. Hunters with such permits were divided into two groups. One group received cash offers, complete with cheques, for their permits and their actual willingness to sell proved to be $63 per permit, on average. The second group received questionnaires and their hypothetical willingness to sell was $101 on average, a value 60 per cent higher than the actual willingness to sell of the comparable first group.

Bohm (1972) divided 605 people into five groups and asked each person the largest sum he or she was willing to pay to watch a specific television programme. The members of one group knew they had to pay nothing; hence, their bids were the hypothetical bids. The other four groups knew they actually had to pay the amounts bid, or some proportion of that amount. There proved to be no statistically significant differences between the means of any pair of groups, indicating a lack of hypothetical bias.

A sequence of three experiments using lotteries was reported by Sinden (1987). In each experiment, students were divided randomly into two equal groups. One group could enter a real lottery while the other could enter a hypothetical, but otherwise identical, lottery. In two of these experiments, there was no statistical difference between the mean hypothetical willingness to pay and mean actual willingness to pay. In the third experiment, mean hypothetical payments were less than mean actual payments. Hypothetical bids should exceed actual payments if respondents free ride and have no expectations of actually having to pay. Thus, this third result does not necessarily indicate the existence of hypothetical bias.

The 15 tests reported in Cummings *et al.* (1986), and that of Bishop and Heberlein (1979), give the overall impression that hypothetical bias of the order of $\pm$ 50 per cent does exist. The five experiments reported by Bennett (1987), Bohm (1972) and Sinden (1987) allow the more hopeful interpretation that bias can be absent, or where it exists hypothetical bids can underestimate true willingness to pay. The present research offers a further 17 tests from a series of four experiments.

Method

In actual decisions, a rational utility-maximising consumer will estimate values for the purpose of comparing alternatives and making choices. The natural test for bias is therefore to determine hypothetical willingness to pay, to observe actual willingness to pay in the same circumstances, and then to compare the two. These steps can rarely be followed so neatly in real life; therefore, they are now modelled in structured experiments with student groups.

The different conditions (changes in information, different sources of money and time lags) are each considered in this comparison, but each requires a slightly different setting of the hypothetical questions. Despite different experiments for each setting, the possibility of comparisons between experiments is maintained through a single questionnaire. The null hypothesis for all tests in all experiments is that no hypothetical bias exists or, more formally, that hypothetical willingness to pay equals actual willingness to pay. Referring to this very hypothesis, Cummings *et al.* (1986, p. 51) state, 'the literature abounds with evidence that suggests that (it) be rejected'.

Concepts and context

In their day-to-day expenditures, individuals compare the worth of a potential purchase to the worth of alternatives. They then allocate their limited budget to acquire the bundles of goods that will give them greatest satisfaction. Economic theorists have much to say about this allocation process. However, all adhere to the notion that relative willingness to pay for a good is an appropriate measure of the expected satisfaction or benefit from that good, relative to other goods, and given the individuals' incomes.

The lengthy literature on this allocation process is matched by the lively current developments in the theory, notably attempts to combine behavioural science with microeconomics. For example, Thaler (1985) has suggested new ways to value the utility of gains and losses, to include the utility of the transaction itself, and to better incorporate the household budgeting process.

Soil conservation provides direct and indirect benefits. In addition to any direct increases in agricultural production and any reductions in off-farm costs, it provides at least three kinds of indirect benefit. *Existence value* is the satisfaction that people themselves derive from the knowledge that the soil is preserved and so still exists. *Bequest value* is the satisfaction to today's people from protection of the soil resource for use by future generations. *Option value* is the benefit to today's people from reserving the option to use the soil resource for food production in the future. The term 'preservation value' can conveniently cover all of these three kinds of benefit.

In principle, these benefits can be valued through the notion of willingness to pay. The most simple, direct question to estimate this preservation value is therefore: what is the maximum you are willing to pay to preserve the soil? With appropriate qualifications on marginality of changes and incidence of benefits and costs, answers should capture all three kinds of preservation value. Several of the major qualifications are covered below in the question construction and wording. The issue of will-

ingness to pay or willingness to sell (Knetsch and Sinden 1984) is 'covered' by assuming the reference point that the community has no prior entitlement to a conserved soil resource. Thus, the community must pay for further conservation rather than be compensated for further erosion.

The benefits from consumption derive from the characteristics of a good or service, and those characteristics can include aspects of the production process according to Lancaster's (1966) theory of consumer choice. The present service is no exception because the benefits from existence, bequest and option derive jointly from characteristics of the soil, and the processes of soil use and soil conservation. Indeed, the information provided in experiments 1 and 4 emphasised both benefits and processes. The inseparability of product and process, and the nature of the particular product and processes, will be constant across each pair of hypothetical and actual bids in each of the tests for bias. Any problems that arise from these features of the consumption activity should therefore be minimal.

The hypothetical questionnaire

Choices are guided by expected benefits and so provide a useful way of assessing both general interest and possible willingness to pay for soil conservation. Soil conservation concerns management of a natural resource involving substantial government agencies, as do forestry and wildlife conservation. Thus, question 1, to determine the level of concern for soil conservation relative to similar natural resources, was framed as a choice between more government investment to stop soil erosion, preserve rainforest or preserve rare kangaroo species. Some households may prefer to support none of these but get the money back as, say, a tax rebate. A fourth choice was therefore offered in question 1, namely receiving the equivalent sum back in tax.

The New South Wales government spent $17m in 1984 and $15m in 1985 to buy private forest as a new source of jobs and production because rainforests elsewhere had been preserved. The preamble to question 1, covering the financial magnitudes, used the sum of $17m which is equivalent to $15 per household. The order of the four choices in question 1 was randomised and the wording is shown in the Appendix.

The chances of reliable answers to willingness to pay questions are improved if the question relates to something familiar to the interviewee. Clearly, the questions must also relate to the problem at hand. New South Wales has extensive wheat-producing areas; wheat is made into bread, and wheat production can often lead to serious soil erosion. A second question was therefore related to a loaf of bread which conveniently cost about $1 at the time of the surveys. The four sub-questions of question 2 provided a simple iterative procedure to elicit the maximum extra willingness to pay to help control erosion.

Despite its relevance to the problem, question 2 implies a general levy on the price of a consumer good. Question 3 uses the contrasting context of a voluntary donation to a fund(s) to assist resource management. Question 3 presumes a fixed sum ($5) in the pocket as the budget from which the donation is made. This constraint is lifted in questions 4, 5 and 6 which together elicit the willingness to pay to such a fund from whatever sum happens to be in the pocket at the time of the survey.

Analysis

The hypothesis, that hypothetical willingness to pay equals actual willingness to pay, was tested by three statistical comparisons of the distributions of bids. The means were compared with the conventional *t*-test. The sequence of four experiments involved 17 sets of bids. In all 17, every observation was within two standard deviations of its respective mean. In 11 sets, two-thirds of all observations were within one standard deviation of their respective means. Thus, while 11 sets closely approximate the normal distribution, six do not do so quite as closely.

The Wilcoxon matched-pairs signed-rank test, a non-parametric alternative to the *t*-test, was used to determine the significance of differences in the distributions. Following a repeated-measurement format, the matched pairs were the hypothetical and actual bids at each monetary value, and the numbers of each kind of bid at each value were the data for the test. The correspondence of the hypothetical distribution of bids to the actual distribution was also examined with a chi-squared, goodness-of-fit test.

The arrangements for each of the four experiments enabled the value of individual bids to be recorded, but to ensure confidentiality, individuals within each group were not identified. The statistical tests were therefore performed on appropriate means and the distributions of group bids.

The three tests imply that actual payments are true reflections of surplus, although there is no guarantee of this. All bids are anonymous so that actual bids should approximate true surplus, *ceteris paribus*. In any event, actual bids would seem a close, and certainly the closest possible, measure of true surplus.

Sequence of Experiments

The experiments were defined and completed in sequence. Each experiment elicits hypothetical and actual willingness to pay, in the context of one or more of the survey questions. The questionnaire for each experiment was a subset of the six questions of the Appendix.

Each experiment in the sequence is now described. The individual results are noted for each, and summarised in Tables 1 to 3. The findings as a whole are reviewed in the results and discussion sections.

Experiment 1

The first experiment attempted to test whether hypothetical willingness to pay equalled actual willingness to pay and whether specific information on soil conservation changed valuations. To test for the latter, responses were required before and after the information was available and so the subjects were surveyed twice. The experiment was completed in June 1985 with 51 students and was carried out as follows.

(a) At a Monday lecture, students were asked to complete a questionnaire containing questions 1, 2 and 3. (See Appendix.)

(b) The next Wednesday lecture began with a short, forceful, talk on the benefits of soil conservation by Dr S. J. Perrens, an authority on soil resources.

(c) The same questionnaire was completed and collected.

(d) Students were each paid $5 as wages for their participation. They were then offered the opportunity to make voluntary, confidential dona-

tions to funds to assist soil conservation and/or control of eucalypt dieback.

(e) The following statistical comparisons were made:
  (i) relative preference (question 1) for soil conservation before and after the information;
  (ii) willingness to pay extra per loaf (question 2) before and after the information;
  (iii) hypothetical willingness to pay from a hypothetical $5 (question 3) before and after the information;
  (iv) actual donation on day 3 versus hypothetical donation (from the comparable question 3) on day 3, for both soil conservation and eucalypt dieback;
  (v) actual donation on day 3 versus hypothetical donation (from question 3) on day 1, for both soil conservation and eucalypt dieback.

TABLE 1

Effect of Information on Preferences for and the Comparative Value of Soil Conservation[a]

	Experiment 1		Experiment 4	
Question 1 Preferences for extra government expenditure	Before information (%)	After	Before information (%)	After
(a) to preserve more rainforest	51	26	17	8
(b) to get $15 more back in tax	5	4	10	3
(c) to stop more soil erosion	39	68	68	86
(d) to preserve more kangaroos	5	2	5	3
Total per cent	100	100	100	100
Question 2 Willingness to pay extra cents per $1 loaf	24.1[b]	24.1[b]	14.6[c]	16.5[c]
Question 3 Hypothetical willingness to pay to a fund to assist soil conservation from a hypothetical $5 in pocket	$1.5[d]	$2.1[d]	$1.0[e]	$1.3[e]
Question 5a and 6[f] Hypothetical willingness to pay to a fund to assist soil conservation from whatever sum happens to be in the pocket			$1.6[g]	$2.1[g]

[a] There were 51 students in experiment 1 and 75 in experiment 4.
[b] The calculated t-statistic of 0.01 indicates that there is no significant difference, at the 1 per cent level, between these means.
[c] As [b], but with a calculated t-statistic of 0.1.
[d] As [b], but with a calculated t-statistic of 0.3.
[e] As [b], but with a calculated t-statistic of 0.2.
[f] This question was not asked in experiment 1.
[g] As [b], but with a calculated t-statistic of 0.1.

Comparisons (i) to (iii) concern information effects and the relevant results are shown in Table 1. Comparisons (iv) and (v) concern hypothetical bias and their results are shown in Table 2.

The results given in Table 1 show that the information greatly shifted the preference for expenditure of government monies toward soil conservation. However, there is no statistical difference between the answers to questions 2 and 3 and so the information did not shift willingness to pay from the students own (hypothetical) monies. The tests for hypothetical bias (Table 2) show that there was no statistical difference between the means or distributions of the hypothetical and actual donations for either resource management service in either comparison [comparisons (iv) and (v) in (e) above, and comparisons 1 and 2 in Table 2].

These results suggest a validity for hypothetical responses, and an information stability for personal hypothetical donations as sources of willingness to pay data. Both these suggestions warrant confirmation; thus, the lack of hypothetical bias was checked through experiment 2, and the

TABLE 2

Comparison of Actual Willingness to Pay (WTP) from $5 Wages Versus Hypothetical WTP

Comparison [a]		Test[b]		
		t	Wilcoxon	chi-squared
Experiment 1				
1. Actual WTP (day 3)	: hypothetical WTP (day 3)			
Soil conservation	(1.86 : 2.06)	0.1(2.4)	16(2)	3.5(7.8)
Eucalypt dieback	(1.12 : 1.06)	0.1(2.4)	8(0)	4.7(9.5)
2. Actual WTP (day 3)	: hypothetical WTP (day 1)			
Soil conservation	(1.86 : 1.43)	0.3(2.4)	22(6)	10.3(11.1)
Eucalypt dieback	(1.12 : 1.22)	0.1(2.4)	8(0)	2.5(11.1)
Experiment 2				
. Actual WTP (day 1)	: hypothetical WTP (day 1)			
Soil conservation	(2.76 : 2.40)	0.1(2.4)	21(6)	5.5(9.5)
Experiment 4				
1. Actual WTP (week 7)	: hypothetical WTP (week 7)			
Soil conservation	(1.28 : 1.10)	0.2(2.4)	22(4)	6.6(11.1)
Eucalypt dieback	(0.92 : 0.70)	0.3(2.4)	6(6)	8.6(9.5)
2. Actual WTP (week 7)	: hypothetical WTP (week 6)			
Soil conservation	(1.28 : 1.36)	0.3(2.4)	18(8)	5.5(11.1)
Eucalypt dieback	(0.92 : 0.84)	0.1(2.4)	10(0)	4.9(11.1)
3. Actual WTP (week 7)	: hypothetical WTP (week 4)			
Soil conservation	(1.28 : 1.30)	0.0(2.4)	26(8)	7.9(11.1)
Eucalypt dieback	(0.92 : 0.79)	0.1(2.4)	12(2)	16.2(11.1)
4. Actual WTP (week 7)	: hypothetical WTP (week 1)			
Soil conservation	(1.28 : 1.01)	0.2(2.4)	16(4)	20.3(11.1)
Eucalypt dieback	(0.92 : 1.27)	0.3(2.4)	13(6)	16.2(11.1)

[a] Actual dollar bids are in parentheses.

[b] The *t*-test is at a 1 per cent level of significance, whereas the other two tests are at the 5 per cent level. Calculated values are listed first, with test levels in parentheses. To retain the null hypothesis (of no difference between hypotheticals and actuals), calculated *t* and chi-squared values must be less than test values whereas calculated Wilcoxon T statistics must exceed test values.

information effect through experiment 4. The students' actual donations came from wages they had just been paid. Thus, the lack of bias must still be checked when their own money is involved, and experiments 2 and 4 attempt to do this.

Experiment 2

The objective of this experiment was to retest for hypothetical bias using question 3, with its donations to soil conservation. This time the subjects were a class of 40 external students, all mature persons taking courses by correspondence but visiting the University for a residential school.

The experiment was completed in August 1985 and proceeded as follows.

(a) The class was asked to complete a questionnaire with questions 1, 2 and 3. The donation option for eucalypt dieback was omitted from question 3.
(b) Questionnaires were collected.
(c) Students were given $5 in cash, as wages for their participation.
(d) They were then given the opportunity to make a voluntary, confidential donation to a fund to assist soil conservation.
(e) As in experiment 1, they placed their donations in unmarked envelopes, to enable individual donations to be recorded but to maintain confidentiality.
(f) The hypothetical payments from question 3 were compared to the actual payments.

Once more, the mean hypothetical donation did not significantly differ from the mean actual donation, and the hypothetical distribution did not differ from the actual distribution (Table 2). Both experiments 1 and 2 were eliciting actual donations from the $5 given to the subjects immediately beforehand; thus, an inevitable question is: would subjects respond in the same way if they made the donation from their own money? This question was explicitly addressed in experiment 3.

Experiment 3

This experiment was the first test for hypothetical bias where the subjects were asked to donate their own money. An internal class of 13 students was used for this purpose.

The experiment was held in October 1985 and proceeded as follows.

(a) The class was asked to complete a questionnaire with questions 1, 2, 4, 5 and 6.
(b) Questionnaires were collected.
(c) Students were then given the opportunity to make a voluntary, confidential donation to soil conservation from their own pockets.
(d) A facility was provided whereby a colleague (a president of a student association) could lend cash if needed.
(e) As they filed from the lecture theatre, students placed their individual donations in unmarked envelopes.
(f) The hypothetical donations (from questions 5a and 6) were compared with the actual donations.

The results given in Table 3 indicate that there is no significant difference between hypothetical and actual mean bids, or between the hypothetical and actual distributions. This result, with actual money, confirms the

TABLE 3

Comparison of Actual Willingness to Pay (WTP) from Own Pocket Versus Hypothetical WTP for Soil Conservation

Comparison[a]	Test[b]		
	t	Wilcoxon	chi-squared
Experiment 3			
Actual WTP (day 1) : hypothetical WTP (day 1) (1.87 : 2.38)	0.2(2.7)	10(0)	4.3(7.8)
Experiment 4			
1. Actual WTP (week 6) : hypothetical WTP (week 6) (1.40 : 1.60)	0.1(2.4)	17(6)	6.3(11.1)
2. Actual WTP (week 6) : hypothetical WTP (week 4) (1.40 : 2.10)	0.2(2.4)	22(6)	6.6(11.1)
3. Actual WTP (week 6) : hypothetical WTP (week 1) (1.40 : 1.60)	0.1(2.4)	13(2)	28.5(11.1)

[a] Actual dollar bids are in parentheses.
[b] The *t*-test is at a 1 per cent level of significance, whereas the other two tests are at a 5 per cent level. Calculated values are listed first, with test levels in parentheses. To retain the null hypothesis (of no difference between hypotheticals and actuals), calculated *t* and chi-squared values must be less than test values whereas the calculated Wilcoxon T statistics must exceed test values.

previous result using the $5 wages. Nevertheless, further tests of this apparent lack of hypothetical bias, using students' own money, seemed warranted and experiment 4 provided these.

Experiment 4

To confirm and expand the results so far, the objectives of experiment 4 were to test whether actual donations from the subjects' own monies differed from hypothetical donations, whether actual donations from $5 wages differed from hypothetical donations, and whether information on soil conservation shifts actual and hypothetical donations.

The arrangements for the experiment also allowed a test for changes in willingness to pay over time. The 75 students of the particular class were surveyed four times between Thursday 6 March and Thursday 17 April, 1986. The experiment was carried out as follows.

(a) In week 1, the class was asked to complete a questionnaire with all six questions. No information was provided and no money changed hands.
(b) In week 4, Dr S. J. Perrens gave his forceful talk on the benefits of soil conservation. Immediately afterwards, the class completed the same questionnaire and no money changed hands.
(c) The third survey with the same questionnaire occurred in week 6. After the questionnaires had been completed and collected, the students were offered a voluntary, confidential opportunity to donate to funds to assist soil conservation and control of eucalypt dieback. The donations had to come from the students' own monies and there were facilities for students to borrow money (anonymously through class presidents) and to change money.

(d) In week 7, the students again completed the questionnaires which were collected immediately. Students were paid $5 wages, for participating, and then offered the opportunity to donate to funds to assist soil conservation and control of eucalypt dieback. Once more, donations were confidential and voluntary.

(e) The following statistical comparisons were made concerning the effects of information:

(i) relative preferences (question 1) for soil conservation before and after the information;
(ii) willingness to pay per loaf (question 2) before and after the information;
(iii) hypothetical willingness to pay to a fund to assist soil conservation from a hypothetical $5 (question 3) before and after the information;
(iv) hypothetical willingness to pay to a fund to assist soil conservation from whatever monies happen to be in the pocket (questions 5 and 6) before and after the information.

The results for these tests are shown in Table 1.

(f) The following statistical comparisons tested for hypothetical bias:

(v) actual donations from $5 wages in week 7 versus the corresponding hypothetical donations (question 3) in weeks 1, 4, 6 and 7, for donations to both soil conservation and control of eucalypt dieback;
(vi) actual donations from own monies in week 6 versus hypothetical donations (questions 5a and 6) in weeks 1, 4 and 6 to a fund to assist soil conservation and/or to a fund to help control eucalypt dieback.

The results for these tests are shown in Tables 2 and 3.

The effect of the information was to shift preferences toward soil conservation when preferences concerned spending government funds (question 1). However, the willingness to pay values from expenditure of the subjects' own monies (questions 2, 3 and 5a and 6) were not changed by the information. These results confirm those of experiment 1 and suggest that individuals are more responsive to information when they 'spend' government income rather than their own. The opportunity costs of using their own monies would usually be obvious to the individuals, and would always be felt directly. The opportunity costs of government use of their tax monies would often be unclear, and would usually be felt less directly. The difference in responsiveness could therefore be due to a widespread perception of differences in opportunity costs.

There were 11 tests for hypothetical bias in this experiment, eight reported in Table 2 and three in Table 3. In each, there was no statistically significant difference between the means of the hypothetical and actual willingness to pay bids. However, the chi-squared test indicated that the hypothetical distribution differed from the actual in four cases, a feature referred to in the next section.

Results

This sequence of four experiments has tested for the existence of hypothetical bias under conditions of changing information, different

opportunities for learning, contrasting sources for the money payments, and time lags between hypothetical and actual donations. In the sequence there were 17 specific tests for bias. The *t*-test and the Wilcoxon test both suggest that the hypothesis, that hypothetical willingness to pay equals actual willingness to pay, be retained in all 17 cases. Using the chi-squared test, the hypothesis would be retained in 13 of the 17 cases but could be rejected in the four cases with the longest time lags (experiment 4, comparisons 3 and 4 in Table 2 and comparison 3 in Table 3). But in these four cases all distributions of bids appeared to follow the normal distribution. The *t*-test result is therefore appropriate in these cases and so the hypothesis of no bias can be retained for them also.

The subjects in each experiment usually comprised three contrasting groups (agricultural economics, natural resources and overseas students) even though all met some other selection criterion. Payments could vary by subgroups due to different knowledge and attitudes, and more relevant perhaps, hypothetical bias could exist amongst such subgroups. Accordingly, the existence of bias was tested by subgroups, for the largest group where payments were made from the subjects' own funds (week 6 and experiment 4). While mean hypothetical bids exceeded mean actual payments, the differences were not statistically significant even at the 1 per cent level. Consistent with all the other results, bias could not be detected within these subgroups.

Discussion

The general lack of hypothetical bias warrants some explanation and four interpretations seem possible. First, the subjects may have remembered their hypothetical responses and then donated similar amounts to maintain their personal integrity or because they felt locked in to that response. In the three related experiments in Sinden (1987), subjects were split into separate groups and each group made either hypothetical or actual payments. This split-sample design avoids interdependence between hypothetical and actual donations, and no hypothetical bias could be detected in these experiments.

There was, of course, little likelihood that donations from any given experiment would influence the level of state investment in soil conservation. A second explanation for the lack of bias could therefore be the lack of incentive for strategic behaviour to influence government investment. Strategic behaviour is the untruthful response to hypothetical questions in an attempt to influence the outcome. The District Soil Conservationist attended all of the third survey session (week 6) and the fourth survey session (week 7) of experiment 4. The wording of question 3, and the verbal instructions that accompanied both sessions, indicated that both hypothetical and actual donations would assist investment in a local soil conservation project. There is no certain way of knowing whether hypothetical bias would be likely to influence any actual investment. However, the opportunity was available for high hypothetical bids to demonstrate to the District Soil Conservationist the need for more local action.

There is further evidence on the issue of strategic behaviour. Students seem a likely social group to indulge in game playing for the fun of it, or even in the hope of gaining better marks in the course. The hypothetical donations were of the order of $0.50 to $2 and so a hypothetical donation of zero is hardly much of a game. Game playing would therefore

require consistently high hypothetical donations relative to actual donations, and as the results of Tables 2 and 3 show, there is little evidence of this. There were several actual donations of about $10 throughout the experiments. Information passed informally to the researcher after the experiments did indicate that these all came from a group of students who had been doubtful of their ability to pass the course. On this basis, the explanation of lack of incentive for strategic behaviour seems hard to sustain.

A third possible explanation presumes that the subjects had already anticipated that the purpose of their experiment was to test for hypothetical bias. Then, for personal reasons, they gave answers to disprove the existence of bias. Experiment 1 was, of course, the first in the series, and so new to the students. Anticipation seems impossible in this experiment and in any case hypothetical bias could not be detected in its results.

All this would seem to leave a fourth explanation for the results. Hypothetical bias may not be inevitable, and indeed may not occur, when a carefully prepared questionnaire is presented carefully to a cooperative group of subjects.

Generalisation from any set of results is difficult. The present findings have been obtained where donations varied between $0 and $10, and were typically between $0.5 and $2. The same kind of result may not occur where payments are substantially larger.

The students who participated in the experiments should have been aware of concepts of willingness to pay, difficulties with surveys and questionnaires, and the possibilities for strategic behaviour. Most should also have been aware of the broader issues of soil conservation and eucalypt dieback. This prior knowledge, and the somewhat artificial environment of the experiments, might mitigate against generalisation of the conclusion that hypothetical bias is not inevitable.

The reasoning leading to the conclusion can also be applied to this question of generalisation. These groups of students did appear to have some incentive to behave strategically, but apparently did not do so. Their knowledge of practical and analytical issues does not appear to have influenced the results, either to induce bias or preclude it, and not least, students seem more likely to play various kinds of games with surveys than any other social group. Overall, despite the main situations for bias appearing to be present, bias was absent. Thus, hopefully the conclusion can be generalised to similar kinds of survey questions for other social groups, and so hypothetical bias may not occur when a carefully prepared questionnaire is presented carefully to a cooperative group of subjects.

APPENDIX

The Questionnaire

1. How would you like to see an extra $17m of government money spent in NSW in 1986? This is about $15 for every family in NSW. Would you prefer that the money be spent,
   (a) to preserve rainforest species?
   (b) to give you $15 back on your tax?
   (c) to prevent soil erosion?
   (d) to preserve rare kangaroo species?

2. The bread we eat is made from wheat, much of which is grown on eroding soils: suppose a loaf of bread produced from this land costs $1.00.
   (a) Are you willing to pay an extra 5¢ per loaf, if all of that extra 5¢ goes to control this erosion? Yes/No .....
   (b) Are you willing to pay an extra 10¢ per loaf, if all of that extra 10¢ goes to control this erosion? Yes/No .....
   (c) Are you willing to pay an extra 15¢ per loaf, if all of that extra 15¢ goes to control this erosion? Yes/No .....
   (d) What is the maximum you are willing to pay extra per loaf if all the extra goes to control soil erosion? .....¢

3. Suppose you had $5 in your pocket and were asked whether you wanted to donate to two funds to assist local resource management problems.
   How much would you: (please use units of 50 cents)
   (a) donate to a fund to help soil conservation? .....
   (b) donate to a fund to help eucalypt dieback control? .....
   (c) keep for your immediate personal expenses? .....
   TOTAL $5.00

4. How much cash do you have in your pocket now? .....
   If you do have cash in your pocket, go to question 5.
   If you do not have cash in your pocket, go to question 6.

5. Some cash in pocket. Suppose you were asked whether you wanted to donate now, to a fund to assist soil conservation in Australia. How much would you
   (a) donate to this fund? .....
   (b) keep for your immediate personal expenses? .....

6. No cash in pocket. Suppose you were asked whether you wanted to donate now, to a fund to assist soil conservation in Australia. And suppose there was a good friend present from whom you could borrow money - providing you pay it back the next day. How much would you donate to this fund to assist soil conservation? .....

References

Bennett, J. W. (1982), 'Valuing the existence of a natural ecosystem', *Search* 13(9/10), 232-5.
—— (1983), 'Validating revealed preferences', *Economic Analysis and Policy* 13(1), 2-17.
—— (1984), 'Using direct questioning to value the existence benefits of preserved natural areas', *Australian Journal of Agricultural Economics* 28(2/3), 136-52.
—— (1987), 'Strategic behaviour: some experimental evidence', *Journal of Public Economics* 32, 355-68.
Bishop, R. C. and Heberlein, T. A. (1979), 'Measuring values of extramarket goods: are indirect measures biased?', *American Journal of Agricultural Economics* 61(5), 926-30.
Bohm, P. (1972), 'Estimating demand for public goods: an experiment', *European Economic Review* 3(2), 111-30.
Brookshire, D. S., Ives, B. and Schulze, W. D. (1976), 'The valuation of aesthetic preferences', *Journal of Environmental Economics and Management* 3(4), 325-46.
Cummings, R. G., Brookshire, D. S. and Schulze, W. D. (1986), *Valuing Environmental Goods: An Assessment of the Contingent Valuation Method*, Rowman and Allenheld, Totowa, New Jersey.

Knetsch, J. L. and Davis, R. K. (1966), 'Comparison of methods for recreation evaluation' in A. V. Kneese and S. C. Smith (eds), *Water Research*, Johns Hopkins University Press, Baltimore, pp. 125–42, for Resources for the Future.

—— 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(3), 507–21.

Lancaster, K. J. (1966), 'A new approach to consumer theory', *Journal of Political Economy* 74(2), 132–57.

Langford, W. A. and Cocheba, D. J. (1978), *The Wildlife Valuation Problem: A Critical Review of Economic Approaches*, Canadian Wildlife Services Occasional Paper No. 37. Ministry of Supply and Services, Ottawa.

Rowe, R. D. and Chestnut, L. G. (1983), 'Valuing environmental commodities: revisited', *Land Economics* 59(4), 404–10.

Russell, C. S. (1982), 'Publicly vested values for fish and wildlife: comment', *Land Economics* 58(4), 559–62.

Schulze, W. D., d'Arge, R. and Brookshire, D. S. (1981), 'Valuing environmental commodities: some recent experiments', *Land Economics* 57(2), 151–72.

Sinden, J. A. (1987), Willingness to pay or willingness to pray for soil conservation: empirical tests of hypothetical bias in consumers' surplus surveys. Paper presented to the 31st Annual Conference of the Australian Agricultural Economics Society, Adelaide, 9–12 February.

Smith, V. L. (1979), 'An experimental comparison of three public good mechanisms', *Scandinavian Journal of Economics* 26, 33–60.

Thaler, R. (1985), 'Mental accounting and consumer choice', *Marketing Science* 4(3), 199–214.