



Winter 1994

CVM Spells Responses to Hypothetical Questions

Peter Bohm

Recommended Citation

Peter Bohm, *CVM Spells Responses to Hypothetical Questions*, 34 NAT. RES. J. 37 (1994).
Available at: <https://digitalrepository.unm.edu/nrj/vol34/iss1/2>

This Article is brought to you for free and open access by the Law Journals at UNM Digital Repository. It has been accepted for inclusion in Natural Resources Journal by an authorized editor of UNM Digital Repository. For more information, please contact disc@unm.edu.

PETER BOHM*

CVM Spells Responses to Hypothetical Questions

ABSTRACT

Part of this note takes a new look at an earlier study of mine (1972), sometimes used—incorrectly—in support of CVM; it is shown that while Cummings and Harrison are right in criticizing this study for not providing such support, they do so for the wrong reasons. Moreover, it is argued here (a) that even if it could be shown that CVM provides acceptable accuracy in trial runs, this property cannot be counted on when CVM is used for actual decisionmaking; (b) that the need for CVM has been exaggerated, since CVM is neither the only approach to estimating non-use values, nor obviously better than using methods that estimate only use-values.

In 1966, a Swedish evening newspaper—*Kvällsposten* in Malmö—conducted an interesting study using the Contingent Valuation Method (CVM). A sample of the population was asked whether they would want the Swedish government to raise its aid to LDCs from 0.25 percent to one percent of GNP. Forty percent said "yes". Immediately afterwards, those who had said yes were asked a second CVM question: "Would you accept this increase in government aid even if taxes would have to be raised" (meaning roughly: "... if it would cost you anything")? Half of those saying "yes" to the first question now said "no".

Somewhat later in 1966, a group of assistant professors at the Economics Department of the University of Stockholm arranged a petition among university employees in Stockholm, urging the Swedish government to increase its aid to LDCs to one percent of GNP. To avoid having the petition flatly dismissed as irrelevant, it was decided that all those who signed the petition would also make a commitment to pay one percent of their after-tax income during one year to unspecified multilateral UNDP assistance to LDC. Sixty-four percent of all university employees signed -and had withdrawals made from their paychecks during the following year. The Swedish prime minister who received the petition said he was impressed by—what may here be called—the non-CVM nature of the petition.¹

* Department of Economics, University of Stockholm. I am indebted to Bill Desvousges, Glenn Harrison, Reed Johnson, Bengt Kriström and Hans Lind for helpful comments on earlier drafts of this paper.

1. It may be added that, some years later, Swedish government aid was in fact raised to one percent of GNP. The role of the petition for this outcome can only be speculated about; most likely it was not large, but probably still much larger than any study based on hypothetical questions.

The first story illustrates the fact—hardly surprising to most of us—that responses to hypothetical questions cannot be trusted to reveal the truth; more specifically, it shows how significant even part of the strategic bias of hypothetical questions can be (here, the part which is revealed when moving from the first to the second—still hypothetical—question). The second story is told because it indicates that the “best available procedure” for public-good (or public-bad) demand revelation need not be CVM, as the Ohio court suggested (see Cummings and Harrison, p. 1). In particular instances, methods involving real payment may be much more convincing (although much more costly, to be sure) for decisionmakers—governments or judges—whose decisions are open to influence. My comments here will refer essentially to these two points. More specifically, I wish to address the questions: (a) are there ways to settle the dispute about the size of strategic bias in CVM *when used for actual decisionmaking*, and is that size robust enough and hence capable of being systematically taken into account; and (b) are there competitors to CVM in the role of “best available” procedures for valuing public goods or bads, at least for certain issues? In dealing with these questions, I will elaborate on the results of two studies of mine which were discussed at some length in Cummings and Harrison (CH)² and which have a further bearing on these issues.

The reader is advised to observe that the comments here are given by a non-United States economist, who has little acquaintance with the United States legal system.

On the relevance of hypothetical questions for decisionmaking

To begin with, it is worthwhile noting that there are important policy questions that can hardly be other than hypothetical. Where the line must be drawn will probably change over time depending on the advances of social science. Thus, for example, few economists would now say categorically that questions about “willingness to pay” (WTP) for all public goods must be hypothetical, which generally seems to have been the belief before the era of Clarke-Groves mechanisms and other recent advances in this field. On the other hand, few would argue that anything close to *all* public goods or bads could be subjected to practicable demand-revealing mechanisms.

In any event, one must ask: what information can we expect to obtain from responses to hypothetical questions such as those used in CVM? As shown by CH, the literature in favor of CVM has not produced any proof that any particular design of this method is known to reveal WTP, or can reveal WTP by any known degree of approxi-

2. R. Cummings & G. Harrison, *Was the Ohio Court Well Informed in their Assessment of the Accuracy of the Contingent Valuation Method?*, 34 Nat. Res. J. Vol. 1.

mation, when CVM is applied to a new issue. Thus, although we know a lot more about CVM now than we did some ten years ago, we have not been able to identify a design that would remove the original doubts about CVM, especially with respect to "strategic bias" and "lack of sincerity". If CVM is deemed an acceptable procedure to calculate natural resource damages in CERCLA litigations, i.e., if responses to hypothetical questions are deemed to provide an adequate basis for decisions which have an impact on CV respondents, it is the extent of strategic bias that is our primary concern.

Why has it not been possible to determine whether or not any particular CVM design is capable of repeatedly revealing actual WTP with a reasonable level of accuracy, and specifically, without any significant strategic bias? As CH show, proponents of CVM have certainly made some efforts in that direction. In that context, the results of an experimental study of mine³ have been referred to extensively in support of the use of CVM. The justification for this interpretation of my results was refuted by CH, and I am the first to agree with that conclusion. In fact, as I shall argue here, the results of my 1972 paper support the view that responses to hypothetical questions give upward-biased estimates of WTP. By taking a new look at this paper, we can also see a specific example of an experimental design which can be used to answer the question posed above.

Taking a second look at Bohm (1972)

Aside from trying to substantiate the claims just made, I will offer some critical comments on CH's interpretation of my 1972 paper. In addition, I will highlight a few characteristics of this study that are important to keep in mind when evaluating its results. Given CH's rather detailed presentation of my paper, I will try to be brief.

(1) The *purpose* of the 1972 paper was to analyze the effects of different incentives on *non-hypothetical* WTP statements. The major part of the experiment was aimed at comparing five approaches (denoted I through V, see CH for details) for estimating average WTP for a specific public good. Each approach had different payment consequences, at least one with incentives to overstate WTP, at least one with incentives to understate WTP, and one or two with mixed incentives. The public good was showing a well-acclaimed stage-show by two famous comedians on closed-circuit TV; the program had not yet been shown on regular TV and would not be for another two months. Thus, it seems fair to say that the public good in question was for most people neither of an unknown quality (or *genre*), nor of trivial value or interest.

3. P. Bohm, *Estimating Demand for Public Goods: An Experiment*, European Econ. Rev., 3, 111-130, June 1972.

Participants were told that stated aggregate WTP would determine whether or not the good would be produced. If it were produced, participants would have to pay according to the rules presented to them. The purpose of the test was explained as one of finding out in more detail, and in a new fashion, about viewers' valuation of TV programs. The study was held on the premises of the TV company.

Moreover, it must be emphasized that various attempts were made to perform the test in a referendum-like manner, in particular with respect to the existing misrepresentation incentives. These incentives were brought to the attention of the participants who were explicitly asked not to give in to the incentives, since this would reduce the value of the study for the TV company. This latter design is in line with the "training" that people in democracies such as Sweden receive, e.g., encouraging them to withstand "petty egoistic" incentives not to participate in general elections (saving time versus having only an insignificant effect on the outcome). Further, it seems unlikely that government decisionmakers would allow important decisions on compensation, conservation, clean-up operations, et cetera, i.e., decisions otherwise made by the politicians themselves, to be determined by voting mechanisms where misrepresentation incentives are present, without openly discussing the existence of these incentives and verbally dealing with their possible implications in an attempt to neutralize them or at least reduce their impact.

(2) The study demonstrated that there were no significant differences the five approaches. In other words, the dominance of certain misrepresentation incentives, especially free-riding incentives, so generally accepted as a fact in the literature at the time, did not show up in the results when the approaches were used in the "referendum-like" manner. Thus, among these different approaches, a strategic bias did not manifest itself.

(3) Since the direction of the incentives differed among two to four of these approaches, with No. I providing clear incentives to *under-report* WTP and No. V clear incentives to *over-report* WTP, the (approximately) similar responses to all of them can hardly be interpreted as reporting anything else than a rough unbiased estimate of the *true* average WTP.⁴

(4) Although the purpose of the experiment was *not* to test responses to hypothetical WTP questions, a fairly large group of subjects was asked to report their *hypothetical* WTP for the good (question VI:1). The average WTP for this group appeared to be higher than those for

4. For reasons that remain unclear to me, CH do not seem to want to accept this interpretation. Rejecting it is crucial, as we shall see, for CH's argument that the study does not say whether CVM overstates, understates, or equals true WP in this particular application.

the five non-hypothetical groups; however, the difference was found to be statistically significant at the 95 percent level with respect to only approach No. III. It is this lack of statistically significant differences for the four other pairs of non-hypothetical/hypothetical approaches that seems to account for the interpretation, so common in the literature, that the CVM approach in this study did not differ from the non-hypothetical approaches.

In the 1972 study, parametric tests were used. As shown in CH's paper, Glenn Harrison recently subjected the raw data of this experiment to a non-parametric test. The results obtained are essentially the same as those reported in the 1972 paper. However, in the Harrison study, approaches I and III were significantly lower than VI: I at the 98 percent level; and approach IV was close to significantly different from VI:I at the 95 percent level. The results of these new computations are presented in Table 1, rows 1 to 6.

Given the current interest in the relation between the hypothetical and non-hypothetical (or real-payment) approaches, and given the fact that none of the non-hypothetical approaches revealed any misrepresentation of preferences (see point (2) above), it is interesting to compare the responses of all the *non-hypothetical* groups (I-V, which contained 23, 29, 29, 37, and 39 subjects, respectively) with those of the *hypothetical* approach (VI:I, 54 subjects). As can be seen from the bottom line of Table 1, the CVM approach (VI:I) gave higher WTP responses than the non-hypothetical ones, at the 95 percent significance level. (The average WTP of group VI:I was 25 percent larger than that for groups I-V.)

Group	I	II	III	IV	V	VI:I
I	-					
II	.36	-				
III	.33	.56-				
IV	.44	.58	.65	-		
V	.60	.99	.68	.87		-
VI:I	.018	.24	.011	.058	.13	-
I-V pooled						.036

Table 1. Probability values for testing the null hypothesis that pairs of approaches originate in the same distribution; a Kolmogoroff-Smirnov test of the Bohm (1972) experiments. (Source: GlenN Harrison)

This result strengthens the conclusion I originally made in passing and with at the time—a sense of stating the obvious, i.e., that responses to hypothetical questions tend to give unreliable estimates of WTP.⁵ Now, with the addition of the comparison between CVM (VI:I) and the pool of the non-hypothetical groups (I-V), the only conclusion

5. In response to CH's concluding query about the reason for this "strong conclusion": it was the difference between VI:I and III (the only significant difference in the material) that led me to conclude in the 1972 paper that my results were compatible with the view that "... that people respond in an 'irresponsible fashion' ... to hypothetical questions." See *supra* note 4.

that seems possible to extract from the 1972 experiment is that the CVM approach in this particular setting tends to produce upward-biased estimates of true WTP.

However, CH seem to interpret the results in quite another way. They conclude that "... depending on which pair of experiments one focuses on, we have shown that one can argue from these results that the CVM understates, overstates, or equals true WTP". The eye-catching part of this statement is, of course, that CVM can be said to understate true WTP in this particular study. Specifically, they argue that VI:2, the multi-unit auction that followed VI:I, understates true WTP and then so does CVM (VI:I), since the results from VI:I and VI:2 are virtually identical.

However, CH's interpretation of VI:2 as providing an underestimate of WTP for the commodity in question is hardly correct.

First, it should be noted that the auction approach of VI:2 is quite different from the other approaches, which either try to elicit non-hypothetical WTP responses under "referendum-like" conditions (I-V), or ask a simple hypothetical question (VI:I) concerning a public good. Furthermore, when subjected to the VI:I question, respondents may get the impression from the setting they are in that there is a chance the program will in fact be shown to them. Therefore, they may be exposed to a strategic-bias incentive to "encourage" the company to produce the good, or praise them in advance for doing so, by overreporting their WTP. By contrast, VI:2 effectively stated to the same respondents: "We are going to show the program, but only to the 10 percent highest bidders!" This may have encouraged competition among the participants by emphatically stating that 90 percent of them would eventually have to leave the room, while the highest bidders would remain there and look forward to seeing the show.

The manner in which people might respond in this new situation is dramatized here for the sake of the present argument. However, in my original paper, I tried to avoid a "sociological" discussion of behavior in this competitive situation, except to suggest briefly the obvious—that the behavior is likely to be the same as when people attend an auction where the winners can be identified by all those participating. In Sweden, auctions of this kind are often a significant social event, where some participants—to continue the dramatization one step further—might "show off", e.g., by trying to indicate that money is no problem for them, or generally behaving in ways which would attract attention. This means that some people place bids in amounts they would not pay for the same objects when offered on a regular market. In my 1972 paper, I simply referred to this as "auction fever".

However, unlike what CH suggests in note 50, such behavior is not necessarily irrational. Instead, it may demonstrate that, when confronted by approach VI:2, people bid on something more than what was involved in approaches I-V or VI:I. That is, in VI:2, the commodity auctioned off is a composite: the good itself plus the social implications of being known to be among the few who will consume the good. As in similar auctions, some bidders seem to attach a value to the latter dimension of this composite commodity. While I certainly agree with CH that the WTP estimates provided by VI:2 may well have been downward-biased, I suggest that what is biased here is the estimate of WTP for *the composite commodity*, not the good itself, which is what approaches I-VI:I refer to.

Thus, it hardly seems meaningful to compare estimates obtained by VI:2 for one commodity with estimates obtained by other approaches with respect to *another commodity*. Specifically, VI:2 cannot be used for an unequivocal interpretation of my results as (also) saying that VI:I, by being virtually identical to VI:2, would understate WTP for the commodity in question.

It should be noted that CH do not seem to want to pay attention to points (1)-(3) stated above, which were the main points of my 1972 paper. These points summarized my finding that the five different non-hypothetical approaches did not yield diverging WTP estimates, which refuted the conventional wisdom at the time. Therefore, the relevant result from my 1972 paper to be highlighted here is: For the particular case of public goods analyzed in this paper, all the non-hypothetical approaches generated good approximations of true WTP, in contrast to CVM, which significantly overstated true WTP.

What proof is needed for using CVM in actual decisionmaking?

The burden of proof is clearly placed on those who argue that CVM can be used for WTP revelation. Given results such as those reported here as well as by Bennett⁶ and CH, there should be no doubt that the required proof is simply absent. However, even if the requisite proof existed, i.e., a series of independent *experimental* studies showing that CVM in each particular context consistently revealed WTP or some stable percentage of WTP, such results could hardly validate the use of CVM for decisionmaking, e.g., by United States courts in CERCLA litigations. New applications of CVM, the results of which now would be used by courts, thereby possibly affecting the well-being of the CVM respondents to a significant extent, would strengthen the misrepresentation incentives. Then, we would need to know the effects of this new level (and possibly kind) of misrepresentation incentives.

6. J. Bennett, *Strategic Behaviour: Some Experimental Evidence*, 32 J. Pub. Econ. 355 (1987).

I have here taken for granted that if CVM is to be used for decisionmaking, this will be made known to the respondents before the questions are posed. If, instead, the purpose of the investigation were successfully kept secret from the respondents, we would have a second, potentially equally serious, "lack-of-sincerity" problem. But then, how would one know for certain whether such efforts were successful? Uncertainty on this point would tend to reduce the usefulness of the exercise, since the two incentive problems have different effects on the respondents. Aside from the fact that a democratic society might want to avoid manipulation of its citizens by hiding the purpose of questions which they are subjected to, making it clear that the responses are to be used for decisionmaking would, in most cases, also make it known in what direction the responses can be biased.

Another implication of accepting the face value of responses to hypothetical questions is that signals would be given to various parties involved in the case at hand, be they firms, organizations, news media or individuals, to urge respondents to use their now, practically speaking, unlimited voting power. Thus, even if it had been repeatedly shown that a certain CVM design so far has given estimates with some small maximum degree of approximation, the next application of this approach to a new situation may not. In other words, a significant strategic bias may well appear, when CVM is used for actual decision-making, if not earlier. And now, of course, there is no non-hypothetical approach with which CVM could be compared. If there were, the non-hypothetical approach would have been the best available method.

Therefore, while I agree with CH's final conclusion that "a good number of . . . studies show that CVM values can overstate real economic commitments, and that these overstatements *can be* quite large" and that the "courts must assess the demonstrated *potential* for such overestimates in their deliberations concerning the legal necessity of results from CVM surveys in CERCLA litigations", it seems necessary to add that this assessment must concern the new situations where it may have been made clear to CVM respondents that their responses might actually influence, or even determine, compensation or other real consequences for themselves.

One way for a method to be helpful in revealing actual WTP would be to incorporate in its design a gauge of the extent of misrepresentation that might occur. (If not, it would be entirely up to the judges to determine what they think is the likely extent of misrepresentation.) It was with that aim in mind that I suggested the "interval method" and had it tested in practice (as mentioned by CH).⁷ This method al-

7. P. Bohm, *Estimating Demand for an Actual Public Good*, 24 J. Pub. Econ. 135 (1984); see also Bennett, *supra* note 7.

lows one to estimate an upper and a lower bound to aggregate WTP, the interval between which would grow with the impact of misrepresentation incentives. The usefulness of this method would be, practically speaking, similar to that of a perfect revelation mechanism if the impact of these incentives were small. In the non-hypothetical test, which determined whether or not a Swedish government agency should produce an actual (excludable) public good, half of the population of consumers (here: local governments) was asked to report their WTP in a setting which confronted the respondents with incentives to over-report but no incentives to under-report, while the other half was asked the same question with the opposite incentive structure.

If the resulting interval of average WTP exceeded the production costs per consumer, the good would be produced and the respondents had to pay according to the rules presented to them. (This is what actually happened.) As already reported in CH, the interval between the upper and the lower bound to WTP turned out to be quite small. It was small enough to determine the discrete output decision in this particular case with an accuracy close to that where an approach had existed that would reveal true WTP. The point to be emphasized here is that this approach—given that the only incentives at work are those referred to above—reveals the extent of WTP misrepresentation to the political decisionmakers and that the extent was small when used in this non-hypothetical application.

Why is this test of the interval method, which uses only non-hypothetical approaches and thus in no way deals with the CVM, relevant here? I turn to that question in the final section.

CVM—still the best available procedure?

Even though CVM does not meet reasonable quality standards in the sense of a known minimum degree of precision, is it nonetheless the “best method available”? At least for certain cases, there may be better alternatives. For example, the value of environmental damage inflicted on one particular group in a certain area may be approximated by the WTP to avoid or eliminate similar damages as expressed by another group in another area, where in fact a demand-revealing process had been used (say, by using the interval method, or the “minimum real WTP” approach mentioned in the second paragraph of the introduction, or, if practicable, some truly incentive-compatible approach). Then, this latter kind of information can hardly in general be considered to be inferior to the information that a CVM study of those already harmed and now to be compensated would provide.

Obviously, it is unlikely that there are a great many litigation cases where the damage is sufficiently similar to the potential damage, or damage avoidance, of an action that can be subjected to such a test.

Nonetheless, there are some such actions and there are methods, such as the interval method, which could be used, provided that the potentially harmed people were given the power to veto such actions by making a sufficient economic commitment. This provision is certainly a considerable one, but can it be excluded out of hand? More important, perhaps, is the general point that all indirect ways through which real economic commitments for avoiding (the effects of) environmental damages do not seem to have been sufficiently investigated, for it to be possible to state that the CVM is generally the "best available" approach to estimating statutory use values, i.e., total values. For example, it may well be that using travel cost and hedonic methods alone—likely to yield estimates of use values only—is a better approach for certain kinds of issues than using CVM as an instrument for estimating total values.

This criticism does not imply that CVM is "useless". First, CVM is clearly a practical instrument for inexpensive "early-warning" investigations of almost any issue. If responses to hypothetical questions reveal "surprisingly strong" reactions, the CVM would serve as an important signal for politicians' "introspection" and for indicating that more needs to be done in terms of empirical investigations or searching for indirect evidence from other sources. Second, the government can use responses to hypothetical questions put to a collective of pollutees as the basis for *claims* when suing a polluter on behalf of this collective, just as claims by a party of one, who sues another for harm inflicted on the plaintiff, are based valuations by the plaintiff. In both cases, it is up to the court to evaluate the *bias* in the estimates of the harm inflicted by those who stand to gain from having the claims accepted.

Scientific analyses of the properties of different CVM designs in practical applications should, of course, continue. But what has come as a surprise to many bystanding economists for quite some time now is not only the considerable extent to which CVM in fact has been relied upon as a provider of information, especially in the United States, but also the doubtful arguments with which many CVM proponents have defended their method. This now seems to have polarized the CVM debate in an unfortunate way, as can be seen from some of the more questionable anti-CVM arguments advanced in the Exxon Conference volume from April 1992.⁸ What quite often seems to be lacking on both sides is a truly scientific—and hence at least cautious—attitude toward the available evidence.

8. Cambridge Economics, Inc., *Contingent Valuation: A Critical Assessment*, (conference held in Washington, D.C. April 2-3, 1992).

In particular, it seems to me that CVM protagonists often have failed to pay sufficient attention to the fact that we are, after all, talking about hypothetical questions here. Terminology could be part of the problem. Contingent valuation and the acronym CVM represent an embellishment for hypothetical valuations, and this choice of terminology may have clouded some minds. Since this terminology will hardly cloud the minds of the defense in CERCLA litigations, one may wonder how long courts would be able to brush aside arguments by the defense that CVM "only provides responses to hypothetical questions". In fact, accepting hypothetical questions as the "best available" method for reliably calculating the values of intangibles may in the end jeopardize the courts' attention to such values.

CONCLUSION

I agree with CH's view that the economics literature has not provided any conclusive evidence that responses to hypothetical questions reveal WTP with a known degree of approximation or without serious bias. Indeed, there is much evidence to the contrary, including the results of my 1972 study, despite its having been referred to *in support* of using hypothetical questions as a sufficiently precise instrument for estimating WTP. It has been emphasized here that once responses to hypothetical questions are accepted for clearcut decision-making, the strategic bias is significantly strengthened and the implications of this effect are unknown.

It is doubtful whether, strictly speaking, the use of hypothetical questions can be seen as a scientific method, let alone the "best available procedure for reliably calculating statutory use values" as suggested by the Ohio court. Potential *indirect* methods for estimating such values, reflecting non-hypothetical economic commitments, cannot be said to be nonexistent. Although perhaps much more costly to use, such methods are likely to be much more precise whenever they are applicable. However, the capacity or role of these methods as "better available procedures" needs to be further evaluated. An important part of this work requires convincing local governments that experimentation with real-payment approaches is needed, or using such methods for decision-making on environmental or conservation issues in the voluntary sector. With respect to valuation issues highlighted by the Exxon Valdez accident or similar catastrophes, such non-hypothetical field experiments may refer to projects such as the construction of protection devices against oil spills in parts of an archipelago, using additional land for bird sanctuaries, or conserving coastal lands that government has declined to protect.⁹

9. Postscript 2.

Postscript I

The above text was written prior to the publication of the NOAA panel report in January 1993.¹⁰ Most of what I have had to say here seems to agree with what is argued in this report (in particular, that CVM tends to give upward-biased WTP estimates, but that this is no reason to consider the method useless). However, there are two points on which I disagree with the conclusions of the NOAA panel report.

First, its recommendation to use "reliably conservative" versions of CVM "as a partial or total offset to the likely tendency to exaggerate willingness to pay" NOAA panel report (p. 4610) is something that is easy to say but hard to operationalize and accomplish. If it were *known* how this could be achieved, all CV analysts would probably have liked to do it all along. But since real-world "conservative" CVM versions can hardly be trusted to be "reliably conservative" (e.g., in that they may now yield an *underestimate* of WTP), such versions tend to be even less useful since we would then not even know "for certain" the sign of the bias.

Second, like many others, the NOAA panelists are attracted by the referendum (or dichotomous-choice) approach—where people are asked whether or not they accept to pay a contribution of a given amount—since this approach lacks an obvious strategic bias. All right, but this approach opens up a Pandora's box of other problems related to—what may be called—a worsening of the hypothetical status of the questions. For example, subjects may respond differently if they know or do not know whether or not others would *have to* pay the amount in question; moreover, they might be confused by knowing (or protest if learning later that they did not know) that others are asked about other amounts, very much unlike a referendum. These problems of interpretation by respondents and/or CV consumers do not confront versions of CVM where the (hypothetical) payments are explicitly allowed to differ. It is due to the existence of these problems that my comments above referred only to open-ended versions of the CVM.

Postscript 2

A brief sketch of an example of a non-hypothetical field experiment of the type indicated at the end of the paper is given here. It concerns a study (by Per Olov Johansson, Bengt Kriström and the present author) of the valuation of wetland conservation in Sweden. This study is now underway, pending financial support from the Swedish EPA.

The Swedish government has made a list of wetland areas for protection and identified which ones it can fund for protection and

10. National Oceanic and Atmospheric Administration, Natural Resource Damage Assessments under the Oil Pollution Act of 1990, 58 Reg. 4601 (Jan. 15, 1993).

which ones it cannot, at least not for the next five years. One suitable area in the second group is chosen for the study. The main points of the study design are as follows:

(1) The area in question can be rented for protection during a period of five years (during which the government has declared it definitely cannot fund this project). The rental is SEK A.

(2) The government (the Swedish EPA) decides more or less arbitrarily that it values protection of this area during the next five years at SEK B ($B < A$), on behalf of the present and future Swedish population, *excluding the present population in a set of counties in the neighborhood of area*. (This amount of SEK B would be financed separately by the Swedish EPA, if the results of the study indicate that the conservation contract shall be purchased.)

(3) The local population (N) is asked, in a "referendum-like manner" as explained below, about their WTP for conserving this area during the five-year period. If stated average WTP is found to exceed average project costs ($(A - B)/N$), see further below), the conservation contract will be purchased and 80 percent of the contract cost will be paid by the local population.

(4) Why "80 percent"? The primary purpose of the study is to compare a CVM estimate with an estimate according to the interval method.

(I) Ten percent of the population (at least 100 subjects) will be asked an ordinary open-ended CVM question. (II) Another ten percent will be asked to state their WTP, again without having to pay anything but now being told that their responses provide part of the information on which the government bases its decision whether the contracts should be purchased or not. This group is confronted with an incentive (if any) to overstate their WTP. (III) 80 percent of the population will be asked to state their WTP and, if the outcome is that the area should be preserved, they would have to pay the percentage of the WTP stated that would cover their share (80 percent) of project costs. If so, the remaining 20 percent of project costs would be covered by the WTP test budget.

(5) The "voters", groups II and III, will be informed essentially as follows: (i) since it is difficult to elicit true WTP statements from people in a simple and direct fashion, due to misrepresentation incentives, different voters will have to be subjected to different payment consequences, some of which may lead to overstatements and others to understatements of WTP; (ii) given the group to which the individual (household) belongs, information will be provided as to what incentives are relevant here (group III, see point (6) below); (iii) if it is found that aggregate WTP exceeds project costs, subjects in group III will have to pay, as stated in their respective instructions (see point (7) below); (iv) the information problem stated in point (i) implies that voters (and

so they are told) are confronted with different, and therefore "unfair", payment consequences; specifically, a minority of 20 percent is given a free ride; a minor consolation for the majority is that the "lucky ones" will be selected at random (in a process that is open to public supervision).

(6) The incentive provided to group III is to understate WTP and this group is so informed. Some observers in the literature have argued that incentives are unclear, when the subject is committed to pay "only" a share (100 percent or less) of stated WTP. This position must be interpreted as saying that there would be an incentive for some to overstate WTP in the state of the world where they need to pay less than 100 percent of their stated WTP. However, if such a state materializes, it would mean that revealed aggregate WTP is higher than project costs, and hence, that the project *will* be carried out. If so, there "cannot" be any rational reason for overstating one's WTP, since the only reason for this would be to increase the chances of having the project accepted. Therefore, it is flatly stated to this group that there is an incentive (if any) to understate preferences in, so to speak, all possible states of the world.

(7) The project will definitely be carried out if average WTP in group III exceeds average project costs (AC). It may be carried out even if this lower bound to true average WTP is below AC, a necessary condition being that the upper bound provided by group II is above AC. This decision is up to the government (the Swedish EPA). If, in this latter case, the government decides to have the project carried out, it will be at an extra cost to the government, since the commitment by group III no longer covers 80 percent of costs.