FINAL REPORT

COOPERATIVE AGREEMENT CR 810224
Budget Period II

THE USE OF CONTINGENT VALUATION DATA FOR BENEFIT/COST ANALYSIS IN WATER POLLUTION CONTROL

Robert Cameron Mitchell
Resources for the Future

and

Richard T. Carson
University of California-San Diego

December, 1986

Submitted by Resources for the Future

to

Dr. Peter Caulkins, Project Officer
Office of Policy Analysis
U. S. Environmental Protection Agency
Washington, D.C. 20460

DISCLAIMER

Although prepared under EPA Cooperative Agreement No. CR810224-02, this report has neither been reviewed nor approved by the U.S. Environmental Protection Agency for publication as an EPA report. The contents do not necessarily reflect the views or policies of the U.S. Environmental Protection Agency, nor does mention of trade names or commercial products constitute endorsement or recommendation for use.

ABSTRACT

This study addresses EPA's needs for data and methods to value the benefits of water pollution control regulations. Previous work under this cooperative agreement has resulted in a national data set which we have used to make aggregate benefit estimates for different levels of national freshwater quality. In chapter 1, we present a summary of this study's findings and which reflects our further thinking on issues such as the nature of the benefits which the study measured and which extends our original analysis to address such topics as the distribution of the benefits and costs of water pollution control.

In chapter 2 we address the issue of how our national data might be used to value local freshwater quality changes. In it we present a valuation function from the national freshwater benefits survey. This function allows the estimation of marginal changes in water quality using the mapping provided by the RFF water quality ladder. It is based on a fairly simple, theoretically plausible, model which provides a good fit to the cross-sectional data from our survey. We then develop a technique to apply the function to estimate local benefits and apply the method to a boatable to fishable improvement in the Monongahela River system in Pennsylvania. This allows us to compare our estimate with the one obtained in a local CV study by Desvousges, Smith, and McGivney (DSM). Our aggregate estimates, based on reasonably plausible multipliers, resulted in aggregate estimates that coincided quite closely with those estimated on the basis of DSM's WTP amounts. In this chapter we also present an analysis of our data to extract a counterpart to recreation day values. The latter exercise is useful only as a further way to judge the plausibility of our findings; our "recreation day" values are not worthwhile in themselves.

In chapter 3 we make the argument that political markets, particularly referenda, are the most suitable model for CV surveys which value pure public goods. We then present the findings of an innovative CV survey of a sample of California residents which attempts: (1) to predict their vote on a water bond referendum and (2) to test a method, using double sampling techniques, to have respondents make discrete choices WTP judgments about proposed tax prices. The results of this study demonstrate the validity of a referendum model in that we were able to predict the actual vote for the proposition in the subsequent election. They also suggest that the new method we developed to obtain WTP amounts for different tax prices shows considerable promise.

TABLE OF CONTENTS

Appendices

TABLES

FIGURES

Chapter 1

THE RESOURCES FOR THE FUTURE NATIONAL WATER BENEFITS STUDY

In 1977 Robert Dorfman, after calling all available estimates of the benefits of national pollution control programs open to serious question, reluctantly turned to a 1969 Gallup survey for the National Wildlife Federation as the best source of data for estimating the benefits of these programs. Since this survey was not designed for this purpose, Dorfman characterized the estimates he derived from this data set as “perhaps suggestive, but certainly untrustworthy” (Dorfman, 1977: 336).

In this chapter, we summarize the methodology and findings of our study of national freshwater quality benefits (hereafter referred to as NWBS) which were the subject of an earlier report to EPA (Mitchell and Carson, 1984). These data, which we believe to be reasonably trustworthy, are the first to be based on a national contingent valuation study. The material presented in this chapter revises our previous report in several respects and extends it. It contains a revision of our typology of water pollution benefits, a discussion of the the distribution of the benefits and costs of water pollution control, a comparison of our estimates with the fishing benefits estimated by Vaughan and Russell (1982), among other things. In the next chapter we present a valuation function based on these data and a possible technique for using this function to value sub-national freshwater quality improvements.

BACKGROUND

Although numerous water benefit studies have been conducted in the past thirty years, they have been of limited use in estimating national water quality benefits.[1] For example, local site benefit estimates derived from travel cost studies generally do not control for water quality (Dwyer, Kelly and Bowes, 1977), nor can they measure nonuse values. Of the handful of studies which directly measure nonuse benefits using contingent valuation surveys (Gramlich, 1977; Oster, 1977; Greenley, Walsh and Young, 1981, 1982; Desvousges, Smith and McGivney, 1983; Blomquist, 1983a), all value the benefits of one particular site or river basin. Recent work by Randall, Hoehn and Tolley (1981) and Hoehn and Randall (1982) demonstrates, theoretically, that independently derived benefit estimates for sites or areas which are potential substitutes for each other can not be aggregated to obtain national benefits in a straightforward manner. Randall, Hoehn, and Tolley provide empirical evidence that performing such an aggregation may result in a gross overestimate of total benefits. We avoid this geographical aggregation problem by using a national probability sample of the American public and by asking them to value a national set of water quality improvements.

Contingent valuation (CV) uses survey research techniques to elicit people's preferences in the form of willingness-to-pay (WTP) monetary amounts.[2]

---

1. See Tihansky (1975) and Freeman (1982). Freeman notes a general tendency for studies which use the appropriate economic theory and suitable estimation techniques, such as Feenberg and Mills (1980), to lack an adequate data base for making national estimates.

2. Some of the methodological issues related to this technique are discussed in the course of this paper. Strategic behavior is not discussed in what follows, because the burden of much laboratory and field research is that it is unlikely to be a significant problem in CV studies such as this one (see Cummings et al., [1986] and Mitchell and Carson [forthcoming] for a review of this literature).

In its standard form, the CV survey describes a detailed hypothetical market in which a specified public good may be purchased and asks respondents how much of their household's current income in dollars they would be willing to give up in exchange for a specified increase in the level of the public good. Thus we ask the respondent for a direct evaluation of his or her household's compensating surplus (CS) from a change in the public good in question which can be represented as:

$$CS = [e(p_o, q_*, q_o, U_o) = Y_o] - [e(p_o, q_*, q_1, U_o) = Y_1], \qquad (1)$$

where e( ) is the expenditure function, p_o the vector of prices for marketed goods, q_* the vector of public goods which remain fixed, q_o and q_1 the initial and subsequent level of the public good being valued and, Y_o and Y_1 the initial and subsequent levels of income associated with each of the two expenditure functions. Usually the valuation question is repeated several times for different levels of the good so that a Hicksian compensated demand curve can be traced out.

Since its initial applications in the 1960s (Davis 1963; Knetsch and Davis 1966; Hammack and Brown, 1974) considerable effort has been devoted to establishing its theoretical basis,[3] developing the methodology,[4] and, where possible, comparing its estimates with those using market demand-based measures.[5] There now appears to be widespread agreement among welfare

---

3. For arguments that CV data are generated in forms consistent with the theory of welfare change measurement see Randall, Ives and Eastman (1974). Freeman (1979); and Just, Hueth and Schmitz (1982).

4. Cummings et al. (1986) comprises a major review of this work. See also Brookshire and Crocker (1981). Schulze, d'Arge and Brookshire (1981), Rowe and Chestnut (1983) and Mitchell and Carson (forthcoming).

5. Studies making travel cost comparisons are Knetsch and Davis (1966), Bishop and Heberlein (1979), Desvousges, Smith and McGivney (1983) and Seller, Stoll and Chavas (1984). Those making hedonic price comparisons are Brookshire, Thayer, Schulze, and d'Arge (1982), Blomquist (1983b). Cummings, Schulze, Gerking, and Brookshire (1986).

economists that the correct measure of benefits is willingness-to-pay and that, provided it can be administered without bias, the contingent valuation method can be used to estimate the Hicksian consumer surplus measures (Freeman, 1979; Just, Hueth, and Schmitz, 1982). If one accepts the compensating, surplus form of WTP as the appropriate welfare measure for a specified improvement in the water quality enjoyed by an individual household, and takes the current distribution of income as given, then a point on the Samuelson-Bradford bid (or benefit) curve (Bradford, 1970; Randall, Ives, and Eastman, 1974) is given by summing all households' WTP amounts for the new level of water quality. Optimal provision of water quality occurs at the point where the aggregate marginal cost and benefit curves cross.

In the presence of uncertainty, contingent valuation obtains estimates of option price, the correct ex ante welfare measure (Graham, 1981). The total value for water quality improvements measured in NWBS includes a number of categories as shown in figure 1-1.[6] Several of these, particularly the categories making up the existence class benefits, show no traces in marketplace transactions. Because of this, many environmentalists and some resource economists have criticized benefit-cost analysis on the grounds that the existence class of benefits tend to be uncounted or undercounted by benefit measurement techniques which use observed marketplace behavior.[7]

---

6. In contrast to our earlier (Mitchell and Carson, 1984) report, we do not include option value as a separate benefit category among the types of benefits measured in this study. Option value is now seen as a correction factor to techniques like travel cost and hedonic pricing which measure expected consumer surplus rather than option price (Chavas, Bishop, and Segerson, 1986; Smith, forthcoming).

7. For an interesting account of a dialog between regulators, environmental group leaders, and economists on this issue, see Swartzman, Liroff and Croke (1982). The conterence volume edited by Peskin and Seskin (1975) contains several papers by economists which explicate the difficulty of measuring nonuse benefits and the problems of conducting benefit-cost assessments if the analyst is unable to quantify these benefits.

Figure 1-1   A TYPOLOGY OF BENEFITS FROM AN IMPROVEMENT IN FRESHWATER QUALITY

Benefit Class	Benefit Category	Benefit Subcategory (Examples)
Use	In-Stream	Recreational (water skiing, fishing, swimming, boating)
		Commercial (fishing, navigation)
	Withdrawal	Municipal (drinking water, waste disposal)
		Agriculture (irrigation)
		Industrial/Commercial (process treatment, waste disposal)
	Aesthetic	Enhanced Near Water Recreation (hiking, picnicking, photography)
		Enhanced Routine Viewing (commuting, office/home views)
	Ecosystem	Enhanced Recreation Support (duck hunting)
		Enhanced General Ecosystem Support (food chain)
Existence	Vicarious Consumption	Significant Others (relatives, close friends)
		Diffuse Others (American public)
	Stewardship	Inherent (preserving remote wetlands)
		Bequest (family, future generations)

There are reasons to believe that these benefits may comprise a sizable portion of water quality benefits (Fisher and Raucher, 1984). Among benefit measurement techniques, the contingent valuation method is uniquely able to measure existence benefits since it can elicit values from both users and nonusers of a given amenity and elicit a the full range of values from the users.

DESIGN CONSIDERATION AND FEATURES

In order to obtain valid responses in a CV survey, respondents must understand the scenario and amenity being valued in the way intended by the researcher, expend the effort necessary to arrive at a considered value for the amenity, be uninfluenced by features of the scenario which are not intended to influence their values, and influenced by those which are. These factors are interrelated. Unmotivated respondents, for example, may not pay attention to important aspects of the scenario such as the description of the amenity's location or its quality level with the result that the values they give are actually for a different amenity than the one being studied. Because the national water quality program potentially is a more abstract amenity than many of the amenities valued by previous CV studies, such as the quality of local lakes or rivers, we conducted extensive questionnaire design research for this study,[8] which led to the design features we describe in this section.[9]

---

8. A first round of work, conducted in 1979-80, developed a precursor to the present instrument. This instrument was administered in 1980 to a national sample of 1576 respondents by appending it to another survey we were conducting (Mitchell and Carson, 1981). An experiment was conducted in this survey to test the payment card elicitation method for bias. The second phase of the development work was conducted by the Research Triangle Institute. During this effort our expanded draft instrument was reworked and pretested in the summer of 1984. The last round of work on the instrument was conducted by the Opinion Research Corporation and the authors on the basis of observed interviews and field testing.

The Clean Water Act of 1972 and its subsequent amendments suggest three levels of minimum national water quality which should be valued: boatable, fishable, and swimmable. Our design research indicated that these concepts were also meaningful descriptions of water quality to respondents. Matching these levels of water quality with physical water quality criteria is no easy task, nor is there complete agreement on how to do this. In our survey instrument, we used a water quality index developed by W.J. Vaughan for our 1980 pilot study which maps the rungs shown in figure 1-2 back into physical water quality parameters such as dissolved oxygen. A description of this index is presented in appendix B. Use of this ladder in the survey as a visual aid greatly facilitated the task of communicating the several quality levels to the respondents.[10] The scenario's wording emphasized the nonuniform distribution of water quality implied by the concept of "minimum" water quality. Respondents were told that although the present minimum level is boatable, most of the nation's freshwater bodies are currently fishable and perhaps 70-80 percent are swimmable. When asked to value the boatable minimum level, respondents were asked how much they would be willing to pay "to keep the nation's freshwater bodies from falling below the boatable (minimum) level where they are now." This established a "below boatable" baseline which represented the minimum level of national water quality which would occur if all present annual expenditures for water pollution control by industry and governmental entities ceased. This is also the baseline from which the U.S. Commerce Department measures water pollution control expenditures. By

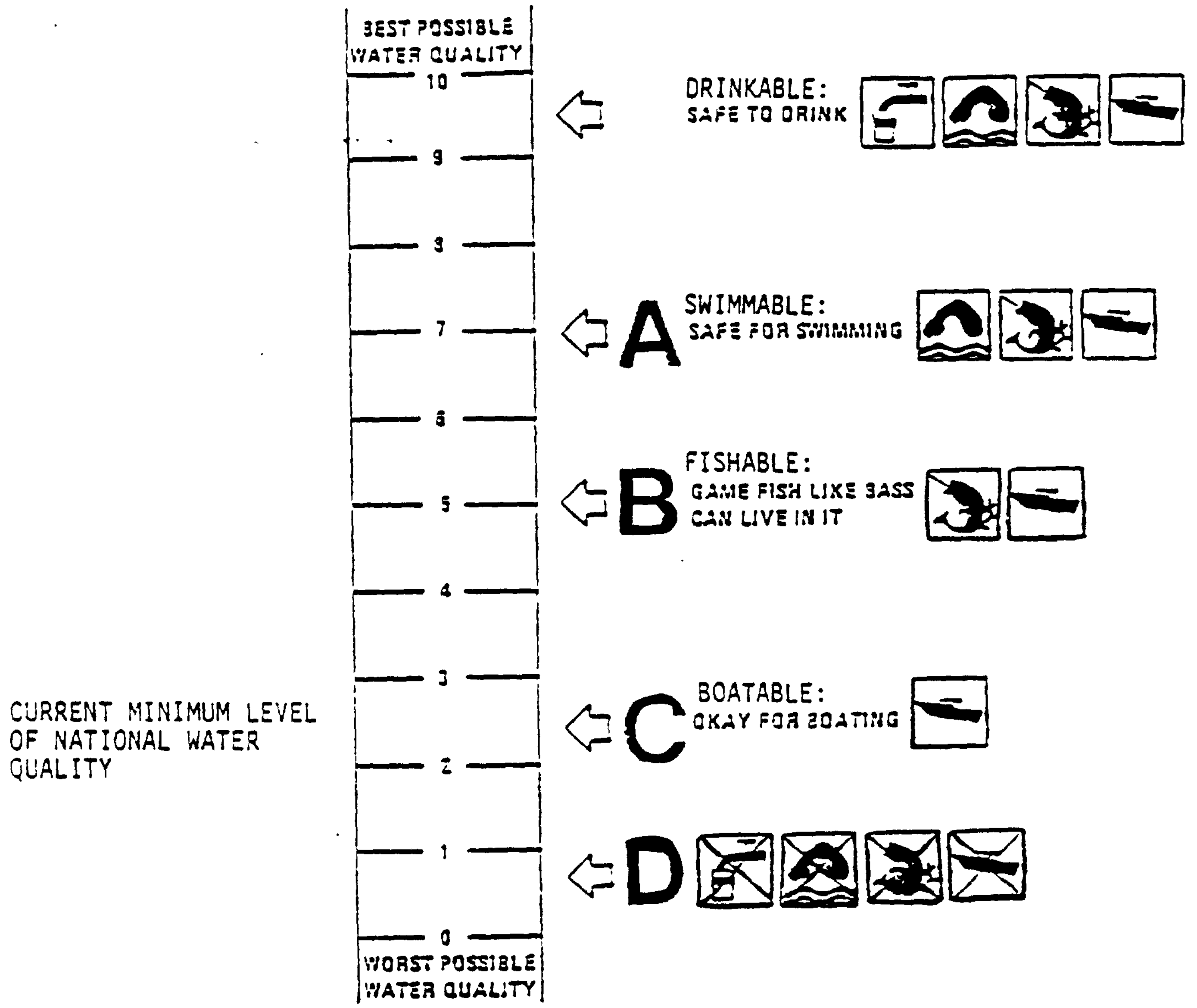
---

9. Appendix A presents the full text of our instrument.

10. Desvousges, Smith, and McGivney (1983) successfully used the ladder for the same purpose in their Monongahela River study.

Figure 1-2 RESOURCES FOR THE FUTURE

WATER QUALITY LADDER



adopting this baseline, respondents were offered the opportunity to purchase water quality improvements in a form which allows the compensating surplus-WTP measure to be used in all the valuation questions.[11]

The payment vehicle used in this study, annual taxes and higher product prices, correspond with the way citizens presently pay for water quality and was accepted by the respondents without protest as appropriate for this purpose. In an effort to avoid the starting point bias associated with the commonly used bidding game method (Randall, Ives, and Eastman, 1974), our elicitation procedure used the grounded payment card format we first developed and tested in our 1980 study. Respondents were divided into five income groups based on their household income and given a payment card containing a large array of amounts. In order to provide a meaningful context for the valuation exercise, five of the amounts on the card were identified as the amounts average households of that income group are currently paying in taxes and higher prices for nonenvironmental public goods such as defense, the space program, and police and fire protection.[12] The willingness-to-pay questions

11. This type of baseline is quite useful and plausible for public goods whose quality deteriorates in the absence of continuous expenditures. The property rights associated with public goods of this type are developed in more detail in Mitchell and Carson (forthcoming).

12. An experiment, conducted as part of our 1980 pilot study (Mitchell and Carson, 1981), tested for possible starting point bias induced by these information points and found the WTP amounts were insensitive to the number and dollar amounts of goods similar to those used in the present study. A second experiment, conducted in 1983 in a pretest for the present study, compared the use of identical grounded and nongrounded payment cards to see if the information points contribute to the quality of the data. Although the findings from this pretest are tentative, owing to the small sample size (N = 93), we found no significant difference at each level of water quality between the mean and median WTPs for the two types of payment cards. However, the interviewers strongly felt the information on the grounded card helped respondents to arrive at meaningful answers, and it appears that this information reduced the <u>unexplainable</u> variance in the WTP responses.

asked respondents to state an amount on the payment card or “any amount in between” they were willing to pay for each of the three levels of national minimum water quality.

Because our development work indicated that some respondents tended to confuse drinking water benefits with freshwater benefits, the scenario was worded to distinguish the two types of benefits. To ensure that all respondents were aware of the full range of appropriate benefits a “values” card (Desvousges, Smith and McGivney, 1983) was used which listed the major reasons why households might value water quality.[13] To promote respondent understanding of the water quality levels, additional descriptions were provided. Regarding the fishable level, for example, respondents were told that “although some kinds offish can live in boatable water, it is only when water gets this clean that game fish like bass can live in it.” The scenario also reminded respondents that they are currently spending part of their income on water pollution control, a condition they needed to understand for us to implement our WTP-compensating surplus questions. Our pretests found that a number of respondents wanted to know how much they were paying for this purpose. This created a potential problem since we could not inform them of this amount prior to eliciting their WTP amounts because of the likelihood that some would base their value on this figure instead of independently determining their maximum WTP amount. By providing this information at a later stage in the interview and by offering respondents the opportunity to revise their original WTP amounts on the basis of this information if they wished, we were able to coax reluctant respondents to give us initial values and to test the effect of providing this information.

---

13. We did not intend the respondents to take any of the commercial in-stream or withdrawal benefits described in figure 1-1 into account and it is unlikely, given the wording of the CV scenario, that they did so.

In order to provide the maximum opportunity for respondents to arrive at a considered value, we solicited a total of four WTP amounts from each respondent for each of the three water quality levels. The first bid (WTP_F) is the amount given for each of the three WTP questions (boatable, fishable and swimmable; Appendix A, questions 24, 26, and 27). The reconsidered (WTP_R) bid is the amount (whether changed or unchanged) offered after their three first amounts were repeated to them, the total was stated, and they were encouraged to make any revisions they wished (question 29). The informed (WTP_I) bid is the amount given after respondents were informed of the range of the amounts households in their income group (question 33) were actually paying for water (and air) quality. Finally, respondents were asked if they would increase their WTP amounts if their bids were not enough to reach any of the three goals, including the boatable water quality goal. The amounts given after this question (35) is the highest (WTP_H) bid.

Given the variety of measures -- three quality levels by four separate measurements -- it is useful at this point to make clear our assumptions about the nature of the amounts each of these bids elicits. We assume are that many or most of the respondents do not have a well formed value when asked in a CV survey to value a good which they are unaccustomed to purchasing. Faced with such a first-time request for such a value, some respondents are unable to offer a value. The interviewers were instructed to avoid putting pressure on these respondents to give what would only amount to meaningless values. The remaining respondents, however, know within a reasonable range where their value for the good may lie and a few may even have a good idea of the actual value. On the assumption that respondents are generally cautious (i.e., risk averse consumers) when faced with sizable purchases, we believe the WTP_R amounts are likely to represent the lower bound of their WTP range. In the case of the WTP_H amounts, where the request for revaluation could have been interpreted by some respondents as implying that they had not given a high enough WTP amount and "should" give more,[14] the WTP responses are likely to represent the range's upper bound.[15] With regards to the boatable, fishable, swimmable quality levels, respondents who attempted the valuation exercise had the least difficulty arriving at a total value for the nation's water pollution control program. Specifying the values of the intermediate goals was likely to be a more difficult task for the respondents.[16] We therefore have the greatest confidence in the accuracy of the total (swimmable) value as representing the benefits of the current program (providing the national minimum level is raised above the fishable level). The boatable and fishable levels require larger confidence intervals.

A major question in valuing water quality improvements is the shape of the benefit curve between the three goals of boatable, fishable and swimmable. If people's willingness to pay is totally contingent upon the attainment of each goal, the function is a step function, and intermediate or partial improvements would provide no additional benefits. Two questions, asked of equivalent subsamples (A and B), explored the respondent's views about water policies which promise partial improvements to one or the other of the fishable and swimmable goals. In the halfway policy question, respondents

---

14. The prospect offered respondents in the highest scenario was quite drastic -- that even the boatable level was threatened if a higher WTPB bid was not forthcoming.

15. This approach is consistent with Bohm's (1984) suggestion that CV questions be asked so as to elicit an upper and lower bound.

16. Evidence supporting this assumption is can be found in the response patterns. A number of respondents gave equal dollar values for each of the three levels, a response strategy consistent with uncertainty about how to value the intermediate improvements.

were asked (version A) if they would still be willing to pay their revised amount for swimmable "if the best we could do was to raise the minimum only halfway from fishable to swimmable." The 95 percent question (version B) asked respondents if they would still be willing to pay the fishable amount if "five percent of the nation's water bodies remain at the boatable level...The lakes, rivers and streams comprising this five percent would all be located in heavily industrial and/or urban locations where a lot of people live." Both of these questions will be useful in examining modifications of the current provisions of the Clean Water Act.

With a public good such as water quality which is unevenly distributed geographically, policy makers are interested to learn the extent to which respondents value provision of the good outside their home area. Pretests showed the most widely understood definition of a home area for a survey such as ours was the respondent's state. After being reminded of their total revised bid ($WTPTOT_R$), respondents were asked how many dollars or what percent of this amount they would give to their state and to the rest of the nation for water improvement.

It is sometimes believed that a respondent who is asked a value for a particular water quality level unwittingly values, instead, a more general package of environmental improvements. To minimize this type of bias[17] we explicitly asked respondents to keep in mind that no matter what amount they give for water pollution control they will also continue to pay for the nation's other environmental programs such as air pollution and "air quality will remain at its present level or improve slightly." To test for the presence of this bias, an experiment was conducted in the present study using

17. Policy-package part-whole bias (Mitchell and Carson, forthcoming, chapter 10).

equivalent subsamples. At the point in the interview where respondents were told what they are actually paying for pollution control, those in subsample A were only told the amount for water quality control whereas respondents receiving treatment B were given the amounts for both water and air pollution control. If part-whole bias was present in the WTP_R bid, we hypothesize that version B respondents would disproportionately reduce their WTP_I amounts to compensate for their previous overspending of their environmental account.

According to the national area probability sampling plan used in this study, personal interviews were conducted at 61 primary sampling points in the contiguous United States by the Opinion Research Corporation using experienced professional interviewers. Each interview took approximately 40 minutes. The response rate was 79 percent of eligible respondents with a total of 813 people being interviewed.[18]

FINDINGS

Of the original 813 interviews, 564 or 70 percent yielded usable WTP amounts. The remainder consisted of 72 don't knows (9%), 18 refusals to answer the WTP questions (2%), 133 protest[19] zeros (17%), 16 inconsistent

---

18. Complete details of the sampling plan and its execution can be found in Mitchell and Carson (1984).

19. Protest zeros are zeros given by respondents who object to some aspect of the scenario, such as paying for the good by the specified vehicle, or who fail to understand the hypothetical market. They were identified by a series of eleven followup questions asked of each respondent who gave a zero bid. If the respondents said they gave a $0 bid because that is what the level of water quality is worth to them or because they lack enough money to pay anything, their WTP amount was coded as $0. Respondents giving protest zeros tend to be of two types: (1) those who are generally hostile towards government taxes and expenditures, and (2) very strong environmentalists who believe it is immoral to place a dollar value on environmental amenities. The first category appears to be somewhat larger than the second.

(too high) responses (2%), and 10 inconsistent (too low) responses (1%).[20] Given the degree of interest and effort involved in answering complex CV scenarios such as the one used in this study, this level of item response in a national sample, while high, is acceptable,[21] provided the estimates are adjusted to take into account the fact that the nonrespondents were not a random subset of the sample. These adjustments are discussed later in the paper.

Unadjusted WTP Amounts

Table 1-1 presents the unadjusted WTP amounts for each of the four series of bids measured in the study. Using the reconsidered series of bids, the respondents who gave usable responses were willing to pay $106 annually for maintaining boatable quality water ($WTPB_R$), $80 more to reach the fishable minimum water quality level ($WTPF_R$), and an additional $89 to move from the

---

20. Responses judged inconsistent (too high) were those which exceeded 5% of the household's income while those judged inconsistent (too low) were WTP amounts of less than $5.00 (usually $1.00) given by respondents with above average to high incomes whose answers to attitude questions showed strong support for water pollution control expenditures. The "too low" responses may be regarded as protest zeros which missed being identified because the token positive amounts given by these respondents removed them from our protest zero screen. One possible explanation of the "too high" values is that they represent strategic behavior. The fact that they were mostly given by respondents with low educational levels suggests thoughtless rather than strategic behavior.

21. The present study's 70 percent response rate to the WTP questions represents a forty percent improvement over the 1981 study's 50 percent rate. This improvement occurred despite the fact that we increased the standards for accepting a WTP answer as valid. Based on an examination of response rates to other difficult questions in national surveys, we speculate that this may be close to the upper limit for the usable response rate in random samples asked to value complex public goods, and that efforts to increase this rate might well result in lower quality data by inducing respondents who lack genuine opinions to give values nonetheless.

Table 1-1 MEAN UNADJUSTED ANNUAL HOUSEHOLD WILLINGNESS TO PAY AMOUNTS FOR DIFFERENT LEVELS OF NATIONAL WATER QUALITY BY TYPE OF BID

Water Quality Level	First Bid (F)	Reconsidered Bid (R)	Informed Bid (I)	Highest Bid (H)
Nonboatable to Boatable (WTPB)	$111 (10;$40)*	$106 (10;40)	$125 (11;48)	$141 (13;50)
Boatable to Fishable (WTPF)	80 (8;30)	80 (8;30)	96 (9;35)	108 (10;50)
Fishable to Swimmable (WTPS)	89 (12;25)	89 (12;25)	102 (12;25)	116 (13;25)
Total WTP (WTPTOT)	280 (25;125)	276 (25;120)	323 (27;150)	366 (29;150)
Number changing their bids at each stage		75	104	136

*(Standard error of the mean; median), N = 564.

fishable minimum quality level to a national minimum of swimmable quality water ($WTPS_R$) for an unadjusted total ($WTPTOT_R$) of $276. An examination of the changes made by the 75 respondents who reconsidered and revised their amounts after giving their WTP_F amounts shows that most of them corrected mistakes caused by misconceptions about the elicitation process. In the WTP_I iteration, 104 or 18 percent of the respondents revised their bids after being informed of the approximate level of their current payments for water quality (water and air quality) improvements. Those changing bids tended to be respondents who discovered that they were actually paying more money than their previous (reconsidered) amount and wished to increase their WTP_R amounts. Of those who discovered that they were actually paying less than they said they were willing to pay, few reduced their earlier bid. Finally, those who changed their amounts in the last iteration -- after being confronted with the assertion that the amount they had previously committed themselves to might not be enough to maintain even the present minimum level of water quality (a strong statement) -- tended to be of two types: (1) respondents whose informed bid was still below their current payments and (2) respondents who already had a reconsidered bid much higher than their current payments. Overall, taking those who made multiple changes into account, approximately 30% of the respondents changed one or more of their WTP amounts. Of these respondents about a third changed more than once.[22]

As noted above, we believe the WTP_R series represents the most valid basis for estimating WTP (after adjustment for nonresponse). The informed

---

22. The mean bids for the first and reconsidered conditions were not significantly different except for WTPB. Each of the other two revision opportunities generally resulted in mean bids which were significantly higher from their predecessors. These tests are given in Mitchell and Carson (1984).

and, in particular, the highest series of WTP questions put a significant amount of social pressure on the respondents to increase their willingness to pay and should be viewed as upper bounds.

Test for Policy-Package Part-Whole Bias

The results of our test to see if respondents erroneously valued a broader pollution control policy which includes air pollution are reassurring. None of the t-tests of the differences in the willingness to pay for any water quality level between subsample A, (who were told what they were paying for water quality) and subsample B, (who were told the amounts they are currently paying for both air and water quality improvements) are greater than .75.

Geographical Allocations of WTP Amounts

Regarding the respondents' geographical allocation of their WTP amounts, they allocated an average of 67 percent of their $WTPTOT_R$ amount for water quality improvements to be spent in their state and 33 percent of this amount to be spent out-of-state. The median in-state percent (70%) was almost identical to the mean. Only one person out of three wanted all of their WTP amount spent in-state. The percentage of in-state benefits was positively correlated with the number of years lived in state and age, and negatively correlated with education, income and recreational use of out-of-state water. While these data show significant state benefits, the level of out-of-state benefits is consistent with a strong federal role in water pollution control.

Existence Benefits

Also of interest here is the determination of what portion of the average respondent's WTP amount is for the existence benefits discussed earlier and displayed in figure 1-1. Fisher and Raucher (1984) suggest that one way of indirectly estimating the lower bound nonuse or existence benefits for water quality, as a percentage of total WTP, is to dividing the sum of WTP by nonusers in the sample amount by the sum of the entire sample's WTP. Doing this, nonuse benefits appear to be substantial. We note, however, that this estimate is sensitive to how nonuse is defined. For example, when nonuse is defined as no instream recreational use of freshwater by the respondent in the past 12 months, existence benefits calculated by this procedure amount to 39 percent of the total WTP amount. When nonuse is extended to include everyone in the respondent's household, existence benefits amount to 30 percent of the total. Finally, if nonuse is defined as no direct or indirect (e.g., picnicking, camping, duck hunting etc. by freshwater) activities by anyone in the household, an absolute lower bound for existence benefits of 19 percent is indicated.

PARTIAL IMPROVEMENTS

According to the answers to the "halfway" and "95 percent" questions, the benefits of partial improvements are considerable. Almost nine out of ten (89 percent) of those who answered the question said the 95 percent improvement from boatable to fishable was worth the same to them as the complete improvement.[23] Those who wished to pay less for the partial improvement were disproportionately residents of large urban areas. This is understandable because the question informed respondents that the "lakes, rivers and streams comprising this five percent would all be located in heavily industrial and/or urban locations where a lot of people live." Each person who was unwilling to

---

23. Which we defined as where 99 percent or virtually all the nation's lakes, streams and rivers would be fishable.

pay the same amount was asked how much he or she was willing to pay for this partial improvement. The WTP amount for raising 95 percent of the nation's water to at least the fishable level is $74, or 8 percent less than the $80 $WTPF_R$ amount for raising 99 percent to at least this level. Turning now to the halfway improvement question which was asked of subsample A, we find a somewhat lesser percent (73 percent) willing to pay the same amount for the halfway improvement from fishable to swimmable as they were for the total improvement to the swimmable level. Because those who were not willing to pay the same amount were willing to pay a somewhat greater percent for the partial improvement than in the 95 percent case, the overall reduction in $WTPS_R$ for swimmable water quality is slightly less (6% of $WTPS_R$).

It is possible to compare these estimates of the benefits of the 95 to 99 percent fishable water partial improvement with a recent estimate made by Vaughan and Russell (1982) using a participation-travel-cost model. Vaughan and Russell valued the benefits accruing to fishermen from improving national freshwater so that all waterbodies are at least at the fishable quality level. This improvement is equivalent to raising three to five percent of the waterbodies from quality levels of less than fishable to fishable quality,[24] an increase quite similar to the 95 vs. 99 percent improvement we asked our respondents to value. It might be expected that our estimates should be somewhat higher than Vaughan and Russell's due to the more inclusive nature of our benefits. On the other hand, it is likely that recreational benefits dominate the benefits of the 95 vs 99 percent improvement in our survey, since

---

24. The definition of fishable water is slightly different in the two studies. The water classified as fishable by Vaughan and Russell included some water which was capable of supporting "rough" fish, such as catfish, but not bass. In our study. fishable water was defined as supporting bass.

the 95 percent level provides for a large number of available substitutes and is likely to fulfill many people’s stewardship needs. Vaughan and Russell’s estimate of the benefits for this improvement range from 200 - 1200 million (1983) dollars with 500 million dollars as the best rough point estimate. Considering the difference in methods and data bases, this amount is quite similar to our 490 million dollar point estimate.[25]

DISTRIBUTION OF BENEFITS AND COSTS

Baumol and Oates (1979) have noted that studies of the distributive effects of environmental policy are still in their infancy despite the crucial importance of the equity issue for environmental policy. Based on their review of the then available literature on distributive benefits, they raise the possibility that the less affluent may believe that environmental improvements come at their expense. Baumol and Oates cite poll data as evidence for a “consistent pattern of disproportionately strong support for environmental programs among higher-income groups” (Baumol and Oates 1979: 184). One of the major advantages of the CV method over other benefit estimation techniques[26] is the ease with which it provides information on the distribution of the benefits for the program being valued, thus permitting the identification of losers and gainers when cost is considered.

Table 1-2 presents the distribution of water quality benefits for five broad income categories. In absolute terms, average willingness to pay for water quality increases sharply with income; the respondents in the highest

---

25. Which is derived by multiplying the average reduction in $WTPF_R$ by the number of 1983 United States census households.

26. Harrison and Rubinfeld (1978) discuss the difficulties of estimating the distribution of benefits using the hedonic pricing approach.

Table 1-2. WTP_R FOR WATER QUALITY UP TO SWIMMABLE LEVEL WATER AND IMPORTANCE OF NATIONAL GOAL OF PROTECTING NATURE AND CONTROLLING POLLUTION BY INCOME GROUP

Household Income	N	Mean	Std. Error	Median	As % of Income	National Goal of Protecting Nature and Controlling Pollution Very Important*
under 10,000	125	$ 61	$ 6	$ 35	.90	60%
10,000-19,999	154	171	16	100	1.18	71
20,000-29,999	130	225	20	150	.92	66
30,000-49,000	97	422	45	270	1.13	63
50,000 and over	41	1154	281	600	1.32	66
All Respondents**	564	276	25	120	1.05	66

* Question wording: "Some national goals are more important to people than others. How important to you personally is a national goal of protecting nature and controlling pollution? Is it very important, somewhat important, or not very important to you."

**Including those who did not give their household's income.

income category are willing to pay almost 19 times as much, on the average, as those in the lowest income category. As a percent of income, however, willingness to pay is quite even across income categories, a finding consistent with the close to unitary income elasticity for water quality. This finding is also consistent with the broad-based pattern of support for environmental goals and the environmental movement which became apparent in numerous public opinion polls in the late 1970s (Mitchell, 1979) and the 1980s (Council on Environmental Quality, 1980: Ladd, 1982; Mitchell, 1984), and in the distribution of responses presented in table 1-2 to a question in our present survey which asked respondents how important to them personally is a "national goal of protecting nature and controlling pollution." As indicated there, at least 60 percent of every income group said such a goal is "very important" to them personally with only modest (and insignificant) differences between the high and low income groups. It thus appears that demand for environmental quality in general, and for improved water quality in particular, is broad based although the monetary benefits are subject to strong income constraints.

Two fairly recent studies of the distribution of water pollution control costs (Lake et al., 1979; Gianessi and Peskin, 1980) found the costs tend to be mildly regressive overall and especially regressive at the lower income levels,[27] because these costs are paid largely through sewer fees and higher prices for a number of basic consumer goods. The inequitable character of

27. Comparisons between the two studies are somewhat difficult because of differences in their baselines and demographic projections. However, both show the lowest income group is paying more than twice the percent of their income toward water pollution control than those in the highest income groups. The regressive impact of water pollution control costs is mitigated somewhat by the federal sewage treatment plant program. Control costs for air pollution are more regressive (Gianessi and Peskin. 1980) owing to the absence of comparable federal programs.

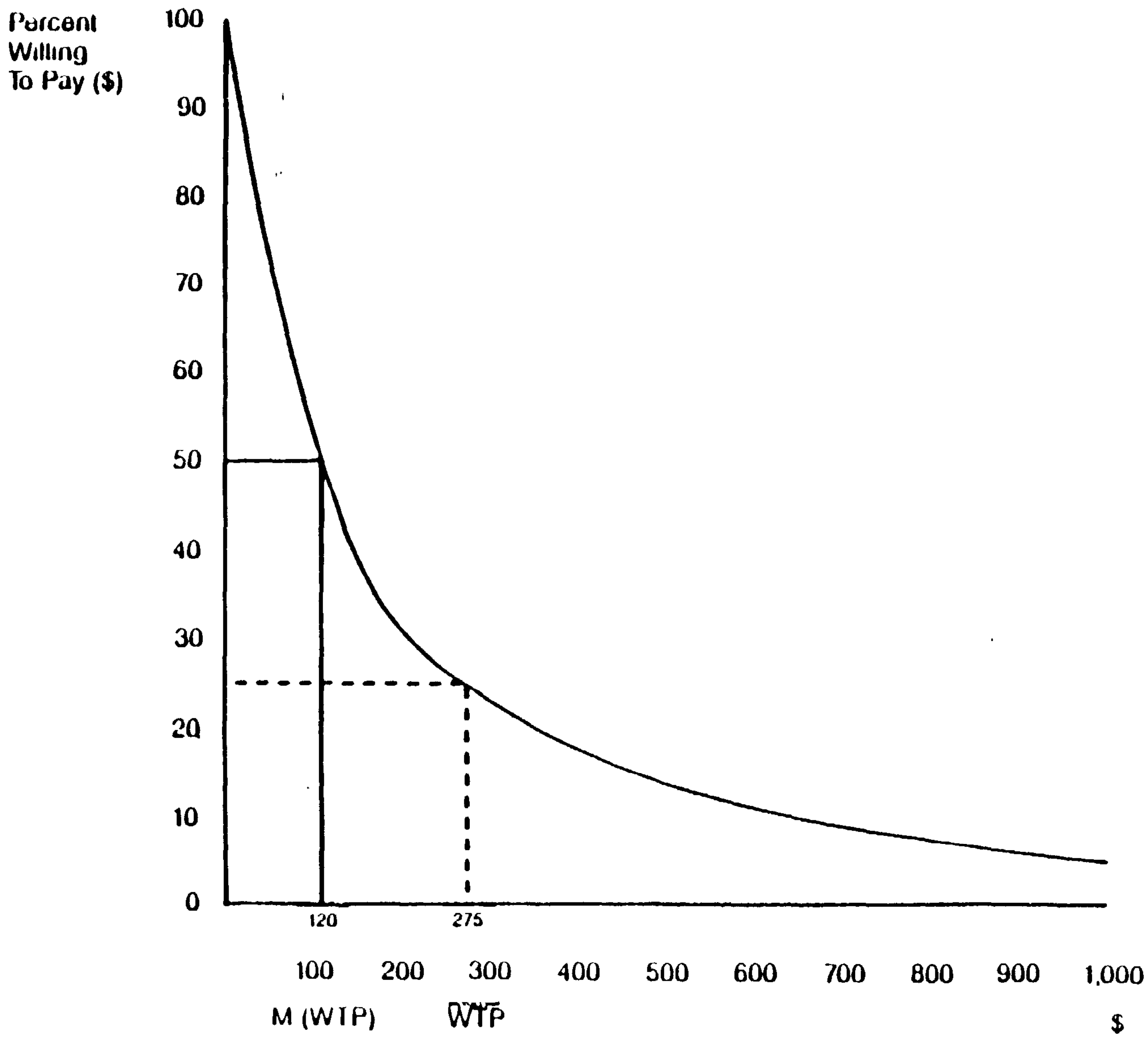
present water pollution control policies is made more apparent by treating our CV survey as an analogue of a voter referendum. If the referendum was on a flat tax, the median voter would rule, and $120 is the maximum annual amount that would be approved by a majority. If the referendum proposal was for a progressive tax, with each of our broad income groups paying the median $WTPTOT_R$ amount for that group, the indicated overall average payment is $164. Both of these amounts are far short of our sample mean of $276, although it should be noted that our income categories are fewer than the income brackets on which differential tax rates are based and therefore may underestimate the amount that would be approved by an ideal referendum on a progressive tax. A less regressive distribution of net benefits would be achieved if a larger (than current) portion of the costs of water pollution control is collected by income or other progressive tax.[28] A more complete convergence of the incidence of costs and benefits toward the Lindahl solution would also require increasing the amount collected for water pollution control from recreational user fees, because water users account for a disproportionate number of those in each income group who are willing to pay more than their group's median WTP amount.

Figure 1-3 shows the distribution of willingness to pay on a graph of dollar amounts by the percent of the public willing to pay that amount or more for the level of the public good in question. This is simply a plot of one minus the cumulative distribution function for the willingness to pay responses, smoothed to eliminate discontinuities at respondents' favorite numbers. It shows visually the sizable difference between median WTP and mean WTP mentioned above.

---

28. Nonpoint source controls and subsidized sewage treatment plants are the primary direct federal expenditures on water pollution control.

Figure 1-3 PERCENT WILLING TO PAY SPECIFIED AMOUNTS FOR A FIXED QUANTITY OF NATIONAL WATER QUALITY



ADJUSTING FOR SAMPLE SELECTION AND ITEM NONRESPONSE

One of the most serious but generally unrecognized problems in CV surveys is the need to adjust the data to compensate for bias introduced by the inevitable failure to interview every person selected for the sample and by the failure of some persons who are interviewed to give valid answers to the WTP questions. Item-nonresponse is a particular problem in studies such as ours which use comparatively elaborate scenarios. To a large degree this is a sample selection problem in which respondents with low education are unable to answer the questions asked of them and in which respondents who are very distrustful of the government refuse to give usable answers.

Before estimating aggregate benefits from our data using the WTP_R (revised) series, we adjusted these data for item nonresponse and sample selection bias in two phases.[29] In the first, which imputed WTP values for the thirty percent of the respondents with missing or invalid WTP values, we assigned each observation with a valid value for $WTPTOT_R$ to one of six categories ordered by $WTPTOT_R$ and used CART, a tree structured classification procedure recently developed by Breiman et al. (1984), to estimate a classification tree. This tree is given in figure 1-4. The square boxes in the tree represent terminal nodes and were used as the imputation classes. Each observation with a missing/invalid value For $WTPTOT_R$ was classified into one of the terminal nodes according to a series of binary splits based on the values of different independent variables. The missing $WTPTOT_R$ values for these observations were imputed by randomly assigning values to these observations taken from that node's pool of valid $WTPTOT_R$ values.

---

29. The procedures used in this section as well as a number of alternative methods are described in Carson (1984).

While CART is a very powerful non-parametric technique which has much to recommend it in situations where economists are currently using logit or probit,[30] the feature which is crucial for our purposes is its surrogate splits. These identify the alternate splits which can be used in place of the optimal split. For example, the first binary split in the CART tree in figure 1-4 shows that households with an income of less than or equal to $15,000 go down the tree to the left and those with a greater income go to the right. What if, as is the case with our data, a large number of the respondents who failed to answer the WTP questions also did not answer the income question? CART solves this problem by estimating the splits on the other variables which best mimic the optimal income split and using these splits for the observations for which data on the optimal variable are missing. In our example, age is the best surrogate split variable and observations with a missing income value are accordingly sent left or right on the basis of age.[31]

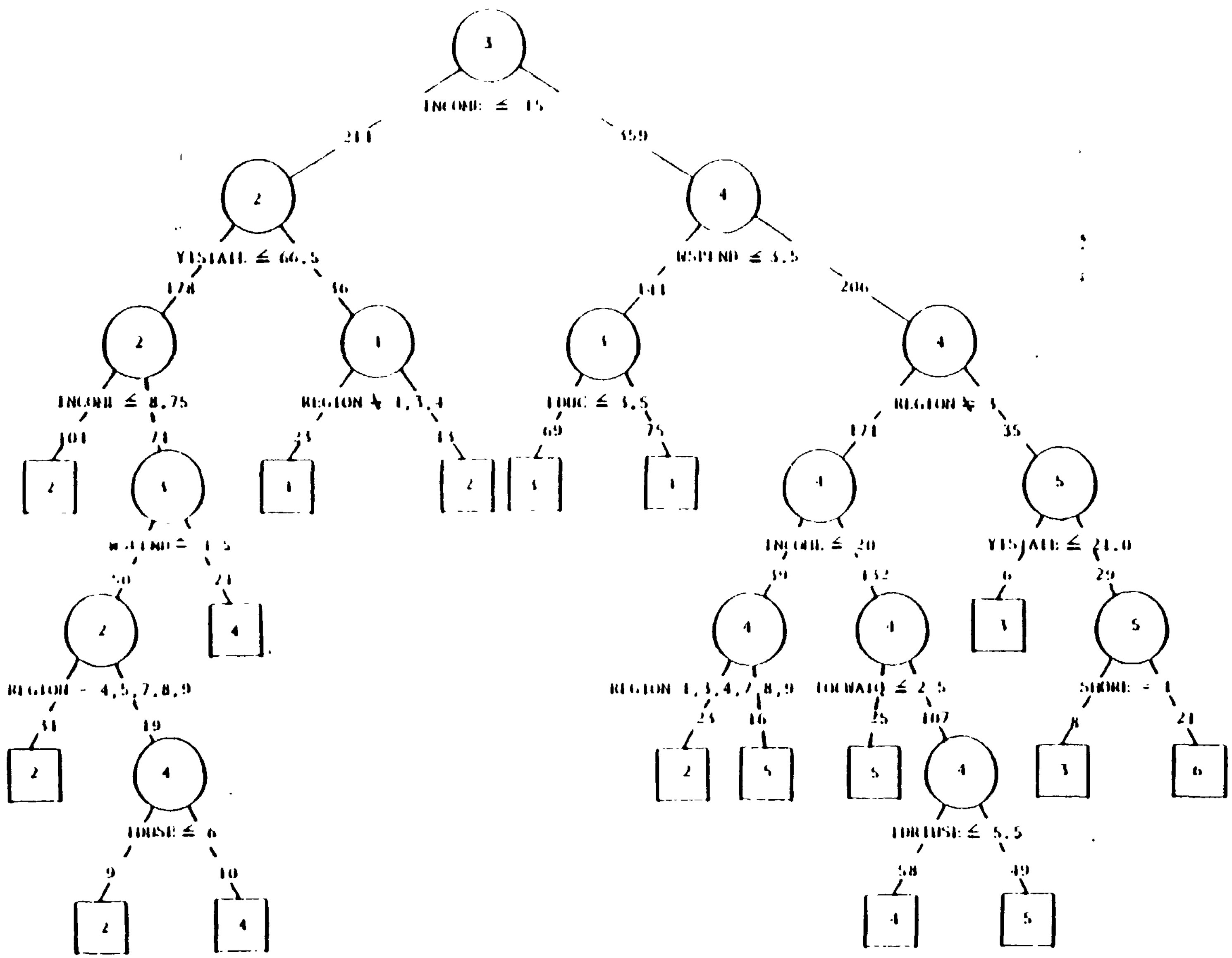
In the second phase, we corrected for response rate bias by using the household weights supplied by the Opinion Research Corporation to weight the observations to make the sample more representative of the Census population.

---

30. CART attempts to minimize a given lose function with respect to a binary split among the values of possible predictor variables. This optimization is carried out by means of an exhaustive search which is made possible by fast sorts and some recent mathematical results which show that the whole search need not be carried out in a number of common special cases. The loss function used in our case was to incur one unit of loss for each category away from the true category an observation was classified. Thus an observation whose true class was four and predicted class was a six incurs two units of loss. A problem with such trees is overfitting through chance patterns as the number of observations becomes small. This problem is solved by using 10-fold cross-validation to prune the tree upward (See Breiman et. al. 1984, for a discussion).

31. CART also provides useful information about the structure of the public's willingness-to-pay. Thus, although the tree in figure 1-4 is in general agreement with our regression results, it suggests complexities which otherwise would not be apparent and would be difficult to model in a regression framework.

Figure 1-4 CART CLASSIFICATION TREE USED TO OBTAIN IMPUTATION CLASSES FOR NATIONAL WATER QUALITY BENEFITS STUDY



The [ ] categories are: 1: WTPTOTR $0-25; 2:$26-74; 3:$75-149; $4:$150-249; 5:$250-499; 6:$500+.

Predicted class label in ○ and [ ] 's

As is typical in national probability sample surveys, women are somewhat overrepresented in our unweighted sample of respondents and young black males are underrepresented. A combination of household weights and imputing the missing values reduces the adjusted $WTPTOT_R$ value by 12 percent with each of the two correction techniques contributing approximately equally to this reduction.[32] This scale factor was applied for consistency to the rest of the WTP_R series as shown in table 1-3.

AGGREGATE BENEFITS AND COSTS

We can now assess the aggregate benefits implied by our data and compare them with current and projected costs of water pollution control programs. Commerce Department estimates (Farber, Dreiling, and Rutledge, 1984) place water pollution control expenditures in 1982 at $22.2 billion (1983 dollars) and project them to be approximately the same in 1983.[33] According to the most recent Council on Environmental Quality forecasts (1979), total pollution control expenditures in 1987 will be 52 billion (1983) dollars. The more

---

32. Two more common methods of imputing missing values, using the mean values based on "hot deck" imputation classes developed from combinations of the demographic variables and maximum likelihood imputation, resulted in very similar values in the adjusted $WTPTOT_R$ -- $246 and $237 respectively. Thus, if the mean value is the primary concern, the choice of how to impute the missing values is not critical. The method of using an _ad hoc_ combination of demographic variables does not use all of the available information in the data set and the EM maximum likelihood procedure is very sensitive to the normality assumption. The non-parametric CART procedure avoids both of these problems and provides an informative picture of the problem's structure.

33. These expenditures are currently purchasing national water quality levels where most lakes, rivers and streams are somewhere between fishable and swimmable in quality. A small number of rivers and lakes, mostly near urban and industrial area, are only of boatable quality.

Table 1-3. ADJUSTED ANNUAL HOUSEHOLD VALUES FOR BEST ESTIMATE OF NATIONAL WATER QUALITY BENEFITS*

	Mean	Standard Error of the Mean	95% Confidence Interval
WTP_R (Boatable)	$93	$8	$77-109
WTP_R (Fishable)	70	6	58-82
WTP_R (Swimmable)	78	9	60-96
$WTPTOT_R$	242	19	205-279

recent but less comprehensive[34] U.S. EPA (1984) estimate for 1987 is 41 billion.[35] These markedly higher projected expenditures are due largely to moving from BPT (best practical technology) to BAT (best available technology) standards and the implementation of nonpoint source controls on agricultural and urban runoff.

We did not attempt to value (figure 1-1) the withdrawal categories, commercial fishing, or marine recreation. These benefit categories are generally believed to be small in relation to those we measured and in most cases are more amenable to valuation using more traditional techniques. Recent estimates of these benefits (Feenberg and Mills, 1980; Freeman, 1982) range from 5 to 25 percent of the total benefits of reaching the swimmable water quality goal.

Our best estimate of the benefits of achieving the national swimmable water quality goal from a baseline of nonboatable water is 20.3 billion dollars a year.[36] Extreme bounds of 17 to 45 billion dollars a year in benefits can be developed by taking the lower 35 percent confidence interval for $WTPTOT_R$ and the upper 95 percent confidence interval for $WTPTOT_H$. We feel the high end of this range substantially overstates the possible benefits and believe a more reasonable range is 19 to 30 billion dollars a year.

---

34. The estimates are less comprehensive in that they do not include some expenditures not required by the 1972 Water Quality Amendments. In comparable terms, the EPA estimates represent a five to eight billion dollar reduction over the older CEQ estimates.

35. The reader should be warned that expenditure estimates such as these are subject to a number of problems chief among which are how to value the services of capital goods used in pollution control and how to assign costs to particular legislation or sets of regulations. Portney (1981) provides a useful discussion of these issues.

36. For the benefits categories measured in our survey. This amount is obtained by multiplying the adjusted $WTPTOT_R$ by the 83,918,000 1983 census households (Bureau of Census, 1984).

DISCUSSION

The findings presented in this chapter indicate the potential benefits of water quality improvements in the nation's lakes, rivers and streams are large and in excess of the current costs of the program. They are not, however, sufficient to cover the anticipated increase in expenditures necessary to improve the remaining waterbodies to the national swimmable water quality goal. Our data show that most of the potential benefits could be realized if the current national water quality goals are modified by designating a few industrial waterways to remain at the boatable quality level and by setting a national minimum water quality goal for the remainder of the nation's freshwater bodies at slightly below the swimmable level (as long as a high percentage of the nation's water is swimmable). In this case, at least, the public appears to understand the nature of marginal benefits.

As the first benefit analysis of a major government program using a national contingent valuation survey, this study also illustrates some of the method's complexities and capabilities. Despite its apparent simplicity, the task of designing and implementing a valid and reliable CV instrument is neither a simple nor inexpensive undertaking. This is especially true of studies such as ours which value a geographically dispersed amenity and use a random sample which includes both users and nonusers. Three examples of its capabilities can be cited. A necessary condition for choosing the optimal policy instrument(s) for pollution control (e.g. standards, effluent charges, and marketable permits) is to have knowledge about the shape of the benefits function.[37] In addition to easily providing such curves, contingent

---

37. Bohm and Russell (1985) provide a review and synthesis of the relative performance and properties of these instruments under different conditions.

valuation is unique in its ability to incorporate intrinsic benefits into such curves and to explore their shape in areas where policies have not previously existed. Elected officials and bureaucrats have been reluctant to move from the *status quo* toward incentive compatible pollution control instruments (Bohm and Russell, 1985) without detailed information about who the gainers and losers are for particular policy changes. The contingent valuation method readily provides this information about the gainers in a form which policymakers understand. Finally, no competitor method can measure benefits under alternative conditions (such as attaining the national swimmable minimum standard vs. attaining that standard for all but a small percent of the nation's freshwater bodies) with the efficiency made possible by the use of multiple scenarios in a single CV survey.