

## ESTIMATING DEMAND FOR PUBLIC GOODS: AN EXPERIMENT

Peter BOHM \*

*University of Stockholm, Sweden*

Received April 1971

The purpose of this paper is to describe a test involving five different approaches to estimating the demand for a public good. The test was conducted in a setting which permitted a real collective choice and in which each subject was committed to actual payments when relevant. The results indicate that the well-known risk for misrepresentation of preferences in this context may have been exaggerated. The test would seem to encourage further work in the field of experimental economics.

The revealing of individual preferences for public goods would produce important guidelines for output decisions and distribution policy. But, as most of the literature on public goods tells us, the odds against achieving a practicable approach which avoids incentives to misrepresentation of preferences are overwhelming.<sup>1</sup> The purpose of this paper is to present an outline and the results of a test involving five different approaches to estimating the demand for a public good in a setting which permitted a real collective choice and which each subject was committed to actual payments when relevant. Four of these approaches, explicitly or implicitly discussed in the literature on public goods, have been shown a priori to provide incentives for the subjects to misrepresent their preferences. In our test, these approaches were

\* I am grateful to members of the Research Department of the Swedish Radio-TV Broadcasting Company for expert advice and to my colleagues at the Department of Economics, University of Stockholm, for helpful comments and assistance. I am particularly indebted to Mr. Hans Bonnevier who was in charge of the experiment presented in this article.

<sup>1</sup> This conclusion, well known from standard theory, remains generally true even when taking into account a recent contribution by Drèze and de la Vallée Poussin [4]. They show that one can avoid providing incentives to misrepresentation of preferences in certain situations (when already in an optimum position and under certain conditions outside optimum when behavior is governed by minimax strategies).

supplied with certain additional arguments which on reasonable grounds could be assumed to mitigate – although to an unknown extent – such incentives. A fifth approach, which could not be assumed a priori to provide any obvious incentives of this kind, was added to the test.

The main result of the test was that none of these five approaches (each applied on a random sample of consumers) gave an average maximum willingness to pay that significantly deviated from that of any other of the approaches. In particular, an approach which provided no incentives for *overstatements*, but possibly the opposite, did not produce an estimate significantly different from that obtained from another approach which provided no incentives for *understatements*, but possibly the opposite.

Another result, “which confirms” preconceived ideas in this area, was that the estimate from a sixth and completely different approach, aiming at the calculation of a *hypothetical* willingness to pay, significantly exceeded the estimate obtained from one of the above-mentioned approaches.

The main importance of this study, however, would seem to be that it provides an experimental approach to tackle the problem of selecting a practicable method for estimating demand for public goods. Moreover, people subjected to our test reacted in a way which is promising for future experiments of this kind.

A general outline of the six approaches is presented in section 1. Section 2 spells out certain crucial requirements of empirical tests in this field. A short description of the experimental situation is given in section 3 (and the appendix). The hypotheses and the results are presented in section 4 and the results are subjected to some tentative interpretations in section 5.

## 1. Six approaches to determine the demand for public goods

### 1.1

We shall here discuss only such situations where a known aggregate willingness to pay which exceeds the costs for a given increase in the output of a public good means that this increase should be made. Then, if we ask the people about their maximum willingness to pay for a given increase in the output of a public good and make them *pay the amount*

*stated*, we would expect to encounter the so-called "free-rider" problem. That is to say, there is a possibility that people will reveal a willingness to pay that is biased downward in the hope of getting "something for nothing". This may be said to be version of the approach discussed (and refuted) by Samuelson [5] and others. (I)

At another extreme, we may ask the people about their maximum willingness to pay when at the same time they are explicitly or implicitly given the impression that they *will not have to pay at all* or will have to pay only a negligible amount. This approach resembles the most common "approach" in practice where people are asked in polls etc. about their interest in, say, more police protection and where they believe that their opinion will have some influence on politicians and that the costs for the individual will be negligible (if any attention is paid to this aspect at all). In these cases, there is the possible tendency of revealing an *upward* biased demand. This approach and the one mentioned above could therefore be expected to present certain limits to the true maximum willingness to pay. (V)

In between these two extremes, other approaches may be defined. They differ from the two just mentioned in the sense that there may be simultaneous tendencies to an upward bias for some people and a downward bias for others. One approach, originally discussed by Bowen [2], and here adapted to the case of a given (or indivisible) increase in a public good, may be seen to have such implications (see also Dorfman [3]). In this approach the people are asked about their maximum willingness to pay where the total cost of a given increase in the public good will be shared equally among all consumers. Thus, given the total cost and the number of consumers, they will in fact be asked if they are willing to *pay a given price identical for all consumers*. In this case it is perhaps likely that those who are willing to pay the given price will exaggerate their stated willingness to pay, thereby increasing the probability of having the increase in the public good at the given price. For those who are not willing to pay this price there is a possible tendency in the opposite direction. Then the net result on aggregate willingness to pay may be biased in any direction, depending on individual preferences in relation to the given price and on their willingness to "cheat". In particular cases, where the price is "low" and statements of negative amounts of willingness to pay are not accepted or are irrelevant, the aggregate amount may be assumed to have an upward bias. (IV)

A different version of the first-mentioned approach, which may also

give rise to understatements from certain people and overstatements from others, is to lay down the rule that people will have to *pay in proportion to their stated maximum willingness to pay* so that total payments exactly cover total costs. Here we would seem to have the same tendency as in the first approach, i.e. that it may pay to state a downward-biased figure. If, however, as is possible, the subject considers that there is a certain probability ( $<1$ ) of paying only a "small" fraction of his stated amount, he may tend – irrationally or not – to exaggerate his willingness to pay. (II)

In still another approach, in fact the one which inspired this study in the first place, it cannot be assumed that the stated maximum willingness to pay will be unbiased or biased upward or downward. In short, people are asked to state their maximum willingness to pay and are told that the decision as to how costs are to be shared among consumers has not yet been made, that it is an issue to be settled by the authorities in light of the *results* of the inquiry and, finally, that it may be *any one of a number of alternatives* explicitly stated, such as (at least) the four alternatives mentioned earlier in this section. I have tried to show in another paper [1] why this situation of uncertainty in fact may be relevant to many cases of political decision making, especially to such cases where information about consumer preferences is most badly needed, as well as why, in these cases, a true state of uncertainty can be made credible to the voters. The important characteristic of this approach in our context is that the voters now face a situation in which incentives to using *simple* strategies, leading to a bias in one direction or another, are absent (cf. [1]). (III)

Finally, one may simply ask people how much they think a certain increase in the output of a public good at most would be worth to them without tying output and/or payment decisions to their responses. This approach, which can be said to involve a hypothetical question, may be the best approximation to the few, direct political approaches or polls actually used in democracies today. (VI)

To sum up, we have briefly presented six different approaches to determining the aggregate maximum willingness to pay for a given increase in the output of a public good, the first five saying that the given increase will be made if the aggregate stated maximum willingness to pay exceeds costs and, if so, each individual will have to pay

- (I) according to his maximum willingness to pay as stated,
- (II) the same fraction of the maximum stated, the fraction being equal

- to costs divided by the stated aggregate maximum willingness to pay,
- (III) according to one of several alternatives, the choice not yet being made,
  - (IV) a given amount, the same for all individuals,
  - (V) nothing.

The sixth approach (VI:1) amounts to asking what the individual thinks the increased volume of the public good would at most be worth to him, not mentioning the conditions under which it would be produced nor what it would cost him if produced.

## 1.2

If any of the six approaches now presented were actually carried out with respect to a certain pure or quasi-public good, say, a bridge, free of toll charge for some reason (legal or other), we are led to believe that at least some of these approaches would lead to significantly biased responses. However, the administrator of such a "referendum" may in light of this presumption want to counteract, if possible, the tendencies to distortions of the kind we have mentioned. In connection with approaches I and II, for example, it may be pointed out that while it is certainly conceivable that it might pay for the individual to understate his maximum willingness to pay, there is a definite risk that such a course may contribute to the good's not being produced, even though he and other potential consumers desire it. For approaches IV and V, the attempt to establish a "counter-strategic" argument may run along quite different lines: It can be explicitly pointed out that while it is certainly remunerative to those who would be willing to pay the price (which is zero in V) to exaggerate their willingness to pay, and vice versa, such behavior would be "morally objectionable", in terms believed to make an impression on the voters. Counter-arguments of this type are less, if at all, relevant for approaches III and VI.

Moreover, granted that approaches IV, V and VI do not require an identification of responses to respondents, a design which excludes anonymity in *all* of the approaches, such that everybody's response is available to all those who may be interested in it, would presumably add significantly to counteracting any distortions.<sup>2</sup> Thus, it can be

<sup>2</sup> For example, in the absence of anonymity it would be hard for anyone known to complain about spending half an hour per day in traffic jams to state that a freeway to avoid these jams would not be worth anything to him.

assumed that an open referendum would actually be preferred by the administrator, even when openness is not strictly necessary.

Using these approaches modified in any given way, what we really want to know, of course, is whether they tend to give identical or different estimates of actual demand, and, if different, which one, if any, tends to reveal the correct aggregate and, if possible, also the correct individual maximum willingness to pay. If, for one reason or another, no correct estimate of the aggregate willingness to pay can be found, the next-best procedure would be to find out which two of these approaches give the narrowest range for the correct aggregate, i.e., which approach gives the greatest underestimate and which gives the least overestimate. This information, especially if the interval is found to be small, may turn out to be as valuable for most purposes as information about an approach to a correct estimation of aggregate willingness to pay.

An experiment could be devised to get some information along these lines. We turn now to a listing of the requirements such an experiment should fulfill in order to be a relevant test case for a real-world referendum on a total or sample basis.

## 2. Requirements for an empirical test of alternative approaches

The task of testing behavior under different specifications of payment consequences by means of an *experiment*, presents us with certain specific problems. We are hardly likely to get interesting results about how much people would like to contribute to the output of a public good if the experimental situation is only hypothetical. In other words, a prime requirement of the experiment is that it should deal with a real decision with respect to a public good, the output of which and the payments for which will actually be carried out under given rules.

Furthermore, the quality of the good in the experiment must be specified in a way similar to what would be expected in connection with a real issue in which, say, the size, design and location of a bridge are known beforehand. For the same reason the costs of producing the good must be announced in advance. Moreover, to approximate a real situation it is necessary to have a large number of people involved. It may, however, not be realistic to have the number specified to an exact figure as in many real-world cases one may expect that the exact num-

ber of beneficiaries (and negatively affected persons, when relevant) is in fact unknown in advance.

A further requirement is that the test should be a *test of the preferences*, and not of the degree of information or analytic capacity of different individuals. Moreover, for the test to be relevant for situations of real referenda (total or sample), people should get information about possible *strategies*, their pros and cons, since these would certainly be topics of discussion before such real referenda. (To take the example of people's being asked to reveal their preferences and being told that they will have to pay according to these, there would definitely be discussions about the advantages to the individual of understating his willingness to pay as well as about "moral obligations" to tell the truth, etc.). All relevant arguments and pieces of information should be spelled out beforehand to meet the last two requirements mentioned. As we saw in the preceding section, this may be achieved in the process of presenting a set of "counter-strategic" arguments.

A remaining difficulty is to avoid giving the impression that the test is devised for purposes other than that of finding out the preferences for the particular public good involved. Thus, to avoid that people for this reason respond to something other than what they have been asked about, the test case should as far as possible be presented as being "important" on its own. The experiment to which we now turn deals with the demand for TV programs from a broadcasting monopoly, for which there is a widely known need for information about "true" preferences.

### 3. An experiment with closed-circuit TV-broadcasting

It is well known to most people that broadcasting companies gather different types of information about the rating of TV programs. Therefore a new type of question in this respect was judged not to come as a surprise. Furthermore, *closed-circuit* television makes it possible to have people subjected to a test, the outcome of which would determine whether a certain program would be shown to them or not and, in the relevant case, whether payments were to be made. Finally, a close approximation to a general case of public goods could be obtained by choosing a program, not yet shown to the public but of such a nature that people would have fairly definite expectations concerning the program. (The program in our experiment was one made by the two,

without doubt, best known humorists in Sweden.) Thus, closed-circuit showing of such a program was taken to meet the requirements stated above.

The test was carried out by the Research Department of the Swedish Radio-TV broadcasting company (SR) in November, 1969. A random sample of 605 persons was drawn from the age group 20 to 70 of the population of Stockholm. They were asked to come to the premises of the broadcasting company to answer some questions about TV programs and were promised a fee of Kr. 50 (\$ 10) for a one-hour "interview". Normally, some 35--50% show up in tests of this kind. On the basis of this experience and on the basis of a simulation of the test carried out at the Department of Economics, University of Stockholm, the sample of 605 (divided into age groups) was randomly divided into six groups such that at least a predetermined minimum number of people in each group could be expected to show up. Each group was divided into two, and each subgroup was given a different type of question, presented by a member of the staff of the SR Research Department. Thus, each type of question was presented on two occasions and with different administrators, A and B (see table 1).

Table 1  
Distribution of persons according to dates and instructions (I--VI).

		Tuesday 25		Wednesday 26		Thursday 27		Sum
		18.45	20.15	18.45	20.15	18.45	20.15	
Total number showing up		38	31	35	35	33	39	211
Adminis- trator	A	VI: 27	III: 14	IV: 20	II: 15	V: 19	I: 12	107
	B	I: 11	IV: 17	III: 15	V: 20	II: 14	VI: 27	104

The persons in each subgroup were placed into a room with two TV-sets and were, for allegedly "practical reasons", immediately given the fees promised them (in four ten-Crown bills, one five-Crown bill and small change to make Kr. 50). The administrator gave an oral presentation of the test which involved a half-hour program by Hasse Alfredsson and Tage Danielsson, not yet shown to the public. The subjects were given the impression that there were many groups of the same size simultaneously being asked the same questions on other rooms elsewhere in the broadcasting company. The responses, given in



writing by the persons in each subgroup, were taken away and said to be added to the statements from other groups. In all relevant cases (which excludes question no. VI) they were told that the aggregate sum of their stated amounts turned out to exceed the "cost" of the closed-circuit televising (presented as rental of a videotape machine and labor costs for operators). Hence, the program was to be shown to them and the payments were actually collected when relevant.

The reader is referred to the appendix for a detailed account of the instructions presented to the different groups. The main part of the instructions given to groups I to V was as follows: Try to estimate in money terms how much you find it worth at a maximum to watch this half-hour program in this room in a little while, i.e. what is the largest sum you are willing to pay to watch it. If the sum of the stated amounts of all the participants covers the costs (Kr. 500) of showing the program on closed-circuit TV, the program will be shown; and you will have to pay

- (to group I) the amount you have stated,
- (to group II) some percentage (as explained) of the amount you have stated,
- (to group III) either the amount you have stated or a percentage (as explained) of this amount, or Kr. 5 or nothing, to be determined later by a lottery you can witness,<sup>3</sup>
- (to group IV) Kr. 5,
- (to group V) nothing. In this case the participants were informed that the costs were to be paid by the SR, i.e. the taxpayers in general.

"Counter-strategic" arguments were added to instructions I, II, IV and V (see appendix).

The subjects in group VI, who received instructions which differed from the instructions to the first five groups, were simply asked how much they found the program to be worth at a maximum. In a second round, these people were asked to give their highest bids for a seat to watch the program and were told that the 10 highest bidders out of an alleged group of some 100 persons were to pay the amount they had bid and see the program (cf. appendix).

<sup>3</sup> This version of approach III meant that the four alternatives had known (and equal) probabilities. Thus, it differs in this respect from the version of approach III mentioned in section 1 and discussed in more detail in [1].

#### 4. Hypotheses and results

The hypothesis to be tested was that approaches I to V would yield identical estimates of aggregate demand for the collective good. If they turned out to give significantly different expected values it would be desirable to know which one gave the best estimate of true demand. Instruction VI was added to the test with the (optimistic) assumption that it could reveal the actual level of aggregate demand, or more exactly, a narrow interval containing true aggregate demand. In particular, it was assumed that question VI:1 would give a slight overestimate of this demand, since it involved no payment obligations on their part (nor did it involve any output decisions), and that being the case, some of the respondents might just be "nice" to the TV company and say the program was worth more than they actually would be willing to pay at most. On the other hand, it was expected that question VI:2 would yield a slight underestimate of the true demand, as people in this case would possibly try to be included among the 10% highest bidders, who would be eligible to watch the program, by stating not their maximum willingness to pay but an amount they believed would be *sufficient* to place them in the group of the 10% highest bidders.

The results are shown in table 2. They imply that no significant differences (at the 5% level) could be found between any pair of instructions I to V. In fact, as can be seen from table 3, these five instructions gave results which were far from significantly different ( $p > 0.20$ ).

With these results, the proposed role of instruction VI was superfluous. This may be seen as a fortunate fact, as instruction VI hardly could have played this role. The results show that instructions III and VI:1 or VI:2 were the *only* pairs that gave significantly different responses (at the 5% level!). This is of course disastrous for the proposed role of approach VI, particularly in view of the presumption that III was the one of approaches I to V in which (simple) strategies for "cheating" were absent.

Available background data of the 211 persons showed that the composition of the groups did not vary very much (see table 4). There were no significant differences of income between any pair of the six groups. The only difference significant at the 5% level concerned groups III and IV with respect to age.

A correlation study of amounts stated and the background variables was carried out for each instruction (see table 5). Out of the 21 esti-

Table 2  
Amounts stated at instructions I–VI: 2.

Kr.	I	II	III	IV	V	VI: 1	VI: 2	
0 – 0.50				1	1	2	5	
0.60– 2.50	2	2		4	3	4	4	
2.60– 4.50	4	5	2	1	4	4		
4.60– 6.50	8	6	15	13	8	10	10	55 out of 70 stating Kr. 5
6.60– 8.50	4	3	2	6	7	3	3	
8.60–10.50	1	7	9	4	8	13	12	All 54 stating Kr.10
10.60–12.50		1		1	1	3 *	1	
12.60–17.50	3	1		6	3	11	12	35 out of 36 stating Kr. 15
17.60–22.50		3		1	1	1	4	All 10 stating Kr. 20
22.60–27.50		1	1		3	2	2	8 out of 9 stating Kr. 25
27.60–32.50	1						1	Both stating Kr. 30
50						1		
Number	23	29	29	37	39	54	54	165 out of 211 stating Kr. 5, 10, 15 etc.
Mean	7.61	8.84	7.29	7.73	8.78	10.19	10.33	
Standard deviation	6.11	5.84	4.11	4.68	6.24	7.79	6.84	
Median	5	7	5	6.50	7	10	10	

Table 3  
Significance tests (*t*-tests) of means with respect to pairs of instructions I–VI: 1.

		II	III	IV	V	VI: 1
Instruction I	<i>t</i>	– 0.74	0.22	– 0.08	– 0.72	– 1.56
	<i>p</i>	> 0.20	> 0.20	> 0.20	> 0.20	> 0.10
Instruction II	<i>t</i>		1.17	0.83	0.04	– 0.89
	<i>p</i>		> 0.20	> 0.20	> 0.20	> 0.20
Instruction III	<i>t</i>			– 0.41	– 1.19	– 2.22*
	<i>p</i>			> 0.20	> 0.20	> 0.02
Instruction IV	<i>t</i>				0.83	– 1.88
	<i>p</i>				> 0.20	> 0.05
Instruction V	<i>t</i>					– 0.97
	<i>p</i>					> 0.20

\* Significant at 5% level.

**Table 4**  
Mean and standard deviation with respect to age, income and sex.

	I	II	III	IV	V	VI	In terms of
Age: mean	41.0	42.0	45.62	38.86	41.78	41.65	years
standard dev.	(12.53)	(14.24)	(13.02)	(13.16)	(12.44)	(13.26)	
Income: mean	18.18	16.69	22.72	17.32	18.97	17.54	taxable
standard dev.	(13.45)	(14.20)	(18.92)	(15.48)	(18.52)	(16.23)	income*
Sex: mean	60.9	58.6	48.3	48.6	48.7	37.0	percent-age men

\* For the national income tax, 1967, Kr. 1000.

**Table 5**  
Correlation coefficients, background variables and amounts stated.

	I	II	III	IV	V	VI: 1	VI: 2	I-V	Positive correlation means that higher amounts tend to be stated by
Age:									
Product-moment corr.*	0.13	-0.36	0.13	0.14	0.07	-0.06	-0.23	-0.04	older persons
(critical value at 5%)	(± 0.42)	(± 0.37)	(± 0.37)	(± 0.33)	(± 0.33)	(± 0.27)		(± 0.16)	
Income:									
Product-moment corr.*	0.36	-0.13	0.45†	-0.14	-0.10	0.15	0.21	0.02	persons with higher income
(critical value at 5%)	(± 0.38)	(± 0.37)	(± 0.37)	(± 0.34)	(± 0.34)	(± 0.27)		± 0.16	
Sex:									
Point biserial corr.	0.30	-0.12	-0.11	-0.37†	-0.23	0.01	0.01	-0.13	men
(critical value at 5%)	± 0.42	± 0.37	± 0.37	(± 0.33)	(± 0.33)	(± 0.27)		(± 0.16)	

\* Pearson product-moment correlation coefficient.

† Significant at 5% level.

mates only two correlations turned out to be significant at the 5% level: In group III there was a positive correlation between income and stated willingness to pay, and in group IV women tended to state higher amounts than men.

These studies of the age, income and sex composition of the groups and the correlation studies thus indicate that the groups were fairly homogeneous in these respects and that the small differences there were need not be taken into account when we now turn to comment on the reactions to the different instructions.<sup>4</sup> Before doing so, we should point out that it is possible, of course, that people who were summoned but did not show up at the experiment may have responded in quite another fashion to the different instructions than the participants did. With the information available, however, there is no way of knowing whether this is so or not.

## 5. Comments on the main result

### 5.1. *The non-significant differences in responses to instructions I-V*

The fact that our results have not revealed any significant differences in the responses to instructions I to V does not, of course, rule out the possibility that such differences actually exist. However, our results make it unlikely that these differences are "very large" – and this in itself may be of considerable practical importance. If two instructions could be singled out, such that one of them were known not to produce an overstatement of the average willingness to pay and the other were known not to produce an understatement, and if, moreover, each of these instructions were to be applied to a large representative sample, we would be able to calculate at some given probability level an interval for the true average willingness to pay that might suffice as a basis for decision making in many real-world situations.

Instructions I and V may be used for this purpose as we have found no reason to believe that they would provide incentives for overstatement and understatements, respectively. We may also note that instruction IV may be modified to produce an upward biased estimate of the average willingness to pay: Those who did not find it worth Kr. 5 to watch the program may have given a downward-biased response. Then, assuming that all those who stated an amount below Kr. 5 in fact reacted in this way, the maximum they would have been willing to pay

<sup>4</sup> A possible exception would be group III where the correlation coefficient between income and amount stated was significant and where the mean income, although it did not differ significantly from the mean income of the other groups, was by far the highest.

could not have exceeded Kr. 4.50, to take the highest "even" figure below Kr. 5. In other words, putting those who gave a response below Kr. 5 all at 4.50, we would end up with an upward biased estimate from instruction IV. This estimate implies an arithmetical mean of 8.16 (4.01) – compare 8.78 (6.24) for instruction V.<sup>5</sup> This suggests that, in certain applications, using instruction IV and modifying it in the way just indicated may be more efficient than using V as an estimate of the upper limit to the average willingness to pay.

It is worth noting in this context the fact that people to a large extent chose "even" amounts such as Kr. 5, 10, 15, 20, etc. (165 out of a total of 211 persons or out of 182 persons mentioning an amount above Kr. 3). This fact should make it relatively easier to detect any general tendency for people to over- or understate their true valuations. For example if there were a general tendency to give correct responses to instruction III, high responses to instruction V (or IV as modified) and low responses to instruction I, the means should have deviated "substantially", because of the fact just mentioned, whereas the variances should not have done so. If, on the other hand, this tendency to overstatements or understatements were partial only, we would have a smaller change in the mean and an increase in the variance. This latter possibility can be seen to be compatible with the figures. And a partial tendency is, of course, reconcilable with the above proposition that the differences between maximum and minimum limits for aggregate demand to be obtained from instructions I and V is not very large.

Our main result, however, is that the reactions received from the different groups are compatible with the possibility of getting *identical* responses to instructions I to V. This possible interpretation would either mean that people do not use available "cheating strategies" even when made explicit or that the "counter-strategic" arguments in I and II and the moral arguments in IV and V, as well as the general feature of non-anonymity, neutralize tendencies to use such strategies. The possibility that instruction I could after all reveal actual preferences for a public good conforms to the view sometimes advocated that people tend to regard their impact on total demand, however small it may be,

<sup>5</sup> Instruction IV actually mentioned only the possible advantages of overstatements and a moral counter-argument, but not the advantages, whenever relevant, of giving a downward-biased response. If one would believe that only strategies about which information were given affected behavior, then instruction IV as it is ( $x = 7.73$ ) would give a non-downward-biased estimate.

as "important". Such behavior may correspond to the fact that people turn up at general elections in spite of the fact that few (if any) of them have ever witnessed an election when a single vote had a decisive influence. This particular example may, however, be more relevant when seeking possible interpretations of an unbiased response to instructions IV and V: People turn up at general elections (and reveal their true responses to instructions IV and V) because they feel morally obliged to do so.

Thus, there are certain reasons why the observed lack of significant differences in the reactions to the different instructions could hold at any sample size much larger than the one used in our experiment. And it seems relevant to draw some attention to these possible reasons, given the outcome of our experiment and considering the long-standing emphasis on a priori arguments in favor of the dominance of "cheating strategies" in the literature on public goods.

As to the practical use of our findings, they seem to support an application of some of the above-mentioned instructions *simultaneously* in a real-world case, say, concerning a local public good such as a bridge (where, e.g. for legal reasons as in Sweden, tolls are not to be charged), a public park, or a recreation area. The inclusion of instruction III in such a study means that this instruction will be tested in its relevant setting (payment distribution decided ex post by the authorities) which was impossible in our experimental situation (where the payment consequences instead were determined by a "public" random process).

## 5.2. Interpretation of responses to instruction VI

As noted above, the responses to instruction VI (both VI:1 and VI:2) deviated significantly from those of instruction III and that instruction only. In view of this result it seems hardly motivated to assume that an instruction of type VI:1 will have any efficient information to provide in tests of this type. The results are of course compatible with the general view that, when no payments and/or formal decisions are involved, people respond in an "irresponsible" fashion. In other words, this result may be seen as still another reason to doubt the usefulness of responses to hypothetical questions, in general, and of ordinary polls to guide political decision making with respect to public goods, in particular.

More surprising, perhaps, is the fact that the average response to VI:1 and to VI:2 hardly deviated at all. Some people (8) reacted by lowering their statements when these turned into actual bids, whereas about the same number of people (10) raised them. Thus, "only" 18 out of 54 changed their statements. If this is considered to be a low figure, it may be explained by the possible feeling people had that a changed bid would imply a confession that they had "lied" in the first round. At any rate, question VI:2 would probably have produced a different result if not preceded by VI:1. The information to be had from such a separate test may, however, be questioned on two accounts. Firstly, this version (VI:2) of the good as a *private* (excludable and rival) good introduces elements of competition, exclusivity and "auction fever" (perhaps, stronger when not preceded by a question of type VI:1) completely absent from the good conceived as a *public* good. Secondly, as has been argued, a person may judge his consumption of a public good as something quite different from consuming the same good as a private one. In the first situation he may be influenced by additional factors, such as his views on the proper arrangement of society. Hence, the transformation of a public good into a private one for purposes of testing demand for the public good may be considered highly inefficient.

### 5.3. *Some direct observations*

A direct observation of great importance for a practical application of any of the approaches discussed here is the fact that people "accepted" the questions as they were posed and, moreover, that almost all of them gave their responses in a matter of a minute or less. In other words, very few acted as if they found the questions difficult or unreasonable. (At the informal discussions after the test, many people seemed to support the view that a valuation of TV programs in terms of money would be an important piece of information to the decision makers.) It should be observed, of course, that people turning up at interviews arranged in this way are special in the sense that they may be considered to be relatively good at giving quick replies and prepared to accept almost any type of questions. The fact that they knew that they would be fairly well paid for their participation (\$ 10) in a test of not more than one hour may to some extent have counteracted this tendency to a biased sample.

To sum up, it would seem fair to say that our results are encouraging



for experimental economics and that a continuation of experiments along these particular lines is warranted. Such future experiments should preferably involve larger samples and larger payments than we dared in this first attempt.

**Appendix: The six sets of instructions (presented orally by the administrators of the experiment)**

*Instructions I – V (General introduction and distribution of fees)*

“My colleagues here and I work for the Research Department of the Swedish Radio-TV Broadcasting Company. Our Department undertakes different kinds of investigations which aim primarily at finding out what viewers think about various TV programs. What we would like to do tonight is to try a new way of doing this. To begin with, we want to warn you that the questions are going to put to you may seem somewhat unusual – but that does not mean that they will be difficult...

The situation is as follows: You are among a large number of people sitting in several rooms like this in this building... We could in a little while show you a half-hour program by Hasse Alfredsson and Tage Danielsson, which hasn't been shown to the public yet. The total costs for a closed-circuit showing of this program would be approximately Kr. 500 (specified). – Now, you have received Kr. 50 for your participation here tonight – and it may turn out that you will spend part of this amount before you leave here. That's because we would like for you to *try* to estimate in *money* terms and *in advance* how much you think it would be worth to you to watch this program. That is to say, we ask you to state the highest “admission fee”, so to speak, that you would be willing to pay to watch the program. If the sum of the amounts stated by you here and all the other participants together exceeds the costs of the program, Kr. 500, the program will be shown to you.

(*Instruction I*) – and you will have to pay the amount you have stated. If the sum falls short of Kr. 500, the program will not be shown and, of course, you won't have to pay anything at all, and you will be free to leave.

To avoid misunderstandings I shall illustrate what I have said by way of example. Each of you will get a form with the question I mentioned

where you are to write your name and amount. Assume now... (an example using the figures Kr. 2:50, 15, 5, 7, 25, 0, 10, etc. as hypothetical amounts stated by the participants).

By stating a small amount, smaller than you are actually willing to pay, you stand the chance of being able to watch the program without paying so much. In other words, it could pay for you to give an understatement of your maximum willingness to pay. But, if *all* or *many* of you behave in this way, the sum won't reach Kr. 500 and the program won't be shown to you.

*(Instruction II)* – for a fee. This fee is to be determined so that you all have to pay a given percentage of the amount you have stated. The percentage will be worked out so that the sum of the fees exactly covers the costs. In other words, all of you together contribute to paying the costs of the program in proportion to what each of you finds the program is worth to you. If the sum falls short of Kr. 500... (see instruction I) ...shown to you.

*(Instruction III)* – for a fee. The size of this fee, however, is not yet determined. It will be determined afterwards and may be any one of the four following alternatives (a prepared table is exposed and explained to the participants): You may be called upon to pay the amount you have stated, that is, the largest amount you consider the program to be worth, *or* a given percentage of this amount (explained as in instruction II above), *or* you will all have to pay the same fee, Kr. 5, *or* you won't have to pay at all, and the costs will be paid by the SR. – None of these alternatives can be regarded as impossible and none can be said to be more likely than any other. What we shall do is arrange a "lottery", here in front of you after you have returned your statements, in order to choose the alternative that will hold. If the sum falls short of Kr. 500... (an example).

*(Instruction IV)* – for a fee of Kr. 5. If the sum falls short of Kr. 500 ... you will be free to leave.

To avoid misunderstandings... (an example). It is easy to see that it would pay for any one of you who really wanted to watch this program to state a much higher amount than he actually would be willing to pay. In this way the total sum of the amounts stated would increase and so would the chances of having the program shown here. But this

would of course make it impossible for us to find out just how much you really think watching this program is worth to you. It could also be said that such an overstatement would indicate a lack of solidarity or respect for the views of your neighbors, who may be called upon to pay for something that is not really desired by all of you together. In other words, it should be seen as something of a "duty" to state the amount you actually find it worth to see the program.

*(Instruction V)* If the sum falls short of Kr. 500... you will be free to leave.

Note that you won't actually have to pay anything. We only want you to state what you think it would be worth to you at most to watch the program.

To avoid misunderstandings... (an example).

It is easy to see it would pay... (see instruction IV) ...a lack of respect for the taxpayers, who... ...it worth to see the program."

All instructions I to V closed with administrative information and a repetition of the question: "...how much you find you would be willing to pay at most...".

#### *Instruction VI*

This instruction begins in the same way as instructions I to V but only up to the point where the costs of the closed-circuit broadcasting is mentioned (the fourth sentence of the third paragraph of the introduction above). As the output decision was not tied to their responses, this cost figure was irrelevant here. Instead instruction VI continued as follows (VI:1):

"We would now like to know how much you would find it worth to watch this program, approximately in the same way as when you make up your mind what a visit to the theater or cinema is worth to you before deciding whether or not to pay the admission fee. In other words, we ask you to *try* to estimate in money what you think it is worth to you to watch this program here; more precisely, the highest admission fee you would be willing to pay, *if* you had been asked to pay an admission fee for watching the program here. To give an example... (same figures mentioned as in the examples in the previous instructions).

(Administrative information and repeated question.)

After the responses to the first question had been collected, a second question (VI:2) was posed to this group:

"We would now like to put a second question to you. It may sound somewhat fresh, but it is most essential for our test. As I stated earlier, you are among a large number of people participating in this test, around 100 persons. We will now make it possible for *ten* persons only to watch the program. These ten persons will be the highest bidders, that is to say, those ten persons of about one hundred who state the highest amounts in this second round. These ten people will actually have to pay the amount they state to watch the show. The others need not pay anything, of course, and are free to leave. Your response to our question, in other words, amounts to the same thing as when you are presenting a maximum offer in writing at an auction. — We might add that you are, of course, free to state the same amount as in the first round, but that you are in no way compelled to state this amount again."

(Administrative information and the question repeated.)

## References

- [1] Bohm, An approach to the problem of estimating the demand for public goods, *Swedish Journal of Economics*, March 1971.
- [2] H. Bowen, The interpretation of voting in the allocation of economic resources, *Quarterly Journal of Economics*, 58, 1943.
- [3] R. Dorfman, General equilibrium with public goods, in: *Public Economics*, ed. Margolis and Guitton, London 1969.
- [4] J. Drèze and D. de la Vallée Poussin, "A tâtonnement process for guiding and financing an efficient production of public goods", paper presented at the Econometric Society Meeting, Brussels 1969.
- [5] P.A. Samuelson, The pure theory of public expenditure, *Review of Economics and Statistics*, 1954.