AN EXPERIMENT IN DETERMINING WILLINGNESS TO PAY FOR NATIONAL WATER QUALITY IMPROVEMENTS

Robert Cameron Mitchell & Richard T. Carson

EPA Report #
June 1981
DRAFT

AN EXPERIMENT IN DETERMINING WILLINGNESS TO PAY FOR NATIONAL WATER QUALITY IMPROVEMENTS

by

Robert Cameron Mitchell
Richard T. Carson
Resources for the Future
1755 Massachusetts Ave., N.W.
Washington, D.C. 20036

(Grant #R 806906010)

Alan P. Carlin
Project Officer
Office of Strategic Assessment
and Special Studies
RD676, West Tower 919
U.S.E.P.A.
401 M St., N.W.
Washington, D.C. 20460

DISCLAIMER

Although prepared with EPA funding, 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.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Executive Summary

The purpose of this research is to develop a methodology for using direct interview survey techniques to estimate national benefits from freshwater water quality improvements. In particular, this study has developed a method for estimating the intrinsic benefits of water quality, a class of benefits which include option, existence, and bequest benefits among others. The method also measures consumer recreational benefits, but does not estimate industrial, commercial or drinking water benefits.

To accomplish our purpose we adapted the contingent valuation or willingness-to-pay (WTP) survey method for use in a national survey. We first developed and pre-tested a research instrument which measures how much people are willing to pay each year in taxes and higher prices for national water quality of three levels which we defined as "boatable," "fishable," and "swimmable" (Chapter 3). This instrument was then further tested in a full scale survey where it was administered by professional interviewers to 1576 people comprising a nationwide probability sample. For experimental purposes, four equivalent sub-samples were asked different versions of the instrument. We performed extensive analysis on the resulting data to determine the extent to which the biases associated with WTP surveys were present (Chapter 4). With one exception, the item nonresponse rate, the results are very favorable.

Because the purpose of our empirical work is to test, validate and further develop the methodology, we do not attempt to derive national estimates from these data. We do, however, develop illustrative estimates for our cases which suggest aggregate benefits within the range of current national expenditures on water pollution control (Chapter 5). Our technique for estimating

intrinsic benefits involves identifying those respondents who do not use water for "in-stream" recreation and using their WTP amounts as surrogate for intrinsic benefits, Our calculations, again illustrative rather than definitive, suggest intrinsic benefits comprise roughly 40-60 percent of the overall WTP benefits (Chapter 5).

On the basis of these empirical tests and our concurrent work on several important theoretical and conceptual issues relevant to water benefits analysis (Chapters 1-2), we conclude that the use of a national survey to measure water benefits (including intrinsic) is a feasible undertaking, We specify the changes in our pilot instrument and its administration which will enable it to perform this task at acceptable levels of reliability and validity (Chapter 6).

The following are some of the major findings of this study in more detail:

- In the course of this project a number of theoretical and conceptual problems inherent in the direct interview survey method were clarified and further developed. In particular, work was done on consumer surplus measures (p.1-13ff), property rights (1-21ff) and the classification of different types of benefits resulting from water quality improvements (1-46ff). A number of conceptual problems arose which were closely integrated with the theoretical issues. These revolved around ill-defined property rights and the unworkability of willingness to accept compensation questions, WTA. Our conclusion was that theoretical considerations and survey considerations must both be considered in the design of WTP

instruments. Thanks to the recent work of Randall and Stoll (1980) and Brookshire, et al. (1980), however, we show that any theoretical impurity resulting from the balancing of these two considerations need not bias the results as the correct theoretical measures are derivable from the appropriate survey measures. Our conclusions on this question are summarized in Table 1.3 on p. 1-23.

- The most innovative aspect of this study is the development of a "macro" WTP approach to benefits estimation. Previously, WTP surveys had been used primarily to assess willingness to pay for locally defined goods ("micro"). For reasons specified in the report, water quality benefits lend themselves to macro WTP estimation at the national level, however. Our macro approach represents the first time, to our knowledge, that a national sample was surveyed for benefits estimation purposes on their willingness to pay for a public good. The development of this macro approach required the design of several specialized research instruments such as the water quality ladder (A-II) and non-localized benefits questions.
One clearly advantageous aspect of the macro approach is that, if correct sampling procedures are used, individual willingness to pay for water quality can be directly and reliably aggregated to the national level, The sampling techniques used to accomplish this aggregation were implemented in the survey used in this project and are described in Chapter 4 (4-22ff) and Appendix V.

- Our pretest showed the traditional bidding game format resulted in respondent fatigue and a serious starting point bias problem. To overcome this problem we developed the anchored payment card (3-14ff). To test for bias induced by the payment card, its format was systematically varied and three versions of the instrument were administered to separate sub-samples. As this experiment showed no evidence of bias, the payment card is a promising technique for WTP studies which wish to avoid the bidding game format.
- Strategic and hypothetic biases are of concern to economists who desire to use benefits derived from willingness-to-pay surveys. Our major conclusion here is that strategic and hypothetic are not opposite sides of the same bias as had commonly been assumed in the WTP literature, but comprise two separate and distinct potential biases. Table 4.3 on p. 4-22 shows the relationship of the two biases and which WTP question characteristics are necessary to minimize their effects. We further suggest and apply to our data several tests for the presence of strategic and hypothetic biases. These tests suggest that strategic bias is not a problem in our study. Our findings with respect to hypothetic are mixed because of an item nonresponse problem. However, regression equations estimated in Chapter 5 strongly suggest that those respondents who did answer the WTP items did not do so in a random fashion; one of the requirements for the absence of hypothetic biases.

- The item nonresponse problem consists of a high level of no response to the WTP questions (38 percent) and a relatively high level of zero amounts (16 percent). This problem may be attributed to the circumstances of the interviews (it was not possible to provide the interviewers with special training or instructions for this test as would be the case in a full scale implementation of the method and the WTP questions were asked after respondents had answered a half hour's worth of questions for another study) and the question wording (a too strong incentive was offered to the respondents to say water quality wasn't worth anything to them) (4-49ff). Improvements in the method,as suggested in Chapter 6, should reduce this problem to manageable proportions. Recommendations are made for weighting procedures (6-6) which can adjust for the remaining missing data.
- In order for WTP benefit estimates to be credible, a theoretically sound predictive model must be constructed and tested, We have hypothesized the primary determinants of willingness to pay amounts for water quality to be: income, water use, and environmental attitudes. To measure these and several secondary determinants, we chose items from the long environmental survey which preceded the WTP survey. Econometric estimation of this model (5-15ff) strongly supports our theory, The estimates are robust and highly significant (Table 5.5, p.5-21). A special test for heteroskedasticity appropriate for equations with both interval and dummy data was developed for this estimation and successfully implemented (Appendix VIII).

Preface

This study represents one product of several which Resources for the Future has prepared under a Cooperative Agreement with the United States Environmental Protection Agency for "Methods Development for Assessing Economic Benefits of Water Pollution Water." The particular methodological approach which we adopt in this report, a macro willingness to pay survey, emerged as we studied the problem. It builds on a tradition of innovative research using the willingness to pay methodology which extends back to the 1960s and which has flourished during the 1970s as economists have grappled with the challenging task of measuring benefits. Our use of the method diverges from this young tradition in several important respects, however, and in this sense is innovative and experimental. In the course of changing our methodology we also have found it necessary to address a number of generic methodological, conceptual and theoretical issues pertaining to benefits estimation. The fruits of our thinking on these issues is also contained in this report. In this area, too, we are building on the work of our predecessors.

To state a truism: benefits estimation is a difficult and challenging enterprise. Several years ago, Robert Haveman, commenting on a paper which analyzed 60 benefit studies, declared: "To me, the situation is ... extremely discouraging, because, in my view, what has passed for benefit estimates in these studies forms a catalog of what not to do in cost-benefit analysis" (Haveman, 1975). In our endeavor to avoid joining this infamous roll of abortive or misguided benefit studies we hewed as close

as possible to the six methodological criteria set forth by A. Myrick Freeman III in The Benefits of Environmental Improvement (1979a;10-12) and to his dicta:

> Part of the art of benefit analysis involves sensitivity to the gap between the ideal and the available and knowing how much confidence to place in the estimates being generated. (1979a;13).

To help the reader to evaluate the extent to which we have succeeded in this task we provide as much information as possible in this report about the possible biases in our method and how we have tried to overcome them. In the case of the major problem which we encountered, item nonresponse, we describe in detail the procedures which we believe can resolve the problem in a future application of the method.

The structure of our report follows from this approach. In the first two chapters we discuss crucial theoretical and conceptual matters. Chapter 3 describes our research instrument. The next chapter describes the potential biases which threaten the validity and reliability of our findings and our success in dealing with the problems they present. Finally, in Chapter 5 we present our findings. Ever mindful that benefit estimates take on a life of their own, however weak their methodological and conceptual basis may be, we offer our findings only for what they are: experimental data to test a method. Our findings are suggestive, but only a full scale application of a revised instrument can produce estimates of sufficient reliability to use for policy purposes. The final chapter discusses the nature of the revisions we propose.

With the necessary disclaimer that we alone are responsible for the work reported here, we wish to gratefully acknowledge the assistance of our colleagues here at RFF and elsewhere. We have benefitted from discussions with Ralph d'Arge and David Brookshire of the University of Wyoming, George Tolley of the University of Chicago, Alan Randall of the University of Kentucky (while on leave at Chicago) and Alan Carlin, our project monitor at EPA. At RFF, Raymond Kopp and Michael Hazilla offered us much useful counsel on statistical and econometric problems as has William J. Vaughan. Vaughan also prepared our water quality ladder, helped us refine our theoretical and conceptual ideas and collaborated with Carson on a much needed constructive test for heteroskedasticity. Clifford S. Russell was extremely generous with his time and skillful, as always in raising awkward questions and wielding an editor's pen.

Robert Cameron Mitchell

Richard T. Carson

June 15, 1981

Chapter 1

THE WILLINGNESS TO PAY METHOD, CONSUMER SURPLUS AND WATER QUALITY BENEFITS

In valuing environmental amenities, benefit-cost analysts try to ascertain what individuals would be willing to pay and/or would have to be paid for a particular public investment in a world where markets were pervasive. In such a world the prices for marketed goods would convey information sufficient to ascertain what "the gainers and losers from some public investment will consider equivalent in value to their respective gains and losses" (Brookshire, et al., 1979:33). Since a world like this does not exist for public goods such as the quality of the nation's freshwater streams, rivers and lakes -- the subject of this report -- the shape of the demand curves for these goods cannot be determined directly and economists have been forced to develop techniques to infer the value of these goods. According to Freeman (1979:4) there are three approaches to determining the values individuals place on improvements in environmental quality when markets fail or are nonexistent: (1) holding a referendum on proposals for alternative provisions of environmental quality, (2) using market data for substitutes or complements of the environmental quality being studied in order to determine the demand curve for the environmental quality, and (3) direct questioning of individuals about the value of environmental quality to them personally. The first method, referendums, have not been used in determining national policy on any environmental quality issues and few legislatures run on platforms of specific provision of an environmental amenity.

The second method is the determination of the demand curve for environmental quality indirectly through its relationship with a market good. This technique has been used extensively, particularly in the area of recreation. Examples of the indirect estimation technique include: (1) the determination of substitutability in household production functions,

(2) the travel cost method which assumes that a complementary position exists between travel cost and enjoyment of environmental quality and (3) hedonic pricing which assumes that environmental quality is not a pure public good and that a consumer can substitute (trade) market goods to obtain more or less of the environmental amenity. (Property values are usually used).

Each of these three methods of using market generated data has limitations[1] which are unique to the method. In addition they all suffer from the common inability to estimate the demand for benefits which are strongly separable in utility functions,[2] a characteristic which severely limits their utility for water benefit estimations. Freeman (1979b), for example, suggests that environmental amenities which are not directly associated with private good consumptions are separable from a utility function standpoint. Existence value certainly meets this criteria and thus is probably a separable component of a consumer's utility function, Cicchetti and Freeman (1971) argue that some forms of option value are probably strongly separable. Hence most of the water pollution control benefits we will later define as "intrinsic" and which are a primary subject of this report are not capable of being estimated by means of these techniques.

---

[1] See Brookshire, et al. (1979), Freeman (1979a) and Feenberg and Mills (1980) For critiques.

[2] Strongly separable utility functions take the form:

$$U = V[U^1(X) + U^2(Y) + U^3(Q)]$$

where X and Y are subsets of marketable goods and Q is the public good. Changes in Q have no effect on the marginal rates of substitution of any of the marketable goods. For a discussion of separability condition in utility functions with respect to public goods see Freeman (1979a:70-78) or Mahler (1974).

The third approach,which is the one employed in this study, uses the direct technique of asking people in surveys what they are willing to pay or to accept for specified levels of the public good. The use of surveys, as Brookshire, et al. have argued at length, allows the analyst to shortcut the problems inherent in the indirect method by "positing a world of pervasive markets in a form totally consistent with theoretical models of valuation for public goods" (1979:28). Most uses of the WTP method,[3] including ours, limit themselves to hypothetical markets where no money or goods actually change hands. In a couple of intriguing instances, however, researchers have used the method in the context of a simulated market. One case involved subjects paying the amount they bid to see a closed circuit TV program (Bohm, 1972); the other one measured hunters' willingness to accept money for Canadian geese hunting permits by paying them the amounts they were willing to accept in exchange for a surrender of the permit (Bishop and Heberlein, 1980). The simulated market technique has little applicability to most environmental goods because it requires exclusion from the benefit (not seeing the TV program; surrendering the hunting license), a situation which is inconsistent with how public goods such as air and water are actually provided or how it is possible to provide them in an experimental situation,

[3]We use WTP for convenience, as the method properly refers both to people's willingness to pay (WTP) for a public good or their willingness to accept (WTA) compensation for the imposition of a public bad.

This Study

The objectives of this study are to design and validate a method which can: 1) measure the national benefits of freshwater water pollution control to consumers and 2) determine what portion of these benefits come from in-stream recreational values (e.g. boating, fishing) and what portion from the intrinsic or non in-stream recreational values (existence, option, aesthetic, etc.). Very little empirical work has been done on the latter objective and no previous study has measured the former using the WTP method. Our method employs a questionnaire to ask a national sample what they are willing to pay for national[4] water quality of specified levels: boatable, fishable and swimmable.

We adopt the willingness to pay method because it is the only one of the three valuation methods which can be used to estimate intrinsic benefits. It has the significant added advantage that willingness to pay results obtained for a national probability sample of respondents may be straightforwardly blown up to give national benefit estimates. Studies using an indirect method, when based on specific sites, present a problem in this regard, for aggregation from single, or even a few, sites to the nation as a whole involves problems of definition and computation. (See, for example, the companion report by Vaughan and Russell under this cooperative agreement.) This method is not without its problems too, which we will discuss at considerable length insubsequent chapters. For

[4] All the previous uses of the WTP method to estimate the benefits of environmental public goods were limited to local or regional studies. For reasons we will discuss in subsequent chapters, the characteristics of national water quality and its benefits are such as to make a national WTP survey a feasible and desirable undertaking.

the moment we should simply note that the methodological requirements for a successful WTP survey are formidable. Not only must the instrument describe the hypothetical market in a manner which meets the requirements of economic theory, it should also be understandable to respondents with less than a high school education. The sampling and field work must meet high standards, and the sample size should be large enough to permit reliable estimates. The fit between the respondent's experiences and the hypothetical situations described in the questions must be close enough to render the situations meaningful to the respondents.

In this chapter we discuss briefly the willingness to pay method of benefits in the context of economic theory and of the types/which accrue from water pollution control. Our purpose is to clarify the theoretical basis for our measurements and to review the literature on intrinsic water benefits.

Benefit-Cost Analysis

The purpose of this study is to estimate certain benefits resulting from raising the ambient level of fresh water quality in the United States. These benefits are one side of a benefit-cost analysis and may be defined in terms of the (Hicks-Kaldor) Pareto optimality conditions (Mishan, 1976a)[5] which allow for the possibility that those who gain in utility by a change in state can compensate those who lose utility as a result of the change. In our case,where water quality is assumed to be a normal good, benefits are the largest amount of the numeraire the individual is willing to pay to obtain a given higher level of water quality, while costs are the smallest amount that those producing the water pollution are willing to accept for reducing their pollution enough to achieve the specified level of water quality. This can be expressed in terms of utility for consumers and producers.

$U(W^{o},Y) = U(W^{+},Y - B)$ for consumers

$U(W^{o},Y) = U(W^{+}, Y + C)$ for producers[6]

where

W^{o} = the initial provision of good W

W^{+} = a higher level provision of good W

Y = income or all other goods (numeraire)

B = the amount of Y consumers are willing to pay to obtain W^{+}

---

[5] Benefit-cost analysis has long recognized that decision makers should consider criteria other than economic efficiency in implementing a policy, in particular distributional issues. These criteria are not considered in this study. For a discussion, see Mishan, 1976a.

[6] Since this is a study of benefits rather than cost we will not consider production cost and producers surplus and their associated problems (See Mishan, 1976a).

C = the smallest amount of compensation that producers are willing to accept to reduce their pollution enough to achieve W+

These definitions can be seen to be those of the Hicksian (1956) compensating measure of consumer surplus, a topic which we will shortly address at greater length. Following Mishan (1976a, 1976b) we assign a minus sign to cost and a plus sign to benefits and aggregate over consumers and producers, The standard benefit-cost equations for a change from one state to another can be expressed in terms of the Hicksian compensation measure as follows:

(1) $\Sigma B^c - \Sigma C^c > 0 \simeq \Sigma CM > 0$

(2) $\Sigma B^c - \Sigma C^c < 0 \simeq \Sigma CM < 0$

(3) $\Sigma B^c - \Sigma C^c = 0 \simeq \Sigma CM = 0$

Where

B^c = Total benefits of the proposed change

C^c = Total costs of the proposed change

CM = Hicksian compensation measure

The discussion thus far has been deceptively simple. We now need to address the complications which arise from the special characteristics of public goods, the nature of public policy, and the limitations of the survey WTP method. These matters are a necessary background to the resolution of the debate over exactly which consumer surplus measure is most appropriate for measuring the benefits of environmental amenities.

Public Goods and Public Policy

Public goods such as national levels of water quality are those which once produced can be supplied to everyone at zero marginal cost and whose enjoyment by one person does not interfere with the enjoyment of another. Furthermore, individuals cannot easily be excluded from enjoying the benefits of the public good, once it is produced. These goods are normally produced as a result of government action, either by government requiring firms or individuals to produce the goods or by government directly subsidizing this production from tax revenues. Once produced, public goods are usually provided free. In the case of water quality Congress declared its intent in the Federal Water Polltuion Control Act (1972) that all freshwater bodies reach fishable and swimmable quality by 1983. Private firms now have to clean their water discharges to meet government regulations, and the federal government subsidizes the major portion of a waste water treatment plant construction program for local governments.

For goods which are provided through markets, individuals are always free to optimize by trading along their budget lines in order to reach the highest indifference curve possible. In this situation, measuring the consumer surplus is a straightforward problem. This is not the case for national water quality, however. First, since "clean water" is a public good, it is provided free to citizen consumers who wish to boat, fish, water ski or simply contemplate it. As such it is available at any given time only at the quality level provided by government policy irrespective of whether some consumers are willing to pay more for higher water quality. In the case of national freshwater this quality level consists of two

factors: a) the ambient quality level (boatable, fishable, swimmable, etc.) and b) the amount of the overall stock of freshwater which is mandated to reach a specified quality level, Thus if the government had set a boatable water quality standard for all freshwater, those who wished to have a higher standard for the nation's water (e.g. fishable) would have no way to obtain it short of changing government policy.[7] Even if this were not the case, it would still make no sense to use survey techniques to ask consumers how much they were willinging to pay for the quantity and quality of national freshwater they regard as personally optimal. Let us say person A might be willing to pay $339 a year for national water of fishable quality and person B $400 for boatable quality water. Once having obtained data in this form, however, it is impossible to aggregate the WTP amounts to get a national benefit estimate for any water quality but the highest/level for which WTP amounts are available. That is, we can reasonably count B's amount for boatable water as the amount which he would also be willing to pay for the higher, fishable, level which A

---

[7] This is an oversimplification, of course. Many public goods, fresh water included, are also available privately at a cost. Naturally, national water quality of a certain level can only be provided by the government. But a consumer faced with the absence of public lakes and streams of fishable water quality in his or her locality may be able to obtain access to private water of that quality for a fee of some kind. The existence of numerous private swimming pools, clubs and beaches attests to the widespread use of privately supplied water for swimming, The availability and desirability of these optional sources of water presumably influence the value people place on the public supply of freshwater.

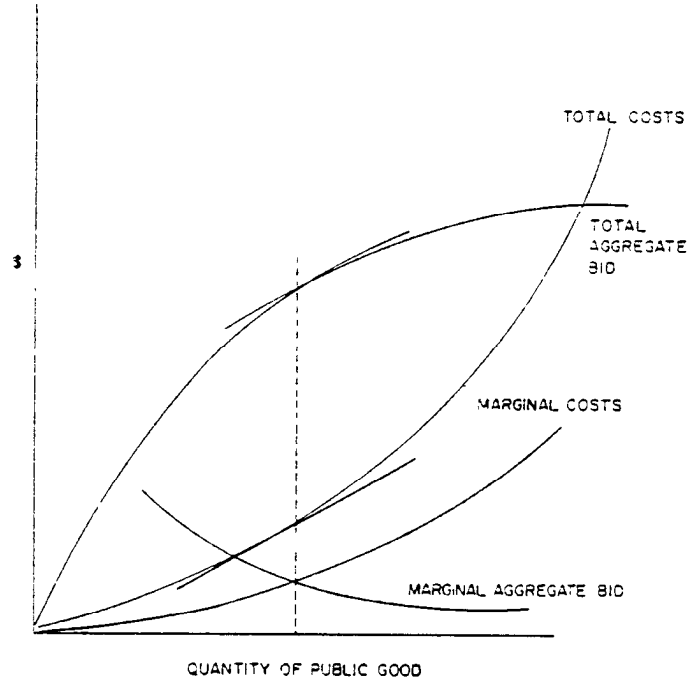
regards as optimal, giving a total WTP amount of $739 for fishable water. We cannot, however, reverse the process and determine what A would be willing to pay for any level of water quality below his optimum. He might be willing to pay most of his $339 for water of boatable quality or he might not be willing to pay anything for water of such inferior quality,

The irrevelance of the consumer's willingness to pay for his or her optimal personal provision of the public good greatly limits the range of consumer surplus measures which are appropriate for the study of national water benefits. David Bradford (1970), in an expansion of Samuelson's (1954) early demonstration that the demand for a public good is the vertical summation of individual demand curves, takes these factors into account in developing his theoretical framework for the valuation of public goods in benefit-cost analysis. This framework and its subsequent expansion by Randall, Ives and Eastman (1974) has been the theoretical basis for most of the WTP surveys. Bradford makes the assumption that individuals choose between various bundles of goods which may differ in quantity and quality and proposes the concept of an aggregate bid/benefit curve (more recently referred to as the total value curve) which he defines as the vertical summation of the individual bid curves. Because of this assumption, Bradford was able to demonstrate that over any relevant range, the aggregate bid curve and its corresponding marginal bid curve (demand curve) need not be continuous or downward sloping. If the aggregate cost is known and the marginal cost curve is derivable, the Bradford framework

resembles the traditional profit maximization framework with the optimal production occurring where the marginal aggregate bid curve and the marginal aggregate cost curve intersect (See Figure 1.1). What is being optimized here is total welfare or utility rather than profits. This intersection can be shown to be the point where the rate of commodity substitution equals the rate of technical substitution which is the traditional welfare economics position necessary for Pareto optimality (Bradford, 1970; Henderson & Quant, 1971). Consumer surplus is usually used as the measure of the aggregate benefit curve.

This caveat should be added. If a unidimensional scale (underlying metric) is unknown or does not exist, it will be impossible to estimate the demand or supply curves for the public good. This means that only specific levels of production can be compared with the initial level or with other specified levels. This is, however, not as serious a problem as it might appear since policymakers almost always choose between a limited number of alternative policies, the benefits of which can be measured in the framework we present.

Figure 1.1 COLLECTIVE OPTIMIZATION OF THE QUANTITY OF PUBLIC GOOD PROVIDED*



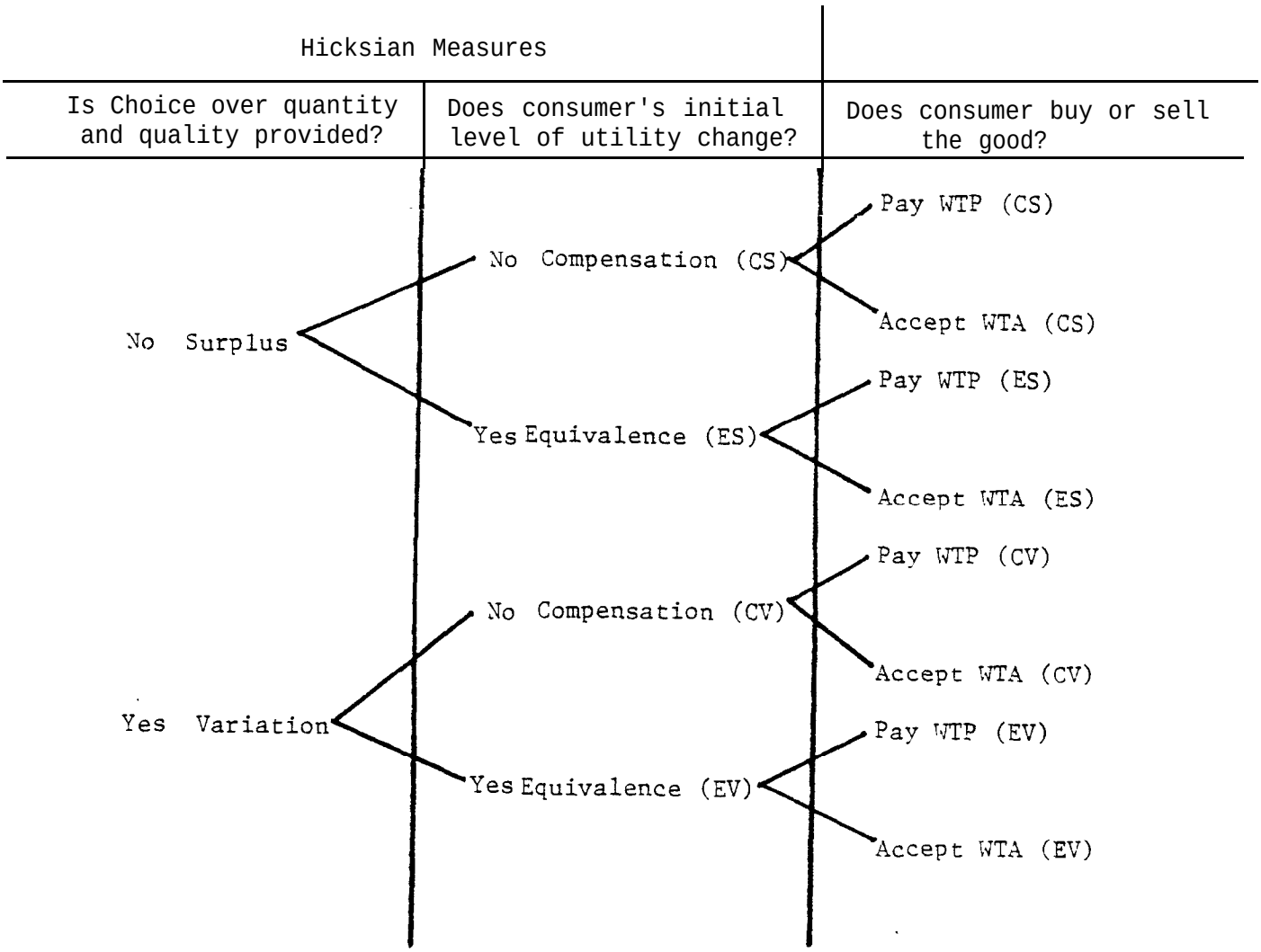
*From Randall, Ives and Eastman (1974).

Consumer Surplus

The concept of consumer surplus has been the subject of considerable debate among economic theorists (Curie, et al., 1971) and among those who use the concept i benefit-cost analysis it has been the subject of some confusion until recently. Consumer surplus was at the center of the welfare economics of Marshall and Pigou. After a period of neglect, it became a point of contention between two eminent theoretical economists; Paul Samuelson and John Hicks. Samuelson (1947) argued that consumer surplus was a subject of "historical and doctrinal interest, with a limited amount of appeal as a mathematical puzzle," a view echoed more recently by Silverberg (1978) who charged that "attempts to use consumer surplus to measure welfare losses are largely the application of the inappropriate to measure the undefinable." Hicks, on the other hand, argued strongly that consumer surplus is useful to welfare economics and his view has come to prevail amongst those who conduct benefit-cost analysis.

Hicks (1941, 1943, 1956) in a series of works beginning with The Revision of Consumer Surplus (1941) and concluding with The General Theory of Demand (1956) redefined the concept in an attempt to overcome the objections to the Marshallian version. He developed four definitions of consumer surplus which become eight when both price increases and decreases are taken into account. These measures are set forth in Table 1.1. The distinction between the surplus or variation measures depends on whether the consumer is allowed in response to price change to adjust his or her purchases to optimize his or her consumption/(variation) or whether the consumer is simply offered fixed quantities of a particular good (surplus), The second set of distinctions depends upon whether the

Table 1.1 TYPES OF CONSUMER SURPLUS MEASURES FOR CONTINGENT VALUATION STUDIES



consumer's reference point is his or her initial level of utility or not. In the compensation type, the individual moves along the indifference curve determined by his or her present utility. In the case of the equivalence type, the individual moves from a point on one indifference curve (his or her initial utility) to a point on another indifference curve, Thus the equivalence measure always represents either a gain or a loss in utility.

Since none of these measures fulfills the need for a single concept to measure welfare loss or gain from various price or quantity changes, analysts have to choose which of them meets the requirements for their particular case. Mishan, for one, in a series of writings (1947, 1960, 1971, 1976a, 1976b) argued that the Hicksian compensation variation measure is the appropriate measure of welfare gain or loss if a potential Pareto improvement is being considered. He further argued that the variation form rather than the surplus form is the correct measure of consumer surplus. Mishan went so far as to drop all discussion of the compensation surplus measure in his later works including his influential book, Cost-Benefit Analysis (1971, 2d ed. 1976a). The choices between surplus and variation, and compensation and equivalence, were much discussed during the 1970's as analysts conducting the WTP surveys tried to determine which consumer surplus measure is most appropriate for the case of non-marketed environmental goods, the property rights for which are ill defined and which are provided to consumers in fixed quantities. The appropriateness of measures involving paying for the good (WTP) versus accepting compensation for it (WTA) was also discussed and tested empirically during this period. We conclude

from our review of these discussions and experiments that the most appropriate measure of consumer surplus for WTP surveys is the compensation-surplus WTP measure and that when methodological considerations preclude the use of questions in this form, the equivalence surplus WTP measure should be used,

Surplus vs. Variation

Let us address the surplus vs. variation choice first. Mishan relegated the surplus form to the dust heap, a position taken implicitly by others, most recently Daniel Feenberg and Edwin S. Mills in their book Measuring the Benefits of Water Pollution Abatement (1980). As we have shown above, however, our case of well defined levels of water quality fits the model of lumpy goods which Randall and Stoll (1980) have shown require the use of Hicksian surplus measures. Since our case is typical of many environmental amenities,the surplus measures are appropriate for most WTP surveys because only they measure people's willingness to pay for fixed quantity/quality bundles of public goods.[8]

The Surplus Measures: Definitions

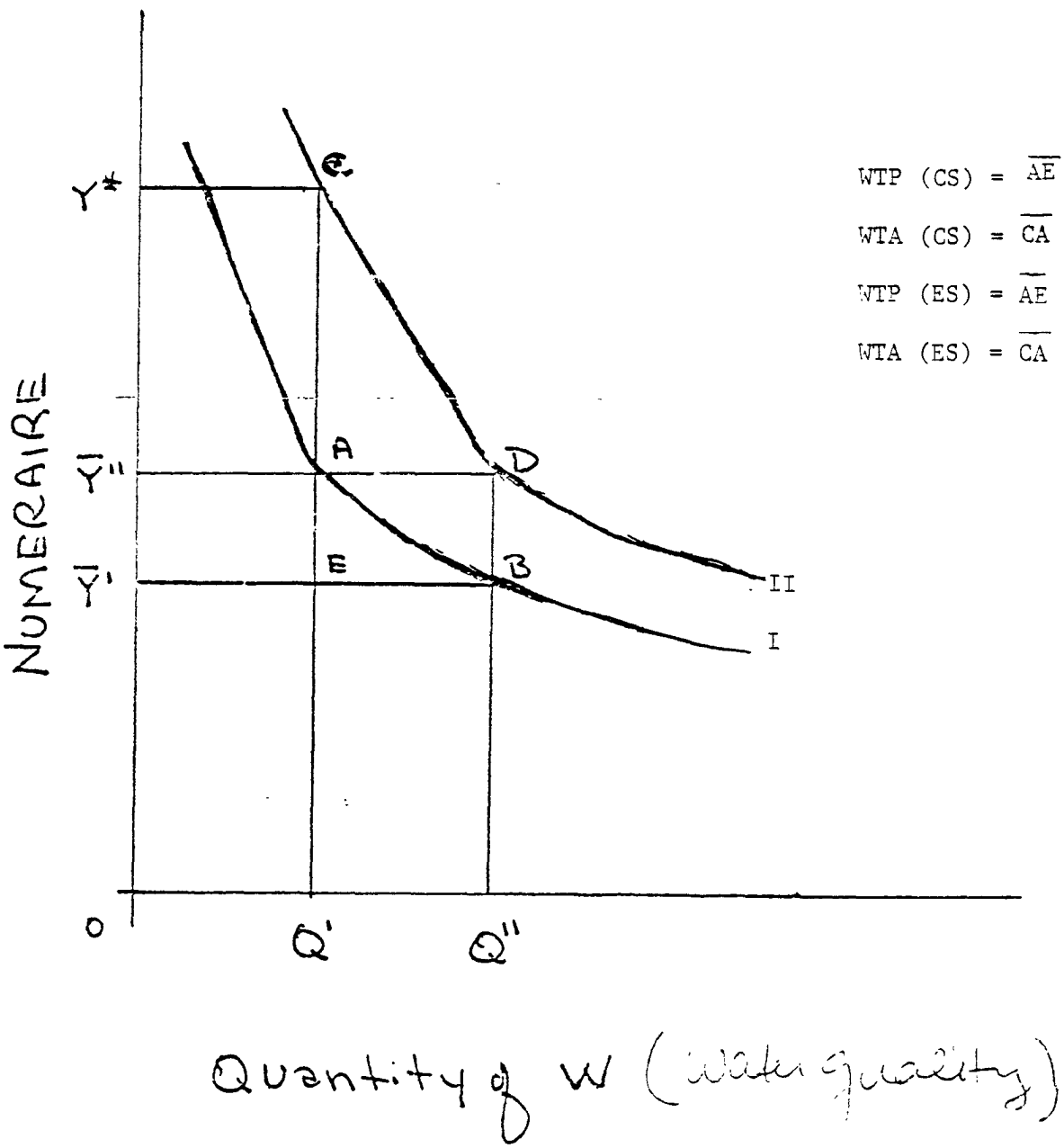
This leaves four measures of consumer surplus as the object of our concern. Before proceeding further let us define these in words and identify them graphically as follows:

[8] Freeman (1979b), after correctly distinguishing the variation and surplus measures according to Hicks' definitions, inexplicably ignores this distinction when he argues that if people are only offered fixed quantities of goods the compensating variation measure is equivalent to the compensating surplus measure and hence one only needs concern himself with the variation measures.

- WTP (CS) --The maximum amount a consumer is willing to pay to obtain a prespecified level of W (e.g. water quality) and have his or her utility remain the same as it was initially,
- WTA (CS) -- The minimum amount a consumer is willing to accept for having W decline to a prespecified level without changing his or her utility.
- WTP (ES) -- The maximum amount a consumer is willing to apy to avoid having W lowered to a prespecified level; either the change in water quality or the payment will lower the consumer's utility.
- WTA (ES) -- The minimum amount a consumer is willing to accept to forego a promised increase to a prespecified higher level of w. Either the payment or a higher level of W will increase the consumer's utility level.

On Figure 1.2, if the initial position is A, and the prespecified improvement is Q' to Q", then WTP^{cs} is the amount of Y represented by the line segment $\overline{AE}$ WTP^{es} is the amount Y represents by the line segment $\overline{AE}$. The reduction in utility is accomplished by moving the consumer from D(Q" on II) to A(Q' on I). The consumer is then indifferent between trading $\overline{AE}$ amount of Y to get back to Q", the original endowment of W. WTA^{cs} is the amount of Y represented by the line segment CA. The consumer in this case is moving along indifference curve II going from Q" to Q' in exchange for AC of Y. WTA^{es} represents an increase in utility, To make this example parallel with the the WTP^{es} measure, the consumer will move from A(Q' on I) to D(Q" on II) and is asked how much Y would it take to move back to his or her original endowment of Q but remain on indifference curve II. That quantity shown on the graph is $\overline{CA}$.

Figure 1.2 HICKSIAN SURPLUS MEASURES



From Figure 1.2 it is readily apparent that the two willingness to pay measures are equal and that the two willingness to accept measures are equal. Further, it is apparent that the WTA measures are not income bounded. Without proof (which may be found in Willig, 1976; Randall and Stoll, 1980; and Brookshire, Randall, and Stoll, 1980) we cite the following useful generalizations about the relationship between the measures: (1) WTP $\leq$ Marshallian Consumer Surplus (M) $\leq$ WTA, (2) for the case of zero income elasticity of income for the public good, all of the Hicksian measures are equal and are also equal to the Marshallian (M) consumer surplus, and (3) when income elasticity (price flexibility of income for the good)[9] is small (generally less than 1) and/or WTP (WTA) is small relative to income (generally 5% or less) the bounds between WTP and WTA have been rigorously defined and are usually less than estimation error. From these findings we may conclude that the two WTP or the two WTA measures may be freely substituted for each other and that these measures will be close to the Marshallian consumer surplus observed from market data and that the WTA measures could be derived from the WTP measure or vice versa. Empirically the bounds between the WTA and WTP measures would be / testable if it were not for respondents' aversion to the WTA measures which we discuss shortly.

---

[9] Price flexibility of income for a good is analogous for the income elasticity for a good except that only specified quantities of the good are supplied (Randall and Stoll, 1980).

Table 1.2 TYPES OF PROPERTY RIGHTS FROM THE CONSUMER'S PERSPECTIVE

Implied Property Rights	Legal Property Rights: Yes (Vested by Law)	Legal Property Rights: No (Not Vested by Law)
Yes (consumer holds)	Legal property *Boatable*	"Squatters Rights"
NO (consumer does not hold)	Hypothetical *Fishable* *Swimmable*	Non property

Criteria for Choosing Between the Hicksian Surplus Measures

Now that we have defined the four types of Hicksian surplus measures let us consider them from the standpoint of measuring consumer surplus in WTP/A surveys.[10] They are formed by combinations of two set of distinctions: equivalence vs. compensation and willingness to pay vs. willingness to accept. To determine which combination is the correct measure for an environmental good being valued in a WTP/A survey we need to compare the property rights posited in the questionnaire with the actual distribution of property rights for that good. Before making our argument we need to distinguish two types of property rights. The usual sense of property right is a right vested by law. In much of what follows we speak of property rights in different sense, as the actual endowment of goods held by a person, to which he or she can add or subtract (Silverberg, 1978). Freeman calls this "implied property rights" (1979b). Table 1.2 shows the relationship between these two types of property rights, names the categories, and locates the boatable, fishable, swimmable levels of water quality.

Speaking now of property rights (implied), the initial endowment or implied property right defines the initial indifference curve that the consumer is on. Additions or subtractions of goods to the consumer's initial bundle of good which are counterbalanced (thereby preserving the same utility level) are Hicksian compensation measures. Changes in the initial endowment or implied property right which are not exactly counter balanced (thereby shifting the consumer to another indifference curve) are equivalence measures. From the standpoint of the individual

[10] At this point we will temporarily refer to these surveys as WTP/A in order to avoid terminological confusion.

consumer, if producers have the right to pollute waterways then consumers must bribe them into not polluting if the consumers desire better water quality. This calls for a WTP measure. In the opposite case, where consumers own the right, producers must bribe the consumers if they wish to pollute and a WTA measure should be used. Compensating surplus measures are appropriate when the contingent situation described to respondents in a WTP/A study uses the same distribution of property rights as actually exists at the time of the study. In this case there is no redistribution implied in the instrument and the potential Pareto-improvement becomes the proper criterion. Where the instrument posits a property right which differs from the existing situation, redistribution is implied and the equivalence surplus measure is called for (Mishan, 1976a). Table 1.3 cross-tabulates the existing and the contingent property rights to show which measures of consumer surplus are theoretically correct for the four combinations. While these distinctions are clear theoretically, in practice they are difficult to apply to WTP/A instruments. We will illustrate this difficulty by discussing our choice of consumer surplus measures and why we believe WTP/A surveys are restricted to the equivalence and compensating WTP measures.

We sought to measure the respondent's consumer surplus for three levels of national water quality: boatable, fishable and swimmable. To identify the theoretically appropriate consumer surplus measure we had to decide what property right (implied) consumers presently have for these environmental amenities. The Federal Water Pollution Control Act (as amended)

Table 1.3 ROLE OF PROPERTY RIGHTS IN DETERMINING THE RELATIONSHIP OF THE WTP, WTA, COMPENSATING AND EQUIVALENCE DIMENSIONS OF CONSUMER SURPLUS MEASURES FOR WTP/A SURVEYS VALUING ENVIRONMENTAL PUBLIC GOODS

Contingent Property Right Specified in the Questionnaire (Implied)	*The Existing Property Right (Implied)*	
	Consumers Own	Consumers Do not own
Consumers own	Compensating WTA	Equivalence WTA
Consumers do not own	Equivalence WTP *RFF Q.82*	Compensating WTP *RFF Qs. 83, 84*

endows the public (individual consumers) with a legal entitlement to fishable/swimmable water nationwide, the goal specified in the Act to be achieved by 1983. Its Congressional architects declared: "This legislation would clearly establish that no one has the right to pollute -- that pollution continues because of technological limits, not because of any inherent right to use the nation's waterways for the purposes of disposing of wastes" (Rosenbaum, 1977:159). Feenberg and Hills (1980), however, contend that in practice property rights to water quality are ill defined and in a state of flux. We agree and think this is particularly the case from the consumer's point of view. Many consumers are personally unaware of the national goal. What they hear about is national freshwater lakes and stream virtually all of which are at the boatable level at the present time, although what they experience locally may be of higher quality. In this context and with regard to the overall national level of water quality which is the public good we are valuing, we believe the implied property right is such that it is appropriate to treat freshwater of boatable quality as if the rights to it are actually owned by consumers and to regard rights to water of higher quality as not (yet) owned by them.

When it comes to deciding how to specify the property right (implied) in our questionnaire theoretical purity gave way, as we believe it must, to methodological realism. In theory the distribution of property rights (implied) for water quality, as specified above, should be replicated in the questionnaire. If we did this the consumer surplus associated with

boatable water over some base (very low quality) level would have to be measured by a compensating WTA question and swimmable and fishable water by a compensating WTP question. While we followed this theoretically desirable practice for the swimmable and fishable levels (Qs. 83, 84), for methodological reasons we measured the boatable level (Q. 82) with an equivalence WTP measure instead of a compensating WTA item.

We made this substitution because the hypothetical market presented in WTP/A instruments must accord sufficiently with the respondents frame of reference, otherwise respondents will give meaningless answers. Clearly, asking our respondents how much they are willing to pay for higher (fishable and swimmable) levels of water quality than they presently enjoy (WTP^{cs}) meets the frame of reference test especially as compared with the alternative of asking them to accept compensation for reductions in levels which they have not yet received (WTA^{cs}). The use of WTP^{cs} is not appropriate for boatable water, however, since the respondents already enjoy national water of that quality. It would be inconsistent to have them pretend that national water quality is non-boatable and to ask them how much they are willing to pay to raise it to the boatable level. The theoretically appropriate measure, WTA^{cs}, also fails the frame of reference test. Analysts who have attempted to ask WTA questions report that an unacceptably large number of respondents respond to WTA questions by either refusing to answer the questions or by saying there is no price they would accept for the loss of environmental quality being valued. In one study of the value people place on visibility in the Pour Corners region 52 and 51 percent of

two samples recorded infinity bids for the WTA questions (Eastman, et al., 1974:581) In another study of the value of hunting to hunters, 54 percent refused to accept any finite amount of compensation (Brookshire, et al., 1980:487). The WTA format places respondents in a situation which is both unfamiliar and which is perceived by many as unfair. People are not accustomed to being offered compensation for environmental goods and apparently some feel offended by the notion. These considerations lead researchers who have experimented with the WTA format to conclude: "We cannot recommend compensation (WTA) games" (Eastman, et al., 1974:583) and "iterative bidding formats for the direct observation of WTA^{cs} do not appear to collect reliable value data" (Brookshire, et al., 1980:488).

Fortunately the empirical consequences of yielding to methodological considerations in the choice of the consumer surplus measure are minor. Randall and Stoll (1980) and Brookshire et al. (1980) have calculated rigorous bounds for the difference between WTP and WTA measures. Using their equations the WTA measures can be derived from the WTP measure and the differences between the two are small. For example, using equations (11) $\frac{M - WTP}{M} \simeq \frac{\xi M}{2\bar{Y}}$ and (13) WTA - WTP $\simeq \frac{\xi M^2}{\bar{Y}}$ of Randall and Stoll (1980), and assuming for illustrative purposes the price flexibility of income (ξ) = .7, household income ($\bar{Y}$) = \$18,000 and WTP = \$250, WTA can be derived from WTP. Equation (11) is solvent for M using a quadratic and then substituted into equation 13. The difference between the WTP and the WTA measures is approximately \$2.50 or 1 percent of WTP.

The Nature of Water Benefits

Water pollution has a wide range of effects on various types of consumers and potential consumers. Insofar as these effects are harmful, they impose "costs." Since the expense of reducing pollution involves another type of "cost" we can avoid unnecessary semantic confusion by calling the losses imposed by a reduction of environmental quality "damages," and the gains associated with reduced pollution "benefits" (Freeman, 1979b).

The basis for determining what is to be regarded as a damage or a benefit is individuals' preferences about the ideal state of the world. We tend to assume a societal consensus about which effects of a given change in pollution should be defined as benefits and which as damages, but such a consensus is not inevitable. If, for example, a significant segment of the population harbored an aesthetic preference for misty landscapes they might regard a reduction in air visibility from 100 miles to 40 miles caused by the operation of large scale coal-fired power plants in the Southwest as a benefit rather than as a damage. Fortunately, a strong consensus does seem to exist as to which environmental changes should be considered benefits and which as damages; otherwise benefit estimation would be even more complex than is currently the case. The consensus does not extend to the amount of the benefit created by a change in an environmental good. Since this varies across individuals, "We define the benefit of an environmental improvement as the sum of the monetary values assigned to these effects by all the individuals directly or indirectly affected by that action" (Freeman, 1979b:3).

As the benefits associated with changes in an environmental media such as water are diverse, any attempt to estimate benefits must specify which benefits are to be measured and which are not. Otherwise certain benefits may be inadvertently left out or others may be overestimated due to double counting. There are several lists of the benefits of improved water quality in the literature (Feenberg and Mills, 1980; Freeman, 1979a), none of which is full satisfactory. Table 1.4 offers our categorization of water benefits. It builds on previous distinctions for the most part, but adds a category of non-direct use benefits which we call "indirect" benefits and assembles all the non-direct use benefits in a single "intrinsic" category.

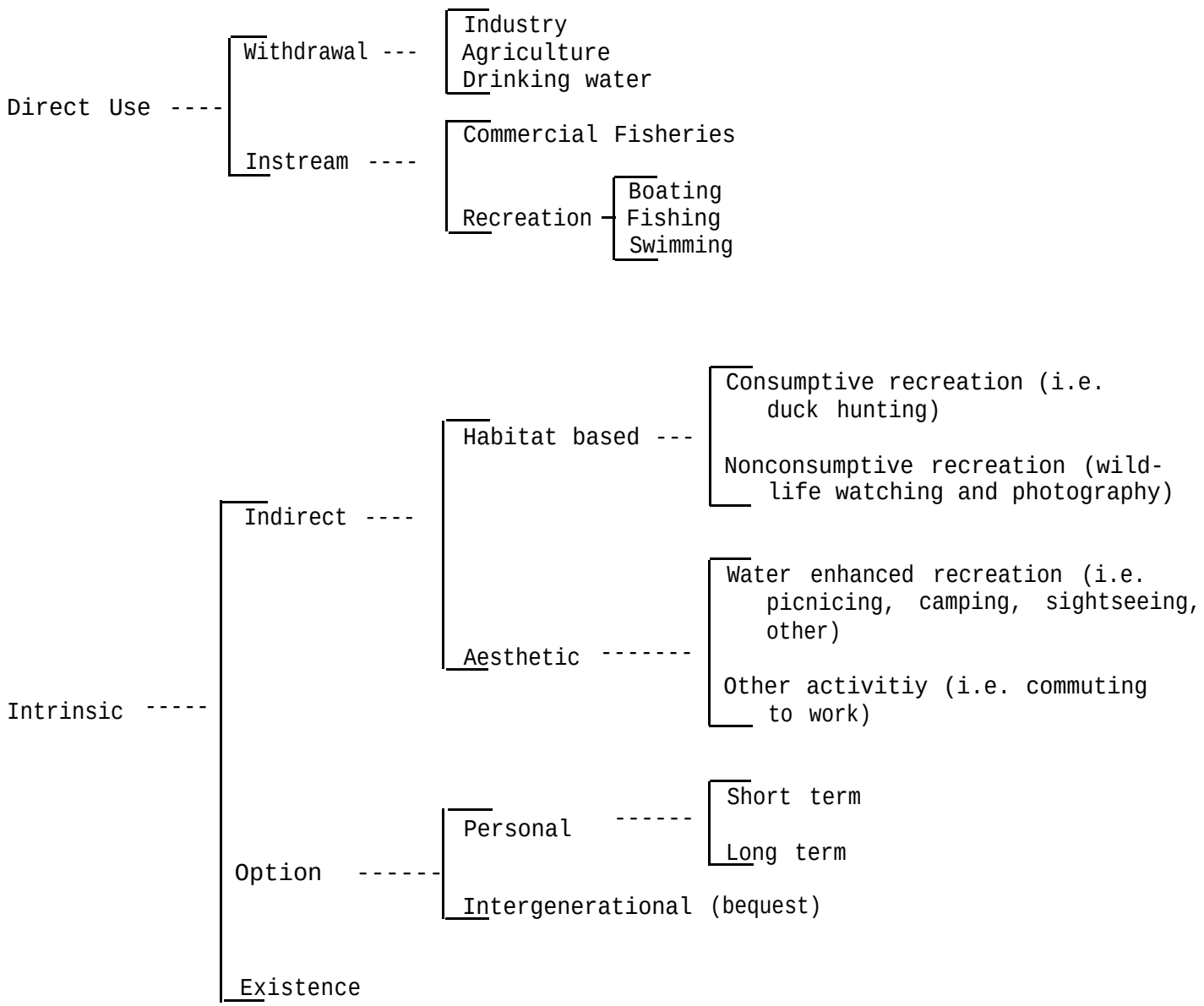
Direct use refers to activities which currently use water either by withdrawal or by instream use (Feenberg and Mills, 1980:8). Improved water quality in freshwater rivers, streams and lakes can result in a variety of withdrawal benefits.[11] Industries which require water of a certain quality for their processes might have lower water purifying costs and less damage to equipment which uses water. Likewise the costs of purifying water for use in washing agricultural produce might be lowered. Drinking water benefits would occur if the improved quality of raw water supply sources lowers the costs of purification and/or reduces the health damage by previously unremoved pollutants. (A companion report to EPA under our cooperative agreement by Mark Sharefkin addresses the question of drinking water benefits.)

Instream use benefits occur in two ways: via increased output or lower costs in commercial fisheries and via the array of activities --

[11] These are comparable to what Freeman (1979a) calls "diversion uses."

NATIONAL BENEFITS OF CLEAN WATER

Table 1.4



fishing, swimming, boating and the like -- by which people use water to recreate. These activities are very popular -- two-thirds of our respondents said they participated in at least one of these activities during the past two years -- and attempts to estimate national direct use recreational benefits assign them a considerable portion of the total benefits of water pollution control. Freeman (1979a), gives a set of best point estimates where recreation accounts for 55% of the total. Most studies of water recreation benefits either use the travel cost or the "participation model" approach, although Gramlich (1977) used the WTP method to estimate the benefits of achieving swimmable quality water in the Charles River Basin of Massachusetts.

Water quality benefits extend beyond its direct use to include various indirect and potential uses. We call this category of values intrinsic since they stem from the inherent characteristics of freshwater bodies. Although important, these benefits are less tangible than the direct use benefits and are consequentially less studied. Freeman's forty page review of recent water benefit estimates devotes a mere two pages to the handful of studies on this topic and concludes: "This is a very tenuous empirical basis from which to estimate national non-user benefits" (1979a:162). Prior to the present research no study attempted to measure

the overall intrinsic benefit for water quality and the existence, intergenerational and option values of water quality were measured by only a single regional study using the WTP method. (This study is described in Walsh, et al., 1978 and Greenley, et al., 1980).

We divide intrinsic benefits into three major sub-categories: indirect, option and existence. Indirect benefits occur when water supports or enhances out-of-stream (non water contact) activities. Duck hunters and bird watchers who enjoy observing acquatic species benefit from the availability of marshes and lakes which provide the necessary habitat for these birds.[12] Fresh water is an aesthetically pleasing setting for such diverse recreational experiences as picnicking by a stream, hiking in wilderness areas, strolling through a New England village located on a river, or visiting the gambling casinos at Lake Tahoe. Aesthetic benefits also accrue to people for whom rivers, lakes or streams serve as a backdrop to their normal activities. Although some would list property values here as a distinct indirect benefit category, we believe property values should be regarded as a surrogate measure of aesthetic and recreational benefits. Adding them to the list would result in double counting (Freeman, 1979b).

---

[12] In a recent paper, Hay and McConnell (1979) review the sparse literature on the value of non-consumptive wildlife recreation and attempt to estimate the reduced form participation model demand for such activities. For comments on statistical procedures see Vaughan and Russell (1981) and Hayward and McConnell (1981).

Weisbrod (1964) first identified option values as an additional form of benefit that must be added to the consumer surplus measure. The essential nature of option value is contained in Greenley, et al.'s definition (1980) of option value as a willingness to pay for the "opportunity to choose from among competing alternative uses of a natural environment in the future." We distinguish between option value based on whether the individual values the future opportunity to choose for his or her personal use (personal option value) or the use of future generations (intergenerational option value).

Let us consider personal option benefits first. These benefits refer to the value people place on a particular environmental amenity on the chance that they personally may wish to use that amenity at some time in the future. Among the three conditions which Weisbrod asserted must be met for determining the presence of option value is that a decision about supplying the amenity in the future is about to be made and should that decision be negative it would be very difficult or impossible to reestablish it (Cicchetti and Freeman, 1971:528). There are two situations where this condition holds and we distinguish between what we call short term and long term individual option value on the basis of these conditions. The first is where present use or failure to protect an amenity will damage it irreversibly.[12a] If the damage can be reversed in the future (at some expense of course) and the individual does not expect to exercise the option in the "near" future, the individual need not make a present choice between the damaging use and non-use to preserve

[12a] Our use of irreversibility extends to situations where the damage could be undone at a future date but at a much greater expense.

his or her use option. We define long term option value, therefore, as the value people place on a good which is regarded as facing possible irreversible damage.

What about the situation where an individual is uncertain about whether or not he or she may wish to use an amenity in the relatively near future? Under certain conditions such an amenity will have option value for a person even when it is not threatened with irreversible damage. We call this short term option value which we define as the price people will pay to have the option to use a good immediately or, in the case where a period of repair (e.g. pollution control) is required to make the good usable, to use the good as soon as possible. Unless the person wishes to use this (non-irreversibly threatened) good as soon as possible, however, it should have no option for him or her. For example, Lake W. is not now swimmable because of seepage from septic tanks but if a sewage treatment plant were constructed it could be made swimmable in five years, It is not threatened with irreversible pollution. If person X wishes to have the option to swim in the lake as soon as possible (e.g. five years from now), he or she has a short term option value for that amenity. If the person has a longer option time frame, however, it would make no sense for the person to express a WTP option value today since the potential to clean the lake up after a five year effort will continue to exist. Put another way, since the damage can be reversed in the future the individual need not make a present choice between the damaging use (continued use of septic tanks) and a cleanup program to preserve his or

her option to swim in the lake. Instead, he or she should use his or her money for other purposes. Table 1.5 summarizes the conditions under which people will hold long and/or short term option values for environmental goods.

Intergenerational or bequest option benefits comprise the willingness of members of the present generation to pay to endow succeeding generations with some natural environment. Some individuals may place a value on preserving such amenities as streams from being essentially destroyed by strip mining operations simply because they would feel better knowing that these streams would still be available for their children or future generations to use if they want to. A parallel argument is made by some that ecosystems and species should be preserved even when they have no present "use" because the reduction of genetic diversity in this manner reduces the possibilities available to future generations to use such species in the ways we are presently unable to imagine. This perspective has become law in the Endangered Species Act and was instrumental in delaying the construction of the Tellico Dam in Tennessee when it was found that the dam threatened an endangered species of minnow, the snail darter.

Table 1.5 CONDITIONS UNDER WHICH PEOPLE MAY HOLD LONG AND SHORT PERSONAL OPTION VALUES FOR ENVIRONMENTAL GOODS

Desire to have option to use good as soon as possible?	Irreversible Threat? Yes	Irreversible Threat? No
Yes	Long and short term	Short term
No	Long term	No option value

As with the long term personal option value, these benefits rest on the assumption that the action taken by the present generation poses an irreversible threat to the environmental good in question. It is worth noting here that the benefit-cost analysis procedures in current use effectively value benefits or costs a generation or more in the future at zero by imposing real discount rates of about 5 percent (Ben-David, et al., 1979:33).

The only empirical study of the option values of water quality is by a team of economists from Colorado State University who designed a WTP instrument on the basis of Henry's (1974) option value mode. Henry's model posits the "preservation of an irreplaceable environmental asset facing an imminent irreversible commitment, until such time that sufficient information becomes available affecting the future option decision of selecting from among alternative uses" (Greenley, et al., 1980:3). The researchers interviewed a sample of two hundred and two residents of Denver and Port Collins. In order to measure the recreation, option, existence and preservation benefits of different levels of water quality in the South Platte River Basin (Northeast Colorado) the respondents were asked a formidable array of willingness to pay questions (twelve in all) using the bidding game format. The personal (short term) option value question posed two alternatives for the Basin. Alternative I featured a large expansion in mining development which would severely pollute, in an irreversible fashion, "many" lakes and streams. Under Alternative II, any decision to expand mining would be postponed

until information became available, sufficient for the respondent to make a decision "with near certainty as to whether it is more beneficial to you to preserve the waterways at level A (the highest level) for your recreational use or to permit mining development" (Greenley, et al., 1980:13). Using an additional fraction of a percent to the region's sales tax as the payment vehicle,[13] an annual mean bid per household of $22.60 to postpone development was reported for the 177 respondents who answered the question. The study also measured intergenerational option benefits by asking the following question:

> Q.28 If it were certain you would not use the South Platte River Basin for water-based recreation [which they defined as including both direct and indirect recreational use], would you be willing to add __ cents on the dollar to present sales taxes every year to ensure that future generations will be able to enjoy clean water at level A? (Walsh, et al., 1978:82).

A bequest value of $16.97 a year per household is reported for a subsample of 24 non-recreationists.[14]

---

[13] They repeated each bidding game using a second bidding vehicle, an additional charge to the respondent's water bill.

[14] The researchers eschew using the intergenerational option benefit amount for the recreators in their sample because they doubt the recreator's ability to leave out their personal recreational considerations when answering this question. Based on our review of the instrument this is the correct decision, but it reduces the sample size so much that the bequest estimate can only be regarded as suggestive (Greenley, et al., 1980:15, 33).

The final type of intrinsic benefit is existence benefits. In 1967 Krutilla wryly commented regarding wilderness that: "There are many persons who obtain satisfaction from mere knowledge that part of wilderness North America remains even though they would be appalled by the prospect of being exposed to it" (1967; see also Krutilla and Fisher, 1975). Existence value is the willingness to pay for the knowledge that a natural environment is preserved (Greenley, et al., 1980:1) quite apart from any use or expectation of use by the respondent or by future generations. The lone attempt to measure the existence benefits of water quality is the above mentioned Colorado State study which uses the following question:

> Q.27 If it were certain you would not use the South Platte River Basin for water-based recreation would you be willing to add ___ cents on the dollar to present sales taxes every year, just to know clean water exists at level A as a natural habitat for plants, fish, wildlife, etc?

They report a mean figure of $24.98 for the 24 non-recreationalists who answered this question.

Since the Colorado State study represents the state of the art in estimating option and existence benefits, a closer examination of its methodology is relevant to our purposes. Three questions will be addressed. Is it methodologically sound? How adequate are their measurements and estimation procedures for option and existence values? How much credence should be placed on their annual benefit estimate for the South Platte River Basin of $61 million of which $26.4 million or 43 percent is attributed to recreation benefits (both direct and, using our terminology, aesthetic) leaving 57 percent attributed to option, existence and bequest benefits?

The study is a useful methodological experiment from which we can learn a great deal thanks to the admirably complete report they wrote for their sponsor (Walsh, et al., 1978) and which is available through NTIS. Unfortunately, the study's flaws are such that the researchers' decision to extrapolate their findings without qualifications or reservations in the form of aggregate point estimates in the report and in a brief journal article (Greenley, et al., 1980) is unwarranted and potentially misleading.

Since we are primarily concerned with the study's approach to measuring intrinsic benefits, we will only briefly mention the more serious of its other methodological problems. These are:

- A low response rate -- only 37 percent of the sampled households which received the letter announcing the intention of the researchers to interview a household member participated. According to sampling theory this low a rate means that the findings cannot be generalized to the total population of those areas which constitute the study's sampling frame.
- Starting point bias. The large difference in results between their two bidding vehicles -- sales tax increase and increase in sewer

bills -- may be attributed to the aggregate yearly payment implied by the starting point for each vehicle.[15] Furthermore, the mean bids for option, bequest, and existence values are very close to the starting point for each vehicle.[16] Since their questionnaire involved so many bidding games, a combination of respondent fatigue and a willingness to please the interviewer possibly may account for a large portion of the bids.

- The payment vehicle, additional taxes at the regional level, is ambiguous. Since water quality actually is paid for in higher prices and federal income taxes for the most part, the respondents are already paying large amounts for this purpose. We have no way of knowing whether the respondents are

---

[15] The starting points for the vehicles were one quarter of a cent increment in sales tax and $.50 a month on the water sewer bill. Prior to bidding the respondents were informed how much additional money they would pay a year for every one quarter cent increment in sales tax. (Walsh, et al., 1979:29). The study report does not say whether an annual amount for the water/sewer fee was calculated for the respondents but even if it wasn't the respondents would be able to calculate this easily themselves. For the entire sample they report an annual recreation value of $18.60 for the water fee vehicle and $56.68 for the sales tax. (Every respondent bid for recreation using each of the vehicles, total N = 174) The only explanation they offer for respondents' willingness to pay only about one-fourth as much in water-sewer fees as in sales tax was that they "may have perceived inequities" in the fees since everyone, including tourists, would be liable for sales taxes (Greenley, et al., 1980:17). However, since the starting points for the two vehicles "generated revenue of $6 per year in water-sewer fees and $25 per year in sales tax for a typical household of four with an average income of $13,500 per year" (Greenley, et al., 1980:11), it is more likely that the difference results from starting point bias.

[16] In Table 1 of Greenley, et al. (1980) they give the mean bid for option, bequest and existence values for each vehicle. In every case, irrespective of vehicle, the bids for these values hover around the starting point. The average difference from the starting point is 17 percent. (It is true, however, that the bequest value lies slightly below the starting point, while the other two have mean bids above the starting point, suggesting that people do value bequest values less than the other two.)

willing to pay these amounts plus the additional amounts elicited in the bidding games or not.

Of direct importance to our present discussion is the method used by the Colorado State researchers to measure the option and existence values. Their approach is additive. They ask separate questions for each of the four benefit categories (recreation, option, bequest and existence) and add the resulting mean bids to get a total WTP figure for the Basin's water quality. Since the additive technique requires each benefit to be measured independently with no overlap, the WTP instrument must ensure that respondents bid on one value at a time and only on that value. Otherwise double counting will occur biasing the total estimate upward and making it impossible to derive reliable estimates for the component values. A close scrutiny of the wording of the recreational and option value questions in the Colorado State instrument raises serious doubts about their independence.[17]

Here is the wording of the question they used to measure recreational value:

---

[17] In the case of the existence and bequest values, however, they recognized after the fact that their survey "did not ask users about (these benefits) in such a way as to permit adding them to user's values" (Walsh, et al., 1978:39). For this reason they restricted their estimates of these benefits to the very small number of non-recreationists. In the discussion which follows we consider only the recreation and option values, both of which they estimated for the full sample, although we believe our criticism also holds for the other two measures.

Suppose a sales tax was collected from the citizens of the South Platte River Basin for the purpose of financing water quality in this Basin. All of the additional tax would be used for water quality improvements to enhance recreational enjoyment. Every Basin resident would pay the tax. All bodies of water in the River Basin would be cleaned up by 1983. Assume that this is the only way to finance water quality improvement.

14. Would you be willing to add ____ cents on the dollar to present sales taxes every year, if that resulted in an improvement from situation C to situation B?

15. Would you be willing to add ____ cents on the dollar to present sales taxes every year, if that resulted in an improvement from situation C to situation A?

The three water quality levels A (best), B (medium) and C (worst) were represented by photographs showing colored water features associated with mine drainage. Although the wording says all the additional tax would be used "to enhance recreational enjoyment" the question does not explicitly ask the respondent to limit his or her answer to recreational benefits nor does it inform the respondent that he or she will be presented with subsequent opportunities to say how much they are willing to pay for other (intrinsic) values. Since the apportioning of water quality values to precise categories is not a familiar undertaking for most people, the form of the question with its emphasis upon the quality shift from C to B and C to A and the use of the pictures which depict aesthetic degradation serve to create the impression that the respondent is being asked about water pollution in general. The bids for the recreation question probably should be regarded as the consumer's total willingness to pay for an increase in water quality in the area from C to A.

The option value question has the same weakness. Although the researchers are careful to specify the option characteristics in accord with their theoretical model, the question is worded in such a way that the respondents could interpret it as asking them to value water quality of level A while bearing in mind the economic tradeoff of foregoing mining activity. (A Further problem with the option question is that the respondents may not believe level C to be irreversible since the recreational questions in the interview told them that level C could be improved to levels B or A.) The option question is worded as follows:

> In the near future, one of two alternatives is likely to occur in the South Platte River Basin. The first alternative is that a large expansion in mining development will soon take place, creating jobs and income for the region. As a consequence, however, many lakes and streams would become severely polluted. It is highly unlikely, as is shown in situation C, that these waterways could ever be returned to their natural condition. They could not be used for recreation. Growing demand could cause all other waterways in the area to be crowded with other recreationists.
>
> The second possible alternative is to postpone any decision to expand mining activities which would irreversibly pollute these waterways. During this time, they would be preserved at level A for your recreational use. Furthermore, information would become available enabling you to preserve the waterways at level A for your recreational use or to permit mining development. Of course, if the first alternative takes place, you could not make this future choice since the waterways would be irreversibly polluted.
>
> 26. Given your chances of future recreational use, would you be willing to add ____ cents on the dollar to present sales taxes every year to postpone mining development? This postponement

> would permit information to become available enabling you to make a decision with near certainty in the future as to which option (recreational use or mining development) would be most beneficial to you?

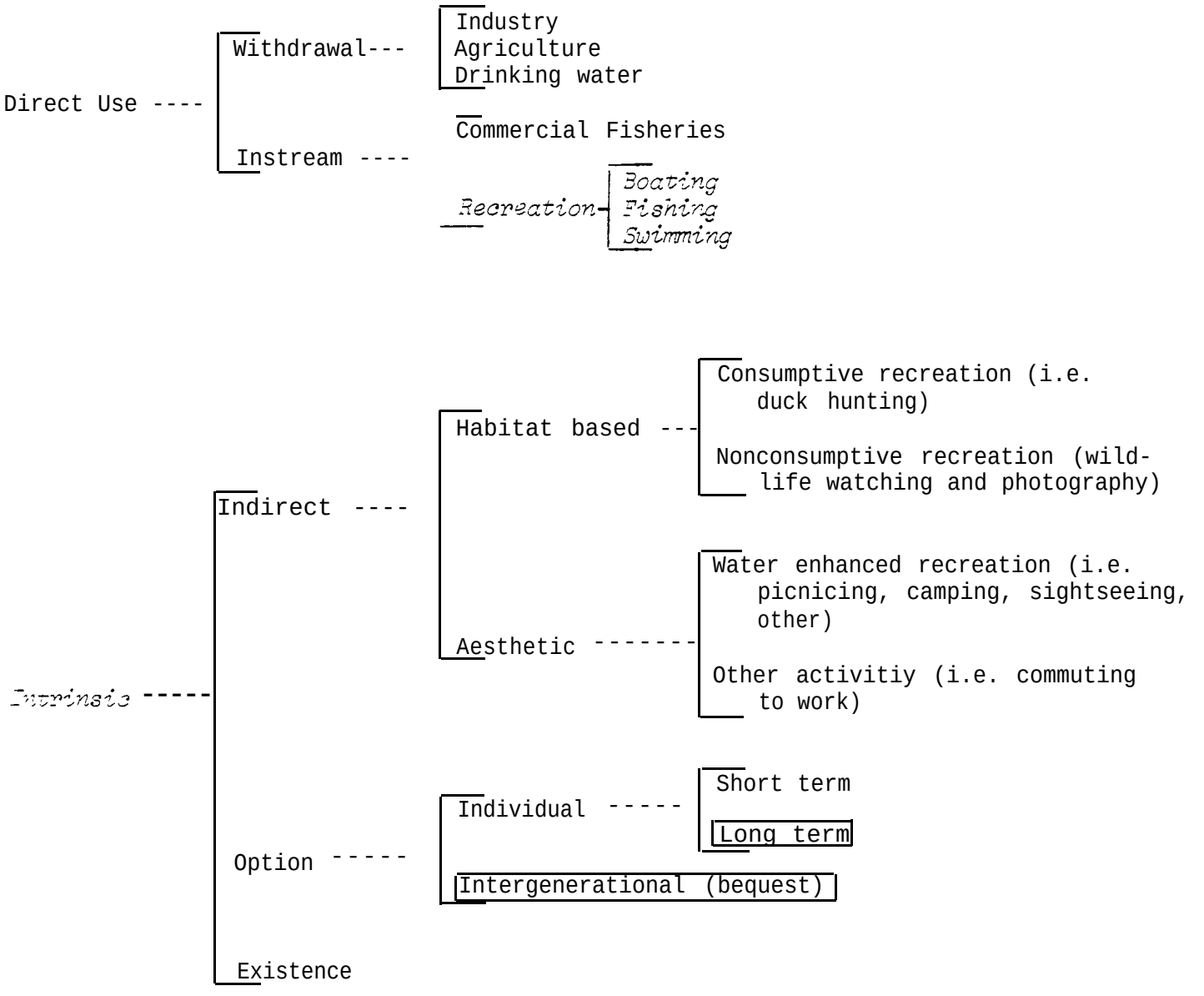
Whereas the "recreational value" questions (14 and 15) ask the respondents to imagine that they are at level C or B and to say how much they are willing to pay to move to level A, the "option value" question asks them to assume that they are at level A and asks how much they are willing to pay to remain at level A instead of moving to level C. In this respect, the question is simply another way of measuring the consumer surplus for level A and we would again anticipate that the respondents' WTP amounts will reflect their total recreational and intrinsic values for water quality rather than just the intended independent (and additive) option value. Of course the bids on this question will be influenced by the additional information conveyed by Question 26, namely: a) the water quality change will be irreversible; b) keeping water quality at level A involves economic tradeoffs (jobs and income) and c) there is pressure for such development. This last point, which is implied rather than stated, might lead the respondent to believe that further mining activity is inevitable,[18] and therefore to give low or zero bids.

---

[18] These factors may explain why the "option" question received a lower mean bid ($23) than the "recreation" question ($57).

In this study we use an approach which contrasts with the Colorado State method in several respects. First, we do not attempt to measure the various sub-categories of intrinsic benefits as they did, although we do obtain separate estimates for the overall intrinsic benefits and for the in-stream recreational benefits. Second, we begin by ascertaining the individuals' total consumer benefits (recreational plus intrinsic) through a sequence of WTP questions. Only then do we apportion these total benefits to the separate recreational and intrinsic categories as the basis of information which we obtained in the interview about the respondents' recreational use or non-use of freshwater. Our process is subtractive rather than additive and uses self-reported behavior rather than answers to specific WTP questions to distinguish recreational from intrinsic benefits. Working backwards from a total benefit figure has the advantage of forcing respondents to consider their budget restraints more realisitically than in the case when they are asked to value a sequence of component benefits without confronting the overall expenditure involved in these separate decisions. Table 1.6 shows which of the benefit categories in Table 1.4 we measure in this study. We present our findings in Chapter 5.

Table 1.6 NATIONAL BENEFITS OF CLEAN WATER MEASURED BY THE RFF SURVEY



Categories in italics are those estimated in this report. The subcategories in the boxes are not included in our intrinsic benefits total because the changes in water quality which we value are defined as irreversible.

CHAPTER 2
THE MACRO APPROACH TO WILLINGNESS TO PAY STUDIES

Our review of studies using the willingness to pay method reveals two distinct research traditions. In one tradition, willingness-to-pay questions are used in national polls as a measure of environmental concern. In the other, the questions are employed by economists to develop benefit estimates for particular environmental goods. We have named these approaches the macro and micro, respectively. Each has advantages and disadvantages for benefit estimation. We have experimented with a new kind of macro approach, one which borrows heavily from methodological innovations developed by practitioners of the micro approach. In this chapter we describe these two approaches and the rationale behind our synthesis.

The Macro Willingness to Pay Approach

Since 1969 at least 8 different surveys have asked questions using the "macro willingness to pay" (macro WTP type). The kind of environmental public goods covered in these surveys range from air pollution devices on new automobiles Viladus, 1973) to the more general category of "cleaning up pollution now" (Gallup, 1971). They also vary in how they ask for the amount. Some questions are open ended, but macro WTP questions usually offer a specific amount or a limited sequence of specific amounts for the respondent's judgment. For example, in 1969 a Harris poll for the National Wildlife Federation asked 1500 adults nationwide:

> You are already sharing in the costs brought to us all by air and water pollution. In order to solve our national problems of air and water pollution the public may have to pay higher taxes and higher prices for some products. To get real clean-up in your natural environment, would you be willing to accept a per-year increase in your family's total expenses of $200?

The question was repeated for the amounts of $100, $50 and $20. Other examples of macro questions include these taken from national surveys.

> Would you be willing to pay an additional $20 per year on your electric bill in order to cut down air pollution caused by power plants? (Federal Energy Administration, 1977; August 1975 survey).
>
> (After asking people the amount of their last electric bill ....) Now suppose that the only way to stop the electric power plants from polluting is to install expensive equipment, and this equipment made your electric bill go up unless you used less electricity than you use now. How much more would you be willing to pay a month to clean up this form of pollution? $_____ (Viladus, May 1973 survey).

The past uses of the macro WTP approach have the following characteristics:

1. Purpose: In these earlier uses, macro WTP questions were not intended to provide the basis for benefit estimates in the strict sense. They were used for the conventional poll takers purpose of measuring public concern about environmental goods. It is assumed that asking people the amount of money they are personally willing to pay for pollution reduction is a more stringent test of people's concern than questions which simply elicit concern without reference to the cost. The relevant audience for these studies are those who normally use public polls on environmental issues.

2. Survey Method: The macro WTP questions were used in social surveys conducted by professional polling organizations. Because the respondents were chosen by modern sampling techniques, with sample sizes ranging from 800 to 1500, the results may be generalized to the appropriate sampling frame within a statistically determinable degree of accuracy. The interviewers are trained adult workers under contract to the polling

organization whose work is subjected to independent checks. In each use of the macro WTP questions they have been just one component of a larger list of questions.

3. Specification of the good and procedure for ascertaining WTP: The nature and geographical distribution of the environmental good is described in general terms. People are asked about "cutting down air pollution," for example, with no mention of where this would happen or how much "cutting down" is involved. No attempt is made to vary the amount of the good, to provide visual aids describing it, to present the parameters of a hypothetical market in the good, or to sepcify the geographical location which would receive the environmental benefit.

4. Test for biases: The standard assumptions about the reliability/ validity of survey research are applied to the macro WTP questions. These assume that a question is reliable if it uses words which are understood by all the respondents, is unambiguous in meaning, is neutral in its wording, and asks about a matter on which respondents may be presumed to have an opinion. Validity is established by judgment of whether or not the description of the environmental good in the question appears to be adequate (face validity). No attempts were made to undertake specific tests for threats to reliability and validity. Data reporting was limited to presentation of the marginal results and cross tabulation by standard background variables.

5. Sampling Frame: The sampling frame for these surveys was a large geographic area. Most were national (the lower 48 states) although macro WTP questions have occasionally been used in state surveys.

We use "macro" as a label for this tradition of WTP questioning because of its focus on national benefits.[1] No matter what environmental good these questions solicit willingness-to-pay amounts for, the money would pay for supply of the good across the country. The micro approach, as we will see, is interested in the benefits for a specific geographic area.[1a]

The Micro Willingness to Pay Approach

Since the Second World War, economists have been increasingly faced with the need to measure the use values associated with natural resources. Insofar as values associated with goods are measured in the market place in terms of price, obtaining dollar estimates for them is relatively straightforward. But natural resources, including the amenities of clean air and water, have characteristics which severely limit the use of exchange to determine their value for society. Because they have the attributes of public goods especially in that it is difficult or impossible to exclude consumers from using them, they are outside conventional market structures. The rather intangible nature of some of the values these resources convey, such as aesthetic and existence values, means that people are likely to have difficulty imagining the good with precision and conceiving of a hypothetical market in those values.

---

[1] If the Grand Canyon has symbolic national value then the location of the benefits is national rather than local.

[1a] Macro need not refer to only national benefits. For instance, the benefit could be global (CO_2, d'Arge et al., 1980) or regional as in a survey of WTP for air quality regulations in California of a random sample of all California (if California only generated and was affected by the air pollution). At the margin the distinction between macro and micro become blurred.

Economists have experimented with ways to overcome these obstacles in order to simulate a market in environmental goods. Among the myriad of techniques developed over the past three decades for this purpose (see Wyckoff, 1971; for an overview as of 1970) is the use of survey research instruments to ask people what they are willing to pay for such goods. Although Ciriacy-Wantrup suggested such a technique -- which he called the "direct interview method" -- as early as 1947 (Wyckoff, 1971:13), it apparently did not come into actual use until the 1960s when Davis (1963) used questionnaires to estimate recreation benefits. Since that time the technique has been used repeatedly by economists to measure such things as recreational benefits (Binkley and Hanemann, 1978; Darling, 1973, McKinney and MacRae, 1978); water quality benefits (Gramlich, 1977; Walsh, et al., 1978) (Davis, 1980); benefits of decreased risk from a nuclear power plant accident (Mulligan, 1978); aesthetic benefits from foregoing a geothermal power plant (Thayer, forthcoming); aesthetic benefits of air visibility (Randall, et al., 1974; Brookshire, et al., 1976; Rowe, et al., 1979a and b); and aesthetic and health benefits of air quality (Brookshire, et al., 1979)

In the course of this research the direct interview technique has been refined and a great deal of study has been given to its possible biases. Much of this work has been undertaken by Randall and colleagues (Randall, et al., 1974) and by d'Arge, Brookshire, Rowe and others from the University of Wyoming in their series of studies on the aesthetic benefits of air pollution reduction. In 1979 the latter group produced a major methodological study of the technique for EPA (Brookshire, et al., 1979).

Figure 2.1 gives the text of a micro WTP question. It was used for a 1975 study of the aesthetic damages of a possible power plant near Lake Powell in Utah and illustrates the essentials of the micro approach. This approach, particularly as used in the air pollution benefits studies, differs from the micro approach in a number of important respects.

1. Purpose: The micro studies are specifically designed to obtain estimates of economic benefits by gathering data which enable the fitting of a demand curve for the value in question. Their designers seek to gather data which will be accepted as valid for this purpose by their fellow economists.

2. Survey Method: The field work for the micro WTP studies is usually conducted by the researchers using student interviewers who are specially trained for the study. The WTP questions are the centerpiece of the survey instrument which is dedicated solely to the benefits measure study. In a number of the past studies sample sizes have been very small by conventional survey research standards; sub-groups which are the focus of extensive analysis sometimes consist of only 20-30 cases. Sometimes the descriptions of the sample frame and procedures are sketchy or lacking entirely[2] in the report so it is difficult to know whether the findings can be generalized reliably to larger populations and what those populations might be. In other case (e.g. Rowe, et al., 1979b:85-89) a representative rather than random sample was used which precludes such generalization.

[2]For example, the interview dates, the response rate, and/or the method of selecting the respondents may be missing.

QUESTIONNAIRE

Bidding Game for Estimation of Recreationists' Demand
for Abatement of Aesthetic Environmental Damage

Good Morning/Afternoon. My name is _____. I'm doing research for the Economics Department at the University of New Mexico, as a part of the Lake Powell Research Project, funded by the John Muir Institute for Environmental Studies.

This research is designed to more closely examine some of the trade-offs between industrial development, recreation and the environment in the Lake Powell area. In connection with these objectives. I would like to ask you a few questions to see how you feel about environmental quality and its future In this area.

1. How many members of your family are here with you? _____persons.

2. What is the expected length of your stay? _____days.

3. Where are you staying? _____(a) local resident. _____(d) developed or semi-developed campground
_____(b) lodge, Page motel
_____(e) remote (specify location)
_____(c) passerby

4. If you don't mind, could you please indicate which of the following brackets your family income falls into:

_____ 0 - 4,999	_____ 20,000 - 26,999
_____ 5,000 - 9,999	_____ 25,000 - 39,999
_____ 10,000 - 14,999	_____ 30,000 - 49,399
_____ 13,000 - 19,999	_____ 50,000 and up

There are plans to construct a large electric generating plant north of Lake Powell. This plant is expected to be at least as large as the Navajo Plant on the south side of the lake.

5. Have you noticed the Navajo Plant or its smokestacks? _____yes _____no

Depending on exactly where and how a new plant is constructed, it could have a significant effect on the quality of the environment. If the plane is built near the lake, it could be visible for many miles up and down the lake. If air pollution is not strictly controlled, visibility in the area may be significantly affected.

These photographs (show) are designed to show how a new powerplant on the north side of the lake might appear. Situation A shows a possible plant site but assumes that the powerplant would be built at some distant location, not visible from the lake area. In Situation B the powerplant is easily seen from the lake, but emits very little smoke; visibility is virtually unaffected. Situation C is intended to show the situation with the greatest impact on the environment of recreationists in the area. It is easily seen from the lake, and the smoke substantially reduces visibility.

Vacationers, of course, spend considerable amounts of money and time and effort to equip themselves with vehicles, boats, camping and fishing gear, and for traveling to the destination of their choice. It is reasonable to assume that the amount of money you are willing to spend for a recreational experience depends, among other things, on the quality of the experience you expect. An improved experience would be expected to be of greater value to you than a degraded one. Since it does cost, money to improve the environment, we would like to get an estimate of how much a better environment is worth to you.

First, let's assume that visitors to GCNRA are to finance environmental improvements by paying an entrance fee to be admitted into the recreation area. This will be the only way to finance such improvements in the area. Let's also assume that all visitors to the area will pay the same daily fee as you, and all the money collected will be used to finance the environmental improvements shown in the photos.

6. Would you be willing to pay a $1.00 per day family fee to prevent Situation C from occurring, thus preserving Situation A? $2.00 per day? (increment by $1.00 per day until a negative response is obtained, then decrease the bid by 25¢ per day until a positive response is obtained, and record the amount.)

7. Would you be willing to pay a $1.00 per day fee to prevent Situation B from occurring, thus preserving Situation A? (Repeat bidding procedure).

8. (Answer only if a zero bid was recorded for question 6 or 7 above.) Did you bid zero because you believe that:

_____ the damage is not significant

_____ it is unfair or immoral to expect the victim of the damage to have to pay the costs of preventing the damage

_____ Other (specify) _____

9. In your opinion, has visibility, depth or color perception in this area been significantly reduced by air pollution?

3. Specifications of the good and procedures for ascertaining WTP: Because of the importance of making the situation as realistic and credible as possible, great attention is given to the description of the environmental good in the micro studies. It is typically described as occurring in a specific locality (usually the locality where the interviewing is taking place); a time frame is specified; and an extensive verbal description of the good is supplemented with pictures or other visual devices. A great deal of care is also given to the procedures for eliciting the WTP amount. The survey instrument describes a hypothetical market with a substantial degree of institutional detail; specific, plausible means of payment are specified; and contingencies relevant to the respondent's valuing the good are described. A common feature of most of these studies is the use of a "bidding game" procedure to ascertain the dollar amount the respondent is willing to pay.

The bidding game works in the following manner: after the hypothetical market is staged by means of preliminary questions, verbal description, and the use of the visual aids, a particular good is identified and the person is asked whether he or she is willing to pay $x for the good. If the starting amount (e.g. one dollar) is agreed to, the interviewer increases it by a set interval (e.g. 50 cents) until the respondent rejects an amount. The study may then require the interviewer to decrease the amount rejected by a smaller amount (e.g. $.25) until the precise maximum amount the individual is willing to pay is reached.[3] This procedure is usually repeated for several levels of the good in question so that the demand curve can be traced out.

4. Test for biases: Because they are explicitly intended to provide benefits estimates for policy purposes, micro studies attempt to obtain as close a surrogate as possible to actual market behavior.

---

[3] Several micro studies also used parallel procedures to ascertain how much respondents were willing to accept (WTA) in return for the loss of the environmental good.

> The efficacy of bidding games used for this purpose [to measure aesthetic environmental improvements] depends on the reliability with which stated hypothetical behavior is converted to action, should the hypothetical situation posited in the game arise in actuality (Randall, et al., 1974:135).

Since many economists are skeptical about the fit between attitudes and behavior, credibility in this regard is crucial. Accordingly, those conducting micro studies have placed a great deal of emphasis upon testing for potential biases. In a number of cases, most notably the studies done by d'Arge, Brookshire, and their colleagues, tests for biases are built into the study design as when comparable samples are offered different dollar amounts as starting points for the same environmental good in order to test for starting point bias. Strategic bias has been examined in a similar manner.

5. Sampling frame: The environmental amenities valued by the micro WTP approach are, as we have seen, location specific. Those interviewed for these studies are generally sampled from people who live or recreate in the particular area. This conjunction of a local good and a local sample is intended to reduce the artificiality inherent in the bidding games since people will be bidding on a good which they can easily comprehend and which is of immediate concern to them. For the South Platte River Basin (Colorado) 202 residents of Denver and Fort Collins were interviewed (Walsh, et al., 1978); for the Glen Canyon Recreation Area the 82 respondents included local residents, motel visitors, developed campgrounds visitors and remote campers (Brookshire, et al., 1976).

Comparison

This brief description of these two ongoing research traditions captures the essential features of each as they existed in 1979 when planning for the RFF experiment began. Each has a major strength and a compensating weakness.

Realism

Of the two approaches, the micro approach has been far superior in its realism. People are asked about a good which they personally have experienced or which they would experience in that location if pollution levels increased. The several values associated with the good (existence, aesthetic, health, etc.) are differentiated and the value chosen for measurement is described in detail both verbally and, if possible, pictorially. The payment vehicle and the hypothetical market are designed to match the respondent's experience as closely as possible. In comparison, the designers of the macro questions have made very little effort to stimulate a market or to describe the environmental goods in detail.

Generalizability

Realism is an important factor in designing reliable and valid measures of WTP. But once reliable and valid benefit estimates have been obtained from a set of respondents, for our purposes it is necessary to aggregate them to obtain overall benefits estimates. The great strength of the macro approach with its use of a national sampling frame is the ease with which the results can be generalized to give a national benefits estimate. In contrast, it is difficult to aggregate micro study findings beyond the location where the study was conducted and it is extremely difficult to make reliable national estimates from a series of micro studies.

Probability Sampling and Aggregation

Survey research has a standard solution to the aggregation problem -- probability sampling. If Gallup wants to predict the national presidential vote, he interviews 1500 people nationwide who are chosen by an elaborate sampling procedure based on statistical principles. Providing his survey takes place immediately before the vote and that his interviewers adhere to the sampling plan, he will be able to predict the vote with an accuracy of $\pm$ 3 percent. Good sampling requires: 1) designation of the appropriate sampling frame for the population to which one wishes to generalize (in the Gallup example this is people living in non-institutionalized settings in the lower 48 states), 2) design of a sampling plan which will give every relevant person (e.g. adult voters) a known probability of inclusion, and 3) strict execution of the sample., Once the sampling frame is chosen, the design and execution of the sample is straightforward, although certain adaptations can be made to a strict probability design in the interests of economy without undue bias resulting (see Sudman, 1976, for a review of these procedures).

The choice of the sampling frame necessarily depends upon the researcher's problem and purpose. For WTP studies, it should be the population for which the researcher wishes to have an aggregrate benefit.[4] There are two separate issues involved which complicate the choice of a sampling frame for WTP studies: a) which groups can be presumed to "have" benefits that should be included in any comprehensive measure and b) what groups are relevant under different equity positions; i.e., do only those who pay get to have their benefits counted? Let us suppose that he or she wishes to

[4] The researcher also needs to define any special sub-populations which are likely to have an especially high value for the good in question, If there are such sub-populations, he or she may need to oversample these people. Otherwise they may be too few in number to enable a reliable estimate to be made of their benefits. For example, one in fifteen men in an area may be fishermen. If 300 people are sampled for a study of water recreation benefits in an area only 20 are likely be to fishermen (0.066 x 150) When benefits are aggregated across the entire sample, the benefits for over-sampled sub-population(s) must be weighted to reflect their proportion of

estimate noise pollution control benefits. In the case of a village which wishes to use WTP techniques to estimate the benefits of ordering quieter garbage trucks,which would be paid for out of village property taxes, the appropriate sampling frame is the residents of the village. If noise regulations are a state matter and their cost is paid for by state taxes, then the state population would be the appropriate frame. In both these cases the selection of the sampling frame is simplified because the same population is affected by and pays for the public good in question.

Choosing the appropriate sampling frame becomes more complex where the two do not coincide. The table below shows the four possible relationships between paying for and using a public good. Using our example of the town contemplating the purchase of garbage trucks, an example of B is visitors

		Pay for the Good: yes	Pay for the Good: no
Use the good	yes	A	B
	no	D	C

to the town who would benefit from quiet garbage trucks although they wouldn't pay for them since they are not subject to town property taxes. Position D would include deaf residents and absentee property owners. Note that by using the sampling frame of the town residents, we include some D's (town population = A + D minus absentee taxpayers). Sampling frames comprised of those who live in political jurisdictions responsible for public goods almost

inevitably include both users and non-users. For example, those who reside in a city with a public school system include the childless, people whose children are too young or old for public school, and those who send their children to private schools. Note also that the use of the town population as the sampling frame leaves out some D's. Presuming that property taxes are the source of the town's revenue, absentee landlords would not be represented in a sample of town residents. A different sampling frame consisting of property tax payers would, of course, include them but it would exclude renters[5]

B is an important category for some benefits estimates. Consider the case of the huge Four Corners power plant at Fruitland, New Mexico in the Southwest (Randall, et al., 1974). Residents of the area and visitors who come to enjoy the scenery use the public good of high air visibility without paying the cost of maintaining it. This cost is (would be) borne by those in Los Angeles (and elsewhere) who purchase their electricity from the utility which owns the plant. Nevertheless, area residents and visitors are a crucial sampling frame for a WTP study of the aesthetic benefits of local air visibility.

A further complication is introduced when we consider the question of intrinsic benefits. It may be worth something to Los Angeles residents (D) who never recreate or intend to recreate in the Four Corners area to know that the extraordinary air visibility in that area is untouched by the emissions of the plants which provide their electricity. Indeed, and here we come to position C, it may be worth something to residents of Ohio as well. A local or even regional sampling frame is inadequate if the researcher wishes to include intrinsic benefits in a national estimate of the benefits of high visibility in the Southwest.

---

[5] Recognizing, of course, that renters eventually pay all or some of the taxes imposed on landlords.

Interrelationship Between Generalizability and Realism

The sampling frame and the realism of the WTP instrument are interrelated. Where users and payers are in the same population (position A), both the description of the good and the payment vehicle can be related to their actual experience and realism is enhanced. People in position B, may be more unrealistic in their WTP estimates than those in A or D because they know they are not paying for the good and are unlikely to think they will have to pay for it in the future. The good may be especially abstract and hard to imagine for those in position D who pay for the good but who do not use it. Thus the potential for measurement bias is reduced when the sampling frame consists primariy of A's. To the extent that respondents anticipate that their answers will affect their level of payment or their level of supply of the public good, B's estimates will tend to overestimate the consumer surplus and D's to underestimate it owing to the effect of strategic bias.

This description of the strengths and weaknesses of the two research traditions as they have been practiced to date is summarized in the following four-fold table.

Generalizability

Realism		High	Low
	High		micro
	Low	macro	

The obvious goal for a study of public good benefits is to move to the box where the data are both realistic and generalizable. This is a difficult task because the two dimensions are somewhat incompatible, necessitating

tradeoffs between degree of realism and degree of generalizability. Thanks to the experimental micro studies of the 1970s, however, we have a much greater knowledge of the properties of willingness to pay measures. For example, micro research has shown us that certain potential problems such as strategic bias are not as much of a problem as some had thought (see Chapter 4). Knowledge such as this gives the researcher greater flexibility in designing a WTP research instrument, flexibility which was essential to our effort to devise a macro instrument which was workable yet sufficiently realistic in its description of water quality to give US valid results. In Chapter 4 we argue the need to jointly minimize the potential for strategic and hypothetical bias.

The RFF Macro Approach

For public goods which are mandated at the national level and are paid for by everyone in higher prices and taxes there is a need to obtain national benefits estimates. The quality of water in the nation's fresh-water bodies is such a public good. In 1972 Congress passed the Federal Water Pollution Control Act Amendments (later amended). In this law:

> Congress has declared its intent "that the discharge of pollutants into the navigable waters be eliminated by 1985" and that "wherever attainable, an interim goal of water quality which provides for the protection and propagation of fish, shellfish, and wildlife and provides for recreation in and on the water be achieved by 1 July 1983." In effect, this amounts to a commitment to make all the nation's navigable waters "swimmable and fishable" by 1983 and wholly free of pollutants in 1985. (Rosenbaum, 1977:158).

The law established a national permit system for all municipal and industrial effluent discharges according to national standards and mandated the use of "best practicable" technology to control water pollution by 1977 and the "best available" technology by 1983. Although it is implemented by the states, the standards and compliance deadlines are set by Washington.

The cost of this mammoth pollution control program is ultimately borne by all U.S. taxpayers and consumers. The federal government provides the construction monies for municipal waste treatment facilities in what is the largest single public works project ever authorized by Congress. Municipal taxes pay to maintain and operate the waste facilities. The expanse of controlling the non-municipal effluents are borne by industry (and ultimately the consumer) and other operators. The reach of the law extends beyond effluent pipes to the many "non-point" sources of water pollution such as fertilizer runoff from farmers' fields.

After a careful consideration of the alternatives, we decided to adopt a macro approach in our study of the intrinsic benefits of water pollution control. A primary impetus for this decision was the national character of control programs. In addition we were influenced by the following considerations

1. The results of the various micro experiments suggested some of the biases involved in the use of surveys would be manageable at the macro level.

Factors mitigating against a micro design:

2. The fact that unlike air pollution, water pollution does not lend itself to the efficient use of site-specific visual aids. This is because: a) perception of water quality is mediated strongly by individual settings; b) the diverse visual values of water include everything from clarity to surface debris; and c) not all visual degradation is due to pollution, making it difficult to distinguish between natural and the human-produced.

3. The diversity of local water bodies in many parts of the country. Lakes, streams and rivers each have different characteristics and even within a particular geographical location they may take many different forms. This diversity poses great problems for micro studies which seek to do more than measure the water quality benefits for a single body of water. Air, in contrast, is a far more homogeneous medium.

Factors favoring a macro design:

4. Both the use of fresh water (for recreation, aesthetic pleasure, etc.) and the payment for the cost of improving its quality occur at the national level. Of course, individual use takes place at the local level, but such use occurs all over the country. Moreover, some people use water in areas far distant from their homes. As noted above, every person pays the cost of improved water quality through a combination of taxes and higher prices and the cost is imposed as a result of national decisions by Congress and EPA.

5. The terms used in the national law mandating the water cleanup to describe the several levels of water quality -- "fishable," "swimmable" -- are readily understood by individual citizens and do not require location specific visual aids.

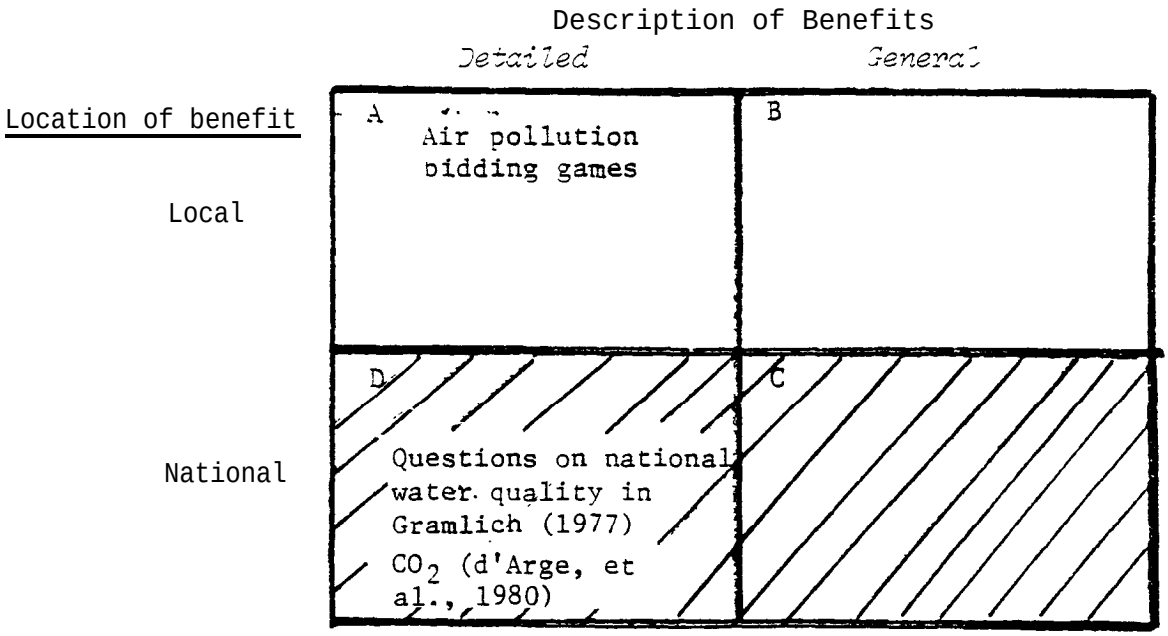
6. That a national survey is particularly suited, for reasons described earlier, for the measurement of the intrinsic value of improved water quality for our special task.

Figure 2 summarizes the major aspects of WTP benefits study design and locates the RFF approach in relation to the other types of approaches which have been used in the past. In contrast to the earlier macro studies, the description of water quality in our instrument is detailed. In contrast to the air pollution bidding games, we use a national sample and measure the benefit for the nation as a whole.

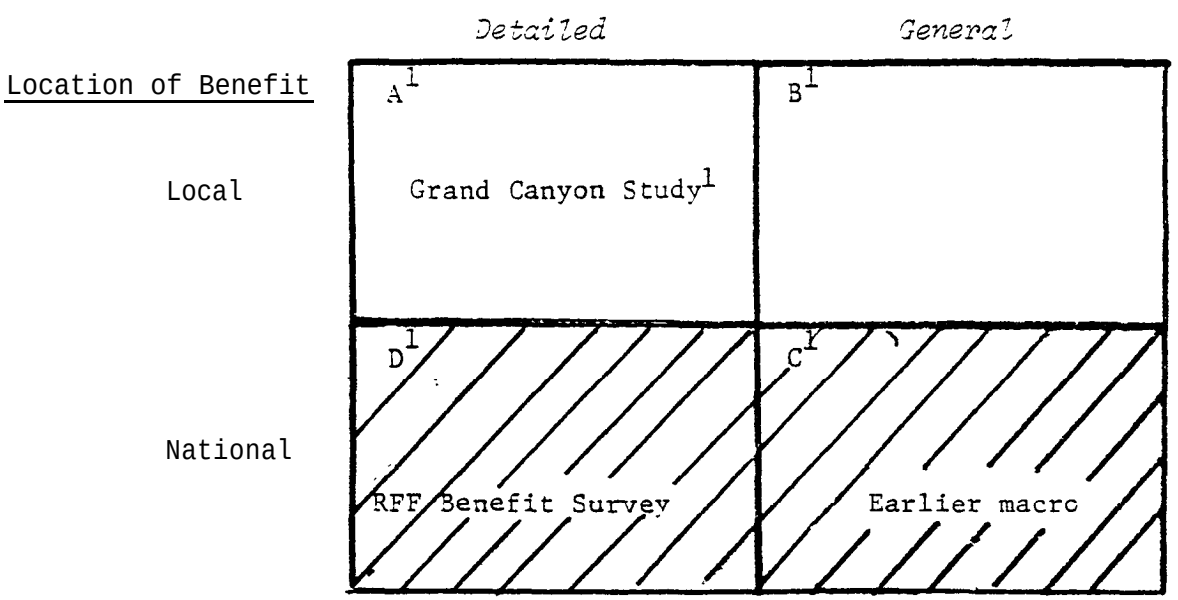
Figure 2

TYPOLOGY OF WILLINGNESS TO PAY STUDIES

Local Sample



National or Non-local Sample



[1] See footnote 1, page 2-4.