

FIELD-TEST ELICITATIONS OF DEMAND FOR PUBLIC GOODS

PETER BOHM

Stockholm University

Market economies would be more capable of establishing consumer sovereignty if the provision of public goods could be based on consumer demand. Ever since [Wicksell \(1896\)](#), as restated in [Samuelson \(1954\)](#), it has been noted that free-rider incentives seem to preclude such an outcome. Theory has suggested an ingenious method for neutralizing these incentives (see, e.g., [Clarke, 1971](#); [Groves and Ledyard, 1975](#); [Green and Laffont, 1979](#)). However, since this method is difficult for ordinary people to understand, governments may hesitate to use it in actual decision making (see, e.g., [Bohm, 1984](#); see also experimentation with the method reported in [Davis and Holt, 1993](#)). Some economists have suggested that the best approach available is to ask well-structured hypothetical questions about the willingness to pay (WTP) for public goods and rely on the responses as good enough approximations, in spite of incentives to distort, or not carefully consider, responses to hypothetical questions.

A large group of experimentalists have analyzed the nature of free riding in the context of so-called public exchanges, a laboratory device for investigating the consequences of various aspects of non-rivalry in consumption related to voluntary public goods (for references, see [Davis and Holt, 1993](#)). However, there have been few attempts to identify and test practicable mechanisms for implementing consumer power over public-good decision making in real-world settings as a means of observing the extent of free riding in such contexts. A couple of exceptions to this rule are reported here.

The following pair of issues are in focus in two field tests of public-good decision making ([Bohm, 1972, 1984](#)):

Do incentives to misrepresent WTP produce significant distortions (strong free-rider behavior) when WTP for public goods is elicited in the field?

If not, can combinations of simple, not incentive-compatible, mechanisms be used to attain verifiable approximations useful for demand estimation in practice?

In test case 1, a random sample of the inhabitants in Stockholm was asked to participate in a new type of TV rating. The test was carried out by the Swedish public TV company (although it was the experimentalist who took the initiative and was responsible for the experimental design). Subjects were asked to assign a money value to a new program they could watch if their aggregate WTP were high enough to cover costs. (This was in 1969, when TV was still exciting in Sweden; in addition, the new program offered could be expected to attract wide interest.) In test case 2, the public good was an information package that could only be made available by the government, while the decision to produce the package was to be determined by the actual consumers.

The two field tests were designed to reflect the elicitation instrument selected, if used for decision making in actual practice concerning a non-trivial public good. In particular, consumers were told that:

- their collective decision would determine whether the public good in question would be produced or not, specified as whether or not their aggregate WTP exceeded production costs;
- if the good were to be produced, the consumers actually had to pay according to the rules for payment given to them; and
- there were incentives for misrepresentation of preferences for the public good; in addition, the participants were informed about various arguments likely to appear in a public WTP elicitation process.

The last point is particularly important since the ambition was to mimic a public decision making process where people would have had time to discuss the issues with others and would be likely to have heard – directly or indirectly – the media analyze the issues and possibly argue in favor of a particular kind of ‘voting’ behavior. Given this objective, attempts were made in the tests to lay bare the principal arguments likely to be exposed in such a process and their implications for information about the incentives confronting the respondents. Although it is difficult to speculate on the nature of an equilibrium state of such information, it was presumed that certain incentives to misrepresent WTP would be widely known and talked about, and that the organizers of the ‘referendum’ would try to counter these incentives by referring to the ‘duties’ of citizens participating in this kind of public process and/or the meaninglessness of conducting ‘referenda’ of this type if voters simply gave in to such incentives.

TEST 1. WTP for closed-circuit preview of a popular TV show (Bohm, 1972).

Subjects summoned to the test rating of a TV program, not previously shown but with well-known features and quality, were divided into six groups. Subjects in the first five groups were told that if their aggregate stated WTP exceeded a certain cost of showing them the program, they would be given the opportunity to watch the program, and prior to which each person would have to pay in group:

- I. the WTP stated,
- II. a percentage (as explained) of the WTP stated,
- III. the WTP stated or a percentage (as explained) of the amount stated or 5 *kronor* or nothing, all four with equal probability,
- IV. 5 *kronor*, the current average price of a cinema ticket,
- V. nothing; taxpayers would foot the bill.

This means that the first two groups, and those in group IV whose WTP fell short of 5 *kronor*, were exposed to incentives to understate WTP, while group V, and those in group IV whose WTP exceeded 5 *kronor*, were exposed to incentives to overstate WTP. With a dominating number of WTP statements above 5 *kronor* in group IV (as turned out to be the case), this group can be taken to offer an overstated mean WTP. There were no clear misrepresentation incentives for group III. Accordingly, in [Figure 1](#), the groups

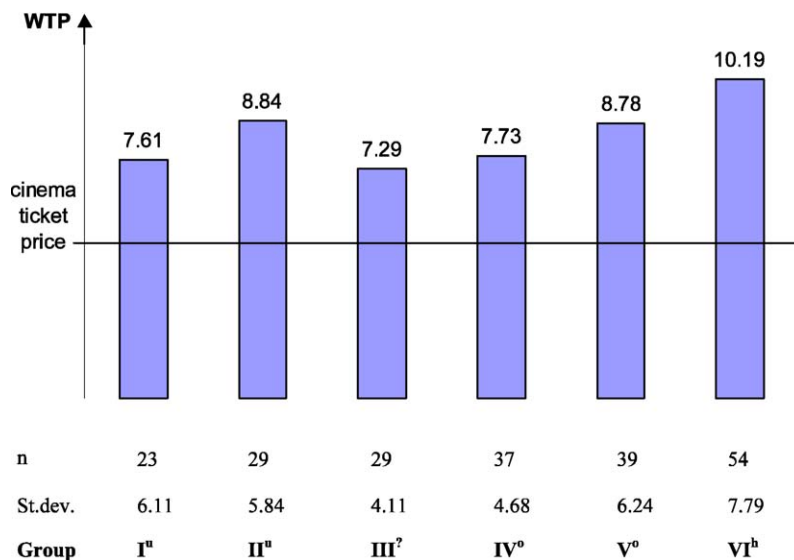


Figure 1. Mean WTP for a closed-circuit preview of a coming TV highlight, shown if (a) aggregate WTP exceeded 'production' costs and (b) subjects pay according to pre-specified rules for groups I–V, all non-hypothetical – two with incentives to understate WTP (u), two with incentives to overstate WTP (o), and one of unknown incentive effects (?); group VI faced a hypothetical question (h). Subjects: A representative sample of 211 inhabitants of Stockholm randomly allocated to the groups. *Results:* No significant differences in mean WTP among the non-hypothetical groups I–V, in spite of the diverging incentives to misrepresent WTP. Hypothetical WTP (VI^h) significantly different from non-hypothetical WTP (aggregate for groups I–V).

are presented as I^u, II^u, III[?], IV^o, and V^o, where u (o) stands for incentives to understate (overstate) WTP. *Result 1:* There were no significant differences between mean WTP in these five groups. Thus, in the particular setting used here there were no signs of strong free-rider behavior; in other words, the same order of magnitude of true aggregate WTP could have been elicited using any of the five approaches.

Subjects in the sixth group (VI^h, h for hypothetical) were simply asked to “estimate in money how much you think watching this program would be worth to you” without indicating whether or not the program would be shown or whether or not their responses could have an effect on such a decision. Comparing mean WTP responses to this hypothetical question with the pool of the responses to the five non-hypothetical questions (Bohm, 1994), a Kolmogoroff–Smirnov test showed a difference at the five percent level of significance, i.e., that *result 2:* the hypothetical version, nowadays called a ‘contingent valuation,’ tended to give an overstatement of the WTP.

TEST 2. WTP elicited for actual decision making concerning a government-provided public good (Bohm, 1984).

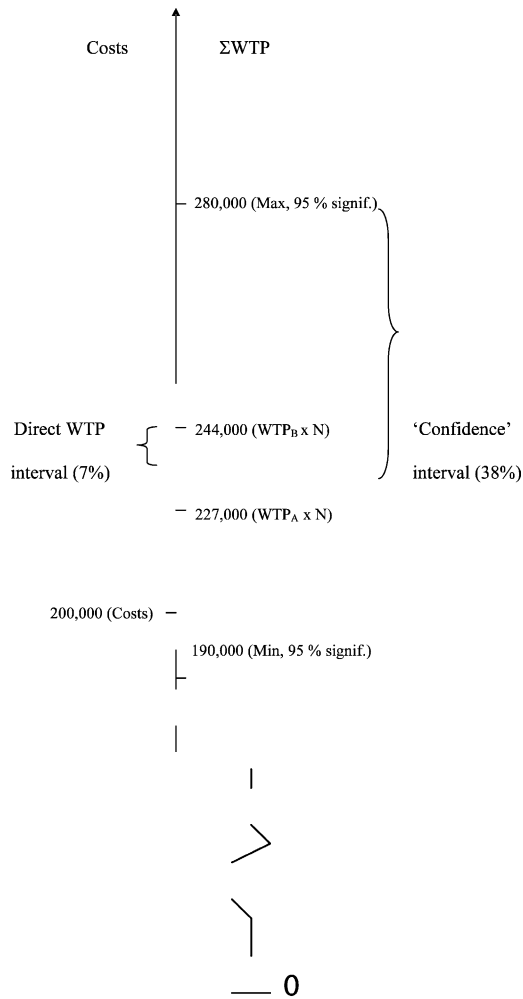


Figure 2. Aggregate effects of misrepresentation incentives using the interval method. Random fifty percent of a population (or sample) exposed to an incentive to understate (Group A) and the rest to an incentive to overstate (Group B) public good demand. Mean stated WTP_i ($i = A, B$) by consumers, the N ($= 274$) Swedish local governments, for a real-world public good (statistics) multiplied by N ; $WTP_A = \text{SEK } 827$ and $WTP_B = \text{SEK } 889$; in SEK. *Results:* Interval between WTP means equals 7 percent of average WTP and exceeds costs. Correction for sampling error: 95 percent confidence interval equals 38 percent of average WTP; 8/9 of interval exceeds costs; central government interpretation: true aggregate WTP likely to exceed costs; therefore, the good was produced and respondents paid according to pre-specified rules.

In 1982, a Swedish government committee investigating potential operational demand criteria for deciding whether or not to produce costly statistics allowed one

particular statistical investment project (providing access to information about certain attributes of housing in Sweden) to be determined by the following "interval method."

Potential consumers were identified as the 279 local governments in Sweden. About half of them were randomly allocated to group A, whose members were asked to report their WTP for access to this type of service according to the principle of group II^a in the test reported above. The remainder of the population (group B) was asked to report their WTP in a way similar to the principle of group IV^o above, but excluding those who did not want to pay the stated price (an excludable public good). More specifically, if the good were to be provided, those in group A would have to pay a percentage of the WTP stated so that the average payment in this group equaled per capita costs of the project, and group B would have to pay a fixed fee of SEK 500 (in 1982, about \$100) if they had stated a WTP of at least SEK 500, while those stating a WTP below SEK 500 would be excluded from access to the service.

The average responses from these two groups gave an estimate of a lower and an upper bound to the true average WTP, since there were incentives to underreport but not overreport (overreport but not underreport) WTP in group A (B). The method could then be used so that the relation between the resulting (average or aggregate) WTP interval and the (average or total) project cost would determine the outcome as follows: If the WTP interval exceeded (fell short of) the cost, the project would (would not) be carried out. If the interval overlapped the costs and (a) the interval was small, i.e., responses to incentives to misrepresent WTP were weak, the project would not have any significant net benefit of the project; hence, it would not be carried out. If instead (b) the interval was large, the government would have to make the decision.

If the government would prefer to use the method only after observing the influence of sampling errors when randomly allocating the population to the two groups, the direct WTP interval just mentioned could be replaced by one including one-sided confidence intervals, as illustrated in Figure 2. The resulting larger interval would then be used instead of the interval directly observed, but in a manner similar to that reported above.

The results from the actual test carried out indicate, as shown in the figure, that regardless of the two alternative designs of the WTP interval, the natural decision was to have the public good produced. (And that was the way the government interpreted the results.) Of the 279 local governments, 274 responded and all of the 130 respondents who qualified for consumption of the good paid the charges. The results reveal a very small direct WTP interval (some 7% of the interval mid-point) and a "95% confidence interval" of 38%, 8/9 of which exceeded the cost figure.

References

- Bohm, P. (1972). "Estimating demand for public goods: An experiment". *European Economic Review* 3, 11–130.
- Bohm, P. (1984). "Revealing demand for an actual public good". *Journal of Public Economics* 24, 135–152.
- Bohm, P. (1994). "CVM spells responses to *hypothetical* questions". *Natural Resources Journal* 34, 37–50.
- Clarke, E. (1971). "Multipart pricing of public goods". *Public Choice* 11, 19–33.

- Davis, D., Holt, C. (1993). "Experimental Economics". Princeton University Press, Princeton, NJ.
- Green, J., Laffont, J.-J. (1979). "Incentives in Public Decision-Making". North-Holland, Amsterdam.
- Groves, T., Ledyard, J., 1975. "An incentive mechanism for efficient resource allocation in general equilibrium with public goods". Center for Mathematical Studies in Economics and Management Sciences, Northwestern University.
- Samuelson, P. (1954). "The pure theory of public expenditures". *Review of Economics and Statistics* 36, 387–389.
- Wicksell, K. (1896). "Finanztheoretische Untersuchungen". Jena; English translation in: Musgrave, R., Peacock, A., "Classics in the Theory of Public Finance". Macmillan, London, 1958.