A CONTINGENT VALUATION STUDY OF LOST PASSIVE USE VALUES RESULTING FROM THE EXXON VALDEZ OIL SPILL

Richard T. Carson

Robert C. Mitchell

W. Michael Hanemann

Raymond J. Kopp

Stanley Presser

Paul A. Ruud

November 10, 1992

A Report to the Attorney General of the State of Alaska

ACE 10917117

A CONTINGENT VALUATION STUDY OF LOST PASSIVE USE VALUES RESULTING FROM THE EXXON VALDEZ OIL SPILL

Richard T. Carson

Robert C. Mitchell

W. Michael Hanemann

Raymond J. Kopp

Stanley Presser

Paul A. Ruud

November 10, 1992

A Report to the Attorney General of the State of Alaska

ACE 10917118

TABLE OF CONTENTS

CHAPTER 1 — INTRODUCTION AND OVERVIEW
§ 1.1 Introduction 1
§ 1.2 The Grounding of the Exxon Valdez 2
§ 1.3 Assessing the Value of the Services Lost 5
§ 1.3.1 The Contingent Valuation Method 5
§ 1.3.2 The Services to be Valued 6
§ 1.4 Development of the Contingent Valuation Study 8
§ 1.5 Estimate of Lost Passive Use Values 11

CHAPTER 2 — DEVELOPMENT OF THE SURVEY INSTRUMENT
§ 2.1 Introduction 12
§ 2.2 Initial Development 12
§ 2.3 Preliminary Design Research 15
§ 2.4 Key Design Issues 17
§ 2.5 Initial Pretesting 22
§ 2.6 Pilot Studies Overview 23
§ 2.7 Pilot I — San Jose, CA 26
§ 2.8 Multiple Year Payments 27
§ 2.9 Pilot II — Toledo and Dayton, OH 30
§ 2.10 Pilot III — Georgia 32
§ 2.11 Pilot IV — Toledo and Dayton, OH 35

CHAPTER 3 — STRUCTURE OF FINAL QUESTIONNAIRE
§ 3.1 Introduction 37
§ 3.2 Section A — Initial Questions 37
§ 3.3 Section A — Description of Scenario 42
§ 3.4 Section A — Valuation Questions 54
§ 3.5 Section B — Perception of Damages and Plan 60
§ 3.6 Section B — Respondent Household 62
§ 3.7 Section C — Demographic Questions 63
§ 3.8 Section C — Strength and Reassessment Questions 64
§ 3.9 Section D — Interviewer Evaluation Questions 65

CHAPTER 4 — SURVEY EXECUTION
§ 4.1 Introduction 68
§ 4.2 Sample Design 68
§ 4.3 Field Enumeration 70
§ 4.4 Interviewer Training 70
§ 4.5 Interviewer Supervision 72
§ 4.6 Quality Control Edits 72
§ 4.7 Validation of Interviews 73
§ 4.8 Interview Characteristics 74
§ 4.9 Sample Completion 75

ACE 10917119

§ 4.10 Selection Bias and Sample Weights 76
§ 4.11 Data Entry 78

CHAPTER 5 — ANALYSIS
§ 5.1 Introduction 80
§ 5.2 Attitudinal, Knowledge, and Demographic Questions 81
§ 5.3 Interviewer Assessment Questions 86
§ 5.4 Depiction of the Spill and Perceptions of Spill Prevention Plan 87
§ 5.5 WTP Questions 91
§ 5.6 Statistical Framework 95
§ 5.7 Univariate Estimation of Willingness to pay 97
§ 5.8 Reasons for WTP Responses 102
§ 5.9 Valuation Function 104
§ 5.9.1 Imputation of Missing Values For Predictor Variables 105
§ 5.9.2 Estimation of a Valuation Function 106
§ 5.9.3 Other Possible Predictor Variables 110
§ 5.10 Adjustments to WTP Responses 111
§ 5.11 Sensitivity of the Median WTP Estimate 112
§ 5.12 Stability and Replicability of Median WTP Estimate 117
§ 5.13 The Measure of Damages 122
§ 5.14 Concluding Remarks 123

REFERENCES

APPENDICES (Separate Volume)

Appendix A — Survey Instrument

Appendix B — Sample Design and Execution
Appendix B.1 — Primary Sampling Units (PSU's) for National Sample
Appendix B.2 — Sample Allocation and Completion Rates by PSU
Appendix B.3 — DiGaetano Memo of 8/12/91
Appendix B.4 — Westat Edit Form
Appendix B.5 — Westat Validation Form
Appendix B.6 — Household Screener

Appendix C — Survey Marginals
Appendix C.1 — Tabulation of Close-Ended Questions
Appendix C.2 — Coding Schemata for Open-Ended Questions
Appendix C.3 — Tabulation of Coded Open-Ended Questions

Appendix D — Question-by-Question Verbatim Responses

ACE 10917120

TABLES

Table 3.1 Visual Aids Used in Survey 44
Table 3.2 Program Cost by Version and Question 59
Table 5.1 Items Most Seriously Damaged by Spill 84
Table 5.2 Perceived Sponsor of Survey 91
Table 5.3 Program Cost by Version and Question 93
Table 5.4 A-15 Response by Version 94
Table 5.5 Questionnaire Version by Type of Response 95
Table 5.6 Turnbull-Kaplan-Meier Estimation Results 97
Table 5.7 Weibull Estimates 100
Table 5.8 Medians and Means for Four Distributions 99
Table 5.9 Reasons Not Willing To Pay Amount 103
Table 5.10 Reasons Not Sure Whether Willing To Pay 104
Table 5.11 Reason For Being Willing To Pay 105
Table 5.12 Prediction of Log Income 107
Table 5.13 Weibull Valuation Function 108
Table 5.14 Summary of Sensitivity Tests 118
Table 5.15 Distribution of Responses Across Surveys 120
Table 5.16 Dollar Amounts Used in Each Survey 121
Table 5.17 Weibull Hazard Model Estimation For Each Survey 122

FIGURES

Figure 5.1 Weibull Estimate of Percent Willing to Pay as a Function of Amount Specified 101
Figure 5.2 Estimated Survival Curves 123
Figure 5.3 95 percent Confidence Intervals for National WTP Survival Curve 124

ACE 10917122

CHAPTER 1 — INTRODUCTION AND OVERVIEW

§ 1.1 Introduction

This report summarizes the development, implementation, and results of a contingent valuation (CV) study designed to measure the loss of passive use values[1] arising from injuries to natural resources caused by the Exxon Valdez oil spill. The study was undertaken for the State of Alaska in connection with the State's action against the Exxon Corporation, Exxon Shipping Company, and Alyeska Pipeline Service Company and its owners.[2]

This report consists of this introduction, the four chapters following it, and appendices. Chapter 2 describes the development of the contingent valuation survey instrument. Chapter 3 presents and discusses the final survey instrument used in assessing the damages.[3] Chapter 4 discusses the technical aspects of the survey's administration and the processing of the survey data. Chapter 5 contains the analysis of the data collected and includes the estimation of damages. This report also contains several appendices related to the survey instrument and the data collected using it.

The core study team for this contingent valuation project was led by Richard T. Carson of the University of California (San Diego) and Robert Cameron Mitchell of Clark University. The other members of the study team were W. Michael Hanemann of the University of California (Berkeley), Raymond J. Kopp of Resources for the Future, Stanley Presser of the

---

[1]Passive use values encompass what economists refer to as option values, existence values, and other nonuse values (Mitchell and Carson, 1989; Kopp and Smith, forthcoming 1993). *See Ohio v. Department of Interior*, 880 F.2d 432 (D.C. Cir. 1989).

[2]*Alaska v. Exxon et al.*, Case No. A92-175 Civil (D. Alaska). Originally filed August 15, 1989, in State Superior Court, Third Judicial District.

[3]Throughout this report, the physical effects of the spill of oil on the natural resources are called *injuries*, while the monetized value of these injuries are called *damages*.

University of Maryland (College Park), and Paul A. Ruud of the University of California (Berkeley).[4] Carson, Hanemann, and Kopp are resource economists; Ruud is an econometrician; and Mitchell and Presser are survey researchers.

Lexecon, Inc. served as project coordinator and special consultant to the state litigation team. Serving in various advisory capacities were Richard C. Bishop of the University of Wisconsin (Madison), Gardner M. Brown of the University of Washington (Seattle), Howard Schuman of the University of Michigan (Ann Arbor), Norbert Schwarz of the *Zentrum fuer Umfragen Methoden und Analysen* (Mannheim, Germany), Paul Slovic of Decision Research (Eugene, Oregon), and Robert M. Solow of the Massachusetts Institute of Technology. Bishop, Brown, and Solow are economists; Schwarz and Slovic are cognitive psychologists; Schuman is a survey researcher. None of these individuals is responsible for any decisions concerning the study or this report; the authors bear sole responsibility for any errors or omissions.

§ 1.2 The Grounding of the Exxon Valdez[5]

Prince William Sound (the Sound) lies near the top of the 850-mile arc of the Gulf of Alaska which extends from the Aleutian islands on the west to the islands of southeast Alaska. It is a remote, rugged area of great natural beauty. Much of this region was pristine before the spill. Prince William Sound is one of the continent's largest tidal estuary systems, a rich environment where rivers meet and mingle with the tides. In terms of water surface alone, the

---

[4]The authors wish to acknowledge Michael Conaway and Kerry Martin of Natural Resource Damage Assessment, Inc., who provided administrative and logistical support to the study team, and Valerie Fraser Ruud who provided editorial assistance.

[5]The discussion of the grounding of the Exxon Valdez and the characterization of Prince William Sound and the resulting spill of oil are taken from the "State/Federal Natural Resource Damage Assessment Plan for the Exxon Valdez Oil Spill: Public Review Draft," published by the Trustee Council, Juneau Alaska, August 1989.

Sound is about the size of Chesapeake Bay. Its many islands, bays, and fiords give it a shoreline more than 2,000 miles long.

The Sound lies within the boundaries of the Chugach National Forest. To the southwest is the Kenai Peninsula, which contains the Kenai Fiords National Park. The western portion of the Sound is within the Nellie Juan-College Fiord Wilderness Area; both the National Forest and National Park are accessible by air and boat from Anchorage, Alaska's major population center, making the area popular with recreationists. State ferries run among the larger communities. In recent years, the number of cruise ships and other tourist visits to the area has steadily increased.

The Kenai Peninsula points southwest to the Kodiak Archipelago and the Alaska Peninsula which are separated by the Shelikof Strait. Along the Alaska Peninsula's coast is Katmai National Park. Southeast of the Strait lies Kodiak Island, once the base of Russia's Alaskan sea otter fur trade which nearly destroyed these native mammals through excessive hunting. Their numbers, coaxed back from the edge of extinction, had grown back to a healthy population throughout the spill-impacted area. The Alaska Peninsula tapers, then scatters into the islands of the Aleutian Chain.

The maritime climate nourishes a lush landscape. Bears, whales, bald eagles, puffins, seals, sea lions, and sea otters are among the wildlife people come to see. Glaciers that carved the intricate fiords still send icebergs floating out to sea. These are the largest glaciers outside Antarctica and Greenland. They descend from permanent ice fields capping the coastal Chugach mountain range.

The Trans-Alaska Pipeline System terminates at the port of Valdez on the northern edge of the Sound. In 1989, the pipeline carried two million barrels a day of oil produced on

ACE 10917125

Alaska's North Slope. Approximately two tankers per day load Trans-Alaska Pipeline System oil at Valdez and transit the Sound.

At 12:04 a.m., March 24, 1989, the tanker Exxon Valdez, carrying more than 50 million gallons of North Slope crude oil, ran aground and ruptured its tanks on Bligh Reef in Alaska's Prince William Sound. The oil spill that followed was the largest tanker spill in U.S. history. Approximately 11 million gallons of crude oil poured into the Prince William Sound in less than five hours. By August 1989, the oil had moved across nearly 10,000 square miles of water in Prince William Sound and the Gulf of Alaska. More than 1,000 miles of shoreline were oiled.

The oil killed thousands of wild animals. Oil and its breakdown products are expected to linger in some areas for years, affecting or potentially affecting:

- Surface water and sediments;
- Land managed by natural resource trustees, including submerged land, wetlands, shoreline, beaches, geologic resources, and other features of the land;
- Marine plants and microorganisms;
- Fish, shellfish, and other marine invertebrates;
- Marine mammals, including sea otters and seals;
- Birds, including seabirds, waterfowl, shorebirds, and raptors.

The State of Alaska filed suit against the Exxon Corporation and other potentially responsible parties claiming compensation for a wide range of natural resource injuries.

Shortly after the Exxon Valdez oil spill, the State of Alaska and the United States undertook a series of joint scientific studies to identify injuries to natural resources resulting from the spill. The state also undertook the economic studies required to quantify certain types of losses. The contingent valuation study discussed in this report was conducted to measure the loss of passive use values.

§ 1.3 Assessing the Value of the Services Lost

Because the resource injuries would give rise to lost passive use values and because the contingent valuation method is the only technique currently available for measurement of such values, the State of Alaska commissioned a state-of-the-art contingent valuation study. The CV team was provided with a description of natural resource injuries caused by the Exxon Valdez oil spill that included the nature and magnitude of the injury and the time frame for recovery. These injuries included: oiled shoreline, bird and mammal deaths, and effects on fish. These injury estimates were understated for the reason that, in January 1991, when the study went into the field, some of the crucial science studies were not yet completed. Hence, lower limits of then current estimates of injuries were used in order to avoid litigation issues relating to what might later prove to be overstatements of provable injuries. Similarly, optimistic restoration or recovery periods were used for the same reason.

§ 1.3.1 The Contingent Valuation Method

The CV method uses survey questions to elicit peoples' values for private or public goods or services by determining what they would be willing to pay for specified changes in the quantity or quality of such goods or services or what they would be willing to accept in compensation for well-specified degradations in the provision of these goods or services.[6] The method attempts to elicit peoples' willingness to pay (WTP) or willingness to accept (WTA) compensation in dollar amounts. The CV method circumvents the absence of markets for services provided by natural resources by presenting consumers with hypothetical markets in which they have the opportunity to buy or sell the services in question. The market in a

[6]Much of the discussion in this section is drawn from Mitchell and Carson (1989) and Carson (1991).

ACE 10917127

contingent valuation study may be modeled after either a private market or a political referendum. Because the elicited values are contingent upon the particular hypothetical market described to the respondent, this approach came to be called the contingent valuation method.

Generally, respondents are presented with survey material which consists of three parts:

1. A detailed description of the services being valued and the hypothetical circumstance under which it is made available to the respondent. The researcher constructs a model market in considerable detail which is communicated to the respondent in the form of a scenario during the course of the interview. The scenario describes the services to be valued, the baseline level of provision, the structure under which the services are to be provided, and the method of payment. All elements of the scenario must be designed to maximize its plausibility.

2. Questions that elicit the respondent's value for the services. These questions are designed to facilitate the valuation process without biasing the elicited dollar amounts.

3. Questions about the respondent's characteristics (*e.g.*, age, income), preferences relevant to the services being valued, and use of the services. This information, some of which is usually elicited preceding and some following the scenario, is used to estimate a valuation function for the services.

§ 1.3.2 The Services to be Valued

The values obtained in this study are almost exclusively passive use values due to two key aspects of the study.[7] First, private services such as commercial fishing, which were being claimed by private parties, were excluded from the injury scenario. Second, with direct use public services, such as recreational fishing, the principal user groups are comprised primarily of Alaskan residents. In the multi-stage sample selection process, no Alaskan households were included in the final sample. As a result of this random selection, the vast majority of recreational users of the area affected by the Valdez Spill had no chance of being selected to be

[7]The contingent valuation technique measures total value, *i.e.*, direct use values and passive use values.

ACE 10917128

interviewed.[8] Therefore, the damage estimates produced by this study are comprised almost entirely of lost passive use values.

The value of services may be measured in terms of willingness to pay or willingness to accept. In the WTP context, individuals are asked the maximum they would pay to obtain an additional quantity or improvement in the quality of some service or group of services; in the WTA context, individuals are asked the minimum amount they would accept for a decreased quantity or degraded quality of some service. If WTP and WTA were the same for most individuals and services, the choice between them would not be a problem for damage estimation; but, as Hanemann (1991) has demonstrated, a substantial difference between the two is possible for services provided by non-marketed resources. Therefore, the choice between WTP and WTA can have important consequences.

Theoretically, the choice of willingness to pay or willingness to accept depends on the assignment of property rights. In the case of Prince William Sound and other affected areas, the rights to the services are held in trust for present and future generations of Americans. Since the public holds the rights to the services, the correct measure of the value of the degradation in those services is the minimum amount of money the American people as a whole would voluntarily agree to accept to suffer the loss or disruption of the services. Thus, willingness to accept compensation is the theoretically correct measure in this case.

Unfortunately, it is very difficult to design a survey that effectively elicits WTA amounts because respondents tend to regard WTA scenarios as implausible.[9] Therefore, in the current damage assessment, we chose willingness to pay as the valuation framework even though this

---

[8] Had these households been interviewed, their willingness-to-pay responses may have been motivated to a substantial extent by direct use considerations.

[9] See Mitchell and Carson (1989) for a detailed discussion of the problems involved in eliciting WTA responses in contingent valuation studies.

ACE 10917129

choice will understate the true value of losses suffered as a result of the spill, other things being equal.

The next issue is the precise nature of the services to be valued. We would like to position individuals immediately prior to the grounding of the Exxon Valdez and elicit from them the maximum amount of money they would be willing to pay to prevent the losses in services about to be caused by the spill. However, this can present methodological problems because it is very difficult for individuals to mentally "travel back in time" to just before the spill and reliably reveal what their preferences would have been. This problem can be overcome by valuing a comparable reduction in services in the future. In the CV study we conducted, respondents were told that if no action is taken over the next 10 years another oil spill will almost certainly cause injuries to Prince William Sound comparable to those of the Exxon Valdez spill. Respondents were then asked their willingness to pay for a realistic program that would prevent with certainty the injuries which would be caused by such a spill.

§ 1.4 Development of the Contingent Valuation Study

The assessment of lost passive use values arising from the injuries to Prince William Sound involved a sequence of activities which are described in more detail in the following chapters. We will briefly introduce the sequence of activities to provide the reader with a "road map" to the CV study. The process began with the identification of the injuries to the Sound, the magnitude and severity of each injury, and the time required for the Sound to naturally recover. As noted above, injury information was provided to the CV team by natural scientists working for the State of Alaska and was updated periodically. The injury data provided the informational basis for the loss of resources and associated services which were to be valued in the CV survey.

ACE 10917130

The contingent valuation design process began with the development of the valuation scenario, the heart of a CV survey. The initial stage of the scenario development used information gained from a series of six focus groups.[10] These groups, which were conducted in the states of Washington, Alaska, Maryland, Virginia, Missouri, and California, allowed us to explore how individuals perceived the spill and its consequences. We also explored the assumptions individuals brought to the valuation process, assumptions which might help or hinder the elicitation of valid and meaningful values for the spill injuries.

Upon completion of the focus groups, a preliminary draft survey incorporating the valuation scenario was developed. This draft was first tested by administering the survey to a series of individuals who were paid to participate in the survey testing. Observing their responses during the interview and debriefing these respondents afterward provided information upon which to base revisions to the survey instrument.

After repeated testing and revision in this manner and also in field interviews, the draft survey instrument was further refined and then tested in a series of four pilot surveys in different parts of the country. These pilot tests were in-person interviews of a relatively small sample of randomly chosen respondents conducted by professional interviewers. After each pilot survey, the data were analyzed, the interviewers debriefed, and revisions were made to the survey instrument. The use of pilot surveys and instrument revision is an effective iterative procedure which can produce a high quality, reliable survey instrument. The process of developing the survey instrument is described in Chapter 2, and the final survey instrument itself is described in Chapter 3.

---

[10]Focus groups are group discussions up to two hours in length which consider topics introduced by a moderator who leads the discussion. Focus groups are used to explore people's beliefs, attitudes, and knowledge about a particular subject.

ACE 10917131

The survey firm retained to administer the surveys was Westat, Inc. of Rockville, Maryland. Westat is one of the country's most respected survey research firms and is often retained by government agencies to conduct their most exacting surveys. Westat conducted intensive interviewer training, provided field supervision, validated the interviews, and exercised quality control over sampling, data collection, and coding.

Once the survey instrument was finalized, a sample of households to be interviewed was drawn by Westat using standard multi-stage area probability sampling techniques to represent all 50 states and the District of Columbia. Using this procedure, a random sample of 1,599 dwelling units was drawn. Visits to each unit established that 176 were vacant, leaving a final sample of 1,423 occupied dwelling units from which the individual respondents were drawn by further sampling at the household level. Professional interviewers then attempted to administer the survey to each selected respondent. In some instances, even after repeated efforts, no one was found at home; in other cases, respondents refused repeated attempts by interviewers to complete the interview; and in other instances, no one in the household spoke English.[11] In all, 1,043 interviews were completed with a resulting response rate of 75 percent. This response rate is comparable to those of the very best academic surveys. As the surveys were completed, they were coded by Westat and sent in batches to Natural Resource Damage Assessment, Inc. (NRDA) where they were independently recoded and checked against the data provided by Westat. Chapter 4 describes the sample design and survey execution. Once all data were verified, the CV team began to analyze the information statistically and to produce damage estimates.

[11]These non-English speaking households were subtracted from the population to which the estimate would later be extrapolated.

ACE 10917132

§ 1.5 Estimate of Lost Passive Use Values

The CV survey revealed that the Exxon Valdez oil spill was spontaneously mentioned by over half the respondents as one of the largest environmental accidents caused by humans anywhere in the world; and over 90 percent of the respondents said they were aware of the spill. The median household willingness to pay for the spill prevention plan was found to be $31. Multiplying this number by an adjusted number of U.S. households results in a damage estimate of $2.8 billion dollars. A number of alternative statistical assumptions tend to result in only fairly small changes to this estimate. In contrast, mean willingness to pay, which is higher than median willingness to pay, is quite dependent on the particular distributional assumption made, and a very wide range of estimates are hence possible. We, therefore, concentrated on the median household willingness to pay in this report. It represents a statistically solid lower bound for the damage estimate.

A valuation function was also estimated to predict willingness to pay as a function of a respondent's characteristics and perception of the plan and the damages it would prevent. This valuation function has significant explanatory power and is consistent with theory and intuition. It can be used to make adjustments for protest responses, for perceptions of damages prevented which are larger or smaller than those of Exxon Valdez spill, and for differences in the perceived effectiveness of the spill prevention plan. The result of these adjustments suggests that the estimate of median household willingness to pay is a conservative estimate.

Two pilot studies and a separate "tracking" study (all in Dayton and Toledo, Ohio) demonstrate that the median willingness-to-pay estimate is stable over the course of a year and several replications.

ACE 10917133

CHAPTER 2 — DEVELOPMENT OF THE SURVEY INSTRUMENT

§ 2.1 Introduction

The survey instrument used for the Exxon Valdez study was developed over 18 months from July 1989 to January 1991, when the final survey was put into the field. The central part of the survey instrument is the valuation scenario that describes the damages caused by the Exxon Valdez oil spill. A referendum market is established in the instrument for eliciting the value the respondent places on preventing a future accident that would cause an equivalent amount of damage in the Prince William Sound area. Other questions preceding and following the scenario ask about the respondent's attitudes, previous awareness of the spill, understanding of the scenario, and personal characteristics. At appropriate places during the in-person interview, display cards, photographs, and maps are shown to the respondent to supplement the information conveyed verbally by the interviewer.

§ 2.2 Initial Development

We conducted an extensive program of instrument development research for this study. In the first stage of instrument development, we conducted exploratory research primarily through focus groups. In the second stage, we produced the first draft questionnaire and revised it during a series of one-on-one interviews followed by informal field testing. The third and final stage involved formal field testing and development work, including a series of four pilot surveys. In the second and third stages, the survey instrument was continually revised on the basis of preceding work. Throughout the process we followed established survey research methodology to ensure the reliability and validity of the final results.

ACE 10917134

The research goal was to develop a valid survey instrument to measure the value of lost passive use values due to the natural resource injuries caused by the Exxon Valdez oil spill. In designing the survey instrument we sought to meet five objectives:

1. valuation of only the injuries defined in the survey;
2. consistency with economic theory;
3. scenario comprehensibility;
4. scenario plausibility; and
5. an overall perception of neutrality by the respondents.

The first objective was to measure only a defined set of injuries. That objective required carefully describing the specific injuries to be valued and the various recovery times for the injured resources and ensuring as much as possible that respondents did not value more extensive or less extensive injuries than intended. The description of the injuries was based on the best available scientific information. Open-ended questions at various points in the valuation scenario and diagnostic questions which followed the valuation scenario were used in the survey instrument to assess our success in meeting this goal. The latter type of question obtained information which could be used to adjust the WTP estimate to compensate for assumptions about the injuries which differed from those we intended.

The second objective was to develop an instrument that is consistent with economic theory. Specifically, the instrument was designed to obtain an approximation to the monetized loss in utility suffered by the respondents as a result of the injuries caused by the spill. The third objective is a basic survey research goal: potential respondents from all educational levels and varied life experiences should be able to comprehend the language, concepts, and questions used in the survey. We undertook an extensive instrument development research program, described in this chapter, to help us reach this and the final two objectives. We also made a

ACE 10917135

special effort to develop visual materials to enhance the communication of the scenario. These included tables, drawings, and a book of photographs.

Plausibility, the fourth objective, requires that a respondent find the scenario and the payment vehicle believable and take the choice situation seriously. To this end, we adopted the referendum format which asks each respondent to make a judgment as to whether they would vote for or against a program that, if adopted, would cost their household a certain, specified amount in addition to what their household already pays for the use of natural resources and other public good amenities.

The fifth objective is neutrality: the wording and information in the instrument should not be perceived by respondents as promoting the interests of any particular party and that the survey is not consistently perceived as sponsored by any particular party.[12] The instrument's wording was reviewed at various stages in its development by outside reviewers to assess our success in meeting this objective. When faced with a decision between two options where a neutral wording choice was not dictated on the basis of theory or solid methodological ground, we endeavored to choose the conservative option.

In addition to the survey design objectives presented above, there are important decisions regarding the description of the natural resource injuries. The injuries must be described in a balanced fashion. Uncertainty regarding the precise extent of some of the injuries was substantial at the time the final CV survey was conducted. The state chose to have the CV team value a conservative representation of the injuries in order to minimize the litigation risk associated with that uncertainty. Therefore, only injury facts of which scientists where

[12]Respondents and interviewers were not told either that the survey was being conducted for litigation or who was sponsoring the survey.

ACE 10917136

reasonably certain as of the fall of 1990 were used.[13] When the best estimate of the actual state of affairs required a range, the conservative end of that range was used; for example, for animals deaths and the extent of the oiling, this rule required that the lower end of the ranges be used.

§ 2.3 Preliminary Design Research

Early in the first stage of our design research we conducted a series of six focus groups in different locations around the United States, which were followed a year later by a seventh group. Focus groups are group discussions, usually two hours in length, that consider topics introduced by a moderator who leads the discussion. Focus groups are held in a facility with an observation room with a one-way mirror so the researchers can discretely observe the discussion. The 8 to 12 participants are typically members of the general public who are recruited by a market research firm and offered a payment for their participation. The focus group is also tape-recorded for further analysis. Increasingly, this type of qualitative research is used by survey researchers in the early stages of designing contingent valuation questionnaires because they are an efficient way to explore people's beliefs, attitudes, and knowledge about the subject matter, *e.g.*, the Exxon Valdez oil spill, and to obtain their reactions to possible CV scenario elements.

The locations and dates of the focus groups conducted for this study are:

[13]The scientific facts were provided in discussions with Robert Spies, the Chief Scientist for the Joint State-Federal Natural Resource Damage Assessment.

ACE 10917137

1.	Seattle, Washington	July 21, 1989
2.	Anchorage, Alaska	July 24, 1989
3.	Baltimore, Maryland	August 6, 1989
4.	Fairfax, Virginia	August 7, 1989
5.	St. Louis, Missouri	August 17, 1989
6.	San Diego, California	August 25, 1989
7.	New Orleans, Louisiana	March 24, 1990

These sites were selected to provide information from people in diverse parts of the country. Robert Mitchell moderated each focus group discussion. The participants were randomly recruited by a local market research firm from the telephone directory in each city. All participants were aged 18 years and older. The recruiters used a screening questionnaire to recruit pre-set quotas of people and to exclude those who had previously taken part in any focus group. In most cases, the quotas ensured that the group included a balanced number of men and women, a range of ages, and a range of educational attainments. The only exception was the St. Louis group, which was restricted to people living in blue collar households in order to advance our understanding of the views of this segment of the population.

To reduce selection bias and to enable us to assess their pre-existing views about the spill, the focus group participants were not told that the discussion would focus on the Exxon Valdez oil spill until after the first part of the group discussion. During recruitment they were told merely that the discussion would be on unspecified "public issues."[14] The identity of the research sponsor was not revealed at any point to the participants or to the market research firms who recruited them.

In the first focus groups, the discussions explored the participants' knowledge of the Exxon Valdez spill, their beliefs about the cause and nature of the damage, and their perception of the plausibility of possible ways of preventing a future spill. Once particular patterns of

[14]Those who agree to participate in a focus group on a particular topic may not be representative of the general population. This effect is known as selection bias.

ACE 10917138

understanding and knowledge were established and confirmed, new topics were introduced in subsequent groups. In later groups, elements of a possible questionnaire were described in more detail to help us understand how the participants understood these elements and how they used them in the valuation process. These included the payment vehicle, the duration of payments, the description of the damages, the description of a plan to prevent future spills, and the use of particular photographs and maps to communicate factual aspects of the scenario.

§ 2.4 Key Design Issues

In addition to the determination of the good to be valued, the designer of a contingent valuation study must make a number of other decisions about key design issues. These include the choice of the elicitation method, the nature of the payment vehicle, the number of years over which payments are collected, and whether the good is valued in a sequence of other goods.

With respect to the elicitation method, we determined early in the process that respondents should be asked a binary discrete choice question (Bishop and Heberlein, 1979). This type of question, often called a take-it-or-leave-it question, requests the respondent give a yes-or-no response to a specific cost. A single take-it-or-leave-it question is incentive-compatible under fairly general conditions; that is, a respondent can do no better than saying "yes" if the policy is actually preferred at the specified cost or by saying "no" if otherwise. We extended the simple binary discrete choice elicitation to the double-bounded dichotomous choice question (Hanemann, Loomis, and Kanninen, 1991) where the respondent is asked to give a yes-or-no response to a second pre-specified higher amount if the response to the initial take-it-or-leave-it question is "yes" and to a pre-specified lower amount if the initial response is "no." Using both the first and second responses substantially increases the statistical power of the WTP estimate, *i.e.*, it tends to produce a much tighter confidence interval for the WTP estimate for any fixed

ACE 10917139

sample size; however, it does so at the expense of a small downward bias in the estimate because the second response is not, in general, incentive-compatible.[15]

There are three natural choices for the payment vehicle: higher oil prices, higher taxes, and higher prices on a wide range of goods. It is also possible to be more specific, *e.g.*, higher gasoline prices, or to combine payment vehicles, *e.g.*, higher prices and taxes. In selecting a payment vehicle, one looks for broad acceptance of that vehicle as a fair method of paying for the good.[16] One also looks for good coverage; that is, one looks for a payment vehicle by which almost all of the respondents could be compelled to pay. A gas tax, for example, may not be relevant to households without a car. Furthermore, the vehicle should be plausible: the payment vehicle should be perceived as a likely way to pay for the good. Finally, one seeks stability: other policies should not be simultaneously causing large changes in revenue collected via the same payment vehicle used in the survey. Sections 2.9 and 2.10 describe the testing of different payment vehicles during our instrument development research.

With respect to the number of years over which payments are collected, there are three major issues. First, longer payment periods mean that budget constraints, particularly for poorer households, are less binding. Second, periodic payments tend to assure respondents that the good will be provided in future years. Third, "out of sight" goods raise the question of how "committed" a respondent is to the stream of multi-year payments. For reasons discussed in Section 2.8, a single year payment vehicle was adopted.

---

[15]This downward bias is suggested by empirical evidence and probably results from expectations formed by the initial cost estimate given to the respondent. Some respondents who vote to pay the first amount might be willing to pay the second (higher) amount but vote against the higher amount when asked because they feel that the government would waste the extra money requested. In addition, some respondents who are not willing to pay the first amount would be willing to pay the second (lower) amount but may vote against the second amount because they believe that either the government will deliver a lower quality good than that first promised or that the probability of the government delivering the good is lower at the lower price. Both of these voting patterns would result in a downward bias. The extent of the bias depends on the degree to which the second amount is perceived by the respondent as an independent cost estimate.

[16]Protest zeros often result from rejection of the payment vehicle as an appropriate means of paying for the good.

Finally, there are two choices related to "embedding." The first is whether to value the good of primary interest by itself or in a sequence of other goods. Here economic theory provides some important guidance for the valuation of natural resource damages.[17] Due to substitution and income effects, the later in a willingness-to-pay sequence a good is valued, the lower its value.[18] The opposite is true of a willingness-to-accept compensation sequence; the later in such a sequence a good is valued, the greater its value.[19] These two propositions can be combined with the fact that willingness-to-accept compensation for a good is greater than or equal to willingness-to-pay for the same good (Hanemann, 1991) to show that valuing a good first (*i.e.*, by itself) in a willingness-to-pay sequence is the closest that one can get to whatever sequence-specific willingness-to-accept compensation measure is desired (short of measuring willingness-to-accept directly, which cannot generally be done).

The second "embedding" choice is methodological: what is the best design to ensure that the respondents do not answer a different question than the one they are asked, whether by forgetting about their budget constraints or by letting Prince William Sound stand for all oil spills or even all environmental damage? To meet this requirement, the scenario *must* present a plausible choice situation describing the good and its method of provision in adequate detail so that the respondents know what they will and what they will not get. The design choice is whether to value multiple goods in a single survey or to value a single good and carefully differentiate it in the instrument from those other goods with which it might be confused. We

---

[17]For discussions, see Hoehn and Randall, 1989; Bishop, 1990; Carson, Flores, and Hanemann, 1992; Randall and Hoehn, 1992.

[18]These two statements are also true for private goods. Randall and Hoehn (1992) show substantial sequencing effects for a common commodity, *i.e.*, rice in an empirical food demand system. They also show how the phenomena of incomplete multi-stage budget optimization tends to increase the magnitude of sequencing effects.

[19]The income effect is assumed to be positive. Also, these conclusions depend upon the assumption that the environmental amenities embedded together are economic substitutes. Complementarity would imply opposite results.

ACE 10917141

decided to use the single good CV survey for two reasons. First, it avoids several difficulties which are introduced by valuing multiple goods. Second, well designed single-good CV surveys have been shown to be capable of eliciting values that are sensitive to the characteristics of the good being valued.

The first of the two major difficulties with the multiple goods approach is that the more different goods that must be valued in a given CV instrument, the less detail that can be devoted to any particular good.[20] Given the amount of information necessary for the Prince William Sound scenario, adding valuation scenarios for additional goods would have required an unmanageably long interview.[21] The second is that the two most common approaches to valuing multiple goods, asking a series of valuation questions which are intended to be independent of each other and asking an allocation question, both involve serious difficulties in interpretation. A sequence of "independent" valuation questions in a single interview makes the questionable assumption that respondents will be able to value each good independently of the others. Respondents will typically have formed some expectation regarding the likely provision of the first good which it will be hard to get them to disregard without emphasizing the hypothetical quality of the choice situation and thereby detracting from the scenario's plausibility. Allocation questions also have problems as the willingness-to-pay questions are typically ambiguous because they do not specify the conditions under which the good in the

[20] The two primary policy-related reasons for valuing multiple goods are: (1) a desire to value a set of goods which will be provided as a package and (2) a desire to trace out the complete benefit curve for a good by obtaining willingness to pay for successive increments to the current level. The cost of doing a large contingent valuation study encourages policy makers to try to value as many different policy options as possible. There is an obvious trade-off between this objective and the quality of the results obtained. This is not generally an issue in a natural resource damage assessment since the set of injuries has been determined exogenously.

[21] The interviews for this study, with one good, required a median length of 40 minutes to administer. Describing an additional related good in sufficient detail to ensure that respondents understood the characteristics of both goods and the valuation context associated with each would have increased the median interview length to over an hour and substantially increased the effort required of the respondent.

ACE 10917142

second question is to be provided and different respondents will make different assumptions about those conditions.[22]

With respect to the single-good CV survey approach, some have argued on the basis of experiments (*e.g.*, Kahneman and Knetsch, 1992) that respondents in such a survey are incapable of sensitivity to the inclusiveness[23] of the good they are being asked to value. This judgment is faulty because most of these experiments do not emulate the type of market and detailed description of the good used in our study and a number of other studies conducted for policy purposes.[24] Other experiments, which do emulate these features, find respondents are capable of responding to the inclusiveness of the good.[25] In addition, there is considerable evidence in the literature that in well-designed contingent valuation surveys, respondents give quite different values for different types of environmental goods that differ considerably in scale. To make an extreme comparison, Carson *et al.* (1992) found that respondents were willing to pay on average less than $1 to improve visibility in the Grand Canyon on ten poor weather days

---

[22]A dramatic but simplistic example of a private good demonstrates this concern. Assume that our respondent's car coasts into the only gas station on a long stretch of desert road with a leaking radiator and out of gas. Ask the well-defined question, "How much are you willing to pay right now for fixing the radiator and a tank of gas?" Now ask the allocation question, "How much of that amount is for the tank of gas?" The respondent's answer should depend on whether the gas station has already fixed the radiator and been paid; and, if not, whether the gas station can fix the radiator; and, if so, what the cost of fixing the radiator is going to be.

[23]By inclusiveness we mean a situation where one good is nested within a larger good. An example frequently used by Kahneman and Knetsch (*e.g.*, 1992) is cleaning up all lakes in Ontario versus cleaning up the lakes in just one region of Ontario.

[24]See Smith (1992) for a discussion of Kahneman and Knetsch's work in this regard. Mitchell and Carson (1989) and Carson and Mitchell (1992) discuss survey design problems which may cause respondents not to value different goods differently.

[25]Carson and Mitchell (1992) show that respondents clearly distinguished between differences in the inclusiveness of goods in split-sample experiments performed in two large contingent valuation surveys which used discrete choice referendum formats. Both surveys involved situations unfamiliar to respondents. In the first survey, which involved predominantly use considerations, respondents valued preventing water shortages of different magnitudes and frequencies in California; while in the second survey, which involved predominantly passive use considerations, respondents valued preventing risks from mining of different magnitudes and geographic extent in a remote but well known national park in Australia.

ACE 10917143

during the winter, while Randall and Kriesel (1990) found that respondents were willing to pay an average of almost $700 for substantial improvements in several national environmental programs.[26]

In constructing the scenario for this study, we took several steps to minimize the possibility of respondent perceptual error in understanding the good they are being asked to value. First, we paid particular attention in the focus groups and in-depth interviews to how people think about the good we offer them. Second, we used this knowledge, in ways that will be described later, to focus the respondents' attention on what they would and would not get if the program was implemented. Third, each time we used the instrument, both during the development process and in the final interview itself, we asked open and close-ended questions to assess how well respondents understood what we were attempting to convey in the survey. This enabled us in the analysis to identify the presence of any remaining perceptual problems and, to the extent that they were present, to determine if and how they affected the results (see Chapter 5).

§ 2.5 Initial Pretesting

In the second stage of our development work, which took place in the fall of 1989, a draft of the questionnaire was developed and used to conduct trial interviews. During these one-on-one interviews, which took place at Westat's office in Rockville, Maryland, the instrument was continually revised to address various problems that became apparent in the interviews or in post-interview discussions with the respondents. Toward the end of this period, the then

[26]Taking a broader view, Walsh, Johnson and McKean (1992) performed a meta-analysis of 129 contingent valuation estimates involving outdoor recreation conducted between 1968 and 1988. They found that these contingent valuation estimates were sensitive to site quality, region of the country, and type of activity.

ACE 10917144

current draft was subjected to preliminary field testing by a few of Westat's most experienced interviewers. After they had administered several personal interviews, these interviewers were debriefed to assess how well the instrument worked and how it might be improved. In December of 1989, a revised version of the instrument was delivered to Westat for the next round of testing.

§ 2.6 Pilot Studies Overview

The third stage of our instrument development research took place from February to November 1990, when Westat interviewers conducted four sequential pilot surveys at sites in different parts of the country. Each pilot was followed by an interval long enough to allow the data to be analyzed and the questionnaire to be revised to reflect the results of the analysis and interviewer debriefings. Through this iterative process, the instrument was revised and improved until we were confident it met our research objectives.

The pilot survey sites were selected to represent three parts of the country with different socioeconomic characteristics. All interviews were conducted by professional interviewers, face-to-face, at the respondent's home. The location, date, and sample size (N) of the pilot surveys are as follows:

Pilot I.	San Jose, California SMSA, February, 1990, N=105
Pilot II.	Toledo & Dayton, Ohio SMSA's, May, 1990, N=195
Pilot III.	Five rural counties in Georgia, September-October, 1990, N=244
Pilot IV.	Toledo & Dayton, Ohio SMSA's, November, 1990, N=176

The respondents for each pilot study were selected in three stages, the sample size depending on the purposes of the particular pilot. First, a small number of census tracts were selected to cover the demographic groups of interest in the pilot site. Second, listing procedures produced representative samples of households within given tracts. Every *n*th address within

ACE 10917145

an assigned tract was listed by listers working block by block through the tract. This created a list of dwelling units that was used to form a sampling frame. Third, interviewers were assigned to dwelling units where, at the household level, they conducted a screening interview to identify all eligible respondents. These were defined as people aged 18 or older who own or rent their home or pay toward the rent or mortgage. The survey respondent for a given household was randomly selected from this list of eligible respondents.

For each pilot, Westat recruited the interviewers, prepared the interview materials based on the instrument we delivered to them, conducted the interviewer training, supervised the production of interviews in the field, and edited and validated the completed questionnaires. With the exception of a small number of senior Westat officials and the study's project manager and field manager, no Westat employee, including the interviewers and field supervisors, was told who was sponsoring the study at any time during the study. This secrecy helped to minimize the chance that the interviewers would consciously or unconsciously bias the findings in favor of the sponsor.

Working with Westat, the CV team helped to prepare the interviewer training materials for the training sessions, which took place in a hotel meeting room located near each site. When the interviewing for each pilot was concluded, as many interviewers and supervisors as possible were brought together by Westat for a debriefing session. The debriefings were designed to discover any problems the interviewers had noticed with the instrument's wording, question sequence, and visual aids. Additionally, any problems the interviewer encountered with other aspects of the field work, such as gaining access to homes and respondents or using the sampling and screening materials, were also discussed. Interviewers were encouraged to mention every problem they encountered, no matter how small. Particular attention was paid to any interviewer comments that suggested that respondents tended to misunderstand some aspect of

ACE 10917146

the questionnaire or that respondents were not giving meaningful and sincere answers to the valuation questions.

In addition to a quantitative data set based on respondent answers to the close-ended questions, each pilot produced two types of qualitative information: (1) the interviewer and supervisor comments described above; and (2) the comments made by respondents during the course of the interview. The latter comments, rendered either spontaneously or in response to open-ended questions in the questionnaire, were recorded verbatim by the interviewers on the questionnaire. All verbatims were transcribed so they could be analyzed by respondent or by question for a given pilot. Both the quantitative data and qualitative information were used to evaluate the instrument's success in addressing potential problem areas and to discover what aspects of the questionnaire deserved further attention. Following each pilot survey, the questionnaire was revised for use in the following survey.

Although the questionnaire wording was revised many times during the pilot phase of the study, the basic structure of the instrument used in the first pilot survey proved to work well and was used in all subsequent versions. This structure included an initial sequence of sections that described Prince William Sound, the effects of the Exxon Valdez oil spill, and the escort ship program to prevent a future oil spill. These sections were followed by the willingness-to-pay questions which were in turn followed by open-ended questions that probed for the assumptions the respondents had in mind when answering the WTP questions. Toward the end of the questionnaire, respondents were given the opportunity to change their answers to the WTP questions. Throughout the scenario, maps, diagrams, and color photographs were used to help convey information about the area, the spill, and its effects on natural resources.

In each pilot, four sets of discrete dollar amount design points were randomly assigned to equivalent subsamples for use in the initial and follow-up take-it-or-leave-it WTP questions.

ACE 10917147

Also, throughout the development of the survey instrument we sought to develop questions to measure respondent attitudes and characteristics that would help us understand and predict the willingness-to-pay responses. For the most part, these conceptual variables were suggested by theory.

§ 2.7 Pilot I — San Jose, CA

This pilot was the first formal test of the questionnaire under field conditions similar to those that would be used in the final survey. San Jose was chosen because it offered the opportunity to interview people in relatively high education and income areas, one of several diverse demographic groups on whom we wished to test the questionnaire and the group most likely to be able to understand the questionnaire even in its early stage of development. This pilot used a higher-prices-for-oil-products payment vehicle to pay for the escort ship plan. Respondents were told that if they voted for the plan, it would cost their households a specified amount in higher prices for oil products each year for the next ten years.

The overall judgment of the interviewers, as expressed during the day-long debriefing we conducted after this pilot (and each of the other pilots), was that the instrument worked fairly well despite the unusually large amount of text to be read compared with other surveys with which they were familiar. The interviewers said the visual aids engaged the respondents' interest and helped communicate the material in the text. In many places they recommended wording changes to make the instrument simpler, and in some places they recommended that the wording be made clearer for the respondents. The interviewers reported that some respondents had difficulty understanding the concept of a second spill. Some interviewers also thought that some respondents did not clearly understand that they would have to pay to prevent the spill each year for the period of ten years and that some respondents may have been confused about exactly

ACE 10917148

what they were being asked to value. As expected, a number of Pilot I respondents reacted negatively to the payment vehicle because they believed it was not their responsibility to pay higher oil prices for this purpose, but that this should be the responsibility of Exxon or "the oil companies."

§ 2.8 Multiple Year Payments

Most comments made by the interviewers at the Pilot I debriefing could be handled in the course of ordinary questionnaire revision without much difficulty. One of the comments, however, was more troublesome: some respondents had not believed that they would have to pay the specified amount every year for ten years, despite language to that effect in the survey instrument.

Our concern about this matter was heightened by a paper by Kahneman and Knetsch (1992) which was then circulating in draft form. That paper argued that people would give the same (yearly) amount irrespective of the number of years they were asked to pay. Kahneman and Knetsch reported a survey question involving toxic waste in British Columbia where respondents appeared to exhibit this behavior. To better understand this phenomenon, we conducted a seventh focus group and a telephone survey.

In the New Orleans focus group in March 1990, we explored how the participants thought about multiple year payments for common consumer durables like refrigerators, automobiles, and houses and for public goods like water treatment facilities. Many participants in the focus group, who were for the most part from the lower and lower-middle income classes, did not accept the commitment entailed by multi-year payments. Some had no actual experience with buying goods on credit or, with the exception of automobiles or houses, had experience with only short financing periods ranging from a few months to three years. Payments for new

ACE 10917149

automobiles or houses tended to be treated as payments for automobile and house services rather than as purchases. Houses, in particular, were considered something that could be sold if mortgage payments could not be met. These findings suggested that the focus group participants did not truly believe they were making long-term commitments when, for example, they were asked to state how much they would pay each year for 10 years.

As to large local public goods, participants believed that governments could, and often would, alter their spending priorities. This belief led the participants to discount the possibility that they had, in fact, committed to make annual payments for a lengthy period (five years or more) of time. Participants also thought that local governments did and should pay for the purchased public goods at the time of purchase.

Thus, the discussion of public goods tended to reinforce our conclusion from the private goods discussion: some people had difficulty accepting long term payment obligations. Some individuals might not feel compelled to pay the annual amount asked for each of the ten years because they felt that they could recontract at some later point if they no longer wanted to continue to receive or pay for the good. We concluded that individuals were committed to making at least the initial payment and generally to paying for two or three additional years, but that any longer payment schedule suffers from the recontracting problem.

Almost simultaneously with the New Orleans focus group, we used a telephone survey in Columbus, Ohio, to explore the issue of a one-time, lump-sum payment versus an annual payment over an extended period of time (twenty years in this telephone survey). The major problem we saw in conducting such a test was finding a good for which making annual payments did not imply an increased likelihood that the good would actually be provided in future years. One good which has this property is a scrubber in a power plant. A scrubber, once installed, would not normally be removed until the end of its useful life, and yet it requires only small

ACE 10917150

annual payments to maintain it in operation. An additional advantage of scrubbers is that they received a fair amount of attention during the acid rain debate, particularly in the Ohio Valley, and, therefore, could be readily described in a telephone survey.[27]

We surveyed 500 people, who were randomly assigned to either the annual 20 year payment vehicle or the lump-sum payment vehicle. We used a double-bounded dichotomous-choice elicitation framework similar to the one in these pilot studies. Fitting a Weibull distribution to this data and including a dummy variable for the payment vehicle treatment, we find the payment vehicle is a significant predictor of willingness to pay (t=2.81).[28] The lump-sum median willingness to pay is almost twice the annual median willingness to pay.

This finding contradicts Kahneman and Knetsch's (1992) finding that people are not sensitive to the number of years they are asked to pay for a public good.[29] However, the difference between the lump-sum payment and 20 years of annual payments appropriately discounted should have been much larger if respondents actually discounted at the 10 percent rate mandated by the Office of Management and Budget (OMB). The difference we found is consistent with discounting at higher discount rates (*e.g.*, Hausman, 1979) or with strong borrowing constraints (*e.g.*, Lawrance, 1991).

There is no obvious *a priori* basis on which to choose between the lump-sum and the annual payment schemes. On the basis of the telephone survey and the results from the New

---

[27]In order to keep the survey simple, we provided respondents with a list of different types of effects of acid rain, but did not go into the actual magnitude of those effects. As a result, what was valued in this survey was the respondents' perceptions of those effects, not the actual effects.

[28]A test based on a non-parametric approach also strongly rejects the hypothesis of no treatment effect.

[29]Kahneman and Knetsch's finding is likely to be an artifact of the good they had their respondents value which was "a toxic waste treatment facility that would safely take care of all chemical and other toxic wastes in British Columbia." The specification of this good is much vaguer than is the norm in contingent valuation studies, and it does not specify the time period during which the plant would provide its services.

ACE 10917151

Orleans focus group, we chose the lump-sum payment. Individuals were committed to making at least the initial payment and generally to paying for two or three additional years, but that any payment schedule longer than that suffers from the recontracting problem. The lump sum payment avoids the recontracting problem. This payment scheme also has the advantage of eliminating the need to determine what rate ought to be applied to discount future payments. However, it has the disadvantage of forcing a much tighter budget constraint on respondents by not allowing them to pay for the spill prevention plan over the course of several years. Hence, estimates using a lump sum payment scheme are likely to be smaller than those under a payment scheme which allows for smaller payments over more years.

§ 2.9 Pilot II — Toledo and Dayton, OH

The site for this pilot was chosen to represent middle America, both geographically and socio-economically. The sample was chosen from selected census tracts in Toledo and Dayton, Ohio. The instrument used in this survey was substantially revised on the basis of our experience in Pilot I.

Having resolved the one time versus multi-year payment issue, the next key design issue involved the choice of a payment vehicle. While there are a large number of potential vehicles, those that respondents will perceive as a plausible way to pay for a particular good are few. The payment vehicle in a contingent valuation scenario must be viewed as appropriate for the good being valued and not subject to waste and fraud. Payment vehicles which diverge from this ideal will generally result in lower stated willingness-to-pay amounts or higher refusal rates.[30]

[30]There are two types of payment vehicles which may actually raise a respondent's stated willingness to pay above their actual willingness to pay for the good. The first is a charitable contribution which may raise willingness to pay amounts because the contribution to the charitable organization is valued in and of itself. (There may be those who get positive utility simply from the act of paying higher taxes but surely such people are small in number.) Stated willingness to pay may also be higher than actual willingness to pay if a payment vehicle is implausible in the sense that the

ACE 10917152

Preliminary research indicated that two vehicles showed sufficient promise to investigate further. One was income taxes and the other was oil prices. Pilot II included a split-sample test to help us make a choice between these two alternatives. One sub-sample of 95 people received the tax payment vehicle, described as a one-time tax on oil company profits and a one-time federal income tax surcharge "on households like yours" to be paid during the first year of the plan. The oil prices payment vehicle was administered to the other sub-sample of 100 people. In this version, there would be a special one-time surcharge on the oil the oil companies take out of Alaska. Respondents were told the surcharge will reduce oil company profits for one year and also "increase the prices consumers like you pay for products that use oil."

The interviewer debriefing, which took place at the end of the field period, indicated that in general the interviewers felt the Pilot II instrument read more smoothly and presented fewer difficulties in administration than the Pilot I version.[31] This perception was confirmed by our analysis of the verbatims, which did not indicate undue respondent confusion. The number of protest responses was reduced from the previous pilot, most likely because various wording changes, including the explicit mention that the oil companies would pay for part of the cost of the escort ship plan (in both payment vehicles) increased the acceptability of the scenario to some people. However, some respondents still felt that the oil companies, and only the oil companies, should pay the cost of preventing future oil spills.

In the split-sample experiment testing the differences between using the household tax and oil prices payment vehicle, there was a statistically significant difference: in this sample,

---

government is unlikely to actually use it for the purpose of providing the good. In this instance, an implausible payment vehicle signals that the amount stated is unlikely to ever be collected but that the amount stated may influence the provision of the good.

[31]This observation is based especially on the reports of those interviewers who took part in both pilot surveys. These interviewers were used by Westat as travellers to augment the locally available interviewers.

ACE 10917153

willingness to pay was substantially higher in the oil price vehicle compared with the tax version. We deferred the decision about which payment vehicle to use in order to get more data from a different sample.

§ 2.10 Pilot III — Georgia

The interviews for the third pilot were conducted in five rural counties in Georgia: Colquitt, Worth, Liberty, Glynn, and Long. This area was selected in the expectation that its lower socioeconomic status, rural nature, and physical distance from Alaska would help us assess whether improvements would be needed to communicate the scenario to this type of respondent. The Georgia sample had much lower educational and income levels than the Ohio sample.

According to interviewer comments during the debriefing, the respondents' ability to comprehend the scenario was good overall, despite their lower educational attainment. The interviewers did recommend several wording changes to simplify the language and clarify that Alaska is one of the 50 states. They also pointed out that some of the respondents in this sample did not have enough income to pay federal income taxes. This disclosure caused us to modify the next version of the questionnaire so we could identify such respondents.

In this pilot, we conducted another split-sample experiment to compare tax and price payment vehicles, using a sample that was substantially different from that of Pilot II. The experiment was identical in design to that conducted in Pilot II except that the oil price payment vehicle was worded somewhat differently. In the Georgia pilot, respondents who received the oil price vehicle were told that: "These price increases will be in addition to any other change in the price of oil related products that may occur during that year." This modification addressed a confusion in the minds of some Pilot II respondents between the price increase to

ACE 10917154

pay for the plan and the fluctuations in oil and gas prices that occur as a result of market forces over the course of the average year. One hundred twenty-five respondents received the tax vehicle, and 119 respondents received the oil price vehicle.

The payment vehicle split-sample experiment showed no significant difference between the WTP distributions of the two versions (t=-0.52); and, therefore, failed to replicate the result of the first payment vehicle experiment in Pilot II. Thus, the two versions, each using a different "reasonable" payment vehicle, produced similar WTP estimates. Analysis of the respondent comments in the verbatims also showed similar amounts of respondent protest to each payment vehicle.

After a consideration of all the information available from these pilots and our other instrument development research, we decided to use the tax vehicle in the final survey for two reasons. First, the price of gasoline, the major type of oil product through which consumers would pay for the plan if we used the oil prices vehicle, had become quite unstable due to Iraq's invasion of Kuwait. It appeared likely that gasoline prices could increase rapidly in the near future when the final survey would be in the field or, perhaps, decrease if the crisis was resolved peacefully. This instability raised the prospect that if we used the oil prices vehicle, the respondents' WTP amounts might be distorted because of factors unrelated to any economic value they held for preventing future damage to Prince William Sound. Second, the two split-sample experiments showed that, if anything, the tax vehicle tended to elicit the same (Pilot III) or lower (Pilot II) amounts than those elicited by the oil prices vehicle.

We conducted a second split-sample experiment in Pilot III by randomly assigning respondents to versions of the questionnaire that included or excluded one item listed in each of the two questions A-1 and A-3. These items asked respondents whether they should spend more, the same, or less money on "protecting the environment" and how important "protecting

ACE 10917155

coastal areas from oil spills" was to the respondent (A-3f). The issue was whether including these items in lists that otherwise involved non-environmental (A-1) or non-oil related (A-3) items would bias subsequent responses in such a way as to be non-conservative. A *t*-test between the two versions of the survey instrument suggests that the inclusion of A-1e and A-3f had no significant effect (t=-0.10) on the WTP responses, and they were retained in subsequent versions of the instrument.

In this pilot, as in the others, we asked respondents to say who they thought sponsored the study. Although most respondents were willing to answer the question, few seemed to have arrived at a clear opinion. People would often say, "maybe X, maybe Y"; still others would give an answer and then confess that, in fact, they did not have an idea one way or the other. Many people mentioned Exxon or oil companies, many mentioned some governmental agency, and a few mentioned environmental groups. No one potential sponsor was mentioned more consistently than the others. The responses to the follow-up question, which asked respondents to give the basis for naming a sponsor, mostly referred to the topic of the survey or to the idea that it made sense for the sponsor named to have an interest in a study on this subject. Very few respondents made comments that suggested they found the wording biased in one direction or another.

A number of the questions in Section B of the questionnaire were designed to check whether the assumptions the respondents actually had in mind when they answered the valuation questions were the same as the assumptions on which the scenario was based. Although these questions were sometimes difficult to communicate to respondents, the evidence from this pilot showed that we had satisfactorily resolved these difficulties with respect to all but one of these questions. The question still requiring further work was "how many large spills like the Exxon Valdez spill" the respondent thought would occur in Prince William Sound without the escort

ACE 10917156

ship program. (The scenario had explicitly informed respondents that in the next ten years there would be one such spill without the escort ship plan.) According to the Pilot III interviewers, some respondents seemed to take the "how many large spills" question as an invitation to engage in speculation about how many spills might occur rather than to report what they had actually assumed about this when they answered the WTP questions earlier in the interview.

§ 2.11 Pilot IV — Toledo and Dayton, OH

The version of the questionnaire used in the fourth and final pilot survey incorporated revised visual aids to address a few problems which we identified in the previous pilots. The main problem involved the map used to show the extent of the spill over time. Some respondents had misinterpreted the shading on the map as indicating that the entire shaded area was covered by oil at a given point in time. Pilot IV also had a number of minor wording changes intended to make the interview more understandable to less-educated respondents and to dissuade respondents from thinking that any other part of the United States would be protected by the Prince William Sound protection plan. Wording changes were made in several of the predictor questions and Section B follow-up questions to improve comprehension. The "how many spills" question in Section B was substantially revised.

We conducted this pilot in Toledo/Dayton where we had previously conducted Pilot II for three reasons. First, comparing Pilot IV with Pilot II would give us an idea about how stable the WTP estimates were across time and help establish whether the estimates could be replicated. Second, it was convenient to interview in this area because the sample listings and trained interviewers were available from Pilot Study II. Third, it would be helpful in assessing the progress the survey instrument had made by using the interviewers from Pilot Study II.

ACE 10917157

The interviewers were very positive in the Pilot IV debriefing about most of the wording changes and about the interview as a whole. Several interviewers mentioned that the survey was now easier to administer because its progression and central purpose were clearer. They also believed that the revised visual aids better conveyed information about spill damage and that the visual aids in general engaged the respondents' interest in the survey. Some interviewers did say that it was difficult to keep their place in the text when they pointed to the visual aids, and some said that they had trouble maintaining eye contact with the respondents because of this. Comments like these helped us design the interviewer training program we used for the main survey.

The number of spills question still presented some problems as some respondents perceived the possibility of small spills in addition to the big one or the possibility of a spill that would not damage the environment very much because it would largely be contained. As a consequence, in the main survey, we decided to ask respondents directly about the amount of damage they expected to occur in the next ten years without the escort ship program. This more straightforward approach, which was pretested prior to inclusion in the main survey, allowed us to determine the effect of any respondent misperceptions in our statistical analysis.

CHAPTER 3 — STRUCTURE OF FINAL QUESTIONNAIRE

§ 3.1 Introduction

In this chapter we discuss the format and wording of the final questionnaire developed as described in the previous chapter and used in the national survey. The survey instrument will be described section by section. All quoted text in this chapter is from the questionnaire unless otherwise indicated. Any questionnaire text in capital letters is an interviewer instruction and is not read to the respondent. The complete survey instrument, including the show cards and reproductions of the photobook exhibits, is provided in Appendix A.

§ 3.2 Section A — Initial Questions

The first part of the survey instrument consists of preliminary questions, most of which were answered by the respondent before being told that the interview was about the Exxon Valdez oil spill. Interviewers were given strict instructions to limit the information they provided to prospective respondents about the subject matter of the survey to saying: "We are talking to people about their opinions on various issues." If the prospective respondent asked for more information about the topic, the interviewer was instructed to say the following, word for word:

> We are conducting interviews for a study of people's views about some current issues, such as crime, education, highway safety, the environment and energy.[32]

[32] "National Opinion Survey: Main Study — Trainer's Manual," Westat, Inc., January (1991). This typeface will identify lengthy direct quotations from the language of the questionnaire or interviewers manual.

ACE 10917159

If the respondent insisted on knowing more, the interviewer was instructed to say:

> The reason I can't tell you more about the topic of this interview before we begin is because I'd like you to form an opinion about it as you see the materials I have to show you.

The respondent was not given any information that would reveal that the topic of the survey concerned oil spills until question A-5. The Exxon Valdez oil spill was not mentioned until question A-6. Withholding this information made it possible to ascertain respondent concern about a list of social problems and awareness of the Exxon Valdez spill before the spill was revealed as the main topic.

The first set of questions asked how much more or how much less money should be spent on solving six social problems.

> A-1. We are faced with many problems in this country, none of which can be solved easily or inexpensively. I am going to name some of these problems, and for each one I'd like you to tell me whether you think we should spend more, the same, or less money than we are spending now. Here is a card that lists the answer categories.
>
> SHOW CARD 1[33]
>
> First, (READ ITEM) . . . do you think we should spend a great deal more money than we are spending now, somewhat more money, the same amount of money, somewhat less money, or a great deal less money on (ITEM)?

The A-1 series of problems (and the A-3 series described below) was intended to encourage the respondent to think about a broad range of current policy issues. Four of the problems are not environmentally related. Two of those, "fighting crime" and "improving public education," are often identified in surveys as subjects of great concern to the public; and a third, "making highways safer," was chosen as a problem with a level of concern likely to lie below that of "fighting crime" and "improving public education." "Giving aid to poor countries" is known

[33]This card lists five answer categories from "great deal more money" to "great deal less money". See Appendix A.

ACE 10917160

to lie at the lower end of public concern. The fifth item, "making sure we have enough energy for homes, cars and businesses," measures concern about energy supply. The last, "protecting the environment," is a general measure of environmental concern. Following standard practice to minimize order effects, the order in which the items were read was rotated according to a predetermined plan.

The next question was the first of a series designed to measure the respondent's awareness of the Exxon Valdez oil spill. This question sought to determine whether respondents spontaneously identified the Valdez spill when asked to identify "major environmental accidents" that caused the "worst harm to the environment" anywhere in the world and "harmed nature the most."

> A-2. Now, I'd like you to think about major environmental accidents caused by humans. Please think about those accidents anywhere in the world that caused the worst harm to the environment. (PAUSE) During your lifetime, which accidents come to mind as having damaged nature the most? (RECORD VERBATIM. PROBE FOR SPECIFIC DETAIL INCLUDING LOCATION.)

This question is the first of a number of questions in this survey instrument that used an open-ended answer format. The interviewers who conducted this study were familiar with verbatim recording as a result of their general training as Westat interviewers. Their instructions were to record on the questionnaire the respondent's comments as closely as possible, asking the respondent to pause, if necessary, so a comment could be completely transcribed. The importance of the verbatims for this study was emphasized in the training and in the Interviewer's Manual (IM); and the interviewers practiced recording verbatims in the training process. For recording the verbatims, as for recording the responses to all questions, the interviewers were instructed to use a ball point pen.

A standard survey practice in asking open-ended questions is to use follow-up probing questions. The interviewers were trained to use specific probes where necessary to clarify the

ACE 10917161

comment (*e.g.*, "What do you mean exactly?" or "Could you please explain that a little? I don't think I quite understand?"), to understand better the specific reference (*e.g.*, "Could you be more specific about that?"), or to better understand its relevance ("I see, Well let me ask you again" followed by the exact question). Another type of permitted probe was used to determine whether the respondent's comment was complete (*e.g.*, "What else?" "What other reasons/things/examples etc.?"). Interviewers were instructed to write "(x)" after every probe to separate the preceding verbatim from the new verbatim elicited by the probe.

In addition to the standard probes, interviewers were sometimes instructed in the Interviewer's Manual to use specific probes for certain questions. In the discussion of the instrument that follows, all instructions of this type will be identified. A-2 is the first question with a special probe. Here the interviewers were instructed to use two types of probes. The first sought completeness:

> ...if the respondent mentions only one major accident, probe by saying, "Can you think of any others?[34]

The second sought specificity:

> IF THE OIL SPILL(S) ARE MENTIONED WITHOUT LOCATION; ASK: Where did (this/these) spill(s) happen?

The next question, A-3, asked respondents to give their opinion about six more social policies. This time they were asked:

> A-3. How <u>important</u> to you <u>personally</u> are each of the following goals?
>
> SHOW CARD 2[35]

As with question A-1, four items were not environmentally related programs. Three of the programs — "expanding drug treatment programs," "providing housing for the homeless,"

---

[34]"National Opinion Survey: Main Study — Interviewer's Manual", Westat, Inc., January (1991), section 4, p. 4-17.

[35]This card lists five answer categories from "extremely important" to "not important at all". See Appendix A.

ACE 10917162

and "reducing taxes" — are widely supported programs, whereas "putting a space station in orbit around the earth" is not. One of the two environmental programs, "reducing air pollution in cities" had nothing to do with oil spills; and the other, "protecting coastal areas from oil spills," is directly related to the survey's subject matter. The oil spill question was expected to be a good predictor of willingness to pay for an oil spill prevention program.[36]

Question A-4 measures people's views about another environmental policy related to the spill area.

SHOW CARD 3[37]

A-4. Over the past twenty years the government has set aside a large amount of public land as wilderness. By law, no development of <u>any</u> kind, including roads and cutting down trees for lumber, is allowed on this land. In the <u>next</u> few years how much <u>more</u> land do you think should be protected in this way -- a very large amount, a large amount, a moderate amount, a small amount, or none?

At this point in the survey a series of questions was asked of those respondents who did not mention the Exxon Valdez oil spill in A-2 to determine whether they had heard of the spill before the interview. The first question, A-5, is open-ended.

A-5. Have you heard or read about large oil spills in any part of the world (other than those you mentioned earlier)?

A-5A. Which spill or spills are these?
(PROBE: Where did it happen?) (LIST NAME OR LOCATION OF SPILLS BELOW)

If the Exxon Valdez oil spill (referred to in the text of the questionnaire as the "Alaskan oil spill" to neutralize any tendencies the respondents might have had to criticize Exxon for causing the spill) was specifically mentioned by the respondent in the verbatim, the interviewer

[36]This proved to be the case. (See Section 5.9.2).

[37]This card lists five answer categories from "very large amount" to "none". See Appendix A.

ACE 10917163

immediately skipped forward to A-6A. Those who did not specifically mention the spill in A-2 or A-5 were asked A-6:

> A-6. A spill occurred in March of 1989 when the Exxon Valdez oil tanker ran aground on a reef in Prince William Sound, Alaska. Part of its cargo, 11 million gallons of crude oil, spilled into the water. Do you remember hearing anything about this spill?

The respondents who had mentioned the spill were given the same information:

> Earlier you mentioned the Alaska oil spill. This spill occurred in March of 1989 when the Exxon Valdez oil tanker ran aground on a reef in Prince William Sound. Part of its cargo, 11 million gallons of crude oil, spilled into the water.[38]

All respondents, except those who said that they had not heard or were not sure they had heard about the Exxon Valdez oil spill, were then asked an open-ended question to determine what assumptions they had about the most serious consequences of the spill for the natural environment in the Prince William Sound area.

> A-6A. What was it about the natural environment around Prince William Sound that you feel was most seriously damaged by the oil spill? (PROBE: Anything else?) (RECORD VERBATIM.)

§ 3.3 Section A — Description of Scenario

The information presented to the respondents in A-6 begins the scenario description in the questionnaire. The scenario presented the elements of the constructed market in which the respondent would later be asked to vote in favor of or against a plan costing the respondent a specific amount. The remaining portion of the scenario conveys information about Prince William Sound, the transport of oil by ship from Valdez, the Exxon Valdez spill and its effects, and the escort ship program to prevent damage from another spill that would have the same effect on the environment as the Valdez spill.

---

[38]See questionnaire, boxes 1 and 2, pp. 4 and 5.

ACE 10917164

At various places during the presentation of this portion of the scenario, the interviewers showed the respondents one of nineteen visual aids — maps, color photographs, and show cards (listed in Table 3.1). These materials were designed and pretested to help the respondents visualize important aspects of the scenario and to understand the material that was being read to them. The maps and photographs were contained in a spiral bound book with plastic coated pages (to protect them from the elements) measuring 10.5 inches by 12.5 inches. The cards were printed on light cardboard stock and were 8.5 inches by 11 inches in size. They were also spiral bound for ease of use by the interviewers.

The interviewer training for this study emphasized helping the interviewers read the narrative material in a way that would maintain respondent interest and enhance comprehension of the material. The interviewer manual summarized this emphasis:

> This questionnaire is different from most questionnaires you have administered because during much of the interview you will read narrative material about the Alaskan oil spill and the escort ship program. The wording has been extensively pretested and should be presented as it appears in the questionnaire; that is, the material is to be read word-for-word. You should not add any explanations of your own at any point in the interview.
>
> Although there is a great deal of material to read, our pretest and pilot study experience shows that respondents' interest can be maintained throughout the interview. Two factors make this possible. First, the maps, photos, and show cards help a great deal as they add a visual dimension to what the respondent is being told. The second factor is the interviewers' mode of presentation. Respondents tire and are prone to distraction if the material is read to them in one or more of the following ways: a monotone voice, a "sing-song" voice, at too fast a pace, or by running one sentence and paragraph into another without natural pauses. Respondents find it much easier to listen to the material when it is presented in a conversational manner by someone with a pleasant, friendly tone, who uses normal inflections, good pacing and frequent eye contact. [39]

At this point, the scenario narrative introduced the purpose of the survey and provided background information about Alaska, its oil, the way it is transported, and the importance of

[39] "National Opinion Survey: Main Study — Interviewer's Manual", Westat, Inc., January (1991), section 1, pp. 3-4.

ACE 10917165

Table 3.1 Visual Aids Used in Survey

ORDER OF PRESENTATION	ITEM	DESCRIPTION
1	Show Card 1	Question A-1: List of Answer Categories 1-5
2	Show Card 2	Question A-3: List of Answer Categories 1-5
3	Show Card 3	Question A-4: List of Answer Categories 1-5
4	Photograph 1	Map 1 — State Of Alaska
5	Photograph 2	Map 2 — Prince William Sound
6	Photograph 3	Photograph A — Port Of Valdez And Valdez Narrows
7	Photograph 4	Photograph B — Columbia Glacier On Prince William Sound
8	Photograph 5	Photograph C — View Of Prince William Sound
9	Photograph 6	Photograph D — Nesting Gulls And Cormorants On Cliff
10	Photograph 7	Photograph E — Murres
11	Photograph 8	Photograph F — Sea Otter
12	Photograph 9	Photograph G — Tanker Sailing Through Prince William Sound
13	Photograph 10	Map 3 — The Alaska Oil Spill Area
14	Photograph 11	Map 4 — The Alaska Oil Spill: Prince William Sound — Direction Of Oil Flow
15	Photograph 12	Photograph H — Heavily Oiled Shore Soon After Spill
16	Photograph 13	Photograph I — Very Heavily Oiled Shore Before Cleanup
17	Photograph 14	Photograph J — Cleanup Operation On Prince William Sound Shore, Summer 1989
18	Show Card 4	Bird Species Affected By The 1989 Alaska Oil Spill
19	Show Card 5	Marine Mammals and the 1989 Alaska Oil Spill
20	Show Card 6	Containment and Oil Recovery System
21	Show Card 7	Number of Large Spills Expected to Cause Damage to the Alaska Spill Area in the Next Ten Years
22	Show Card 8	Likely Damage to This Part of Alaska in the Next Ten Years Without the Escort Ship Program
23	Show Card 9	Total Yearly Income For Your Household Before Taxes in 1990
24	Show Card 10	Question C-7: List of Answer Categories 1-4

ACE 10917166

this oil for the U.S. supply.

A-6B. I'd like to describe a plan to protect this part of Alaska from the effects of another large oil spill. First, I need to give you some background.

SHOW MAP 1[40]

Here is a map of the state of Alaska. (PAUSE)

In the upper right corner (POINT) is a smaller map showing Alaska on the rest of the United States. As you can see, Alaska is very large compared to the other states.

(As you may know,) in 1967 a large oil field was discovered in Prudhoe Bay on the North Slope of Alaska here (POINT).

In 1977, the Trans-Alaska Pipeline opened to take the crude oil from Prudhoe Bay (TRACE ROUTE ON MAP) down to Valdez, a port on Prince William Sound.

This area in blue is Prince William Sound (POINT).

In Valdez, the oil is piped onto tankers which sail down to ports in the lower part of the United States. There the oil is refined into various products including heating oil, gasoline, and fuel for electric power plants.

About one fourth of the oil produced in the U.S. comes from Alaska.

Here and elsewhere in the narrative, questions are asked to help involve the respondent in the interview and to obtain information useful to the study. Questions A-7 through A-10 probe whether the respondent or anyone else in the household has visited Alaska. The answers to the first questions in this sequence determined which questions were asked subsequently. Interviewers were given specific instructions in the instrument as to whether they should proceed with the next question or skip to a later question.[41]

A-7. Have you ever been to Alaska?

A-7A. Has anyone else living in your household ever been to Alaska?

---

[40]This map shows State of Alaska and the features as discussed in the narrative. See Appendix A.

[41]The many skip patterns used in this study can be examined by reviewing the final survey instrument in Appendix A.

ACE 10917167

A-8. How many times have you been there?

A-9. What year were you (last) there? **(RECORD YEAR OR APPROXIMATE YEAR.)**

A-10. Did you ever visit the Prince William Sound area?

The next part of the narrative described Prince William Sound.

A-10A. SHOW MAP 2[42]

This map shows Prince William Sound. (PAUSE) It is an enlargement of the area shown in blue on Map 1 (SHOW). The Sound is a body of salt water, a little over one hundred miles wide. As you can see, it has many islands and inlets, so its coastline is several hundred miles long (TRACE OUT PORTION OF COAST).

From Valdez (POINT) this is the route the tankers use to the Gulf of Alaska (TRACE ROUTE), a journey of 75 miles.

They leave Prince William Sound for the open sea here. (POINT AT PLACE WHERE THE TANKERS ENTER THE GULF OF ALASKA)

Photographs A - C show various features of the Sound including the Columbia Glacier.

SHOW PHOTO A

This photograph shows Valdez from the air. This is the town (POINT)

and across from the town is the terminal where the oil is piped onto tankers (POINT). These are some tankers (POINT).

The tankers go through the narrows here (POINT) into Prince William Sound. The Exxon Valdez tanker went aground on an underwater reef about here (POINT).

This whole area (POINT) is Prince William Sound.

SHOW PHOTO B

The next photo shows a view of part of the Sound.

As you can see, it is ringed with high mountains. In many areas there are glaciers that break up and produce small icebergs. This photo shows the Columbia Glacier which is more than 100 feet high (POINT TO GLACIER WALL). Icebergs from this glacier sometimes float into the shipping lanes.

---

[42]This map shows Prince William Sound. See Appendix A.

ACE 10917168

SHOW PHOTO C

As you can see in the next photo, the area is largely undeveloped.

Most of the land has been set aside as national forest and state parks. People use the area for fishing, boating, camping and other recreation. In the whole area there are only a few small towns. (PAUSE)

The description then turned to wildlife; the photographs showed respondents living examples of some of the wildlife that was killed by the spill. We did not use any photographs of specific animals that had been harmed or killed by the spill in this study.

This part of Alaska is also home to a great deal of wildlife.

A number of different types of birds, including sea ducks, bald eagles, grebes, and murres live in the area.

SHOW PHOTO D

The next photo shows sea gulls (POINT) and cormorants (POINT) at a nesting site on a cliff. (PAUSE)

SHOW PHOTO E

The next photo shows a group of murres. (PAUSE)

In addition to the birds, animals such as sea otters and seals live around the Sound.

SHOW PHOTO F

Here is a sea otter floating on the water. (PAUSE)

The next section of the scenario described the spill and its impact on the shoreline. After a photograph of a tanker in the sound, the narrative focused on the Exxon Valdez spill.

SHOW PHOTO G

The next photo shows a tanker sailing through the Sound. (PAUSE)

About two tankers a day or over 700 tankers a year make this journey. Many are supertankers which are as long as three football fields.

The supertanker Exxon Valdez was carrying slightly more than 53 million gallons of Alaskan crude oil when it ran aground on an underwater reef.

ACE 10917169

The 11 million gallons that spilled made it the largest oil tanker spill to occur in United States waters. Winds and tides spread the oil over a large part of Prince William Sound and part of the Alaskan coastline outside the Sound.

The following questions interrupted the narrative at this point to keep the respondent involved in the survey.

A-11. At the time this happened, would you say you followed radio, TV, newspaper or magazine reports about the spill ... [very closely, somewhat closely, not too closely, or not at all?]

A-12. Did you get most of your information about the spill from newspaper, from television or from both?

A-12A. (As you may remember from the coverage,) some of the spilled oil evaporated in the first few days after the spill, but much of it stayed in the water and ended up on shore.

Now I would like to tell you how the shore was affected. This map shows the overall extent of the spill.

At this point the interviewer presented another map which conveyed the farthest extent of the spill and the time it took to reach this far.

SHOW MAP 3 (PAUSE)

Here is where the spill occurred (POINT).

The currents floated the oil from Prince William Sound. The blue-green color shows the spill area where some oil spread. The farthest point it reached is here (POINT)

about 425 miles from where the- tanker ran aground.

Altogether, about 1,000 miles of shoreline inside and outside the Sound were affected in some way.

Specific attention was called to the fact that the impact of the oil on the shoreline varied and that the oiling was heaviest in Prince William Sound.

Because of the wind and currents, some shore was heavily oiled, some lightly oiled, and much was not affected at all. The oiling was heaviest in Prince William Sound.

Most of the affected shore outside Prince William Sound was only very lightly oiled. (POINT)

ACE 10917170

SHOW MAP 4

This map shows how the oil spread in Prince William Sound. (PAUSE) The red color shows where the shore was more heavily affected (POINT) and the purple where the effects were lighter. You can also see that many areas of shore were not affected by the spill (POINT).

SHOW PHOTO H

The next photo shows a heavily oiled shore soon after the spill. As you can see, the oil covered the rocks near the water (POINT).

SHOW PHOTO I

The next photo is a close-up view of a very heavily oiled shore in Prince William Sound before the cleanup. (PAUSE)

Attention was then called to the cleanup effort.

As you may know, Exxon made a large effort to clean up the oil on the beaches.

SHOW PHOTO J

The next photo shows some of the cleanup activity that took place in the summer after the spill. One of the cleanup techniques was to wash as much of the oil as possible off the shore into the water where it was scooped up by special equipment and taken away. It was not possible to remove all the oil from the rocky beaches in this way because some had already soaked into the ground and couldn't be washed out. Scientists believe that natural processes will remove almost all the remaining oil from the beaches within a few years after the spill. (PAUSE)

The next portion of the scenario described the effect of the spill on wildlife. Information was provided on Card 4 about the total bird population before the spill to provide a perspective on the number of bird deaths (as measured by the number of recovered bodies) that occurred as a result of the spill. For example, although 16,600 murres were found dead, the total population of murres was described as 350,000. The text called attention to the fact that large kills can occur naturally. The respondents were told that the numbers of dead birds shown on the cards are limited to those that were recovered and that the actual toll is estimated to be three to six times higher. Assurance that none of these species was threatened with extinction was included

ACE 10917171

in the instrument because focus groups showed that this aspect of the spill injuries was important to respondents.

> Now I would like to tell you how the spill affected wildlife in this part of Alaska.
>
> SHOW CARD 4[43]
>
> During the period of the spill there were about one and a half million seabirds and sea ducks of various species in the spill area inside and outside Prince William Sound. (POINT)
>
> As you can see from this card, 22,600 dead birds were found. (POINT)
>
> The actual number of birds killed by the oil was larger because not all the bodies were recovered. Scientists estimate that the total number of birds killed by the spill was between 75,000 and 150,000.
>
> About three-fourths of the dead birds found were murres, the black and white bird I showed you earlier. This is shown on the first line of the card. (POINT)
>
> Because an estimated 350,000-murres live in the spill area, this death toll, though high, does not threaten the species.
>
> One hundred of the area's approximately 5,000 bald eagles were also found dead from the oil.
>
> The spill did not threaten any of the Alaskan bird species, including the eagles, with extinction. (PAUSE)
>
> Bird populations occasionally suffer large losses from disease or other natural causes. Based on this experience, scientists expect the populations of all these Alaskan birds to recover within 3 to 5 years after the spill. (PAUSE)

The mammal deaths were described in a table on Card 5. As with birds, total populations were provided in addition to kill estimates. Three species for which no kills were reported were also listed on the card because in our pretests some respondents assumed there were also injuries to these mammalian species.

> SHOW CARD 5[44]

---

[43]This card lists the number of dead birds recovered and the estimated population before the spill for 12 named species and an "other" category. See Appendix A.

[44]This card lists the number of marine mammals estimated to be in Prince William Sound before the spill and the number estimated to be killed by the spill.

ACE 10917172

The only mammals killed by the spill were sea otters and harbor seals. This card shows information about what happened in Prince William Sound. According to scientific studies, about 580 otters and 100 seals in the Sound were killed by the spill. Scientists expect the population size of these two species will return to normal within a couple of years after the spill.

Many species of fish live in these waters. Because most of the oil floated on the surface of the water, the spill harmed few fish. Scientific studies indicate there will be no long-term harm to any of the fish populations.

Another question interrupted the narrative at this point to give respondents a chance to react to the material.

A-13. I've been telling you a lot about this part of Alaska and the effects of the oil spill. Did anything I said surprise you?

Those who said "yes," were asked:

A-13A. What surprised you? (RECORD VERBATIM.)

After recording the answer, the interviewers were instructed to probe: "Anything else?"

The next section of the scenario introduced the concept of a possible second spill like the first one and described how the escort ship plan would prevent such a spill if the plan were put into operation. It was important for eliciting household willingness to pay that the program be perceived as feasible, as effective, and as requiring the amount of money asked about. To avoid overburdening the respondents with information, only information that our pretesting showed to be essential to communicating a plausible choice situation was included in the narrative. The material on double-hulled tankers was included because during our pretests, some respondents were interested to know whether a switch to double-hulled tankers would accomplish the goal of stopping such a second spill and because the introduction of double-hulled tankers helped to sharply define the ten year period during which the escort ship would be in operation.

A-13B. In the little over ten years that the Alaska pipeline has operated, the Exxon Valdez spill has been the only oil spill in Prince William Sound that has harmed the environment.

ACE 10917173

Some precautions have already been taken to avoid another spill like this. These include checking tanker crews and officers to see if they have been drinking, keeping a supply of containment equipment in Valdez, putting trained cleanup crews on 24 hour alert, and improving the Coast Guard radar.

Congress has also recently required all new tankers to have two hulls instead of one. The Exxon Valdez, like most other tankers, had only a single hull. Double hulls provide more protection against oil leaking after an accident.

However, it will take ten years before all the single hulled tankers can be replaced. Scientists warn that during this ten year period another large spill can be expected to occur in Prince William Sound with the same effect on the beaches and the wildlife as the first spill.

In order to prevent damage to the area's natural environment from another spill, a special safety program has been proposed.

We are conducting this survey to find out whether this special program is worth anything to your household.

Here's how the program would work.

Two large Coast Guard ships specially designed for Alaskan waters will escort each tanker from Valdez all the way through Prince William Sound until they get to the open sea. These escort ships will do two things.

First, they will help prevent an accident in the Sound by making it very unlikely that a tanker will stray into dangerous waters. (PAUSE)

Second, if an accident does occur, the escort ships will carry the trained crew and special equipment necessary to keep even a very large spill from spreading beyond the tanker. (PAUSE)

This drawing shows how this would be done. (PAUSE)

SHOW CARD 6[45]

Escort ship crew would immediately place a boom that stands four feet above the water and five feet below the water, called a Norwegian sea fence, around the entire area of the spill. (POINT IF NECESSARY) Because oil floats on the water, in the first days of a spill, the sea fence will keep it from floating away. The oil trapped by the sea fence would be scooped up by skimmers, and pumped into storage tanks on the escort ships. Within hours, an emergency rescue tanker would come to the scene to aid in the oil recovery and transport the oil back to Valdez.

This system has been used successfully in the North Sea by the Norwegians.

[45]This card displayed a line drawing of an escort ship recovering oil at an oil spill.

ACE 10917174

The drawing on Card 6 proved to be extremely helpful in the pilot studies in communicating the way that the escort program would work. The following wording was used at this point to reinforce the concept of what the program would prevent and that it would be effective.

SHOW CARD 7[46]

This card summarizes what the program would prevent in the next ten years. Without the program (POINT) scientists expect that despite any other precautions there will be another large oil spill that will cause the same amount of damage to this part of Alaska as the last one. (PAUSE)

With the program they are virtually certain there will be no large oil spill that will cause damage to this area.

The next question gave the respondents a chance to say whether they would like to know anything more about the plan. It had an open-ended format.

A-14. Is there anything more you would like to know about how a spill could be contained in this way?

Respondents who said "yes" were asked:

A-14A. What is this? (PROBE: Anything else?) (LIST RESPONDENT QUESTIONS BELOW)

The questions asked by the respondents were recorded verbatim by the interviewers and provided useful information about respondent concerns. The interviewers were instructed to answer only those questions that could be answered by referring back to previous material in the narrative. Otherwise they were told to say they didn't know the answer. If a respondent wanted to know why the interviewer was recording questions but not providing answers, the interviewer was instructed to say:

[46]This card indicated that without the program there would be one spill; with the program no spills.

ACE 10917175

The researchers are interested in knowing whether there is more information about spill containment that needs to be given to the public. This is why I need to ask this question.[47]

The next portion of the narrative described the magnitude of the plan and reinforced its effectiveness while noting that it would not protect from spills outside Prince William Sound.

A-14B. Because two tankers usually sail from Valdez each day, the Coast Guard would have to maintain a fleet of escort ships, skimmers, and an emergency tanker, along with several hundred Coast Guard crew to run them.

Although the cost would be high, the escort ship program makes it virtually certain there would be no damage to Prince William Sound's environment from another large oil spill during the ten years it will take all the old tankers to be replaced by double-hulled tankers.

It is important to note that this program would not prevent damage from a spill anywhere else in the United States because the escort ships could only be used in Prince William Sound.

§ 3.4 Section A — Valuation Questions

At this point in the scenario, respondents were asked to state whether they were willing to pay specified amounts to prevent the damage from a future large oil spill in Prince William Sound. The narrative first informed respondents that the program would be funded by a one-time federal tax payment that would go into a Prince William Sound Protection Fund.

If the program was approved, here is how it would be paid for.

All the oil companies that take oil out of Alaska would pay a special one time tax which will reduce their profits. Households like yours would also pay a special one time charge that would be added to their federal taxes in the first year and only the first year of the program.

This money will go into a Prince William Sound Protection Fund. The one time tax will provide the Fund with enough money to pay for the equipment and ships and all the yearly costs of running the program for the next ten years until the double hulled tanker plan takes full effect. By law, no additional tax payment could be required.

[47]National Opinion Survey: Main Study — Interviewer's Manual, section 4, p. 4-47.

ACE 10917176

Respondents were then given the opportunity to state any questions they have about this method of payment.

> A-14C. Do you have any questions about how the program would be paid for?
>
> A-14C-1. What is this? (PROBE: "Anything else?") (LIST RESPONDENT QUESTIONS BELOW.)

Our pretests had showed that some respondents criticized the notion that citizens should share in paying the cost of the plan. Because this could lead respondents to reject the premise of the scenario — that they should make a judgment about what the plan is worth to them — we included a special instruction in the instrument requesting the interviewer to check a box if the respondent expressed the view that Exxon or the oil companies should pay. The interviewers were instructed to say the following to those who expressed this concern in an attempt to persuade them that the oil companies <u>would</u> pay a share:

> If the program is approved, the oil companies that bring oil through the Alaska pipeline (including Exxon) <u>will</u> have to pay part of the cost by a special tax on their corporate profits.

The next portion of the narrative presented information intended to reassure respondents who might not be willing to pay for the program that a "no" vote is socially acceptable. The reasons presented here for voting against the program were given by respondents during the pretest research for this study.

> A-14E. Because everyone would bear <u>part</u> of the cost, we are using this survey to ask people how they would vote if they had the chance to vote on the program.
>
> We have found some people would vote <u>for</u> the program and others would vote <u>against</u> it. Both have good reasons for why they would vote that way.
>
> Those who vote <u>for</u> say it is worth money to them to prevent the damage from another large spill in Prince William Sound.
>
> Those who vote <u>against</u> mention concerns like the following.

ACE 10917177

Some mention that it won't protect any other part of the country except the area around Prince William Sound.

Some say that if they pay for this program they would have less money to use for other things that are more important to them.

And some say the money they would have to pay for the program is more than they can afford.

Question A-15 used a discrete-choice elicitation format in the context of a referendum model to ask whether the respondent would vote for the program if it cost a specified amount that would be paid by a one-time federal tax payment. In order to obtain responses to a range of amounts, four different versions (A through D) of the instrument were administered by the interviewers to equivalent subsamples. Each version used a different set of dollar amounts in questions A-15 to A-17, each set consisting of a single initial amount and two follow-up amounts. Every respondent who said they would vote for the program at the initial amount was asked whether they would also vote for the program if the cost to their household was a specified second amount higher than the initial amount. Those who said they would not vote for the program at the initial amount and those who were unsure were asked whether they would vote for the program if it cost a specified second amount lower than the initial amount.

A-15. Of course whether people would vote for or against the escort ship program depends on how much it will cost <u>their household</u>.

<u>At present</u>, government officials estimate the program will cost <u>your</u> household a total of $[specified amount here]. You would pay this in a special one time charge in addition to your regular federal taxes. This money would <u>only</u> be used for the program to prevent damage from another large oil spill in Prince William Sound. (PAUSE)

<u>If</u> the program cost your household a total of $(amount) would you vote for the program or against it?

The interviewers received special instructions about how to ask the willingness-to-pay questions and how to handle respondent queries in a neutral manner. The following material

ACE 10917178

comes from the part of the Interviewer's Manual for questions A-15 through A-20. Italics are in the original.

> An important goal of this survey is to find out how people really feel about the escort ship program and how much, if anything, they would be willing to pay for the program to protect the spill area from another oil spill. It is especially important, therefore, that these questions (A-15 through A-20) be asked in a *neutral tone* and that the respondents be given *as much time as he/she wants* to think about these questions. *Do not hurry the respondent in any way.*[48]

The Manual told the interviewers that some respondents may look to them at this point in the interview for cues as to how they should answer, perhaps because the respondent is fearful of appearing cheap or of appearing to be naive and a spendthrift to the interviewer. The Manual then declared:

> In fact, it doesn't matter at all whether people vote "for" or vote "against" the program; what does matter is that their answers represent their own best judgment about their actual willingness to pay based on the information provided to them in the interview and their preferences about how their household should spend its money. This is why you should use a neutral tone and an unhurried manner.[49]

Three responses were provided to the interviewers to use if they were asked certain types of questions at the point where the respondent was deciding how to respond to the willingness-to-pay question. The interviewers were also requested to record these questions and any other comments the respondent made while giving their answer to question A-15 in a space provided for this purpose on the instrument.[50]

> [Respondent] "Gee, I'm not sure, what do you think?"
> ANSWER: "We want to know what you think. Take as much time as you want to answer this question. (PAUSE) We find that some people say they would vote for, some against;
>
> "I'm not sure..." or any other expression of uncertainty.

---

[48]National Opinion Survey: Main Study — Interviewer's Manual, section 4, p. 4-55.

[49]National Opinion Survey: Main Study — Interviewer's Manual, section 4, p. 4-55.

[50]National Opinion Survey: Main Study — Interviewer's Manual, section 4, p. 4-59.

ACE 10917179

ANSWER: "Take as much time as you want to answer this question. (PAUSE) We find that some people say they would vote for, some against; which way would you vote if the program cost your household a total of $____?"

"I don't think the program would really cost this much."
ANSWER: "This is the amount it has been calculated it would cost your household. If further planning shows that it will cost less than this, the amount you would pay would be decreased because the money cannot be used for any other purpose."[51]

In the text of the instrument, interviewers were also instructed to say the following if the respondent expressed the view that Exxon or the oil companies should pay:

(As I said earlier) the oil companies that bring oil through the Alaska pipeline (including Exxon) will pay part of the cost by a special tax on their corporate profits.

A follow-up amount was presented to every respondent. If the respondent said she would vote for the program at the given price in A-15, she was then asked:

A-16. What if the final cost estimates showed that the program would cost your household a total of $(amount)? Would you vote for or against the program?

The amount in A-16 was a preset amount higher than the initial amount. Those who said they would not vote for the program in A-15 or were unsure about this were asked:

A-17. What if the final cost estimates showed that the program would cost your household a total of $(amount)? Would you vote for or against the program?

The preset amount presented to these respondents was lower than the initial amount they were asked in A-15. Table 3.2 displays the amounts used for questions A-15, A-16, and A-17 for each of the subsamples. Chosen on the basis of information obtained from the distribution of the public's willingness to pay for our contingent valuation scenario in the pilot studies, these dollar amounts provide reasonable efficiency in estimating the key statistics, such as the median, while providing some robustness with respect to observing a substantially different willingness-

[51]National Opinion Survey: Main Study — Interviewer's Manual, section 4, p. 4-57.

ACE 10917180

Table 3.2 Program Cost by Version and Question

Version	A-15	A-16	A-17
A	$10	$30	$5
B	$30	$60	$10
C	$60	$120	$30
D	$120	$250	$60

to-pay distribution in the main study.[52]

The remainder of Section A is devoted to follow-up questions designed to provide more information about the reasons for the answers the respondents gave to the valuation questions. Those who voted against the program in both A-15 and A-17 were asked:

> A-18. Did you vote against the program because you can't afford it, because it isn't worth that much money to you, or because of some other reason?
>
> CAN'T AFFORD IT 1
> ISN'T WORTH THAT MUCH . . . 2
> WILL ONLY PROTECT PRINCE WILLIAM SOUND AREA/ NOT ELSEWHERE 3
> OTHER REASON (SPECIFY) 4
>
> _____
>
> _____

The pre-coded answers were identified as common responses in our pretesting. The "only protect Prince William Sound area" answer category was not read to the respondent. Any reason other than those offered in categories 1-3 was recorded verbatim by the interviewer in the provided space. The answer "Exxon or oil companies should pay" was not included as an unread response so that the interviewers would record the complete statement made by the respondent on this matter.

[52]See Alberini and Carson (1990) for a discussion of these design issues.

ACE 10917181

Those who said they were not sure whether they would vote for the program at any of the offered amounts were asked the following open-ended question:

> A-19. Could you tell me why you aren't sure? (PROBE AND RECORD VERBATIM)

Those who said they would vote for the program at either of the offered amounts were asked what it was about the program that made them willing to pay for it.

> A-20. What was it about the program that made you willing to pay something for it? (RECORD VERBATIM)

After a space to record the answer to A-20, the following probe instruction appeared, also with a space in which to write comments verbatim.

> IF NECESSARY PROBE FOR SPECIFIC EFFECT. FOR EXAMPLE, IF R REFERS TO "THE ENVIRONMENT" SAY: How did you think the environment would be affected by the program?

This probe was included as a reminder to the interviewers to probe the respondent's answer to this important question. In the pilot surveys, respondents who expressed seemingly general answers such as to "help the environment" frequently had in mind the Prince William Sound environment that had just been described to them in detail by the interviewer.

§ 3.5 Section B — Perception of Damages and Plan

This section contains a number of questions to assess the beliefs respondents held about key parts of the scenario when they answered the willingness-to-pay questions. Although this type of assessment is difficult to make, as noted in Chapter 2, it can be very helpful in checking whether respondents understood the scenario and accepted its basic features.

The first question in this series, B-1, and its follow-ups, B-2 and B-3, asked about the amount of damage the respondent assumed would happen without the plan.

ACE 10917182

B-1. The first question is about what would happen if the escort ship program is not put into effect. (PAUSE)

SHOW CARD 8[53]

Earlier I told you that without the escort ship program, scientists expect that sometime in the next ten years there would be another large oil spill in Prince William Sound causing the same amount of damage as the Exxon Valdez spill. (PAUSE)

When you decided how to vote, how much damage did you think there would be in the next ten years without the program — about the same amount of damage as caused by the Valdez spill, or more damage, or less damage?

Depending on whether the respondent thought there would be more or less damage, she was asked B-2 or B-3.

B-2. Did you think the damage would be a little more, somewhat more, or a great deal more than that caused by the Exxon Valdez spill?

B-3. Did you think the damage would be a little less than the damage caused by the Exxon Valdez spill, a lot less, or did you think there would be no damage at all?

Everyone who answered "more" or "less" was asked the reasons in an open-ended question (B-4).

B-5, also with an open-ended follow-up, asked whether the respondent thought the plan would cover a greater geographic area than that described in the scenario.

B-5. Next, did you think the area around Prince William Sound would be the only place directly protected by the escort ships or did you think this particular program would also provide protection against a spill in another part of the U.S. at the same time?

B-6. How would it protect another part of the U.S. at the same time? (PROBE: What other parts would it protect?)

The perceived efficacy of the plan was another important dimension assessed.

B-7. If the escort ship program were put into operation, did you think it would be completely effective in preventing damage from another large oil spill?

---

[53]Card 8 contained the three answer categories as to the likely damage to this part of Alaska in the next ten years without the escort ship program such as "About the same damage as the Exxon Valdez spill."

ACE 10917183

Those who said "no" or "not sure" were asked:

> B-8. Did you think the program would reduce the damage from a large spill a great deal, a moderate amount, a little, or not at all?

The final two questions in this sequence assessed other types of beliefs.

> B-9. When you answered the question about how you would vote on the program did you think you would actually have to pay extra taxes for the program for <u>one</u> year or for <u>more</u> than one year?

> B-10. Before we began this interview, did you think the damage caused by the Exxon Valdez oil spill was more serious than I described to you, less serious, or about the same as I described?

§ 3.6 Section B — Respondent Household

The remainder of the questions in Section B measured attributes of the respondent or members of the household which might affect their preferences for protecting the Prince William Sound environment from the effects of another oil spill.

> B-11. How likely is it that someone living in your household will visit Alaska sometime in the future? Is it very likely, somewhat likely, somewhat unlikely, very unlikely, or no chance at all?

> B-12. Does anyone living in your household fish as a recreational activity?

> B-13. Is anyone living in your household a birdwatcher?

> B-14. Is anyone living in your household a backpacker?

> B-15. Have you or anyone else living in your household ever visited the Grand Canyon, Yosemite, or Yellowstone National Parks?

> B-16. Do you think of yourself as an environmentalist or not?

Respondents who indicated that they were environmentalists were asked:

> B-17. Do you think of yourself as an environmentalist very strongly, strongly, somewhat strongly, or not strongly at all?

The final question in this section was:

> B-18. Do you watch television programs about animals and birds in the wild very frequently, frequently, some of the time, rarely, or never?

ACE 10917184

§ 3.7 Section C — Demographic Questions

These questions supplement the demographic information obtained from answers to the household screener questionnaire which the interviewer administered to select the respondent. The first three demographic questions asked in this part of the survey measured age, education level, and number of children under 18 in the household.

Now, I have just a few questions about your background.

C-1. First, in what month and year were you born?

C-2. What is the last grade of formal education you have completed? No high school, some high school, high school graduate, some college, bachelor's degree, postgraduate (master's, law degree, doctorate, etc.)?

C-3. How many children or young people under 18 live in this household?

The last demographic question measured the respondent's household income. The interviewer used the standard device of having the respondent report his or her income category from categories listed on a card. Two follow-up questions were asked of low income people, defined as those with reported incomes of under $10,000, to determine if they paid income taxes.

C-4. This card shows amounts of yearly incomes. Which letter best describes the total income from all members of your household before taxes for the year 1990? Please include all sources such as wages, salaries, income from business, interest on savings accounts, social security or other retirement benefits, child support, public assistance, and so forth.

SHOW CARD 9[54]

If the respondent said "letter A" the following was asked:

C-5. Did (you/anyone in your household) have any taxes withheld from a paycheck or other earnings last year?

C-6. Did anyone living in this household file a Federal income tax form last year?

[54]This card listed 11 income categories, the highest of which was "$100,000 or more" and the lowest of which was "Under $10,000."

ACE 10917185

§ 3.8 Section C — Strength and Reassessment Questions

Respondents who had voted for one or more of the amounts asked about in the willingness-to-pay questions were asked C-7 to measure how strongly they favored the escort ship program:

> C-7. Now that we're at the end of the interview and you have had the chance to see the kinds of questions I wanted to ask you, I'd like to give you a chance to review your answers to the voting questions.
>
> You said you would vote for the escort ship program to protect Prince William Sound from another large oil spill during the next ten years if it cost your household a one time tax payment of $(highest amount the respondent agreed to).
>
> How strongly do you favor the program if it cost your household this much money? Would you say ...
>
> SHOW CARD 10[55]
>
> ... very strongly, strongly, not too strongly, or not at all strongly?

In addition to the four answer categories and "NOT SURE," the interviewers were also instructed to place respondents in a category "DOESN'T FAVOR THE PLAN" if their remarks indicated that this was the case. Those respondents who answered "not too strongly" or "not at all strongly" to C-7 were given the opportunity to change their vote to "against."

> C-8. All things considered, would you like to change your vote on the program if it cost your household $(amount stated in C-7) from a vote for the program to a vote against?

Those who said "yes" or indicated that they were not sure were asked:

> C-9. Why is that? (PROBE: "Anything else?")

The interviewers had received special instructions for this series of questions:

> When you are asking this question (C-7) and the remaining questions in Section C, it is important that you do not give the respondent the impression that you

[55]This card listed the four answer categories.

ACE 10917186

are challenging his/her answers. Therefore, read these questions in a matter of fact way using a neutral voice.[56]

Everyone who was originally willing to pay for the program and had not changed his vote was asked C-10.

> C-10. If it became necessary in future years would you be willing to pay any more money beyond the one time payment to keep the escort ship program in operation?

All respondents, whether or not they were willing to pay anything for the program, were asked an open-ended question:

> C-11. Who do you think employed my company to do this study? (IF NECESSARY, PROBE: "What is your best guess?" "Could you be more specific?")

Respondents were also asked a follow-up question to understand why they thought this.

> C-12. What made you think that?

The last question in the interview was asked for information to use in verifying the interview at a later time.

> C-13. In case my supervisor wants to check my work, I need to ask you for your full name and telephone number.

§ 3.9 Section D — Interviewer Evaluation Questions

All the questions in this section were answered by the interviewers after they left the presence of the respondent. The interviewers were told "we want your frank opinion about these questions" (IM p. 4-91). The first four concerned various aspects of the respondent and his or her attitudes.

> D-1. How informed did the respondent seem to be about the Alaskan oil spill? [Answer categories: Very well informed, somewhat, not very well, not at all informed.]

[56]National Opinion Survey: Main Study — Interviewer's Manual, section 4, p. 4-83.

ACE 10917187

D-2. How interested did the respondent seem to be in the effects of the Alaskan oil spill? [Answer categories: Very interested, somewhat, not very, not interested at all.]

D-3. How cooperative/hospitable was the respondent at the beginning of the study? [Answer categories: Very cooperative/hospitable, somewhat cooperative/hospitable, not very cooperative/hospitable, not cooperative/hospitable at all.]

D-4. How cooperative/hospitable was the respondent at the end of the study? [Answer categories: Very cooperative/hospitable, somewhat cooperative/hospitable, not very cooperative/hospitable, not cooperative/hospitable at all.]

A series of three questions asked the interviewer to assess whether anyone besides the respondent and the interviewer were present during the interview and, if so, how much effect this had on the respondent's answers.

D-5. Not counting you and the respondent, was anyone else present during the interview?

D-6. Did any other person who was present while you administered the survey ask questions or offer answers during the interview?

D-7. How much effect on the respondent's answers do you think the other person(s) had?

The next question asked about the respondent's state of mind when the scenario narrative was presented:

D-8. What was the reaction of the respondent as you read through the material beginning with A6B and ending at A15?[57]

The interviewers rated each of the following three items as "extremely," "very," "somewhat," "slightly," or "not at all." They could also say whether they were not sure.

a. How distracted was the respondent?

b. How interested was the respondent?

c. How bored was the respondent?

[57]This is the descriptive material including the maps and photographs.

ACE 10917188

The next questions concerned only the voting questions.

The next items refer only to the questions about the respondent's vote on the escort ship program (A-15 – A-17).

D-9. Did the respondent have any difficulty understanding these vote questions?

D-10. Describe the difficulties [open-ended].

D-11. How serious was the consideration the respondent gave to the vote questions? Answer categories: Extremely serious, very serious, somewhat serious, slightly serious, not at all serious, not sure.

The last question invited the interviewers to make any other comments they wished to about the interview and the respondent:

D-12. Do you have any other comments about this interview?

In the pilot studies, interviewers varied greatly in the degree to which they took advantage of this opportunity. Some felt moved to say something about every interview, including their personal reactions to the respondent. Others wrote rarely or not at all.

ACE 10917189

CHAPTER 4 — SURVEY EXECUTION

§ 4.1 Introduction

The execution of this large national in-person survey had several distinct steps. A random sample of blocks was drawn in two stages, the individual dwelling units in those blocks were enumerated, and a random sample of the enumerated dwelling units was drawn.

With the sample drawn, attention shifted to the interviewing step. A detailed interviewer training manual was prepared, and Westat's professional interviewers were flown to a two-day training session to ensure the consistent administration of the survey instrument. While the survey was in the field, interviewers were supervised by three regional field supervisors. Interviews underwent quality control edits by those supervisors, as well as by the Westat home office staff.

After the interviews were completed, three characteristics of the interviewing process were examined: the effort required to complete the interviews, the distribution of interview lengths, and the completion rates in each block. This last characteristic is important in determining the sample weights used to make the completed sample representative of the population of U.S. households.

The final aspect of survey execution was the rendering of the data into a form suitable for analysis. Data sets containing the responses to both close-ended and open-ended questions were created.

§ 4.2 Sample Design

The survey was conducted using a multi-stage area probability sample of residential dwelling units drawn from the 50 United States and the District of Columbia. In the first stage

ACE 10917190

of selection, 61 counties or county groups were drawn. Within these selected counties, about 330 blocks (or block groups) were chosen. In the third stage, approximately 1,600 dwelling units were drawn from the selected blocks.

The 61 first-stage selections consisted of Westat's National Master Sample of 60 PSU's (primary sampling units) which were drawn from the continental United States and the Honolulu SMSA which was drawn from the states of Alaska and Hawaii.

Westat's Master Sample of 60 PSU's was selected from a list that grouped the 3,111 counties and independent cities in the continental United States in 1980 into 1,179 PSU's, each consisting of one or more adjacent counties.[58] Before the selection was made, the 1,179 PSU's were stratified by the following 1980 Decennial Census characteristics:

Region of the country;
SMSA versus non-SMSA;
Rate of population change between 1970 and 1980;
Percent living on a farm (for non-SMSA PSU's);
Percent employed in manufacturing;
Percent white;
Percent urban; and
Percent over age 65.

Selection from strata typically increases the precision of the survey results compared to unstratified selection.[59] The 60 PSU selections were then drawn with probabilities proportionate to their population counts.

Because Alaska and Hawaii were excluded from Westat's original sampling list, a new stratum was created consisting of those two states. A random selection of PSU's from this stratum yielded the Honolulu SMSA.

---

[58]The 1980 census was used as results from the 1990 census were not available at the time the sample was drawn.

[59]For a discussion of the comparative advantages of stratified selection, see Kish (1965) or Sudman (1976).

ACE 10

Within each of the 61 PSU's, the second-stage selections were drawn from a list of all the Census blocks in the PSU. The lists were stratified by two block characteristics: percent of the population that was black and a weighted average of the value of owner-occupied housing and the rent of renter-occupied housing. The 334 secondary selections were then drawn with probabilities proportionate to their total population counts.

§ 4.3 Field Enumeration

During 1990, trained field workers listed all the dwelling units (DU's) they found on these blocks (or block groups). (On blocks with a very large number of DU's, only a randomly chosen part of the block was listed.) A random selection from the listed DU's was then drawn, yielding 1,554 dwelling units.[60]

As a check for DU's missed by the listers (as well as to account for units constructed after the listing was conducted), interviewers followed a prescribed procedure at the beginning of the interviewing period to look for DU's that did not appear on the original listing sheets. This produced 45 additional DU's that were selected. Thus, the total sample consisted of 1,599 dwelling units.

§ 4.4 Interviewer Training

All of the professional interviewers Westat used on this study attended one of two two-day training sessions in January 1991. Both sessions were conducted by the study's Project Director, assisted by the Field Director and the three Regional Supervisors. To ensure

[60]Entry for listing purposes could not be obtained on three blocks: two on military bases and the third in a closed community. To adjust for the first two cases, Westat increased the number of housing units selected from the one other sampled block that was on a military base (to which entry was gained). No special measure was taken in the case of the block in the closed community; poststratification (described in a later section) served to adjust for this nonresponse.

ACE 10917192

comparability across sessions, they were run in accordance with a detailed script prepared in advance.[61] Interviewers had read an initial set of study materials before attending the training. The training sessions were a blend of lectures, exercises, and role-playing in pairs (one trainee taking the role of the interviewer, the other playing the respondent).

After general introductions, the first morning began with an overview of the survey, the survey materials, and the roles the interviewer would play. The various aspects of the Screener were then discussed, followed by role-playing and exercises using the Screener.

After a break for lunch, the afternoon of "day one" was devoted to the Main Interview. A complete demonstration interview was conducted to give interviewers a sense of the way the interview was to be administered. The key features of the interview were then highlighted with a special emphasis on the use of the visual aids and the reading of the narrative material. Question objectives were then reviewed, and the remainder of the day was spent role-playing with the Main Interview.

The morning of "day two" was devoted to additional Main Interview role-playing, followed by exercises on probing. After lunch there were two round-robin interviews involving the entire group of trainees. This allowed everyone to hear feedback given to each member of the group. The remainder of the afternoon was then spent on administrative and reporting issues.

After returning home from training, interviewers were required to complete two practice interviews before beginning their actual assignments. These interviews were conducted with households that had not been selected from the sampled blocks; the respondents were not aware

[61]See Westat's "National Opinion Survey Main Study Trainer's Manual."

ACE 10917193

that the interviews were being conducted for practice. The completed questionnaires were mailed to supervisors for review and feedback.

§ 4.5 Interviewer Supervision

All interviewers reported to one of the three regional field supervisors (each of whom had an office assistant), who in turn reported to the field director. Supervisors were responsible for conferring with interviewers on a regular basis, reporting on and managing progress, performing quality control edits, and validating interviews.

Interviewers reported to their supervisor by telephone according to a schedule: twice a week at the outset of the study and at least once a week thereafter. The discussion included general comments, a case by case review, feedback on quality and production, and planning strategy for the remaining assignment.

Supervisors or their office assistants entered all data on interviewing production, time, and expenses into a machine-readable file that generated status reports. Supervisors reported to the field director during a weekly telephone discussion. In addition to survey progress, other matters discussed included case reassignment and refusal conversion strategies.

§ 4.6 Quality Control Edits

Interviewers sent questionnaires to their supervisor as they were completed. Upon receipt, the supervisors were responsible for a comprehensive edit of the questionnaires before sending them to the home office for coding. (The 100 percent edit rule was lifted during the last few days of the field period to allow for quicker turnaround of the final cases.[62]) The edit

[62]The Westat home office staff was responsible for the edits on these few surveys.

ACE 10917194

for completeness and accuracy used the form shown in Appendix B.4. It covered respondent selection, skip patterns, probing, verbatim recording, and other administrative matters. Results of the edits were discussed, as needed, with the interviewers.

Only two problems worth noting emerged. The edits uncovered 37 cases in which respondent selection within the household was carried out improperly. In 32 of these instances, the mistake was clearly a haphazard one that would not be a potential source of bias (*e.g.*, the Family Sampling Table was used in place of the Person Sampling Table, or the line numbers from the enumeration table were used instead of those from Box 4 of the Screener). In two instances, the error was clearly a motivated one (#'s 1508 and 1509); and in three cases it was hard to tell whether the mistake was made for the sake of convenience (#'s 1510-1512). In addition, in one other interview, the proper respondent was selected but broke off the interview at question A-7A; her husband was the respondent for the remainder of the interview (# 1513).

The edits also revealed 50 cases in which data on the household's income was lost through interviewer misunderstanding of the manner in which it was to be entered in the Questionnaire.[63] Four interviewers accounted for about three quarters of these cases. Supervisors were able to re-contact most of these households and recover this information.

§ 4.7 Validation of Interviews

Supervisors validated at least a 10 percent random sample of each interviewer's assignment. These cases were preselected for validation at the home office in advance of the

[63]In response to C-4 the respondent was to indicate which of the income categories (A-K) on CARD 9 best described household income, and the interviewer was to record the category in a blank provided for that purpose. Under that blank, the interviewer was to mark one of four discrete choice responses indicating whether the respondent's answer was in income category A, in the group of income categories B-K, was a refusal, or was a not sure. In 50 cases, the interviewer marked only the discrete choice answer for categories B-K and failed to record the letter designating the exact income category. See Questionnaire in Appendix A.

ACE 10917195

field work. Thus, both interviews and non-interviews were validated. Supervisors sometimes supplemented the preselected cases with additional cases to be validated (if, for example, a traveling interviewer was visiting a PSU).

Most validations were performed by telephone using the form shown in Appendix B.5. Validations on cases without telephone numbers were attempted by mail or in-person. In the 26 instances where validation could not be carried out (because, *e.g.*, no validation questionnaire was returned by a household that had refused to participate in the survey), another case from the appropriate interviewer's assignment was selected for validation (except for a few cases from interviewers who already had at least 10 percent of their assignments validated). Of the 180 cases that could be checked, all were successfully validated.

§ 4.8 Interview Characteristics

The mean interview length was 42 minutes, and the median length was 40 minutes. Ninety-five percent of the interviews took between 25 and 70 minutes to complete. The shortest interview was 19 minutes and the longest was 2.5 hours.[64]

At the beginning of the interviewing period, 4.8 hours of field work were required to complete an interview. By the time the survey was completed, an average of 8 hours of field work was required to obtain each interview. This reflected the large effort put into locating difficult-to-find respondents and converting refusals. The field cost, exclusive of out-of-town travel and supervision, rose from about $50 per completed interview to over $600 per completed interview toward the end of the interview period. PSU's varied widely in the degree of effort

[64]A random sample of the entire population always contains a few respondents who are either extremely talkative or have great difficulty coping with the survey task.

ACE 10917196

required to complete an interview; the average time required ranged from just over three hours in Grand Rapids, Michigan to over twenty hours in Miami, Florida.

§ 4.9 Sample Completion

Visits to each of the 1,599 sampled DU's established that 176 were vacant. At the remaining 1,423 DU's, interviewers attempted to complete a Screener (to collect information on household composition and select a respondent for the Main Interview), succeeding in 1,198 cases. The 225 non-responses to the Screener were distributed as follows:

166	Screener Refusals
2	Language Barrier
7	Physical or Mental Handicap
34	Never Reached
16	Other Screener Non-responses
225	Total Screener Non-responses.

The results from the 1,198 DU's where a Screener was completed and a respondent selected for the Main Interview were as follows:

1,043	Main Interview Completions
91	Main Interview Refusals
34	Language Barrier
13	Physical or Mental Handicap
11	Never Reached
6	Other Non-interview
1,198	Total Screener Completions.

The overall response rate was 75.2 percent: 1,043 / [1,599 - (176 + 2 + 34)]. In calculating the response rate, the thirty-six non-English speaking households (2 Screener Non-responses + 34 Main Interview Non-responses) were ineligible for the survey and were removed

ACE 10917197

from the denominator as were the 176 vacant DU's.[65] Our 75 percent response rate compares favorably with the best academic surveys such as the University of Michigan's American National Election Surveys and the University of Chicago's General Social Survey.

As is typically the case in nationwide in-person surveys, the response rate was lower in large urban areas than in the rest of the country; however, the difference was smaller than that experienced in many comparable surveys. The response rate was about 8 percentage points lower in the nation's 17 biggest metropolitan areas than elsewhere (69.6 percent versus 77.8 percent).[66]

§ 4.10 Selection Bias and Sample Weights

As information about the survey topic was not provided to individuals until the interview proper, willingness to pay for the Prince William Sound Program could not have directly affected whether or not a household responded. It is possible, however, that other characteristics (*e.g.*, household size or, as noted above, residence in large urban areas) were related to responding/non-responding status. Thus, the composition of the interviewed sample could differ from that of the total random sample initially chosen. In addition, the composition of the total sample might have differed from that of the total population because of errors made during block listing.

To correct for these potential problems, sample weights were constructed that incorporated both nonresponse adjustment and poststratification to household totals from the

[65]This calculation ignores the one block that was in a closed community (see footnote 60). As that block was not listed, we don't know exactly how many DU's would have been sampled from it. We can, however, estimate its impact on the response rate by multiplying the response rate reported in the text by 331/332 (the proportion of sampled blocks contributing to the sample of DU's), which yields 75.0 percent.

[66]The response rate for each PSU is provided in Appendix B.2.

ACE 10917198

1990 Decennial Census. The variables used were region, age, race, household size and household type (married couple versus other).[67] Respondents from the western states, older respondents, black respondents, and single households tended to be assigned higher weights.

We have not made any additional corrections to the data set beyond those implied by the weighting scheme described above. Doing so is equivalent to assuming that after weighting, dwelling units chosen for our sample but not interviewed are missing at random with respect to their willingness-to-pay values. To a large degree, this is a plausible assumption because a household's decision to participate or not participate in our survey was independent of our survey's subject matter since it was not revealed to them before participating.[68] It is possible that households who are very difficult to find at home or who generally refuse to be interviewed have systematically different willingness-to-pay values, but it is unclear whether they might be higher or lower. In any event, our response rate is sufficiently high that any sample selection effects should be reasonably small.

Due primarily to logistical and cost considerations, no foreign language versions of the questionnaire were developed.[69] As a result, non-English speaking households were not eligible to be interviewed. Thus, we reduced the 1990 Census estimate of the number of U.S. households (93,347,000) by 2.7 percent, our survey's estimate of the proportion of U.S.

---

[67]For details, see Ralph DiGaetano's August 12, 1991 memo in Appendix B.3.

[68]This is in contrast to mail surveys where respondents may read all of the questions before deciding whether to participate.

[69]A non-English version would have presented administration problems since the multi-lingual interviewers would need to visit widely separated locations in order to adequately represent that population. Any non-English version of the questionnaire would have also required separate testing. These considerations would have led to dramatically escalated survey costs. In addition, although some pockets of particular non-English speaking groups are easily identifiable, *e.g.*, Hispanics in Texas or Vietnamese in California, the possible bias from selection of non-English speakers only in those areas would prevent straightforward generalization to the entire non-English speaking American population.

ACE 10917199

households that were non-English speaking.[70] This yields a population estimate of 90,838,000 English speaking households to which our results may be extrapolated.

§ 4.11 Data Entry

As the questionnaires returned from the field, the numeric responses and the verbatim responses were entered by Westat's data entry department. The numeric data from each questionnaire was entered, to the extent possible, as it appeared on the questionnaire; the data entry incorporated no provision for enforcing skip patterns in the data. The data were entered in batches, and consistency checks were performed on those batches. When data entry activities for a batch of questionnaires was complete, that batch was sent to Natural Resource Damage Assessment, Inc. (NRDA). When the data entry was completed, Westat sent an ASCII dataset to NRDA.

Questionnaires arriving at NRDA were logged and filed and the numeric data were re-entered at NRDA. When Westat produced a dataset, that dataset was compared with the dataset generated at NRDA. For each case, a direct comparison was made of the two values for each variable. Differences were reconciled by an examination of the source questionnaire; and a dataset was constructed incorporating the reconciled values of the two data sets. Tabulations from this dataset, weighted and unweighted, are found in Appendix C.1.

Before sending each batch of questionnaires, Westat also entered the verbatim responses to the open-ended questions. When the questionnaires arrived at NRDA, these verbatim responses were entered again. The two data sets were compared at NRDA by visually comparing the entries for each question. Inconsistencies were resolved by reference to the

[70]The survey's estimate of non-English speaking households was used since the Census Bureau does not provide this information.

ACE 10917200

source questionnaires, and a dataset was constructed incorporating the reconciled responses of the two compared data sets. That dataset is listed in Appendix D.

The dataset of reconciled verbatim responses was used to construct a coding schema for each of the open-ended questions. These coding schemata, provided in Appendix C.2, were used to code the verbatim responses. The coded values were then entered into a numeric dataset. These new data were checked for consistency, and any inconsistencies were resolved by examining the source questionnaire and the coding instructions for the variable in question. These values are tabulated in Appendix C.3.

ACF 10817301

CHAPTER 5 — ANALYSIS

§ 5.1 Introduction

In this chapter the responses of the national sample to the final survey instrument are analyzed.[71] In Section 5.2, the responses to the initial attitudinal questions about different government policy programs, questions about the Exxon Valdez oil spill, and questions about household attributes, including demographic questions, are discussed. In Section 5.3, the questions asked of the interviewers for assessing the quality of the interviews are discussed. In Section 5.4, the questions regarding how the spill and the plan to prevent a future spill were perceived by respondents are examined. In Section 5.5, the responses to the willingness-to-pay (WTP) questions A-15, A-16, A-17, C-7, and C-8 are examined. In Section 5.6, the statistical framework for this analysis is introduced. In Section 5.7, the univariate estimates of our sample's willingness to pay to prevent an oil spill similar to the Exxon Valdez oil spill are presented. In Section 5.8, the reasons given by respondents for their WTP responses are examined. In Section 5.9, a valuation function which predicts a household's willingness to pay from the characteristics of that household is described. In Section 5.10, various adjustments to the willingness-to-pay amounts are made. In Section 5.11, the effect of some alternative adjustments to the median WTP estimate are discussed. In Section 5.12, the replicability and stability of the median willingness-to-pay estimate over time is explored. In Section 5.13, possible ways to approximate more closely mean willingness to accept (WTA) compensation are explored. Finally, in Section 5.14, concluding remarks are presented.

[71]The final survey instrument may be found in Appendix A. Details of the sampling plan and survey administration by Westat were described in Chapter 4.

ACE 10917202

§ 5.2 Attitudinal, Knowledge, and Demographic Questions

The first series of questions (A-1a to A-1f) in the survey instrument asks respondents: "Do you think we should spend a great deal more money than we are spending now, somewhat more money, the same amount of money, somewhat less money, or a great deal less money," on six items: (a) foreign aid to poor countries, (b) making sure we have enough energy for homes, cars, and businesses, (c) fighting crime, (d) making highways safer, (e) improving public education, and (f) protecting the environment. The order in which these questions were asked was randomly rotated. Responses ranged from 49 percent in favor of spending a great deal more money on improving education to 3 percent who thought a great deal more money should be spent on giving foreign aid to poor countries. Thirty-nine percent were in favor of spending a great deal of money to protect the environment; this item ranked third after education and fighting crime (42 percent). A complete breakdown of the responses to these and other questions is contained in Appendix C.1.[72]

Similarly, the A-3 series of questions (A-3a to A-3f) asked respondents: "How important to you personally are each of the following goals? . . . is that extremely important to you, very important, somewhat important, not too important, or not important at all?" The goals were: (a) expanding drug treatment programs, (b) reducing air pollution in cities, (c) providing housing for the homeless, (d) reducing taxes, (e) putting a space station in orbit around the earth, and (f) protecting coastal areas from oil spills. Again the items were rotated. Responses of "extremely important" ranged from 36 percent of respondents who felt that protecting coastal areas from oil spills was extremely important to 4 percent who thought that putting a space station in orbit around the earth was extremely important. A composite category of extremely

[72]Appendix C contains both the actual and weighted counts and the actual and weighted percentages for each closed-ended question in the survey instrument.

ACE 10917203

important and very important categories ranged from 81 percent in favor of protecting coastal areas from oil spills to 15 percent for the space station. In the next question (A-4), the public is roughly split on how much more land the government should set aside as wilderness areas, 56 percent saying a very large or large amount and the rest of the sample indicating a moderate amount to no amount.

Question A-2 began the process of narrowing the scope of the interview to its primary focus: "Now I'd like you to think about major environmental accidents caused by humans. Please think about those accidents anywhere in the world that caused the worst harm to the environment. During your lifetime which accidents come to mind as having damaged nature the most?" The response to this question shows the Exxon Valdez spill to be one of the most salient environmental accidents to have occurred. About two years after the Exxon Valdez spill, over 53 percent of our sample spontaneously named the Exxon Valdez in response to this question. Only two other accidents were named by more than 20 percent of the sample: the oil spills in the Persian Gulf during the war with Iraq (25 percent), and the Chernobyl nuclear reactor accident (20 percent). Nine percent named Three Mile Island.

Another 26 percent of the respondents named the Exxon Valdez in response to the more specific open-ended question A-5: "Have you heard or read about large oil spills in any part of the world (other than those you mentioned earlier)?" Of the 21 percent in our sample who had not mentioned the Exxon Valdez oil spill in response to A-2 or A-5, 74 percent said that they had heard of it when asked A-6.[73] When all three responses are considered, less than 6 percent of the sample said that they had not heard of the Exxon Valdez spill or did not know whether they had heard of it. The significance of this six percent is put into perspective by Carpini and

[73]Until A-6 no oil spill or location was specifically mentioned by the questionnaire. The questionnaire narrowed its focus from "major environmental disasters" in A-2 to "large oil spills" in A-5 to the Valdez spill in A-6.

ACE 10917204

Keeter (1991). They asked a national sample of American adults: "Will you tell me who the Vice President of the United States is?" Twenty-six percent said that either they did not know who the Vice President was or named someone other than Dan Quayle.

From this point onward in the questionnaire the focus is on the Exxon Valdez oil spill. In A-6a, respondents were asked the open-ended question: "What was it about the natural environment around Prince William Sound that you feel was most seriously damaged by the oil spill?" Table 5.1 displays a coded version of these responses.[74] Over 90 percent of those answering this question saw some aspect of the ecosystem (the first nine categories in the table) as seriously damaged. A small percentage of respondents named other injuries such as commercial fishing or recreation. These responses were usually given after one of the more common responses, such as wildlife or birds.

The next block of questions, A-7 through A-10a, asked households whether they had visited Alaska and Prince William Sound in the past. Less than 10 percent of our sample households had visited Alaska and less than 2 percent of our sample households had visited Prince William Sound. Most of those who had been to Alaska had only been there once, on average 14 years ago.

Questions A-11 and A-12 asked respondents about how closely they had followed the Exxon Valdez spill and about their news sources. Twenty-three percent of respondents said they followed the spill "very closely," and 51 percent said "somewhat closely." For respondents who followed news about the spill, television was the primary source. Forty-five percent of respondents said they got most of their information about the spill from television; another 45

[74]Multiple responses were encouraged via the interviewer probe: "Anything else?". The percentaging base is the number of respondents answering this question. Since many respondents gave multiple responses, the percentages total more than 100 percent.

ACE 10917205

Table 5.1 Items Most Seriously Damaged by Spill

A-6a: What was it about the natural environment around Prince William Sound that you feel was most seriously damaged by the oil spill?	N=981
Response Category	Percentage
Wildlife	43%
Sea Life	37%
Birds	34%
Fish/Shell Fish	31%
Mammals	30%
Water	13%
Ecosystem	10%
Commercial Fishing	8%
Economy	6%
Plants	6%
Natural Beauty	3%
Health	3%
Natives	1%
Recreation	1%
Other	7%

percent said they got most of their information from the combination of television and newspapers. Six percent of respondents said they got most of their news about the spill from newspapers, and four percent volunteered another primary source for their news, typically radio or magazines.

The remainder of the questions in Section A of the survey instrument describe the Exxon Valdez oil spill and assess willingness to pay to prevent a similar spill in the future. These questions will be taken up in the next section. The first ten questions in Section B of the survey instrument deal with the way respondents perceived the Exxon Valdez spill and the plan to

ACE 10917206

prevent another similar spill. These questions will also be taken up in the next section.

Questions B-10 through C-6 concern household attributes. Fifteen percent of the sample thought it very likely that they would visit Alaska at some time in the future; and 18 percent thought it somewhat likely (B-10). Forty-eight percent of the households have someone who engaged in recreational fishing (B-12); 31 percent have someone who bird watches (B-13); and 17 percent have someone who backpacked (B-14). In answer to B-15, 44 percent said that someone in the household had visited either the Grand Canyon, Yosemite, or Yellowstone National Parks. In B-16, 60 percent thought of themselves as environmentalists; and of those, 16 percent considered themselves very strong environmentalists which represented about 10 percent of the sample as a whole (B-17). In B-18, 19 percent of the respondents said "very frequently" and another 26 percent said "frequently" when asked if they watched television shows about animals and birds in the wild.

Sample demographics were collected via questions C-1 through C-6. The median age of our respondents was 41, and the mean age was 45. The youngest person in our sample was 18; and the oldest 88. In response to the question regarding education (C-2), 7 percent of our sample had no high school education; 12 percent had some high school education; 34 percent had a complete high school education; 24 percent had some college education; 13 percent had a bachelor's degree; and 8 percent had post-graduate education. Forty-two percent had children and 1 percent had more than four children (C-3). Twenty-seven percent were single; and 15 percent lived in households with more than two adults. Sixty-three percent lived in single family homes. The median household income was in the $20,000-30,000 category. Ninety-four percent of our sample said that someone in their household paid federal income taxes.

ACE 10917207

§ 5.3 Interviewer Assessment Questions

Questions in Section D asked the interviewer to assess different aspects of the interview. D-1 asked interviewers: "How informed did the respondent seem to be about the Alaskan oil spill?" The interviewers believed 33 percent of the respondents to be "very well informed," 40 percent to be "somewhat well informed," 17 percent to be "not very well informed," and 8 percent to be "not at all informed." With respect to interest in the effects of the Alaskan oil spill (D-2), 53 percent appeared to the interviewers to be "very interested" and another 33 percent, to be "somewhat interested." They reported 10 percent to be "not very interested," and 2 percent to be "not at all interested." Questions D-3 and D-4 asked about how cooperative and hospitable the respondent had been at the beginning and at the end of the interview. The interviewers felt that 71 percent had been very "cooperative/hospitable" at the beginning of the interview and that 81 percent had been very "cooperative/hospitable" at the end of the interview. At the other end of the scale, 7 percent of respondents started out not very "cooperative/hospitable" or not "cooperative/hospitable" at all at the beginning of the interview; this percentage had fallen to less than 4 percent by the end of the interview. In about 40 percent of the interviews, another person was present (D-5); but in most of these cases (77%) the other people present did not ask questions or offer answers (D-6). In 80 percent of the cases in which other people did make remarks, interviewers believed that the remarks had little or no effect on the respondents' answers.

Interviewers said that when describing the plan to prevent another Exxon Valdez type oil spill, only 3 percent of the respondents were "extremely" or "very" distracted (D-8b), 2 percent were "not at all interested" (D-8c), 7 percent were only "slightly" interested (D-8c), and less than 3 percent of the respondents were "extremely" or "very" bored during the interview (D-8c). Four percent of the respondents had some difficulty understanding the WTP voting questions (D-

ACE 10917208

9). An examination of the descriptions of these difficulties recorded in the open-ended question D-10 shows that 68 percent of these 39 respondents had difficulties such as difficulty in understanding, not being motivated to pay attention, and language problems. The other interviewers mentioned problems such as the respondent's being a Jehovah's witness and unable to vote, the respondent's having not much money, and the respondent's complaining that this was Exxon's responsibility. Finally, less than 1 percent of the respondents were reported to have taken the voting question "not at all seriously," and another 4 percent were reported to have taken the voting question only "slightly seriously" (D-11).

§ 5.4 Depiction of the Spill and Perceptions of Spill Prevention Plan

The survey instrument contained a number of questions interspersed in the scenario description which were designed to discover how respondents perceived the description of injuries from the Exxon Valdez oil spill and the plan proposed to prevent a similar spill in the future. Question A-13 asked, "I've been telling you a lot about this part of Alaska and the effects of the oil spill. Did anything I said surprise you?" About two-thirds of respondents did not express surprise at the information given to them. Of those who did express surprise, most thought that the effects of the spill, as described in the survey, were less severe than they had assumed prior to the interview.[75] Some respondents said that before hearing the detailed description presented in the survey, they had thought that the recovery period was likely to be longer and that there had been harm to fish and land mammals.

The sequence of questions beginning with A-14 focuses on the plan. A-14a asked: "Is there anything more you would like to know about how a spill could be contained in this way?"

[75]Our focus group and pilot study work had shown that people accepted the spill facts provided in the survey.

ACE 10917209

Only 10 percent of the respondents replied that they had questions with regard to how a spill would be contained. These questions exhibit no pattern; the most common questions asked about the cost of the program or expressed doubts about the effectiveness of the escort ship plan. In response to A-14c, about 20 percent of the respondents said they had questions about how the program would be financed. These respondents (A-14c-1) tended to ask how much the program would cost, to express concern that the money would actually be collected for more than one year, to note that the plan was a good idea, or to argue that the oil companies should be paying all the costs.

This line of questioning resumed after the valuation questions. The questions at the beginning of Section B were to ascertain what assumptions a respondent might have made about certain issues when deciding whether to vote for or against the spill prevention program. Questions B-1 through B-4 assessed the degree of damage the respondent thought would be prevented by the spill prevention plan; the damage caused by the Exxon Valdez spill was the reference point. Question B-1 asked: "When you decided how to vote, how much damage did you think there would be in the next ten years <u>without</u> the [escort] program about the <u>same</u> amount of damage as caused by the Valdez spill, or <u>more</u> damage, or <u>less</u> damage?" Forty-three percent thought the same amount of damage would occur without the program and another 10 percent were not sure. Respondents replying that the damage would be more or less were asked a follow-up question regarding how much more or less and why.

Of the 22 percent who thought there would be more damage, B-2 asked whether it would be a little more (18 percent), somewhat more (42 percent), or a great deal more (32 percent). Respondents offered two common reasons: first, that the prior occurrence of the Exxon Valdez spill might make the damages from the second spill worse, and second, that more oil would be

shipped from Alaska. The other responses tended to be vague, running along the lines of "things are just getting worse" or "there is a potential to kill more wildlife."

Of the 25 percent who thought there would be less damage, B-3 asked whether it would be a little less (44 percent), a lot less (41 percent), or no damage at all (10 percent). These respondents gave one major reason: the first spill would make the second less harmful, usually because people would be more cautious or better prepared. Others thought that there would be more double-hulled ships, that the first accident was a fluke, or they were vague about the reasons why the damage would be less.

The next two questions examined whether respondents thought they were buying protection for a larger area. B-5 asked the respondents: "Did you think the area around Prince William Sound would be the only place directly protected by the escort ships or did you think this particular program would also provide protection against a spill in another part of the U.S. at the same time?" Eighty-four percent of respondents believed only Prince William Sound would be protected, 10 percent said that another part of the U.S. would be protected, and 6 percent were unsure. Those respondents who said that some other part of the United States would be protected were asked "How?" in question B-6. The responses to B-6 showed no distinct patterns. Some thought that the oil would escape the Sound and affect a larger area; some thought that the plan would set a precedent or provide useful experience; others thought that better inspections in Valdez might be beneficial to wherever the final destination of the tanker was; a few respondents named distant locations that they thought might be protected. Many of these responses suggest that those who said another part of the U.S. would be protected were simply trying to "guess" how the plan might have broader impacts rather than relating what they actually thought at the time of answering the WTP questions.

We turn next to the issue of the effectiveness of the escort ship plan. B-7 asked: "If the escort ship program were put into operation, did you think it would be completely effective in preventing damage from another large oil spill?" Forty percent believed that the escort ship plan would be completely effective. Those who did not were asked B-8: "Did you think the program would reduce the damage from a large spill a great deal (45 percent), a moderate amount (32 percent), a little (12 percent), or not at all (3 percent)." Over two-thirds of the respondents were convinced that the escort ship plan would be largely successful in preventing damages from another Exxon Valdez type spill; another 19 percent believed that the plan would prevent some non-trivial amount of damage.

B-9 checked whether the respondent had accepted statements about the period the tax would be in effect: "When you answered the questions about how you would vote on the program, did you think you would actually have to pay extra taxes for the program for <u>one</u> year or for <u>more</u> than one year?" Seventy-one percent said one year, 23 percent said more than one year, and 6 percent were not sure.

B-10 asked respondents for a comparison of their prior beliefs about the damages caused by the Exxon Valdez spill with the description of the damages given in the survey instrument: "Before we began this interview, did you think the damage caused by the Exxon Valdez oil spill was more serious than I described to you, less serious, or about the same as I described to you?" A little over half said that they believed that the damages were about the same. Those thinking that the damages were more serious before the interview out-numbered those who thought they were less severe.

We now jump from B-11 which began a series of demographic questions to question C-11 which asked respondents: "Who do you think employed my company to do this study?" The responses to this sponsorship question are given in Table 5.2 below. These responses suggest

ACE 10917212

that the survey was quite balanced. Twenty-nine percent of the respondents guessed that Exxon sponsored the study; another 13 percent thought that another oil company or "the oil companies" sponsored the survey; 23 percent thought the government (typically the federal government or

Table 5.2 Perceived Sponsor of Survey

C-11: Who do you think employed my company to do this study? (IF NECESSARY, PROBE: "What is your best guess?" "Could you be more specific?")	N=1041
Exxon	29%
Oil Company(s)	13%
Government	23%
Environmental Group(s)	9%
Multiple (Conflicting) Responses	11%
Other	3%
Not Answered/Not Sure	11%

some specific federal agency like the EPA) sponsored the study; 9 percent thought an environmental group or groups sponsored the study; 11 percent gave multiple conflicting responses (*e.g.*, Exxon or an environmental group); 3 percent gave other answers such as Westat or a newspaper; and 11 percent did not venture a guess.[76]

§ 5.5 WTP Questions

The survey instrument used a double-bounded dichotomous-choice elicitation framework (Carson and Steinberg, 1990; Hanemann, Loomis, and Kanninen, 1991) to obtain information about respondents' willingness to pay to prevent another Exxon Valdez type oil spill. In this

[76]An examination of the additional comments made on C-11 and the response to C-12: "What made you think that?" also suggests that the survey was fairly well-balanced as many of the respondents indicated that they were uncertain or could at most point to a few weak indicators to support their sponsorship belief.

ACE 10917213

framework, an initial binary discrete question (A-15) asks how the respondent would vote on the prevention plan if it cost their household $____. If the respondent said "for," he was asked in question A-16 how he would vote if the program cost a higher amount. If the respondent said "against" or "not sure" in A-15, the respondent was asked in A-17 how he would vote if the program cost a lower amount.

The four versions of the survey questionnaire differed only in the amounts used in A-15, A-16, and A-17. These amounts are given in Table 5.3. All cases in the sample were randomly assigned to one of these four versions. Since respondents were randomly assigned to questionnaire versions, no correlation between responses and the version of the questionnaire should be expected except for the WTP questions (A-15, A-16, A-17).[77] A correlation should exist between WTP responses and questionnaire version since the amount respondents were asked to pay differed systematically with the version of the questionnaire.

Turning to the actual responses to the discrete choice WTP questions, Table 5.4 shows the frequencies of each response to question A-15.[78] As expected, the percentage responding with a "yes" or "for" vote declines as the amount the respondent is asked to pay increases, dropping from 67 percent in favor at $10 to 34 percent at $120. The WTP distribution appears to be fairly flat in the range from $30 (version B) to $60 (version C). An examination of the "no" or "against" responses and the "not sure" responses suggests that "not sure" responses are being replaced by "no" responses as the amount the respondents are asked to pay increases from

[77]This statement is true, asymptotically, *i.e.*, as the sample size gets very large.

[78]The frequencies for A-16 are: version A (67 percent yes, 22 percent no, 4 percent not sure), version B (50 percent yes, 39 percent no, 11 percent not sure), version C (42 percent yes, 49 percent no, 9 percent not sure), version D (40 percent yes, 45 percent no, 15 percent not sure). The frequencies for A-17 are: version A (9 percent yes, 85 percent no, 6 percent not sure), version B (24 percent yes, 65 percent no, 9 percent not sure), version C (20 percent yes, 70 percent no, 10 percent not sure), version D (18 percent yes, 70 percent no, 11 percent not sure). It is important to note that a respondent was asked either A-16 or A-17 conditional on the response given to A-15 and not both questions.

ACE 10917214

Table 5.3 Program Cost by Version and Question

Version	A-15	A-16	A-17
A	$10	$30	$5
B	$30	$60	$10
C	$60	$120	$30
D	$120	$250	$60

$30 to $60.

These data could be analyzed with a binary discrete choice model, such as a logit or a probit, but that model would not efficiently use the information in the data set. To use all information in the data set efficiently, the A-15 responses should be combined with the A-16 and A-17 responses. Treating the "not-sure's" as "no" responses results in four response types.[79] These are presented by questionnaire version in Table 5.5.

The yes-yes and no-no responses are the easiest to interpret because we would expect the yes-yes responses to fall as the dollar amount the respondent is asked to pay goes from $30 in version A (*i.e.*, 45 percent say yes to $30) to $250 in version D (*i.e.*, 14 percent say yes to $250). We would also expect the no-no responses to increase as we move from version A (*i.e.*, 30 percent say no to $5) to version D (*i.e.*, 54 percent say no to $60). The no-no responses to version A define the upper bound on the percentage of respondents who may not care about preventing an Exxon Valdez type oil spill. It should be noted, though, that this group of respondents is also likely to include those who do not think that the escort ship plan will work

[79]For most of the respondents giving "not-sure" answers, this interpretation seems to be appropriate. Some respondents gave a "not sure" answer to A-15 and subsequently gave a "yes" answer to the substantially lower amount in A-17. Similarly, some respondents gave "yes" responses to A-15 and "not sure" responses to the higher amount in A-16. A likely interpretation is that these "not sure" responses represent respondents who were reasonably close to their indifference thresholds. Of the 141 respondents who gave one or more "not sure" responses, 111 followed this pattern. The other 30 gave "not sure" responses to both A-15 and A-17; these respondents may not have been capable of answering the WTP questions. We have also treated them as no-no responses, which, again is the conservative course.

ACE 10917215

Table 5.4 A-15 Response by Version

Version	Yes	No	Not sure
A	67.42%	29.92%	2.65%
B	51.69%	39.33%	8.99%
C	50.59%	43.53%	5.88%
D	34.24%	59.14%	6.61%

or who believe that the oil companies should pay the entire cost of the plan.

The data gathered using the double-bounded dichotomous choice elicitation method is sometimes referred to as interval-censored survival data (Nelson, 1982). A yes-yes response indicates that the respondent's maximum willingness to pay lies between the A-16 amount and infinity. A yes-no response, *i.e.*, yes to A-15 and no to A-16, indicates that the respondent's maximum WTP amount lies between the amount asked in A-15 and the amount asked in A-16. A no-yes response indicates that the respondent's maximum WTP response lies between the amount asked in A-15 and the amount asked in A-17. A no-no response indicates that the respondent's maximum willingness to pay lies between zero and the amount asked in A-17.[80] Thus, a respondent's willingness-to-pay response can be shown to lie in one of the following intervals depending on the particular response pattern and questionnaire version:

Version A	0 - 5	5 - 10	10 - 30	30 - ∞
Version B	0 - 10	10 - 30	30 - 60	60 - ∞
Version C	0 - 30	30 - 60	60 - 120	120 - ∞
Version D	0 - 60	60 - 120	120 - 250	250 - ∞ .

One additional consideration affects the categorization of respondents into intervals. In C-7 and C-8, we gave respondents who said "yes" to A-15 or A-17 the opportunity to change

[80] If the amenity being valued is "bad" to the respondent, then the lower bound on the interval is negative infinity rather than zero. This situation is possible with some public goods, but it is unlikely that anyone views an Exxon Valdez type oil spill as something desirable.

ACE 10917216

Table 5.5 Questionnaire Version by Type of Response

Version	Yes-Yes	Yes-No	No-Yes	No, No
A	45.08%	23.35%	3.03%	29.55%
B	25.84%	25.84%	11.61%	36.70%
C	21.26%	29.13%	9.84%	39.70%
D	13.62%	20.62%	11.67%	54.09%

their vote to "no." In C-7, respondents were reminded of the highest amount to which they had said "yes" and asked how strongly they favored the plan if it cost their household that amount. Twenty-four percent said they favored the program "very strongly," 52 percent said "strongly," 20 percent said "not too strongly," 3 percent said "not at all strongly," and three respondents volunteered that they no longer favored the plan. Those respondents who did not say "very strongly" or "strongly" were asked in C-8: "All things considered would you like to change your vote on the program if it cost your household $____ from a vote for the program to a vote against." The WTP interval of the respondents who indicated that they wanted to change their votes (3 respondents in C-7 and 8 in C-8) was set from zero to the highest amount to which they had previously said they would vote "for."[81]

§ 5.6 Statistical Framework

The general statistical framework for survival analysis with interval-censored data (Nelson, 1982) is straightforward. First we obtain a sample containing $i=1, 2, \ldots, n$ agents (*e.g.*, survey respondents) with statistically independent log life-times y_i (*e.g.*, maximum willingness to pay) from a cumulative distribution function (CDF),

[81]In addition, four respondents who did not answer the second WTP question (A-16 or A-17) had their WTP intervals based only on their response to A-15.

ACE 10917217

$$F(y)=\Phi[(y-\mu)/\sigma],$$

where μ and σ are the true values of the unconditional population location and scale parameters.[82] Inspection of the i_{th} unit occurs j times ($j = 1, 2, \ldots, J$) along the non-negative real line $[0,+\infty]$. The first inspection occurs at η_1 and the last inspection occurs at η_j. In the interval, $[\eta_{j-1},\eta_j]$, a unit can be found to be either working or failed. If a unit has failed, then it is interval-censored because it is known that $n_{j-1} \le y_i < n_j$. A unit that has not failed by η_j will be treated as right-censored, because it is only known that $y_i > n_j$.

If η_j is independent of y_i (conditional on y_i having not failed by η_{j-1}), then the likelihood function can be written as,

$$Log\ L=\sum_i \ln\left[\Phi\left(\frac{\eta_{ij}-\mu}{\sigma}\right)-\Phi\left(\frac{\eta_{ij-1}-\mu}{\sigma}\right)\right].$$

This is because the unit i always fails in some interval since η_j can always take on the value $+\infty$ if the unit has not failed sooner. One can maximize this likelihood function by assuming a particular distribution for Φ, such as the Weibull or log-normal (Nelson, 1982); or it can be fit nonparametrically by using a modification of the Kaplan-Meier estimator proposed by Turnbull (1976).

[82]The location parameter, μ, is often parameterized in terms of observed covariates.

ACE 10917218

§ 5.7 Univariate Estimation of Willingness to pay

The Turnbull-Kaplan-Meier nonparametric approach makes no assumptions about the shape of the underlying WTP distribution.[83] As a result, this technique is only capable of estimating how much of the density falls into the intervals defined by the dollar thresholds used in the different versions of A-15, A-16, and A-17. This technique can not estimate mean willingness to pay; and it can not give a point estimate of the median, but only the interval in which median willingness to pay falls. In Table 5.6, as estimated by this approach, 30 percent

Table 5.6 Turnbull-Kaplan-Meier Estimation Results

Lower Bound of Interval	Upper Bound of Interval	Probability of Being Greater Than Upper Bound	Change In Density
0	5	.696	.304
5	10	.660	.036
10	30	.504	.157
30	60	.384	.119
60	120	.236	.148
120	250	.111	.125
250	∞	.000	.111
Log-Likelihood	-1362.942		

of the respondents fall into the interval $0 to $5, 11 percent are willing to pay over $250, and the median falls into the interval $30-$60.

To get a point estimate of the mean or median, WTP must be assumed to have a particular underlying distribution. The most frequently used distribution for survival data is the

[83]From this point on we will use the household weights provided by Westat in performing any estimations. The differences between the weighted and unweighted estimates are almost always quite small, the weighted estimates being slightly lower than the unweighted estimates.

ACE 10817219

Weibull. The Weibull is a two parameter $[\alpha,\beta]$ distribution where $\alpha > 0$ is known as the location parameter and $\beta > 0$ as the scale parameter. The CDF for the Weibull is

$$F(y)=1-EXP[-(y/\alpha)^{\beta}],\ y > 0,$$

and the density is

$$f(y)=(\beta/\alpha)(y/\alpha)^{\beta-1}\ EXP[-(y/\alpha)^{\beta}].$$

Sometimes the accelerated life parameterization, $\lambda=1/\alpha^{\beta}$ and $\theta=1/\lambda=\alpha^{\beta}$, rather than the proportional hazard parameterization, is used. The mean of a Weibull is $E(Y)= \alpha\Gamma[1 + (1/\beta)]$.

The Weibull survivor function,

$$S(y) = 1-F(y) = EXP[-(y/\alpha)^{\beta}],$$

is the demand curve for the public good in question, and the Weibull hazard function, $[f(y)/S(y)]$, is given by

$$h(y) = (\beta/\alpha)(y/\alpha)^{\beta-1},$$

which is closely related to the elasticity of demand, $-yh(y)$. For $h(y)$ constant, we have close to a linear demand curve; and for $h(y)$ proportional to $1/p$, we have close to a constant elasticity demand curve. The 100_{th} percentile for the Weibull distribution can be found by manipulating the CDF and is given by

$$y_p = \alpha[-\ln(1-P)]^{1/\beta}.$$

The median can be found by setting P equal to .5. The Weibull is the simplest distribution that allows either an increasing, decreasing, or constant hazard function. The Weibull is also flexible enough to approximate several other commonly used survival distributions. If $\beta = 1$, then the Weibull reduces to the exponential distribution (the constant

ACE 10917220

hazard case); $\beta = 2$ gives the Rayleigh distribution; β between 3 and 4 is close to the normal distribution; and β greater than 10 produces results close to the smallest extreme value distribution.

Maximizing the likelihood function for our double-bounded WTP data under the assumption of a Weibull distribution yields the estimates in Table 5.7: estimates of $31 for the median and $94 for the mean. The standard errors and accompanying asymptotic t-values indicate that the parameters are estimated precisely. This precision is reflected in the 95 percent confidence intervals for the mean and median. Figure 5.1 is the estimated Weibull survival curve.

Several distributions other than the Weibull can be fitted to our WTP data to illuminate the sensitivity of the estimates to the particular distribution assumed. Table 5.8 shows the mean

Table 5.8 Medians and Means for Four Distributions

Distribution	Median	95% Confidence Interval	Mean	95% Confidence Interval	Log Likelihood
Weibull	30.91	[26.85-35.59]	94.47	[83.45-105.19]	-1345.298
Exponential	46.29	[43.07-49.75]	66.78	[62.73-70.83]	-1464.547
Log-Normal	27.32	[23.67-31.52]	220.43	[113.31-327.55]	-1363.208
Log-Logistic	28.74	[24.91-33.16]	∞	----	-1365.307

and median estimates for the Weibull and three other common survival distributions: the exponential, the log-normal, and the log-logistic.

The median estimates of the Weibull, log-normal, and log-logistic distribution are all quite close and their 95 percent confidence intervals overlap. The median for the very restrictive exponential distribution is about 50 percent larger than those for the other three distributions. All four estimates of the median are consistent with respect to the $30-$60 interval obtained with

ACE 10917221

Table 5.7 Weibull Estimates

Parameter	Estimate	Standard Error	Asymptotic T-Value
Location*	58.771	3.820	15.37
Scale*	0.571	0.024	23.73
Log-Likelihood	-1345.298		
Median	30.91	[26.85-35.59]**	
Mean	94.41	[83.75-105.19]**	

* Proportional Hazard Parameterization
** 95% Confidence Interval

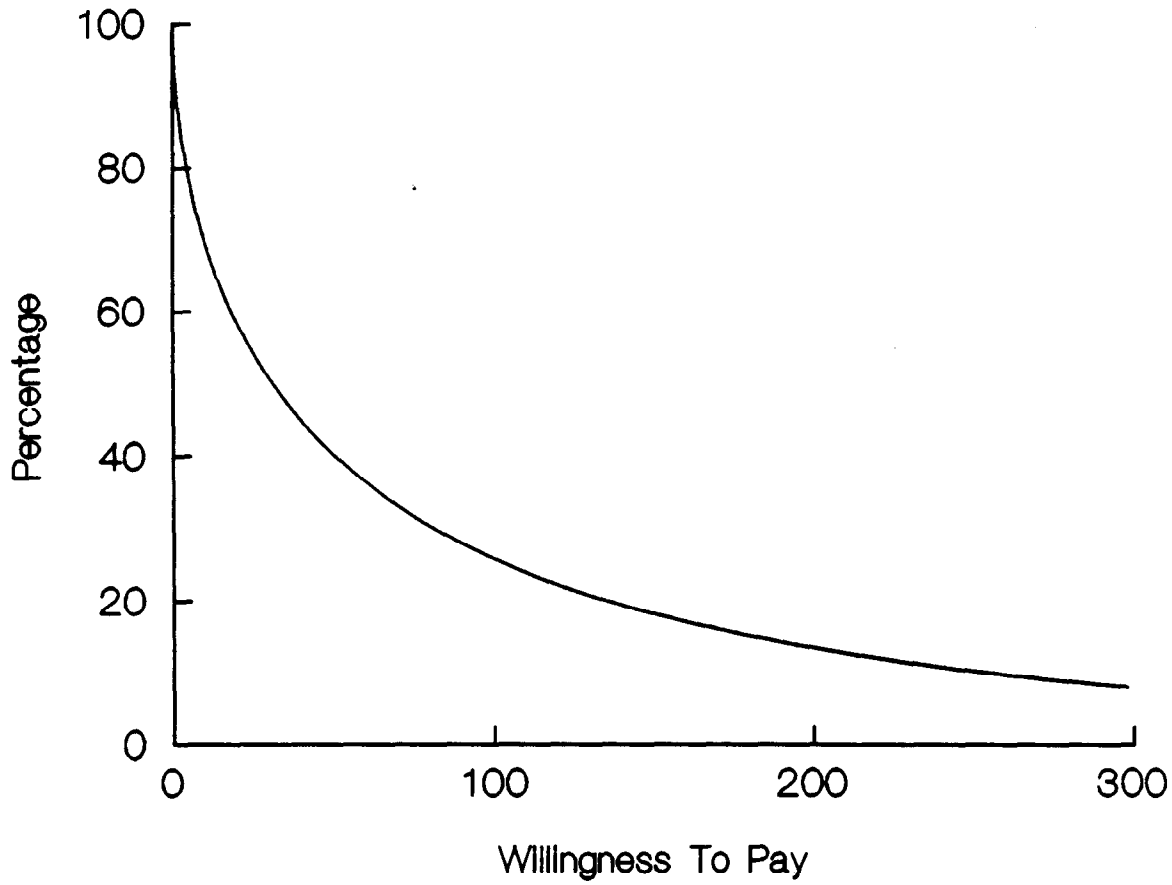
the nonparametric estimator in Table 5.6. The mean estimates are larger than the median estimates and vary greatly. The mean for the exponential distribution is about 30 percent less than that obtained under the Weibull distributional assumption; the log-normal mean is over 100 percent larger than that of the Weibull; and the mean of the log-logistic distribution does not exist.

How can we chose between these distributions? For the Weibull and the exponential, this choice is straightforward since the Weibull distribution collapses to the exponential distribution if the scale parameter is 1. Whether the scale parameter is equal to 1 can be tested by using a likelihood ratio test. This test dictates the rejection of the exponential distribution in favor of the Weibull distribution.[84] It is more difficult to test between the Weibull and the log-normal or log-logistic because these distributions are not nested with the Weibull as is the case with the exponential. In addition, the log-likelihoods of the log-normal distribution (-1363.208) and the log-logistic distribution (-1365.307) are not a lot smaller than that of the Weibull (-1345.298);

[84]The likelihood ratio test statistic equals twice the difference between the unrestricted and restricted log-likelihoods. When the restriction on the scale parameter is correct, this statistic has a $\chi^2_{(1)}$ distribution. Its value was 238.5, which greatly exceeds the 95 percent reference level of 3.84, dictating the rejection of the exponential distribution.

ACE 10917222

Figure 5.1 Weibull Estimate of Percent Willing to Pay as a Function of Amount Specified



whereas the log likelihood of the exponential was quite a bit smaller (-1464.547).[85]

As illustrated by the mean column of Table 5.8, the shape of the right tail of the chosen distribution,[86] rather than the actual data, is the primary determinant of the estimate of the mean. Because the mean can not be reliably estimated and the median can be reliably estimated, we will concentrate on the median in the next several sections. A strategy for obtaining a reliable estimate of mean willingness to pay is discussed in section 5.13.

[85]A non-nested J-test suggests the rejection of the log-logistic in favor of the Weibull. Neither the Weibull or the log-normal clearly dominates on this type of test.

[86]The right tail corresponds to respondents with a very high willingness to pay.

ACE 10917223

§ 5.8 Reasons for WTP Responses

In this section the reasons respondents were willing to pay or not pay for the plan to prevent another Exxon Valdez type oil spill are examined. This examination involves the responses to A-18, A-19, and A-20, most of which are open-ended responses that have been coded into categories.

Those respondents who were not willing to pay either amount in A-15 and A-17 were asked their reasons in A-18. The responses to this question have been placed into the six categories given in Table 5.9.[87] About a third of these respondents said that they could not afford the amount specified or that the program was not worth that much to them. Another third said the oil companies or Exxon should pay. Almost ten percent did not favor the program because they felt it should be protecting other areas, frequently areas near the respondent instead of, or in addition to, Prince William Sound. About 20 percent had some type of complaint about the government. Some of these complaints indicated that the respondent did not think that the spill prevention plan was very important; and in other instances the government was simply deemed incapable of doing things right. In still other instances, respondents indicated that taxes should not be raised for this purpose. The variety of reasons classified as other ranged from simple not-sure's, to being unable to vote because of being a Jehovah's witness, to requiring more information about the plan before being willing to vote yes.

The 47 respondents who said "not sure" to A-17 were asked their reasons in A-19. Table 5.10[88] displays the answers to this question, using the same response categories as in Table 5.9. These not-sure respondents look much like the no-no respondents except for the

---

[87]Because some respondents gave multiple answers, percentages add to more than 100 percent. Both closed-ended responses and open-ended responses to A-18 were coded into these response categories.

[88]The open-ended responses to A-19 were coded into these response categories. Because some respondents gave multiple answers, the percentages add to more than 100 percent.

ACE 10917224

Table 5.9 Reasons Not Willing To Pay Amount

A-18: "Did you vote against the program because you can't afford it, because it isn't worth that much money to you, or because of some other reason?"	N=377
Response Category	Percentage
Can't Afford	24.7%
Not Worth That Much	8.2%
Should Protect Other Areas	9.3%
Oil Companies/Should Pay	33.2%
Government Should Not Pay, Can't Run Well, Anti-Tax	19.2%
Other Reasons Including Not Sure	21.7%

much higher percentage of "other" responses; these responses were varied and not easily coded into a few distinct categories.

Those respondents who were willing to pay at least one of the two amounts specified were asked in A-20 for their reasons. Table 5.11[89] indicates that over two-thirds of the respondents named particular aspects of Prince William Sound that they wished to protect, such as birds, sea otters, or beaches. Twenty-six percent of the respondents made general reference to the Prince William Sound environment. Eight percent of the respondents mentioned people who use Prince William Sound. Twenty-six percent commented that the plan was feasible, well-conceived, effective, or important to implement. Another 16 percent said they supported the plan because its cost was reasonable or affordable given what it would accomplish. Thirteen percent saw the plan as necessary if oil was to be shipped out of Alaska or saw prevention being more cost effective than clean-up. Three percent said that the oil companies should be paying the cost. Six percent gave a variety of general environmental reasons, and 11 percent gave a

[89]The open-ended responses to A-20 were coded into these response categories. Since some respondents gave multiple answers, the percentages add to more than 100 percent.

ACE 10917225

Table 5.10 Reasons Not Sure Whether Willing To Pay

A-19: Can you tell me why you aren't sure?	N=47
Response Category	Percentage
Can't Afford	14.3%
Not Worth That Much	8.2%
Should Protect Other Areas	10.2%
Oil Companies Should Pay	18.4%
Government Shouldn't Pay, Can't Run Well, Anti-Tax	10.2%
Other Reasons Including Not Sure	65.3%

variety of other reasons including not sure.

§ 5.9 Valuation Function

A valuation function is a statistical way to relate respondents' willingness-to-pay to their characteristics. Valuation functions are often developed to demonstrate the construct validity of the estimate from a contingent valuation study. In the simplest sense, the respondent's willingness to pay or an indicator of that willingness to pay is regressed on respondent characteristics such as income and on preferences relevant to the good being valued.

A valuation function is estimated in several steps. First, for those observations with missing values in a possible predictor variable, either those values must be imputed, or the observations must be dropped from any estimation using that variable, a generally undesirable option. Next, which variables to include in the valuation function must be determined. Some variables should clearly be in the valuation function; for other variables the choice is less clear. Finally, the valuation function may be used to make adjustments to WTP estimates for such things as protest responses.

ACE 10917226

Table 5.11 Reason For Being Willing To Pay

A-20: What was it about the program that made you willing to pay something for it?	N=627
Response Category	Percentage
Protect Prince William Sound Area/Environment	26.0%
Protect Specific Aspects/Resources of Prince William Sound	68.1%
Protect People Including Recreators and Visitors	8.3%
Program Feasible/Effective/Important	26.0%
Cost of Program Is Affordable/Reasonable	15.5%
Protect Oil Supply/Prevention Better	13.2%
Oil Companies Should Be Paying	2.9%
General Protect Environment	6.2%
Other Reasons Including Not Sure	10.8%

§ 5.9.1 Imputation of Missing Values For Predictor Variables

A large survey of the general population always has some missing data. For the predictor variables, no approach is conservative by design so we must either impute the missing values using some statistical technique or find ways of operationalizing the variables used in order to avoid missing value problems. It may be useful to first look at the magnitude of the problem. For many of the attitude variables, missing values or not-sure observations are few; for the income variable, about 15 percent are missing values which is typical of large national surveys. We operationalize the attitude variables as dummy variables so that if a respondent did not answer or said "not sure," the condition making the dummy variable equal to one is assumed not to apply. This effectively sets to zero the not-sure responses and the missing values. Missing values for the 12 respondents who did not give their age were set to the median age of 41, and those with missing educational responses were set to the median educational level which was high school graduate.

ACE 10917227

Because most of the missing values are on income, we have estimated an equation to predict the log of income.[90] The estimated coefficients for this equation, which is based largely on demographic characteristics, are displayed in Table 5.12 below. All of the variables have the expected sign, and the equation predicts quite well for a cross-section equation as evidenced by a $\bar{R}^2$ of .46.

§ 5.9.2 Estimation of a Valuation Function

A large number of possible predictors are available for use in the valuation function we wish to estimate. A few, such as income, are obvious choices. Another obvious choice is concern about the environment; different survey questions which tap this dimension can be used to operationalize this variable in different ways. Other good candidates for predictor variables include the likelihood of visiting Alaska and answers to questions which elicit the respondent's perceptions of the characteristics of the oil spill prevention plan. Also, a strong candidate is some indicator of protest responses; this indicator could be parameterized in many ways.

We present our preferred valuation function in Table 5.13.[91] The first two parameters are the scale and location parameters based on the assumption of a Weibull survival distribution. The scale parameter is a little larger than that estimated in Table 5.7. The location parameter is quite different because we are parameterizing the original location variable as a function of the various covariates included in the equation. The first four variables, GMORE, MORE, LESS, and NODAM, are dummy variables indicating which respondents believed that the damage likely to occur in the absence of the escort ship plan would be different from that of the

[90]Three respondents gave income values which seemed implausibly high given their ages and educational attainments. Those income values were set to missing.

[91]Alternative specifications were considered and are discussed in sections 5.9.3 and 5.11 below.

ACE 10917228

Table 5.12 Prediction of Log Income

Variable	Parameter Estimate	White's Standard Error	Asymptotic T-Value
intercept	4.68144	0.96293	4.86
single	-0.37515	0.03991	-9.40
woman	-0.14114	0.03523	-4.01
white	0.15957	0.0493	3.90
age	0.04565	0.00602	7.58
age squared	-0.00049	0.00006	-8.22
live in single family home	0.12405	0.03730	3.33
no high school	-0.51808	0.06134	-8.45
some high school	-0.33948	0.05431	-6.25
some college	0.13040	0.04614	2.83
college	0.36718	0.05930	6.19
graduate school	0.57804	0.06927	8.34
have children	0.06572	0.04055	1.62
5 + children	-0.24115	0.14847	-1.62
spill news mostly from newspaper	0.11266	0.03554	3.17
log median PSU household income	0.46126	0.09897	4.66

Exxon Valdez spill. The coefficients on all four of these variables are significant at the .10 level and follow the expected rank ordering. Those respondents who think that there would be a great deal more damage, GMORE, are willing to pay quite a bit more money than the average respondent. Those who think that there will be somewhat less, but still more damage, MORE, are willing to pay less than the GMORE respondents, but still quite a bit more than the average respondent. Those who think that there would be less damage, LESS, are willing to pay less than the average respondent; and those who think that there would likely be no damage,

ACE 10917229

Table 5.13 Weibull Valuation Function

Parameter	Estimate	Standard Error	Asymptotic T-Value	Covariate Mean
Location	1.684	1.66	1.01	----
Scale	0.670	0.029	22.98	----
GMORE	0.859	0.279	3.08	0.072
MORE	0.664	0.162	4.11	0.162
LESS	-0.270	0.143	-1.88	0.228
NODAM	-0.783	0.426	-1.84	0.028
MWORK	-0.855	0.129	-6.62	0.265
NWORK	-1.735	0.196	-8.85	0.073
NAME	0.202	0.132	1.53	0.520
COASTAL	0.408	0.141	2.90	0.803
WILD	0.259	0.117	2.21	0.556
STENV	0.468	0.226	2.08	0.098
LIKVIS	0.238	0.136	1.76	0.335
LINC	0.282	0.098	2.88	10.227
WHITE	0.418	0.148	2.82	0.784
PROTEST	-1.214	0.143	-8.50	0.179
Log-Likelihood	-1198.793			

NODAM, are willing to pay a lot less.

The next two variables, MWORK and NWORK, indicate respondents who think that the plan will prevent less than a great deal of the damage, MWORK indicating those who think that the plan will prevent some of the damage and NWORK indicating those who think that the plan will not reduce the damage at all. Again, both variables are significant and of the expected negative sign. The NWORK coefficient is about twice the size of the MWORK coefficient in absolute value.

ACE 10917230

NAME is a dummy variable for those respondents who spontaneously named the Exxon Valdez spill in A-2 as one of the major environmental accidents caused by humans. As expected, this variable, which measures salience, has a positive influence on a respondent's willingness to pay.[92] COASTAL, which is a dummy variable indicating which respondents said that protecting coastal areas from oil spills was "extremely important" or "very important" in A-3f, has a large and highly significant positive influence on a respondent's willingness to pay. Likewise, WILD, which is a dummy variable for saying that the government should set aside a "very large amount" or "large amount" of new land as wilderness in A-4, has a positive effect on a respondent's willingness to pay. STENV, identification of oneself as a strong environmentalist (B-17), and LIKVIS, a dummy variable for indicating in B-11 that ones' household was "very likely" or "somewhat likely" to visit Alaska in the future, also predict that a respondent's willingness to pay will be higher.

Respondents with higher incomes, LINC, are strongly associated with having a higher willingness to pay to prevent another Exxon Valdez type spill as is being WHITE. LINC is even more strongly associated with willingness to pay using the subset of respondents who did not have their income values imputed. Respondents who spontaneously protested (PROTEST) in A-14D or A-15A that Exxon should be paying all the cost of the escort ship plan (before they were asked why they were not willing to pay in A-18) were on average willing to pay substantially less than those respondents with the same characteristics who did not protest (that Exxon should pay) by this point in the questionnaire.

Depending on a respondent's characteristics, the median willingness to pay predicted by the valuation function varies widely; the lowest predicted value for a respondent in our sample

[92]Inclusion of a dummy variable for naming the Exxon Valdez as a large oil spill in A-5a makes NAME much more significant, although the positive coefficient on the dummy for A-5a is only significant at about the 15 percent level.

ACE 10917231

is less than $1 and the highest is $441. A restriction on the valuation function that none of our respondents is willing to pay more than 10 percent of their income can not be rejected using a likelihood ratio test at the .05 level.

§ 5.9.3 Other Possible Predictor Variables

A number of other possible predictor variables might be included in the valuation function presented in Table 5.13. Many of these variables measure different aspects of the same underlying trait so that multicollinearity prevents some combinations of variables from being significant in the same equation. Still it is worth commenting on some of these other possible predictor variables. The variables relating to the damage from another spill (GMORE, MORE, LESS, and NODAM), the variables relating to the effectiveness of the spill cleanup (MWORK, NWORK), and PROTEST should always be in the model. The general question A-1f, which asks a respondent about how much money should be spent protecting the environment, is a highly significant predictor of willingness to pay until the more specific variable COASTAL (A-3f), protecting coastal areas from oil spills, is included in the equation. Those respondents believing that reducing taxes is important (A-3d) tend to be less willing to pay for the escort ship plan, although this variable is not quite significant. Paying close attention to the Exxon Valdez spill in the news (A-11) is positively related to willingness to pay but becomes insignificant when NAME is included in the equation. In B-10, those respondents who initially thought the damages from the Exxon Valdez oil spill were more serious than the damages described in the scenario were not willing to pay significantly more than those who believed the damages were about the same as described. Similarly, those respondents who initially thought the damages were less than that described were not willing to pay significantly less than those who believed the damages were about the same. Those who frequently watch TV shows about animals and

ACE 10917232

birds (B-18) are willing to pay significantly more, although this significance does not hold up when the variables in Table 5.13 are also included. Having a backpacker in the household (B-13) and having visited one of the three major national parks (B-14) both predict increased WTP amounts, as does engaging in bird watching, although only B-14 is significant at the 10 percent level. Fishing activities by the household (B-12) appear to have no influence, nor do previous trips to Alaska (although expected visits in the future do). Almost any definition of environmentalist predicts higher WTP amounts, as do most definitions of awareness of the Exxon Valdez spill. After adding income, education is still positively related to willingness to pay although the coefficient is not quite significant. Living on the West Coast is positively related to willingness to pay; but again, the coefficient is not quite significant and declines further when LIKVIS is added to the equation. Age has little effect after income is added to the equation.

§ 5.10 Adjustments to WTP Responses

The valuation function estimated above allows us to examine the effect that various adjustments would have on our median WTP estimate. The first type of adjustment corrects for respondent assumptions inconsistent with three important features of the scenario. Our information about these inconsistent assumptions comes from the respondents' answers to questions in Section B about what they had in mind when they answered the WTP questions. Ideally, respondents would have based their WTP amounts on preventing damages of the same magnitude as those caused by the Exxon Valdez spill. For those respondents who did not, one of four dummy variables in our valuation function has a value of one to represent the particular deviation from this desired perception of the same damage: GMORE, MORE, LESS, and NODAM. Setting the value of these dummy variables to zero effectively forces the perceptions

ACE 10917233

to the same damages. This adjustment reduces the estimate of the median household willingness to pay from $31 to $28.

Another possible adjustment is that for the perceived effectiveness of the escort ship plan. Ideally, all respondents would have perceived the plan as being completely effective. One of two dummy variables in the valuation function have a value of one if a respondent indicated that the plan was not completely effective: MWORK and NWORK. Setting both of these dummy variables to zero forces the perception that the plan was completely effective. This adjustment changes the estimate of the median willingness to pay from $31 to $43.

A third adjustment is that for protest responses. The problem here is how to exactly define a protest response. The most conservative definition is the one used in the variable PROTEST in the valuation function. This indicator variable takes the value of one if the respondent volunteered that Exxon or the oil companies should pay before the respondent was asked why he was against the plan (A-18) and takes the value zero otherwise. Setting PROTEST to zero forces out that consideration and changes the estimate of the median from $31 to $38.

Making all three adjustments simultaneously yields an estimate of $49 for the median household willingness to pay to prevent an Exxon Valdez type oil spill.[93]

§ 5.11 Sensitivity of the Median WTP Estimate

In this section we address the sensitivity of our median WTP amount of $31 to prevent an Exxon Valdez type oil spill to several plausible alternative ways of treating the data. We first examine what would happen to the median WTP amount if one or more of nine categories of

[93]The 95 percent confidence interval for this estimate ($48.97) is [40.71-58.90].

ACE 10917234

respondents were dropped from the estimation. We then describe how two changes in statistical procedures would affect the median WTP amount.

The first category of respondents we will look at are the 31 not-sure/not-sure WTP responses to A-15/A-17. In the previous section, these were treated as no-no responses. Such treatment is consistent with a conservative definition of protest responses. In many contingent valuation studies, these observations would have simply been dropped from the estimation. Dropping these observations raises the estimate of the median from $31 to a little more than $33.

Dropping those respondents who may have had problems in handling the survey instrument and the WTP questions in particular is not uncommon. The interviewer assessment questions in Section D can be used to identify these respondents. The most obvious group to drop are those respondents who the interviewer said gave the voting questions "not at all serious" consideration or "only slightly serious" consideration (D-11). Dropping this 5 percent of the sample raises the estimate of the median about $2. A more expansive definition also drops those who were judged to be "not cooperative" (D-4) by the interviewer, those for whom another person present during the interview had "a lot" of effect on the respondent's answers (D-7), those who were "extremely" distracted (D-8a) during the scenario presentation, those who were "not at all" interested in it (D-8b), those who were "extremely" bored by it (D-8c), or those who had difficulty understanding the WTP questions (D-9). This definition now includes a little less than 10 percent of the respondents. Dropping this group increases the estimate of the median by about $3. An even larger group of respondents can be defined by also including those "not at all informed" about the Alaska spill (D-1) and those "very distracted" (D-8b) or "slightly" bored (D-8c) during the scenario presentation. This group now includes about 18 percent of the sample and dropping them increases the median WTP estimate by almost $7. This analysis suggests that those who did not take the exercise seriously, who were distracted,

ACE 10917235

uninterested, uninformed, uncooperative, or who had difficulties understanding tended, on average, to vote against the amounts they were asked more often than the other respondents in the sample. *A priori*, one would expect these respondents to have a lower value for the good, an expectation that is confirmed by the data. For that reason, these respondents should probably not be dropped.[94]

Another group of respondents who are frequently dropped from the analysis of contingent valuation data are those who "protest" some aspect of the scenario, typically the payment mechanism. In the estimation of the valuation function, we employed a more restrictive definition of "protest" responses than is often used in contingent valuation. We counted as protest responses only those respondents who said that Exxon or the oil companies should be paying for the damage before they were asked questions A-18 or A-19, concerning why they were unwilling to pay the lowest amount asked for the spill prevention program. If we define the protest variable to include all of the respondents who said the oil companies should pay in A-18 or A-19 as well as in A-14D and A-15A, the percentage of protesters rises from 18 percent to 24 percent.[95] Adjusting for this broader definition of protest responses results in an increase in the estimate of median willingness to pay to $44 from $31, as opposed to the increase to $38 seen with the more conservative definition used in the previous section. This adjustment improves the fitted ML likelihood equation; the significance of most of the other predictor variables, income in particular, increases. An even more inclusive definition of protest responses (26 percent of the sample) includes those who are opposed to any taxes, those who

[94]The danger with respondents who are not paying much attention or who have difficulties understanding is that they may give random responses. If we had estimated that these respondents were willing to pay more on average than respondents who were interested in the good, then there would be grounds for concern.

[95]Not all respondents who volunteered that the oil companies should pay gave no-no responses. Of those classified as protest responses on the basis of A-14D and A-15A, 28 percent gave a yes response to A-15 or A-17.

ACE 10917236

think that the money will be wasted, and those who did not understand the program. This definition of protest results in an adjusted WTP estimate of $47.

Some respondents may have thought they were protecting a larger area from another Exxon Valdez type oil spill than we had intended in the scenario. Dropping the 15 percent of the sample who did not say in B-5 that the proposed plan was only protecting Prince William Sound lowers our median estimate by less than $1. This supports our analysis of the open-ended responses to B-6 which suggested that some respondents were trying to "guess" what other benefits the proposed plan might have, benefits which they did not take into consideration when giving their WTP responses.

The possibility of respondents giving an implausible fraction of their income to pay for the good being valued has long been of concern to contingent valuation researchers. A substantial fraction of the sample exhibiting such behavior is usually taken as a sign that some respondents did not take their budget constraint seriously. Often a rule of thumb, such as 5 percent of income, is used as a cut-off point; respondents willing to pay more than that amount are dropped from the sample. Such a rule is easy to implement when a respondent's actual willingness to pay is elicited. It is less obvious how to implement such a rule when the researcher has the interval within which a respondent's willingness-to-pay amount lies. Taking the ratio to income of the lower bound on the interval where the respondent's willingness-to-pay amount lies, we find that no respondents violate a 5 or 10 percent of income rule. Only three respondents violated a strict 2 percent of income rule. Dropping these respondents results in a median WTP estimate a few cents lower. Seventeen respondents violated a very strict one percent rule. Dropping these respondents results in an estimate of median willingness to pay of a little over a dollar lower.

ACE 10917237

The next issue is related to the previous one: whether our estimates of median willingness to pay are sensitive to the imposition of various upper bounds on the interval in which the willingness to pay of a respondent lies. Theoretically, willingness to pay is bounded by income.[96] Our estimation technique treats yes-yes responses as being right-censored, and most of the distributions considered allow for the possibility of infinite WTP values for right-censored intervals. Replacing the upper bound on these right-censored observations with the respondent's income results in virtually no change in the estimated median or mean willingness to pay. Indeed, a likelihood ratio test using the model in Table 5.13 does not reject, at the 10 percent level, a constraint that the upper bound on the WTP interval is 10 percent of the respondent's income. This constraint results in only a few cents difference in the median estimate and an estimate of the mean only a couple of dollars lower. Much stronger constraints, such as upper bounds on the willingness-to-pay interval of 5 percent or 2 percent of household income, also result in only a few cents change in the median; however, the estimate of the mean drops noticeably, *e.g.*, by 25 percent with the 2 percent constraint. This drop reflects, in part, the sensitivity of the mean to the distributional assumption. The median WTP estimate is, as expected, quite robust.

Finally, since a single binary discrete choice question is incentive-compatible, a logit or probit model can be fitted to the first WTP response (A-15). Fitting a probit using the log of the A-15 dollar amounts as the stimulus variable yields a constant of 1.186 (t=7.28) and a slope parameter of -.318 (t=-7.35). The resulting estimate of the median of $41.44 has a 95 percent confidence interval of [32.37-53.66]. This confidence interval overlaps with that of the confidence interval [26.85-35.59] for the Weibull median. This overlap lends support to a belief

[96]More specifically, current income plus borrowing capacity minus existing commitments and subsistence needs.

ACE 10917238

that the double-bounded dichotomous choice approach produces a small downward bias in the estimate of the median or mean in exchange for a large decrease in the size of their confidence intervals.[97]

Table 5.14 summarizes the effects of these sensitivity tests on the $31 median WTP estimate. Each change either increases the median WTP amount, sometimes substantially, or has virtually no effect on it. Thus our $31 median WTP estimate appears to be a robust lower bound.

§ 5.12 Stability and Replicability of Median WTP Estimate

The stability of the estimates of economic quantities over time is often questioned. The work for this study is a unique opportunity to look at this issue. Pilot Studies II and IV were both conducted in Dayton/Toledo, Ohio, as was a "tracking" survey conducted at the same time as the national survey. We thus have three roughly equivalent surveys spanning about a year (May 1990 - March 1991).[98] In addition, we can also compare these numbers to those from the Georgia Pilot III and the national survey.

First, let us examine the possibility that all five of these surveys yielded indistinguishable responses. The dollar amounts respondents were asked to pay differed across the five surveys. If the responses are affected by the dollar amounts, then one should find differences in the

[97]This downward bias is suggested by empirical evidence and probably results from expectations formed by the initial cost estimate given to the respondent. Some respondents who vote to pay the first amount might be willing to pay the second (higher) amount but vote against the higher amount when asked because they feel that the government would waste the extra money requested. In addition, some respondents who are not willing to pay the first amount would be willing to pay the second (lower) amount but may vote against the second amount because they believe that either the government will deliver a lower quality good than that first promised or that the probability of the government delivering the good is lower at the lower price. Both of these voting patterns would result in a downward bias. The extent of the bias depends on the degree to which the second amount is perceived by the respondent as an independent cost estimate.

[98]Only the tax payment vehicle version of the Pilot II survey is used; in that pilot the oil price payment vehicle produced significantly higher WTP estimates than the tax payment vehicle.

ACE 10917239

Table 5.14 Summary of Sensitivity Tests

A. Dropping Respondents	Percent of Sample Dropped	Change in $31 Median WTP Amount
not sure/not sure to vote questions	3.0%	+$2
not at all/only slightly serious consideration of vote questions	5.0%	+$2
negative evaluation by interviewer on one or more of six key indices (includes also respondents in previous category)	10.0%	+$3
not informed/distracted/bored (includes also respondents in previous category)	18.0%	+$7
protested some aspect of scenario (broad definition)	24.0%	+$13
protested some aspect of scenario (slightly broader definition)	26.0%	+$16
plan protecting more than PWS/not sure	15.0%	-$1
WTP more than 2% of income	0.3%	-$0
WTP more than 1% of income	2.0%	-$1
B. Statistical Adjustments		
replacing right-censored observations with respondent's income		-$0
probit using first amount		+$10

responses across the surveys. The first and simplest test for differences is whether the distributions of the responses are statistically indistinguishable across the five surveys. All of the surveys used the same sequence: two questions about whether the respondent would vote yes or no at a specified dollar amount, the amount in the second question depending on the first response. The three possible answers were the same in both questions: yes, no, and not sure. Thus, each survey yielded six possible outcomes, the distributions of which can be compared

ACE 10917240

across surveys. Frequencies for each response type appear in Table 5.15, the last column in the table giving the weighted average for the five surveys. If the response patterns are the same across surveys, then the entries should be similar across the columns.

A casual look at the five surveys suggests that the responses are quite different. This tentative conclusion is confirmed by a statistical test. The likelihood ratio test statistic for the hypothesis that the distributions of responses for the five surveys are the same is 48.73. If the null hypothesis is that the distributions are the same, this statistic is drawn from a $\chi^2_{(20)}$ distribution. Since the .01 critical value for a $\chi^2_{(20)}$ variable is 37.37, the null hypothesis of equivalent responses to the five surveys is rejected at any conventional significance level.[99]

These five surveys differed in several ways. Most obvious is the difference in the dollar amounts used in the WTP questions. This difference is summarized in Table 5.16. Note that Pilot IV and the tracking survey have the same dollar amount patterns. Using the same test procedure as above, we can test whether Pilot IV and the tracking survey have similar response patterns. The χ^2 test statistic value is 8.92 which is not significant at the 10 percent level ($\chi^2_{(5)} = 9.41$), the lowest conventional level of significance. Thus, our testing method supports the null hypothesis of equivalent distributions in a situation where it should.

The differences among the five surveys should be re-examined after accounting for the difference in dollar amounts illustrated in Table 5.16. In general, the percentages of votes for the program track closely the dollar amounts specified. This was tested more formally by estimating Weibull survival models for each of the survey data sets and then testing the null hypothesis that the distributions of willingness to pay implied by these estimates are the same

[99]Ten percent is the lowest significance level customarily used; 1 percent is the highest; 5 percent is the most frequently used.

ACE 10917241

Table 5.15 Distribution of Responses Across Surveys

	PILOT II	PILOT III	PILOT IV	TRACKING	NATIONAL	AVERAGE
A16 YES	0.3053	0.2459	0.1648	0.1962	0.2656	0.2467
A16 NO	0.1368	0.2295	0.2557	0.2679	0.1879	0.2071
A16 NS	0.0632	0.0533	0.0341	0.0670	0.0575	0.0560
A17 YES	0.0737	0.1025	0.2102	0.1196	0.0901	0.1064
A17 NO	0.4000	0.3115	0.3068	0.2967	0.3490	0.3362
A17 NS	0.0211	0.0574	0.0284	0.0526	0.0499	0.0475
TOTAL	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

across the five surveys. A summary of the estimation results for the individual surveys appears in Table 5.17.[100]

Given the variation in the estimated models across surveys, are these differences also statistically significant? The likelihood ratio test makes a comparison between the numbers in Table 5.17 and those for the pooled sample. The likelihood ratio statistic for identical WTP distributions across all surveys is 5.85. For the null hypothesis, this is a realization of a $\chi^2_{(8)}$ random variable. The probability of a value greater than this statistic is approximately 65 percent; thus the equivalence of the WTP distributions across surveys can not be rejected. In other words, the variation in the distributions of the surveys probably results from sampling variation. A more sophisticated analysis would include the covariates of Table 5.13 for each sample. Unfortunately, those covariates are not available for each sample; but a crude analysis suggested a fairly stable valuation function across the different surveys. For example, income in the rural Georgia sample (Pilot III), which has the lowest median WTP estimate, is lower than

[100]To maintain consistency, the national estimates do not include the downward reconsiderations made in C-7 and are, for that reason, higher than reported earlier in this chapter.

Table 5.16 Dollar Amounts Used in Each Survey

	PILOT II	PILOT III	PILOT IV	TRACKING	NATIONAL
A15 [A]	10	10	10	10	10
A15 [B]	30	25	30	30	30
A15 [C]	60	45	70	70	60
A15 [D]	100	90	150	150	120
A16 [A]	30	40	70	70	30
A16 [B]	60	60	100	100	60
A16 [C]	100	90	150	150	120
A16 [D]	250	150	250	250	250
A17 [A]	5	5	5	5	5
A17 [B]	10	10	10	10	10
A17 [C]	30	15	20	20	30
A17 [D]	60	25	30	30	60

income in the rest of the samples. This type of evidence further supports a conclusion that the five surveys produced consistent WTP estimates. This consistency implies that the results can be replicated and that they are stable over the time period considered.

A visual way to examine the differences and similarities between the willingness to pay distributions estimated from the five surveys is to compare the estimated survival (*i.e.*, demand) curves shown in Figure 5.2. The five curves are quite close to each other. The curve for the national survey lies in the center, the tracking survey slightly above, and the Pilot Studies IV, II and III surveys slightly below.

Figure 5.3 displays the survival curve for the national survey flanked by the upper and lower bounds of the 95 percent confidence interval. These bounds are quite close together suggesting that we have achieved reasonable precision in our estimate.

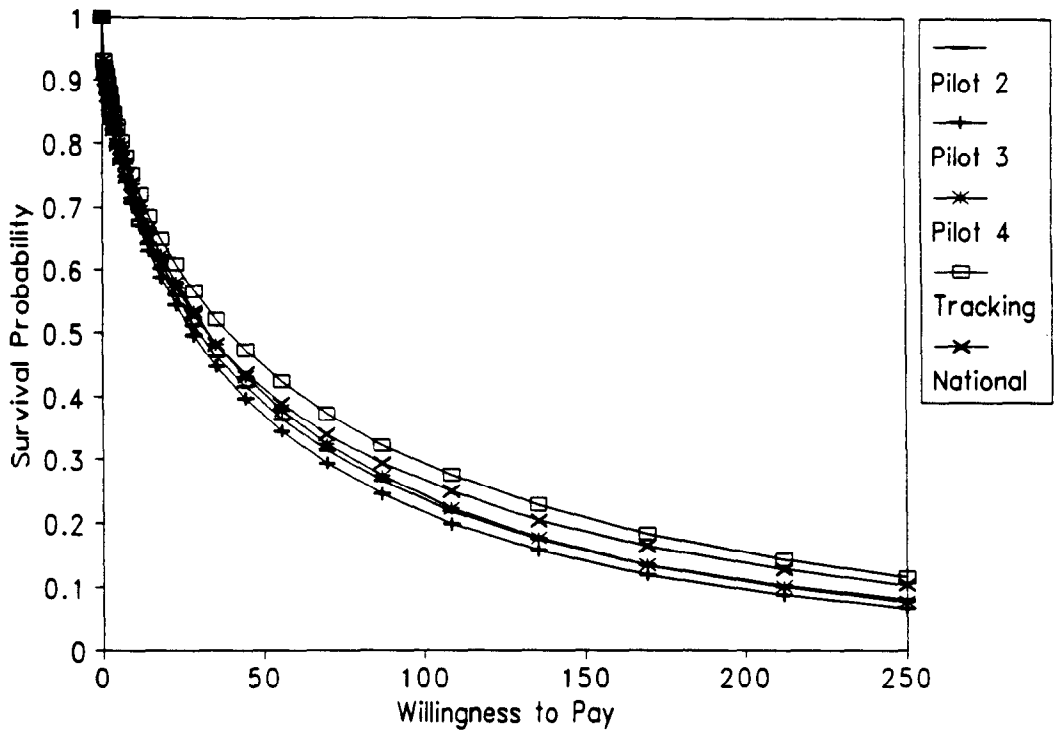
Table 5.17 Weibull Hazard Model Estimation For Each Survey

	PILOT II	PILOT III	PILOT IV	TRACKING	NATIONAL
N	95	244	175	209	1043
MEDIAN	30.11	27.97	32.87	39.20	32.69
95% CONFIDENCE INTERVAL	19.42-46.70	21.61-36.19	24.53-44.04	29.54-52.01	28.49-37.51

§ 5.13 The Measure of Damages

From a theoretical perspective, mean willingness to accept (WTA) compensation is the most appropriate measure of the services lost or disrupted by the Exxon Valdez oil spill. Median willingness to pay represents a very solid lower bound on that quantity. We are currently pursuing methods to get closer to the mean WTA. One line of attack for future research is the use of robust regression analogues for survival data which are less severe in their downweighting of extreme observations than the simple median is but which are still resistant to a small percentage of gross outliers. For example, the sum of conditional medians is likely to be a closer estimate of total willingness to pay than an estimate based on the simple median, yet still very insensitive to outliers. A second line of attack for future research is the development of a semi-parametric estimator for double-bounded interval survival data. This approach would allow us to estimate mean willingness to pay without making strong assumptions about the shape of the underlying WTP distribution. A third line of attack for future research is to adapt the theoretical formulation in Hanemann (1991) and to empirically estimate willingness to accept compensation from a WTP valuation function that includes income. The coefficient of the income variable is related to the ratio of the income elasticity to the Hicksian gross substitution elasticity. This ratio governs the difference between willingness to pay and willingness to accept. An estimate of this ratio would allow us to make inferences about mean

Figure 5.2 Estimated Survival Curves



willingness to accept.

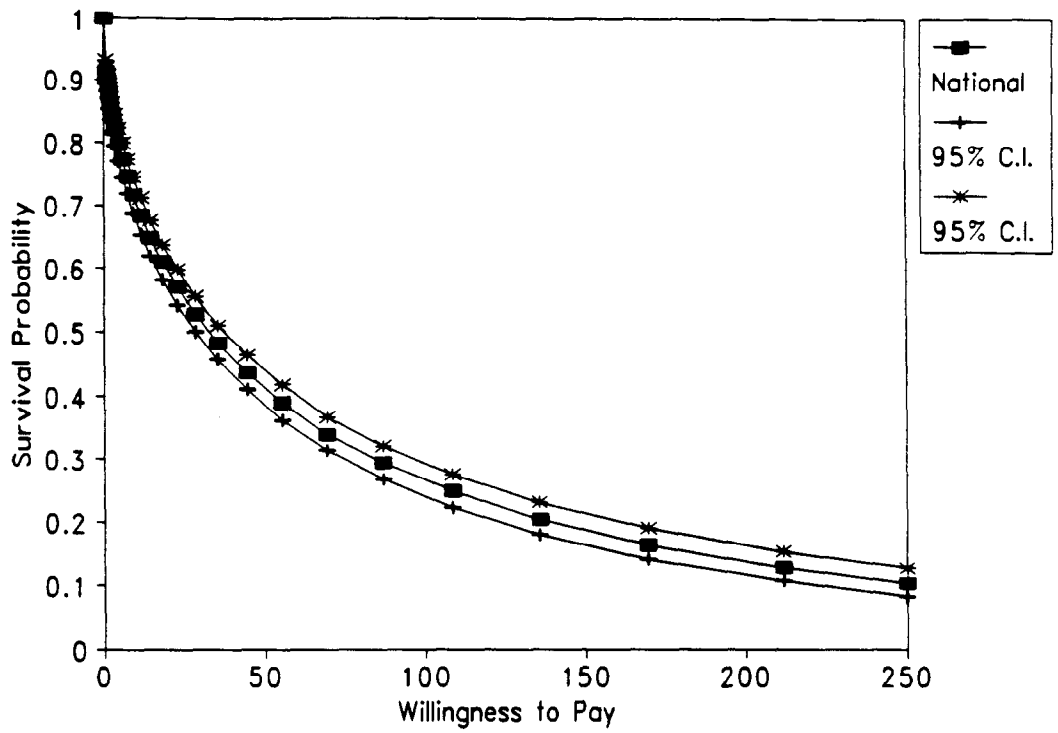
§ 5.14 Concluding Remarks

Our estimate of the lost passive use value as a result of the Exxon Valdez oil spill is 2.8 billion dollars.[101] This estimate should be regarded as a lower bound on these damages. This amount is the public's median willingness to pay to prevent another Exxon Valdez type oil spill given the scenario posed in our survey instrument. Adjusting the actual median WTP estimate for protest responses, perceptions of damages larger or smaller than the Exxon Valdez spill, and

[101]This number is obtained by multiplying the median WTP estimate of $31 by the number of English-speaking U.S. households (90,838,000). The 95 percent confidence interval for this estimate is 2.4 to 3.2 billion dollars.

ACE 10917245

Figure 5.3 95 percent Confidence Intervals for National WTP Survival Curve



for perceptions that the proposed plan would be less than completely effective results in a higher estimate.

The willingness-to-pay responses obtained in our contingent valuation surveys have been shown to be responsive to changes in the dollar thresholds used; and our results have been replicated in several independent studies during the course of a year. Furthermore, they are predicted quite well by respondent characteristics, such as income, concern about coastal oil spills, and self-identification as a strong environmentalist. The sensitivity of our damage estimate to a number of alternative ways to treat the data has been examined at some length. These alternatives either increased the damage estimate or resulted in only very small reductions.

REFERENCES

Alberini, Anna and Richard T. Carson. "Choice of Thresholds For Efficient Binary Discrete Choice Estimation" Discussion Paper 90-34, Department of Economics, University of California, San Diego, August 1990.

Bishop, Richard C. "Embedding as a Potential Problem in Measuring and Interpreting Contingent Existence Values." Unpublished paper. Department of Agricultural Economics, University of Wisconsin, Madison, 1990.

Bishop, Richard C., and Thomas A. Heberlein. "Measuring Values of Extra-Market Goods: Are Indirect Measures Biased?," *American Journal of Agricultural Economics*, vol. 61, pp. 926-930, 1979.

Carpini, Michael and S. Keeter. "Stability and Change in the U.S. Public's Knowledge of Politics," *Public Opinion Quarterly*, vol. 55, pp. 583-612, 1991.

Carson, Richard T. "Measuring the Demand for Environmental Quality," in Braden, J. and C. Kolstad, ed., (Amsterdam: North Holland, 1991).

Carson, Richard T., Mitchell, Robert Cameron. "Sequencing and Nesting in Contingent Valuation Survey." Unpublished paper. Department of Economics, University of California, San Diego, October 1992.

Carson, Richard T., M.B. Conaway, W.M. Hanemann, R.J. Kopp, K. Martin, R.C. Mitchell, S. Presser, and P.A. Ruud. "Valuing Winter Visibility Improvement in the Grand Canyon." Paper presented at the annual Association of Environmental and Resource Economists meeting, New Orleans, LA, January, 1992.

Carson, Richard T., Flores, Nicholas E., and W. Michael Hanemann. "On the Creation and Destruction of Public Goods: The Matter of Sequencing," paper presented at the Meeting of the European Association of Environmental and Resource Economists, Cracow, Poland, June 1992.

Carson, Richard T. and Dan Steinberg. "Experimental Design for Discrete Choice Voter Preference Surveys," in 1989 *Proceeding of the Survey Methodology Section of the American Statistical Association*, (Washington: American Statistical Association, 1990.)

Freeman, A, M. "Nonuse Values in Natural Resource Damage Assessment," in Kopp and Smith (eds.) *Valuing Natural Assets: The Economics of Natural Resource Damage Assessment*, (Washington D.C.: Resources for the Future, forthcoming 1993).

Hanemann, W. M. "Willingness to Pay and Willingness to Accept: How Much Can They Differ?," *American Economic Review*, Vol. 81, pp. 635-647, 1991.

ACE 10917247

Hanemann, W. Michael, John Loomis and Barbara Kanninen. "Statistical Efficiency of Double-Bounded Dichotomous Choice Contingent Valuation," *American Journal of Agricultural Economics*, vol. 73, pp. 1255-1263, 1991.

Hausman, Jerry A. "Individual Discount Rates and the Purchase and Utilization of Energy-Using Durables," *The Bell Journal of Economics*, vol. 10, pp. 33-54, 1979.

Hoehn, John P. and Alan Randall. "Too Many Proposals Pass the Benefit Cost Test," *The American Economic Review*, vol. 79, pp. 544-551, 1989.

Kahneman, Daniel and Jack L. Knetsch. "Valuing Public Goods: The Purchase of Moral Satisfaction," *Journal of Environmental Economics and Management*, vol. 22, pp. 57-70, 1992.

Kish, Leslie. *Survey Sampling*, (New York: John Wiley, 1965).

Kopp, R. J. and V. K. Smith. *Valuing Natural Assets: The Economics of Natural Resource Damage Assessment*, (Washington: Resources for the Future, forthcoming 1993).

Lawrance, Emily C. "Poverty and the Rate of Time Preference — Evidence from Panel Data," *Journal of Political Economy*, vol. 99, pp. 54-77, 1991.

Mitchell, R. C. and R. T. Carson. *Using Surveys to Value Public Goods: The Contingent Valuation Method*, (Washington D.C.: Resources for the Future, 1989).

"National Opinion Survey: Main Study — Interviewer's Manual," (Westat, Inc., January 1991).

"National Opinion Survey: Main Study — Trainer's Manual," (Westat, Inc., January, 1991).

Nelson, Wayne. *Applied Life Analysis*, (New York: John Wiley, 1982).

Randall, Alan, and John P. Hoehn. "Embedding Effects in Contingent Valuation: Implications for Natural Resource Damage Assessment." Staff paper no. 92-14. Department of Agricultural Economics, Michigan State University, 17 July, 1992.

Randall, Alan, and Warren Kriesel. "Evaluating National Policy Proposals by Contingent Valuation." In *Economic Valuation of Natural Resources: Issues, Theory, and Applications*, edited by Rebecca L. Johnson and Gary V. Johnson, 153-176. Boulder: Westview Press, 1990.

Smith, V. Kerry. "Arbitrary Values, Good Causes, and Premature Verdicts." *Journal of Environmental Economics and Management* 22(1992): 71-89.

State/Federal Natural Resource Damage Assessment Plan for the Exxon Valdez Oil Spill: Public Review Draft, report by the Trustee Council, Juneau Alaska, August 1989.

Sudman, Seymour. *Applied Sampling*, (New York: Academic, 1976)

ACE 10917248

Turnbull, Bruce W. "The Empirical Distribution Function with Arbitrarily Grouped, Censored and Truncated Data," *Journal of the Royal Statistical Society*, B38, pp. 290-295, 1976.

Walsh, Richard G., Donn M. Johnson, and John R. McKean. "Benefits Transfer of Outdoor Recreation Demand Studies: 1968-1988." *Water Resources Research* 28, no. 3 (1992): 707-713.

White, H. "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity," *Econometrics*, pp. 817-838, 1980.

ACE 10917249