ESTIMATING WILLINGNESS TO PAY:
WHY AND HOW?

Peter Bohm

University of Stockholm, Stockholm, Sweden

Abstract

The main purpose of this article is to advance a set of conditions which demand-revealing mechanisms must pass in order to be politically acceptable for real-world applications and—to begin with—for real-world experiments. Without such non-laboratory experiments, real progress seems unlikely to take place in this field. So far, there are few indications that these conditions can be met with respect to the proposals made in the literature on public goods. One possible example of a mechanism that meets the "acceptability" conditions is given here. In addition, we present some comments as to why demand-revealing mechanisms constitute an important economic problem, a view which has recently been questioned.

I. Introduction

Recent technical developments in two-way communication, e.g. using cable TV, have made it possible to conceive of people participating directly in public choice decisions in the near future. Whatever one's attitude to using present and potential arrangements of this kind, it no longer seems to be the purely technical or administrative problems that hold us back from involving people in individual policy or managerial issues in the public sector. The main stumbling blocks are—aside from the ideological aspects—the theoretical and empirical problems connected with eliciting people's true preferences.

During the last decade considerable research efforts in public economics have been devoted to problems of estimating demand for public goods. A major part of this research activity has concerned finding a mechanism for direct questioning of consumers whereby incentives to misrepresent preferences would be absent. The existence of such incentives as a main obstacle to revealing consumer preferences for public goods has been part of conventional wisdom since Paul Samuelson's well-known articles on the subject in the mid-fifties (1954). However, a presentation of the problem had already been made by Wicksell (1896).

In this article an attempt is made to find out, from a more practical view-

Scand. J. of Economics 1979
point, where the intensive discussion of demand-revealing methods for public goods during the last ten years has taken us. A main problem seems to be that the perspective of the economists participating in this discussion has been influenced too little by political scientists or observers of practical policy for useful instruments to emerge from this research work. A set of requirements for demand-revealing mechanisms in order for such mechanisms to be adopted in practice are suggested here. A particular example of how these requirements could be met is also presented.

Prior to a discussion about how to estimate consumer demand—or willingness to pay—for public goods we shall say something about why such estimates are desirable. One reason for doing so is that the relevance of this kind of research has recently been questioned. Another is that—as will be argued below—the "how" issue is in practice closely linked to why this kind of estimation techniques are worth looking for.

II. Why?

Johansen (1977) has recently argued that the emphasis on the problem of misrepresentation of preferences in public good theory is misplaced. His main argument for this view is that he does not "know of many historical records or other empirical evidence which show convincingly that the problem of correct revelation of preferences has been of any practical significance" (p. 147) in governmental decision-making. "What matters for the actual decision is the revelation of preferences by the politicians in the actual decision-making process" (p. 149). Politicians, Johansen goes on to say, try to persuade others; thus, they inevitably reveal the preferences embodied in their argumentation. And to succeed and survive as politicians, they must show their constituents by the stands they take that they work hard in the interest, i.e. for the preferences, of their constituents. Thus, politicians have to act in a way that tends to reveal these preferences. In fact, "the two-tier system of electors and representatives tends to diminish the significance and relevance of the theoretical problem of unwillingness to reveal preferences for public goods" (p. 151).

Aside from Johansen's conclusion stated in the first sentence of the preceding paragraph, few economists would disagree, I think, with the general implication of these views. The revelation issue has no particular interest given the present institutions of the two-tier democratic system. But that, as I see it, is beside the point. Why economists have been busy trying to find an approach to estimating individual preferences for public goods is because they want to explore ways of implementing a process of "direct democracy" for certain issues, i.e. have consumers telling their government exactly what their views are about specific public goods instead of letting output decisions rely on the blunt interpretations of these views by politicians. And if consumers were given the opportunity to state their willingness to pay for a specific public

Scand. J. of Economics 1979
good—to be produced or not depending on the information provided by these direct statements—then the possibility that people might misrepresent their willingness to pay would constitute a real problem. The reason is to be found of course in the presence of economic incentives to understate preferences when individual payments are related to the response (the "free-rider" incentive emphasized by Samuelson and others) and to overstate preferences when individual payments are not related to the response.

In passing, we may question whether in fact present institutions are all that free from misrepresentation problems. Johansen admits that joint undertakings by more than one municipality may run into problems of the "free-rider" type (p. 151). We may add similar misrepresentation problems—but of the opposite sign—concerning municipalities in particular in cases where the national government offers large subsidies to public-good type investments in roads, etc. This latter situation is, moreover, just one of many examples where the allocation of a given budget among projects (for road construction or other more or less specific kinds of expenditure) is separated from the question of determining the size of the budget or of a specific kind of earmarked government revenue, thus creating incentives for misrepresentation of preferences. That is to say, politicians will vote for high public expenditure to the benefit of their constituents in one context and for a low fiscal burden on the same group in another context. Whenever such a separation and such "sub-optimization" behavior are made possible in the democratic two-tier process, it will produce instances of revelation problems.

Integrating the financing and the output decisions in the present political process may or may not reduce the risk for the emergence of such revelation problems. If the voters want more of their favorite kind of public expenditure and want to pay less to the government, it is hard to know exactly what set of preferences their representatives will reveal. So, in order to know what trade-offs people actually do want to make, the present two-tier system may not always offer an ideal solution. Other solutions need to be analyzed.

Now, for general or local institutional changes towards more direct consumer influence to take place in a non-revolutionary manner, approval from the present system of government is required. In this perspective, it may be questioned whether recent emphasis in public good theory has been adequate. A considerable part of the contributions to this theory has dealt with what must appear to non-economists as highly complex, technical solutions to the misrepresentation problem. Such devices will hardly convince laymen in government in many countries about the superiority of alternatives to the present system for public decision-making. If I am correct in this view and to the extent that these theoretical contributions cannot be seen as work in progress with a promise of quite another—and more intuitive—end-product in the future, there seems to be a risk for misplaced emphasis here.

There is another possible reason, though, a more cynical one perhaps, why

Scand. J. of Economics 1979
it may be argued that the revelation problem has been overemphasized. It may simply be difficult to convince politicians that any method of direct consumer "power", however obvious its superiority to outsiders, could or should replace the present decision-makers. In other words, politicians may hesitate to abstain from their present influence even in those few and specific public good issues for which a "referendum" approach would be applicable at all. Although politics in democracies is probably competitive enough to prevent a situation where the introduction of new instruments for public decision-making (if "people" want them) were permanently and effectively blocked, this hesitation or direct opposition from a possibly large number of politicians would reduce the chances of governments adopting or even experimenting with new instruments of this type. And this seems to reinforce the requirement that a proposed alternative procedure must contain a realistic and compelling promise to the general public of being "obviously better" than the procedure it is supposed to replace and must for that purpose be quite easy to grasp (pass what we shall call an "intelligibility" requirement) to stand a chance of ever being used in practice.

Improvement in the quality of information about demand for public goods is one reason why it is worthwhile to explore demand revelation procedures. Another is that estimation of option values and consumer's surplus for decreasing-cost industries, cf. Bohm (1977), would require similar mechanisms. Still other reasons exist and may in fact turn out to carry more weight from the politician's point of view. To be able to pursue an effective distribution policy, we need to know how preferences vary among individuals or social groups. If individual demand could be estimated, the distributional impact of the production of specific public goods and of the financial decisions involved would be better known. Even more important for today's politicians may be that information about willingness to pay for specific activities in the public sector could expand the financial capacity for this sector in a way that would not be possible by using general taxes. The basis for this belief is of course that the payments implied here would to a large extent appear as fees for specific services rather than taxes which may be regarded primarily as losses of private purchasing power.

It is against this background of the ability of demand-revealing methods to influence the basis for financial, distributional and allocational decisions in large parts of the public sector that we now turn to some central aspects of the question of how to estimate individual willingness to pay (WTP) for public goods. Before doing so we should emphasize, however, that we do not under any circumstances envisage use of such methods in everyday political decision-making. They should rather be relevant to use now and then when politicians find themselves (or are found to be, in some constitutionally specified way) incapable of reaching a "well-founded" decision, i.e. in highly controversial cases and/or in cases which involve important matters of principle.

Scand. J. of Economics 1979
III. How?

Most of the discussion about estimating WTP for public goods in the literature has focused on finding a mechanism where revelation of true preferences would be a dominant strategy given a rather narrow pay-off definition. In this section we use another perspective and discuss what amounts to a necessary condition for demand-revealing mechanisms to be politically acceptable. A concrete example of how this condition can be met is given in the next section.

A relevant point of departure is to ask whether actual misrepresentation of individual preferences (say, in a local referendum, perhaps on a sample basis) is a fact or a myth. That certain incentives to conceal true preferences may be present in a particular approach to investigating such preferences does not prove that preferences will actually be misrepresented. There may be other—counteracting or dominating—kinds of incentives present as well which would produce a net result of revealed preferences. For example, it has been observed for a long time that the fact that a majority of people participate in general elections violates the "free-rider" hypothesis. Applying this hypothesis to voting behavior indicates that the voter would abstain from voting as he has, practically speaking, no influence on the outcome whereas he has to bear all the costs of voting, i.e. the trouble of getting to the ballot box despite travel expenses, bad weather, etc. That many people actually do vote in general elections—from some 50% in the US to some 80% in Northern Europe, for example—indicates that other aspects, i.e. aspects outside the free-rider argument, influence behavior.

It is not clear how actual voting behavior should be explained. One possibility is of course that the presence of important moral issues, democratic obligations, etc. "force" people to behave as they do. On the other hand, it is possible to point to instances where "moral pressure" has not been able to produce similar results. Sweden—where a conscience of international solidarity is assumed to exist among a relatively large part of the population—may provide a case in point. In a poll in Malmö some years ago, people were asked (1) whether they would like to increase Swedish government aid to less-developed countries and, as a separate question, (2) whether they would like this to take place even if taxes would be raised in proportion. In spite of the fact that question no. 2 was asked right after the answer had been given to question no. 1 and that this was just another opinion poll, half of the supporters of increased aid in the first round vanished in the second round—40% said "yes" to question no. 1, whereas only 20% said "yes" to question no. 2. Here, economic incentives obviously affected people's reported preferences for a public good (in the form of "solidarity" or income transfer to the poor).

Given, as these observations indicate, that economic incentives may or may not influence responses with respect to issues of so to speak high moral content (concerning representation in a democracy, international solidarity, etc.), it can

Scand. J. of Economics 1979
also be observed that economic incentives need not dominate the responses on issues of “lesser dignity”. In the non-hypothetical experiment with collective decision-making in Bohm (1972) where people reported a WTP for, and—if relevant—paid for, a closed-circuit TV broadcasting of a show, the stated average WTP did not differ significantly among groups whose payment consequences varied a great deal. In other words, incentives to overstate and understate demand, with actual requirements to pay when relevant, did not result in significant differences in behavior (average WTP stated) by representative samples of the population (in Stockholm, 1969) concerning a “trivial” public-good type of service.

Already on the basis of these observations it seems fair to draw the conclusion that in a particular case we cannot foresee whether the existence of economic incentives will dominate or be dominated by other incentives that may be present. The odds may change one way or the other due to the issue involved, the investigation design, etc. but the uncertainty will basically remain. Even if we could prove time and again that a given set of economic incentives exists but does not dominate in a particular demand estimating mechanism, we cannot be sure what will happen the next time (and) when the mechanism is used for actual public choice decisions. In fact, once the results of such a mechanism are relied upon, it will be particularly profitable for individuals or groups to “bluff” and try to take “free rides”.

For these reasons and given that direct demand information is asked for, it will continue to be important to look for demand-revealing mechanisms with no incentives for misrepresentation of preferences. But would the invention of such a mechanism meet all needs and guarantee satisfaction to politicians and voters? Probably not. It seems unlikely that satisfaction can ever be guaranteed in a situation like this. A mechanism that convinces economists (as experts) that incentives for misrepresentation are absent in the sense that it does not pay to be dishonest, or that it pays to be honest, may still fail to convince (enough) politicians. And, to put the two groups on a more equal footing, it might not have convinced enough economists either, if they had to stake a substantial part of their future income or professional status on this belief. In practice, things may go wrong as politicians, administrators or voters could misunderstand the rules of the referendum or just because something has been overlooked. And somewhere along the line up to the decision to press the “go” button, someone may have to rely on someone else—say, politicians may have to rely on research carried out by economists—and such confidence may simply not materialize or, once there, it may quickly dissipate.

The point here is not to raise doubts about the feasibility of demand revela-

---

1 More specifically, in some groups individual payment consequences were related to actual responses; thus, there were incentives to understate preferences. In other groups, e.g. where no individual payments were required, there were incentives to overstate preferences. And in still others, there were no clearcut incentives in any direction.
tion in practice. The point is instead to suggest that actual application of methods of direct decision-making for public goods will require that the method can be checked for systematic misrepresentation. That is to say, it must be possible to judge whether in fact a bias is absent in a particular application, not only in principle or in a laboratory setting. Otherwise, I submit, it will be hard to convince voters, politicians and governments to use the method as a public choice instrument on a more permanent basis and/or to make them abide by a counter-intuitive outcome in a particular case.

To satisfy this second requirement—counting the “intelligibility” requirement mentioned in the preceding section as a first requirement—may turn out to be quite difficult. So far the discussion in the literature has focused on finding a single demand-revealing approach capable of avoiding most or all incentives to misrepresentation, without being subjected to a test against another approach to the estimation problem.¹

To indicate how the second “acceptability” requirement—let us call it the “control” requirement—may be met, we shall discuss two possible alternatives here. One is the interval method suggested in Bohm (1972) and (1977) which will be dealt with in some detail in the next section. The other is simply that of using two different methods, each of which is a priori believed to avoid incentives to misrepresent preferences. The idea here is that the “control” requirement would be satisfied by one a priori unbiased method checking another, each used on a separate representative sample of the population. Thus, if both methods produced the same average WTP, in general as well as in relevant strata (income groups or other), there would be enough assurance—by assumption—that correct revelation of preferences has occurred.

As this approach to meeting the “control” requirement calls for not one but two different methods for correct revelation of demand, it does not look all that simple to achieve, at least not at the present time. So far, all existing proposals have been subject to criticism:

(1) Several proposals are based on the same general idea, first developed for use with respect to public goods by Clarke (1971), but discussed also in papers by Groves, Loeb, Ledyard, Green, Laffont and others.² The main point here is that a person i whose reported WTP, influences the outcome—i.e. whether an indivisible project should be accepted (Σ WTP, > costs) or when the optimal volume of a divisible public good is changed—should pay whatever net costs this influence imposes on others. Although it can be shown that people would not have any incentives to misrepresent their preferences under fairly general conditions, there are a number of situations in which misrepresentation may still occur. For example, problems arise (a) when income elasticities for public goods differ from zero or (b) when the payments just mentioned reach non-

¹ See e.g. the mechanisms proposed by Clarke (1971), Groves-Ledyard (1977a) and others with similar basic ideas.
² An example of this proposal is used by Vernon Smith in his article in this issue.

Scand. J. of Economics 1979
trivial aggregate amounts or (c) when these payments are likely to affect very few people, possibly giving rise to careless information gathering and hence poor decisions. Moreover, voters may gain by forming coalitions. Although some of these problems are reduced by using random samples instead of participation by the population as a whole, others remain and new problems can arise. In addition, it seems unlikely that this approach will pass the "intelligibility" requirement. According to the method, each individual may pay a specific amount (in addition to his general or specific contribution to financing the production of the good) which is derived in a complex fashion and based on a rather subtle argument. The appendix to Vernon Smith's article in this issue presents a case in point where it seems unlikely that economists are able to convince ordinary people—and their political representatives—why and how the method works.

(2) Aside from this main group of proposals, others have been suggested, most of which have not been able to survive for any substantial period of time. Thus, the methods suggested, e.g., by Drèze & de la Vallée Poussin (1971) and Malinvaud (1971), require that people follow a minimax strategy although it is far from certain that they would actually do so. Other methods such as the one proposed by Kurz (1973), simply shift the "bluffing" incentives from one stage of the procedure to another, achieving only an increase in complexity. Thus, to sum up, there does not seem to exist any fool-proof method assuring incentives to reveal true preferences, let alone any method that can be understood by ordinary people and by "hostile" politicians. Thus, at the present time at least, it can hardly be possible to have two different demand-revealing methods check one another in order to fulfill the control requirement which we have assumed to be essential for adoption of such methods. Let us therefore turn to another possible solution to this problem.

IV. An Example

Given that we do not seem to be able to use a set of methods which meet even the requirements traditionally made by economists—i.e. the presence of incentives to reveal true preferences—we may try to set out from another angle. Assume that we had two methods both of which were simple to understand, and in both of which there were economic incentives to misrepresent preferences. If, moreover, we knew that the direction of these incentives differed between the two and that, in addition, there were other factors acting as incentives to reveal preferences, a fulfillment of the "control" requirement is conceivable and would be all we needed for the system to be operational.

1 For a summary of most of these problems, see Groves & Ledyard (1977b) p. 139.
2 See e.g. Green & Laffont (1976).
3 See comments in Newbery (1974) and Bohm (1977).

Scand. J. of Economics 1979
An example of this approach is discussed in some detail in Bohm (1972, p. 123) and (1977). Here it is assumed that each of two large samples of the population is confronted with the task of revealing their true WTP, \( \geq 0 \) for a given public good project (or a limited set of alternative volumes of a divisible good). If the project were to be carried out, people in the first sample would pay an amount equal to or at least related to their stated WTP, whereas people in the second sample would pay a given, symbolic amount only or nothing at all. As the relevant information is given to each sample in advance, there would exist economic incentives, in the first sample, to understate WTP and in the second sample, to overstate WTP (at least for all with a true WTP \( > 0 \)). That these incentives would exist is assumed here to be possible to make obvious to everybody. If the average WTP reported by the two samples coincided, we would have to assume that incentives to misrepresent WTP have been dominated by incentives to reveal true preferences and thus that this average WTP value would be the true one.

The reason why this approach is at all conceivable for producing the desired result is based on the following. There are some indications, cf. Bohm (1972), that the existence of incentives to misrepresent preferences may not be sufficient for preferences actually to be distorted when at the same time, there are motives for giving truthful answers ("help guide public decisions by actual preferences" and other moral arguments). Moreover, such motives may be reinforced, e.g., by making responses non-anonymous.

If reported WTP would differ between the two samples, there is still the possibility that useful information about true preferences could be obtained.

---

1 In the experiment in Bohm (1972) this was explained to participants in each sample before they made their binding WTP statements.

Scand. J. of Economics 1979
in this way. As the first sample would reveal at most a value equal to the true average WTP and the second sample would reveal at least a value equal to the true average WTP, we would end up with an interval containing the true WTP. Thus, with an indivisible public good, we could still determine in some cases that aggregate WTP exceeds (falls short of) project costs. This is shown by volume $X_1(X_2)$ in Fig. 1, assuming that this volume illustrates the indivisible good. If, on the other hand, $X = X_2$, we remain uncertain about whether or not aggregate WTP exceeds costs. But if the interval of the aggregate WTP is narrow, we may interpret this as a case of indifference. With a divisible good for which there are WTP estimates given for $X = X_0$, $X_1$, $X_2$ and $X_3$ and where intermediate values are derived by interpolation, we would have to assume that optimal $X$ is somewhere between $X_1$ and $X_2$ in Fig. 1 (where minimum and maximum marginal aggregate WTP, respectively, equals marginal costs). Thus, depending on the interval obtained, an approximation of the optimal volume could be determined.\footnote{To be more specific, the mechanism suggested here could work as follows. For simplicity, let us assume a discrete public good the production of which will be carried out if average WTP exceeds total costs per capita (or part thereof for a good with considerable durability). Production will be financed by the beneficiaries in some approximative fashion, determined by responses from people in the two samples and related to characteristics such as income, family size, location, etc. The payments from the total population of beneficiaries may be adjusted to reduce collection costs and to reach given distribution policy goals at minimum costs; cf. Bohm (1971). Two sufficiently large stratified samples are drawn from the local (or national) population affected by the project. Information about the project as well as about the "referendum"—including information about possible incentives for and against misrepresentation—is provided to everybody and in particular to people in the two samples. An open discussion can be expected in the mass media up to the date of the "representative referendum". Those who have been sampled to participate are required to do so on the same grounds as people who are subpoenaed as witnesses in ordinary court proceedings. Even the compensation for expenses, etc. could be the same as that relevant for court witnesses.}

To sum up, the important things to note about the "interval method" suggested here are

(a) that even if the control requirement is not met, the results of the two tests may be useful by providing a sufficiently narrow interval for average WTP.
WTP; this is not possible for mechanisms which—in contrast to this method—simply try to achieve non-distorted responses from each of the two separate tests involved and for which deviations between the two cannot be interpreted as an interval containing the true WTP (as, for example, both may produce understatements);

(b) that the existence of the control requirement—the comparison between the two tests—will act as an (additional) incentive not to distort responses, as that would contribute to making the tests less valuable as a decision instrument, at worst even useless; at first sight, it may look like this incentive would be present for any pair of tests, i.e. that is would be a result of the introduction of the control requirement per se; but it will hardly be convincing (i) to suggest that two non-distorting methods are used and (ii) to emphasize at the same time that distortions will be checked, revealed and make the whole process useless by the force of the control requirement; and

(c) that the “intelligibility” requirement is met in perhaps the most direct and simplest way possible, i.e. with no attempts to “manipulate” the information problem in order to make telling the truth the dominant strategy at the expense of the simplicity of the question to be posed.

As we have already pointed out, the method suggested here is not to be interpreted as a proposed “final” solution to the problem of revealing demand for public goods in practice and in cases where the traditional political process does not seem to be sufficient or satisfactory. But it has been proposed as a method that can be used as a starting-point for experimental applications in the real world, say for local public good decisions (environmental quality, availability of public transportation facilities, etc.). Without such real-world experiments little can be accomplished by way of real progress in the field of public choice mechanisms. However, for such experiments and other applications of direct decision-making in society to be acceptable to government, the methods proposed must meet certain minimum standards of “intelligibility” and “quality control”, as has been emphasized here. If not, it will be difficult to convince politicians and ordinary people that public choice by direct “demand-revealing” methods will represent an improvement. And regardless of how well a particular method performs in laboratories and with non-representative samples, the laboratory tests—however indispensable they are initially—can never substitute for practical, real-world experiments. Thus, failure to meet requirements of practical politics—some of which I hope we have touched upon here—may keep public choice mechanisms in the laboratories forever.
References


