THE VALUE OF PREVENTING OIL SPILL INJURIES TO NATURAL RESOURCES ALONG CALIFORNIA'S CENTRAL COAST

VOLUME I

March 31, 1996

Natural Resource Damage Assessment, Inc.
P.O. Box 910178
San Diego, CA 92191-0178
(619) 457-3330

THE VALUE OF PREVENTING OIL SPILL INJURIES TO NATURAL RESOURCES ALONG CALIFORNIA'S CENTRAL COAST

VOLUME I

Richard T. Carson

Michael B. Conaway

W. Michael Hanemann

Jon A. Krosnick

Kerry M. Martin

Donald R. McCubbin

Robert Cameron Mitchell

Stanley Presser

March 31, 1996

Table of Contents

PREFACE i

EXECUTIVE SUMMARY iv

§ 1 Introduction and Overview 1-1
- § 1.1 Policy Context 1-1
- § 1.2 Overview of the Contingent Valuation Method 1-2
- § 1.3 Organization of Report 1-4

§ 2 Scenario Identification and Survey Design 2-1
- § 2.1 Introduction 2-1
- § 2.2 Objectives of Survey Instrument Development 2-3
- § 2.3 Basic Design Features 2-5
- § 2.4 Survey Development Work 2-7
  - § 2.4.1 Phase I—Focus Groups 2-8
  - § 2.4.2 Survey Design Issues 2-11
    - § 2.4.2.1 Payment Vehicle 2-11
    - § 2.4.2.2 Number of Scenarios 2-12
    - § 2.4.2.3 Quantity and Type of Information 2-13
    - § 2.4.2.4 Types of Visual Aids 2-14
  - § 2.4.3 Phase II—In-Depth Pretest Interviews 2-15
  - § 2.4.4 Phase III—Pilot Studies 2-16
    - § 2.4.4.1 Design Points 2-18
    - § 2.4.4.2 Pilot Scenarios and Revisions 2-18
  - § 2.4.5 Main Study Survey Instrument 2-23

§ 3 Structure of the Main Study Survey Instrument 3-1
- § 3.1 Introduction 3-1
- § 3.2 Section A—Initial Questions 3-1
- § 3.3 Section A—Description of the Scenario 3-3
- § 3.4 Section B—Choice Questions 3-16
- § 3.5 Section C—Perception of Expected Harm and Program 3-19
- § 3.6 Section D—Respondent Household and Demographic Questions 3-21
- § 3.7 Section D—Reconsideration and Miscellaneous Questions 3-24
- § 3.8 Section E—Interviewer-Evaluation Questions 3-25

§ 4 Implementation of the Main Study Survey 4-1
- § 4.1 Introduction 4-1
- § 4.2 Sample Design 4-1
- § 4.3 Interviewer Training 4-2
- § 4.4 Interviewer Supervision 4-4
- § 4.5 Quality Control 4-4
- § 4.6 Validation of Interviews 4-5
- § 4.7 Sample Completion 4-5

§ 4.8 Sample Weights 4-7
§ 4.9 Data Entry 4-8

§ 5 Evaluation of Open-Ended, Vote-Assumption, Reconsideration, and Interviewer-Evaluation Questions 5-1
§ 5.1 Introduction 5-1
§ 5.2 Open-Ended Questions 5-1
§ 5.2.1 Coding of Open-Ended Questions 5-2
§ 5.2.2 Questions During Presentation of the Scenario 5-2
§ 5.2.3 Vote-Motivation Questions 5-5
§ 5.2.3.1 Interpreting Introspective Questions 5-6
§ 5.2.3.2 Reasons for Choosing to Vote *For* the Program 5-10
§ 5.2.3.3 Reasons for Choosing to Vote *Against* the Program . . 5-13
§ 5.2.3.4 Reasons for Uncertainty about Program Vote 5-15
§ 5.3 Vote-Assumption Questions 5-16
§ 5.3.1 Respondent Assumptions Regarding Expected Harm 5-17
§ 5.3.2 Respondent Assumptions Regarding Program Effectiveness . . . 5-19
§ 5.3.3 Respondent Assumptions Regarding Length of Payment 5-20
§ 5.3.4 Did Respondents Feel Pressured to Vote One Way or Another? 5-20
§ 5.4 Reconsideration Questions 5-22
§ 5.5 Interviewer-Evaluation Questions 5-24
§ 5.5.1 Were Respondents' Choices Influenced by Others? 5-27
§ 5.6 Summary 5-28

§ 6 Analysis of Choice Questions 6-1
§ 6.1 Introduction 6-1
§ 6.2 Definition of Choice Measures 6-1
§ 6.3 Statistical Framework for Analysis 6-4
§ 6.4 Turnbull Estimate of the Lower Bound on the Sample Mean 6-6
§ 6.4.1 Sensitivity of the Turnbull Estimate of the Lower Bound 6-9
§ 6.4.2 Section 6.4.2 Tests for Nay-sayers and Yea-sayers 6-12
§ 6.5 Bivariate Relationships Including Cross-Tabulations Recommended by NOAA Panel 6-16
§ 6.6 Examination of Construct Validity Using a Multivariate Approach 6-23
§ 6.6.1 Definition and Interpretation of Covariates in Choice Function 6-24
§ 6.6.2 A Cluster Analysis Interpretation 6-30
§ 6.7 Sensitivity of WTP Estimate 6-37
§ 6.8 Correction for Non-Taxpayers 6-39
§ 6.9 Summary 6-41

REFERENCES 7-1

APPENDICES [B-E contained in Volume II]

Appendix A—Main Study Survey Instrument
Appendix B—Sample Design and Execution
- Appendix B.1—Household Screener, Missed DU Procedure, and Screener Handcards
- Appendix B.2—Initial Contact Questions and Answers and Supervisor Memorandum
- Appendix B.3—Edit Form
- Appendix B.4—Validation Form
- Appendix B.5—Advance Letter
- Appendix B.6—Refusal Conversion Letters
- Appendix B.7—"Sorry I Missed You" and "No Hablo Español" Cards
- Appendix B.8—Sample Allocation and Completion Rates by PSU

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

Appendix D—Question-by-Question Verbatim Responses
Appendix E—Supplemental Statistical Discussion and Analysis
- Appendix E.1—Description of the Non-Parametric Turnbull Estimator
- Appendix E.2—Comparative Analysis of the COS Survey and EVOS Survey
- Appendix E.3—Recommended NOAA Panel Cross-Tabulations
- Appendix E.4—Supplemental Analysis Tables

List of Tables

Table 2.1 Sequence of Survey Components 2-6
Table 2.2 Phase III—Pilot Studies 2-17
Table 5.1 Respondents Who Asked Questions at A-10A as a Percentage of Respondents Who Answered A-10 5-3
Table 5.2 Respondents Who Asked Questions at A-12A as a Percentage of Respondents Who Answered A-12 5-5
Table 5.3 Reasons for Choosing to Vote *For* the Program 5-11
Table 5.4 Reasons for Choosing to Vote *Against* the Program 5-14
Table 5.5 Reasons Why *Not Sure* About Program Vote 5-16
Table 5.6 Respondents' Perceptions About Expected Harm 5-18
Table 5.7 Respondents' Perceptions About Effectiveness of Program 5-19
Table 5.8 Voting Patterns by Direction Felt Pushed 5-22
Table 5.9 Interviewer Evaluation of Respondent Reaction to Choice Elements 5-25
Table 5.10 Interviewer Evaluation of the Seriousness of Respondent Consideration of the Voting Decisions 5-27
Table 6.1a B-1 Response by B1AMT 6-2
Table 6.1b B1 Choice Measure by B1AMT 6-3
Table 6.2 B1CH Choice Measure by B1AMT 6-4
Table 6.3 Turnbull Estimate of WTP Distribution and Lower Bound on the Sample Mean: B1CH Choice Measure 6-8
Table 6.4 Summary of Sensitivity Tests 6-10
Table 6.5 Description of Sources of Information for Cross-Tabulations 6-17
Table 6.6 Cross-Tabulation Summary 6-19
Table 6.7 Multivariate Analysis of Construct Validity: Probit Estimates for B1CH Choice Valuation Function 6-25
Table 6.8 Cluster Centers 6-33
Table 6.9 Summary Statistics for Sample and Cluster 6-34
Table 6.10 Box-Cox Model Using Cluster Indicators 6-36
Table 6.11 B1CHNT Choice Measure by B1AMT 6-40
Table 6.12 Turnbull Estimate of WTP Distribution and Lower Bound on the Sample Mean: B1CHNT Choice Measure [N=1,085] 6-40

List of Figures

Figure 6.1 Illustrative Distribution With Nay-Sayers and Yea-Sayers 6-12

PREFACE

Study Background

On April 22–23, 1988, approximately 432,000 gallons of San Joaquin Valley crude oil spilled from an above-ground storage tank at a Shell Oil Company refinery into the surrounding environment including the northern reaches of the San Francisco Bay. Pursuant to the settlement ending the resulting litigation (*United States/California v. Shell Oil Co.*, No. C 89-4220, (N.D. Cal 1990)), Shell Oil Company provided funding for, among other things, studies to improve future response strategies to oil spills and ensure better restoration of resources and services affected by such spills.

Funding

The California Oil Spill (COS) Contingent Valuation Study was one of the studies funded by the settlement (Contract between the State of California, Department of Justice, and W. Michael Hanemann, Contract Number 89-2126; Sara Russell, Project Coordinator). Supplemental funding for the COS study was provided by the Office of Oil Spill Prevention and Response (OSPR) of the California Department of Fish and Game (CDFG) (Contract between the State of California, Department of Fish and Game and the University of California, San Diego, Contract Number FG 3499 OS; Pierre duVair, Project Manager) and the Damage Assessment Center of the National Oceanic and Atmospheric Administration (NOAA) (Subcontract between Industrial Economics, Inc. and Westat, Inc., NOAA Contract No. 50-DGNC-1-00007, Task Order No. 56-DGNC-4-50081; Norman Meade, Task Order Manager).

Acknowledgements

The principal investigators on the COS study team were W. Michael Hanemann of the University of California, Berkeley and Richard T. Carson of the University of California, San Diego). Robert C. Mitchell of Clark University was responsible for survey design. Other

members of the core study team included Michael B. Conaway of Natural Resource Damage Assessment, Inc., Jon A. Krosnick of Ohio State University, and Stanley Presser of the University of Maryland, College Park. Paul A. Ruud of the University of California, Berkeley and Thomas Wegge of Jones and Stokes Associates provided support for various aspects of the study. Raymond J. Kopp of Resources for the Future and John B. Loomis of Colorado State University, Fort Collins were internal reviewers. The study's project manager was Kerry M. Martin of Natural Resource Damage Assessment, Inc.

Westat, Inc. of Rockville, Maryland (Martha Berlin, program manager) administered the pilot surveys and the main survey; and the Survey Research Center, University of Maryland (Johnny Blair, program manager) developed the sample weights. Jones and Stokes Associates and Natural Resource Damage Assessment, Inc. (NRDA) provided administrative and logistical support. Patricia Wynne designed the saltwater ecosystem graphics. Philip Lau and Donald R. McCubbin of the University of California, San Diego, Craig Mohn of the University of California, Berkeley, Nick Flores of the University of Colorado, Boulder (formerly of the University of California, San Diego), and David Chapman of NOAA (formerly of the University of California, Berkeley) provided assistance with various survey design, statistical, and programming issues.

Technical assistance was provided by Larry Espinosa and Pierre duVair of CDFG and Norman Meade of NOAA. They also served both as reviewers in their own right and as conduits for review by state and federal scientists. The peer reviewers of the COS study were Richard C. Bishop of the University of Wisconsin, Madison, John P. Hoehn of Michigan State University, and Norbert Schwarz of the University of Michigan, Ann Arbor. Bishop and Hoehn are resource economists; Schwarz is a psychologist and survey researcher.

EXECUTIVE SUMMARY

The purpose of the California Oil Spill (COS) Contingent Valuation (CV) Study was "to execute and document a contingent valuation study of natural resource damages associated with offshore, coastal, or estuarine oil spills in California" (State of California, Department of Justice, Contract Number 89-2126). The COS CV study developed an estimate of per household *ex ante* economic value for a program to prevent a specified set of natural resource injuries to those species of birds and intertidal life that are consistently affected by oil spills along California's Central Coast.

The COS study team designed and implemented a CV survey following best-available practices for survey design and administration. Respondents were given the opportunity to vote *for* or *against* a government program financed by a one-time income tax surcharge on California households. The program would prevent, over the next decade, natural resource injuries from oil spills that harm wildlife and shoreline along California's Central Coast.

The per household sample estimate of total *ex ante* value obtained from the study is $76.45 (with a standard error of $3.78). The statistical approach used to obtain this estimate is a non-parametric maximum likelihood procedure developed by Turnbull (1976) which yields a lower bound on the sample mean. The above estimate includes an adjustment for respondents who did not pay California taxes; that adjustment treats the votes of non-taxpaying respondents *for* the program as votes *against* the program.

The CV survey on which this estimate is based is the culmination of an extensive program of instrument development including focus groups, in-depth pretest interviews, and pilot studies. For the main survey, Westat Inc. completed 1,085 in-person interviews with a random sample of English-speaking California households, achieving a response rate of 74.4 percent.

The qualitative and quantitative responses to the main survey were analyzed in order to

assess the validity and reliability of the measure of value and, as this measure is constructed from respondents' choices, the meaningfulness of those choices. These analyses support the validity and reliability of the results reported above.

Qualitative survey data provided evidence that respondents paid attention to the survey and took their choice seriously and that their choices reflected their perceptions of and preferences for the program. Further, responses to open-ended questions which asked respondents about their choices suggested a good understanding on the part of the respondent of what the program would accomplish and what the program would cost.

The quantitative survey data were used to examine the relationship between respondents' choices and the variables which economic theory suggests should be related to those choices. For both the pairwise approach recommended by the NOAA Panel on Contingent Valuation and a more general multivariate approach, the variables hypothesized to be positively or negatively associated with the probability that respondents voted for the program were found to be consistent with expectations. Further, those variables with clear, expected relationships to respondents' preferences were statistically significant determinants of their choices thus supporting the construct validity of the results.

In addition to its main purpose, this study has several ancillary benefits. First, it can provide a valuable input to decision-makers considering public policies related to oil spill prevention. Second, it can provide a valuable input to those considering the appropriate compensation for a future oil spill along California's coast. Third, it can serve as a valuable input to researchers considering undertaking a contingent valuation study focused on oil spill prevention.

§ 1 Introduction and Overview

§ 1.1 Policy Context

Prompted by the largest tanker spill in U.S. waters—the spillage of over 11 million gallons of crude oil by the Exxon Valdez in 1989—the United States Congress passed the Oil Pollution Act (OPA) in 1990 to help prevent the pollution of coastal waters and seas by oil.[1] At this same time, California, the country's largest producer and consumer of petrochemical products, enacted the Lempert-Keene-Seastrand Oil Spill Prevention and Response Act (OSPRA) to protect the State's 1,000 miles of coastline from oil spills.[2]

California has a long history of oil spills; while the majority of these spills were small, much of the total oil was released from a few large spills.[3] California's largest oil spill occurred in 1969 when an oil platform blow-out in the Santa Barbara Channel spilled two million gallons of crude oil; and, the most recent very large spill occurred in 1990 when the *American Trader*'s anchor punctured its hull, releasing over 390,000 gallons of oil into the coastal waters off Huntington Beach. Given the potentially devastating environmental and economic impacts of an oil spill, it is useful to know how much oil spill prevention is worth to Californians.

To determine worth (*i.e.*, monetary value), economists typically look to information about the public's preferences. The most commonly consulted source of such information is market price. In this instance, however, prevention of the harm caused by oil spills is not a commodity that an individual can readily buy or sell in the marketplace; oil spill prevention is

---

[1] OPA provisions include the establishment of tanker-free zones in environmentally sensitive areas, the use of tug escorts in certain busy tanker lanes, and the requirement that all tankers operating in U.S. waters be equipped with double-hulls by the year 2015. The act also addresses liability issues.

[2] OSPRA expanded the authority, responsibility, and duties of the Department of Fish and Game for marine oil spills, emphasizing oil spill prevention as well as contingency planning, enforcement, and response. It also created the Office of Oil Spill Prevention and Response (OSPR) within the Department of Fish and Game.

[3] The U.S. Coast Guard maintains several extensive databases on oil spills occurring in U.S. waters. *See* also the 1993 report by Mercer Management Consulting for detailed information on California's marine facilities and tank vessel traffic, oil spills, spill clean-up and damage costs, and coastline characteristics.

a non-marketed good, as it is called by economists. At present, the most commonly used approach for valuing non-marketed goods is contingent valuation.

§ 1.2 Overview of the Contingent Valuation Method

Contingent valuation (CV) is a survey-based, economic methodology that can be used to obtain data from which economic values for a wide array of economic goods, including non-marketed goods such as improved air and water quality, may be constructed.[4] CV has been used for both policy purposes and litigation by numerous state and federal government agencies. For example, CV was used by the U.S. Environmental Protection Agency to value the benefits of the Clean Water Act (Carson and Mitchell, 1993a), and by the State of Alaska in estimating the passive use losses resulting from the Exxon Valdez oil spill (Carson, *et al.*, 1992).

The concept of CV was first formally proposed by Ciriacy-Wantrup in 1947. The first reported academic application was Davis (1963), which valued recreation in Maine.[5] Since 1963, the number of published CV studies has grown rapidly, reflecting applications not only valuing environmental goods but also other types of non-marketed goods.[6] A recent bibliography lists over 2,000 CV papers and studies from over 40 countries.[7] A large part of the growth in CV can be attributed to its use by government agencies, such as the U.S. Forest Service and the U.S. Environmental Protection Agency, and by international organizations, such as the World Bank and the Inter-American Development Bank.

The theoretical foundation of CV is the same as that underlying all economic valuation,

---

[4] *See* Mitchell and Carson (1989) for a review of the theoretical and empirical basis of CV.

[5] The first reported application of this methodology appears to be a 1957 study commissioned by the National Park Service (*see* Audience Research, Inc., 1958).

[6] *See* Hanemann (1992) for a brief review of the history of CV.

[7] *See* Carson, Wright, *et al.*, 1995.

regardless of whether the valuation is based on market transactions or other non-market techniques (*e.g.*, the travel cost method used to value recreational activities). In all forms of economic valuation, the analyst constructs an economic value from an observed choice and from knowledge of the circumstances of that choice. Unlike other valuation methods, CV gives an analyst control over the choice presented and over the circumstances by which the choice is framed. In contrast, other valuation methods usually rely on recorded past choices, and the analyst must make assumptions about the features of the choice outside his or her knowledge and control. Furthermore, of hedonic pricing (*e.g.*, property value and wage models), household production function (*e.g.*, travel cost analysis and averting behavior), and CV, the three basic non-market valuation methodologies (Freeman, 1993), CV is the only methodology capable of including passive use value[8] in its estimate of total economic value.

The use of CV has been the subject of an on-going debate in the academic literature and in various policy forums.[9] A recent review of CV by an independent government panel chaired by Nobel Prize winners Kenneth Arrow and Robert Solow provides recommendations for conducting CV surveys and concludes that "CV studies can produce estimates reliable enough to be the starting point for a judicial or administrative determination of natural resource damages—including passive use values" (Arrow, *et al.*, 1993).[10]

---

[8] The term passive use value was first used in *Ohio v. U.S. Department of the Interior*, 880 F.2d 432 (D.C. Cir. 1989), and is synonymous with or inclusive of a number of other terms which have been used in the economic literature including non-use values, existence values, stewardship values, bequest values and option values. *See* Carson, Flores, and Mitchell (forthcoming) for a review of the theoretical and empirical issues related to passive use values.

[9] *See, e.g.*, the recent series of articles in the American Agricultural Economic Association's journal *Choices* (Carson, Meade, and Smith, 1993; Desvousges, *et al.*, 1993; Randall, 1993) and in the American Economic Association's Journal of Economic Perspectives (Portney, 1994; Diamond and Hausman, 1994; Hanemann, 1994).

[10] In this report, we have not directly engaged the critics of contingent valuation (*e.g.*, Hausman, 1993). However, in a number of recent papers we do address various issues raised by critics; see, *e.g.*, Carson (forthcoming), Carson, Flores, and Meade (1995), Carson, Flores, and Hanemann (1995), Carson and Mitchell (1993b; 1995), Carson, *et al.*, (1995a; 1995b), Carson, *et al.*, (forthcoming), Flores and Carson (1995), Hanemann (1994; forthcoming), and Mitchell and Carson (1995).

§ 1.3 Organization of Report

The validity and reliability of a survey depends on the quality of its design and administration. The design and administration of the COS CV survey were guided by many considerations, including those raised in Arrow, *et al.*, 1993, those derived from experience with past natural resource damage assessments and past public policy evaluations involving non-marketed goods, and other research conducted by the principal investigators and other members of the study team. These considerations are discussed in Chapters 2, 3, and 4 of this report: Chapter 2 outlines the design and development phases of the survey instrument; Chapter 3 describes the wording, format, and sequence of the final survey instrument; and Chapter 4 discusses the administration of the main study survey, including the sample design, interviewer training and supervision, quality control, completion rates, sample weights, and data entry. Chapter 5 evaluates the responses to questions pertaining to respondents' choices and perceptions of the scenario and responses to interviewer-evaluation questions. Chapter 6, the final chapter, presents the statistical framework for the analysis and the quantitative results, including the estimate of total *ex ante* economic value and its sensitivity to alternative ways of treating the data and a construct validity model which relates willingness to pay (WTP) to various respondent characteristics.

§ 2 Scenario Identification and Survey Design

§ 2.1 Introduction

The COS study team undertook this research effort in order to construct a monetary measure of the total *ex ante* economic value for preventing a specified set of natural resource injuries. There are two standard (Hicksian) monetary welfare measures used by economists: minimum willingness to accept compensation (WTA) to voluntarily give up a good; and, maximum willingness to pay (WTP) to obtain a good. These are defined in relation to an economic agent, for us, the public. Which of these two is the appropriate measure depends on who holds the relevant property rights in a particular good. If the public wishes to prevent oil spills along the coast and the oil companies have a right to spill oil along the coast, the public must purchase from the oil companies their rights to spill oil; and the maximum WTP of the public is the appropriate measure of how much the prevention of oil spills along the coast is worth to the public. But if the public has the right to an unoiled coastline which the oil companies must purchase in order to spill oil, the minimum WTA compensation of the public is the appropriate measure of how much the prevention of oil spills is worth to the public. Since oil companies do not have the right to spill oil along the coast and the public holds the property right to California's tidelands, submerged lands, and natural resources, WTA is the appropriate measure of economic value. However, using CV to measure WTA has design implications that have usually proven difficult to implement successfully;[11] hence, a choice measure based on WTP was adopted instead. The WTP measure used here represents a lower bound on the desired WTA measure, the sample mean.[12]

In CV studies, choices are posed to people in a survey; analysts then use the responses

---

[11] *See* Mitchell and Carson (1989) for a comprehensive discussion of this issue.

[12] For a theoretical discussion of WTP as a lower bound on the desired WTA measure, *see* Hanemann, 1991, and Carson, Flores, and Hanemann, 1995.

to construct monetary measures of value. Two interrelated decisions must be made in the course of designing the survey questionnaire: how to characterize the object of the choice and how to structure the context in which the choice is presented. The object of choice in CV studies consists of a change in the level of provision of a public good, such as water quality. The context of the choice is the particular sequence of words and illustrations used to convey the essential information about the choice.

The central part of this context, referred to here as the scenario, contains the information relevant to the choice respondents are asked to make. The object of choice is described in detail sufficient for respondents to understand the baseline situation and what would change and what would not change. Frequently, a plausible program to accomplish the change is described. The respondent is told how much the program would cost his or her household, how the money would be collected, and how the money would be used to effect the change. Then the respondent is given the opportunity to choose whether to pay a specified dollar amount and obtain the change in the good, or not to pay and continue with the baseline provision of the good.

In the survey instrument for this study,[13] the object of choice is characterized as the prevention of injuries to wildlife and rocky intertidal and sandy beach shorelines from oil spills along California's Central Coast over the next ten years. The context in which the choice is presented includes: a description of the cumulative harm that is expected to be caused by oil spills that affect wildlife along the Central Coast over the next decade; the presentation of a plausible program which would prevent this harm; and, an explanation of a payment mechanism whereby taxpayers would pay a one-time California income tax surcharge to set up the program and the oil companies would pay all of the costs associated with operating the program for the

[13] A reproduction of the main study survey questionnaire and graphics booklet can be found in Appendix A.

next ten years. A referendum format was used to elicit respondents' choices: respondents are asked how they would vote if an election were being held today and the program would cost their household a specified dollar amount.[14] Other questions preceding and following this choice question ask about respondent attitudes, familiarity with the affected natural resources, understanding of the assumptions underlying the scenario, and personal/household characteristics. During the interview, showcards, maps, and drawings are shown to respondents to reinforce the information presented verbally by the interviewers.

In this chapter, we discuss the development of the main study survey instrument, focusing on the development of the scenario. Section 2.2 outlines the basic objectives that guided the development process. Section 2.3 presents the basic design features adopted at the outset of this study. Section 2.4 describes the three phases of instrument development and, within this context, the resolution of several survey design issues. In the following chapter, the main study survey instrument is described in more detail along with the rationales for key aspects of the final design.

§ 2.2 Objectives of Survey Instrument Development

Throughout the development process, we were guided by the following objectives: the final instrument should be 1) consistent with economic theory, 2) comprehensible to respondents, 3) focused on the defined set of injuries, 4) plausible in regard to the scenario and choice mechanism, and 5) perceived by respondents as neutral.

The first objective was to obtain a measure of damages with a known relationship to the ideal measure suggested by economic theory.[15] Specifically, the survey instrument was designed

---

[14] The referendum elicitation format was recommended by the NOAA Panel (Arrow, *et al.*, 1993, p. 4608).

[15] *See* Mitchell and Carson (1989) for an overview of the economic concepts underlying monetary measures of value.

to enable a monetary measure of economic value to be constructed from a well-defined choice regarding a specified set of natural resource injuries.

The second objective was to use language, concepts, and questions in the survey that respondents from all educational levels and varied life experiences would comprehend. One of the primary purposes of pretesting and pilot work is to test whether or not the wording in the survey instrument meets this standard.

The third objective was to focus respondents on the described set of injuries only. This objective required carefully describing the specific set of injuries in such a way as to minimize the possibility that respondents would envision a more extensive or less extensive set of injuries. In this regard, the *ex ante* nature of the survey scenario is very important: respondents are asked to make a choice concerning a program to prevent *future* oil spills. Given the nature of oil spills and oil spill prevention programs, it is reasonable for some respondents to expect more or less injuries in the absence of the program than the injuries presented in the scenario and to expect that the prevention program may be less effective than portrayed in the scenario. Open-ended and close-ended questions were used to monitor these divergences and to assess their impact on the results.

Our fourth objective was to design a realistic choice context, *i.e.*, a plausible scenario and choice mechanism.[16] Even if respondents understand the choice, they may not consider it plausible. As noted above, we used the referendum format to elicit the respondent's choice. A large number of other design decisions made to enhance plausibility will be noted in this and the following chapter. For example, describing the State as the survey's sponsor helped enhance

[16] Many problems with contingent valuation surveys arise because respondents are asked to make choices in implausible contexts about goods that are too vaguely defined with the result that respondents may perceive their answers as unlikely to influence either provision of or payment for the good. *See* Mitchell and Carson (1989) for a general discussion of CV survey design issues and Mitchell and Carson (1995) for an overview of current CV survey design issues.

the referendum's realism, particularly since a one-time increase in State income taxes was the payment vehicle used in the survey. Further, the State's intent in conducting the survey was explained in such a way that respondents would find it reasonable to be asked about how they would vote on the program.[17]

Perceived neutrality was the fifth goal: respondents should *not* perceive the purpose of the interview as the State's promotion of a particular choice. To this end, we took care to avoid bias in the wording and sequence of material; and, we encouraged respondents to consider a number of reasons why they might want to vote for the program and why they might want to vote against. Again, we used follow-up questions to monitor our success.[18]

§ 2.3 Basic Design Features

Four basic design features were adopted at the outset of this study. The first is the use of in-person interviews.[19] In-person survey administration offers several important advantages over the standard alternatives, telephone and mail surveys. The presence of an interviewer helps to maintain respondent motivation for the approximately half-hour interview that is needed to present a sufficiently detailed scenario. The interviewer is able to pace the narrative to accommodate the respondent's needs and is able to punctuate the narrative with visual aids to more effectively communicate scenario information and maintain respondent interest.

The second design feature is the use of a questionnaire framework consisting of the sequence of *basic* survey components shown in Table 2.1.[20] In previous studies (*e.g.*, Carson,

---

[17] *See* section 3.3.

[18] *See* section 5.3.4.

[19] This mode of survey administration was recommended by the NOAA Panel (Arrow, *et al.*, 1993, p. 4608).

[20] The rationale for this sequence is discussed in Chapter 3.

et al., 1992), this framework has worked smoothly for both respondents and interviewers and has yielded reliable WTP estimates. The first component of this questionnaire framework is a series of questions that measure respondent attitudes towards a variety of government-provided public goods. Following next are background information, a description of the scenario, a vote question, vote-motivation questions, a reconsideration question, vote assumption questions, demographic and other background questions, and a second reconsideration question. The final component is a series of debriefing questions which the interviewer completes after leaving the respondent.

Table 2.1 Sequence of Survey Components

1.	Attitudinal questions
2.	Background information
3.	Description of the scenario including: • natural resource injuries, • the program to prevent some or all of the natural resource injuries, • how the program would be paid for, • reasons to vote for or against the program, and • the cost of the program to the respondent's household.
4.	Vote question
5.	Vote-motivation questions
6.	First vote-reconsideration question
7.	Vote-assumption questions
8.	Demographic and other background questions
9.	Second vote-reconsideration question
10.	Interviewer debriefing questions

For the third design feature—a prevention mechanism—we employed an escort ship program similar to that used previously in the Exxon Valdez survey (Carson, *et al.*, 1992).[21]

[21] The escort ship program described in the Exxon Valdez survey was in fact later set up in Prince William Sound, the site of the Exxon Valdez spill, and subsequently prevented a supertanker, which had lost power, from drifting into the rocks; see, *e.g.*, L.A. Times, March 26, 1993.

Our experience with this program showed that respondents found the escort ship program a plausible way to prevent harm from oil spills in a particular location and only in that location. Respondents also believed the program would be expensive to implement, a belief which helped make credible whichever of the range of tax surcharge amounts respondents received.

The fourth design feature is the referendum elicitation format: the respondent is asked whether he or she would vote for or against a program at a given tax amount. Respondents find being asked about how they would vote on such a matter plausible and are reluctant to vote to tax themselves unless they are convinced that what they would get is worth that amount. In this study, respondents are also asked to explain their vote and to answer questions about their perceptions of various aspects of the scenario.[22] Respondents are also given the opportunity to reconsider their initial vote.

In addition to these four basic design features, throughout the survey development process we used a conservative design strategy wherever competing design decisions were available.[23] Whenever the relevant facts, theory, or methodological considerations did not dictate one correct design decision, we adopted an alternative that would tend to *reduce* the likelihood of a vote *for* the program and therefore reduce the estimated value of the program.

§ 2.4 Survey Development Work

The first stage of development consisted of exploratory work, primarily through focus groups, to discern people's attitudes about oil spills in general, their beliefs about what specific effects oil spills have on the environment, their perceptions about oil spills in particular coastal areas, and their reactions to different scenario features. On the basis of early focus groups and

---

[22] Vote motivation questions and scenario debriefing questions are both recommended by the NOAA Panel (Arrow, *et al.*, 1993, pp. 4608-4609).

[23] This strategy was also one of the NOAA Panel's recommendations (Arrow, *et al.*, 1993, p. 4608).

the basic set of design features summarized above, we drafted a working survey instrument during the latter part of the focus group work.

In the next stage of in-depth pretest interviews, the initial working draft of the survey instrument was continually revised. Our aims during this second stage were to confirm our focus group findings and to see whether the verbal presentation flowed smoothly, whether respondents understood the wording and visual aids, and whether respondents regarded the choice they were asked to make as a credible one.

During the third and final development stage, we conducted a series of formal pilot studies. Pilot studies typically use more formal sampling techniques and larger samples than pretests. A larger number of interviewers are involved and the longer field period makes it possible to reach a more diverse sample. As a result, pilot studies provide a more detailed basis for evaluating how well the interview works in the field.

Peer reviewers in resource economics, psychology, and survey research reviewed the working survey instrument at each development stage.

§ 2.4.1 Phase I—Focus Groups

The design work for this study began with a series of five focus groups conducted in different locations throughout California.[24] Focus groups usually have eight to twelve participants who are led in discussion by a moderator for about two hours. The give and take in focus group discussions is an efficient way to explore what people know and think about a topic and how they might react to an interview about it. Although the people who participate are not usually a random selection from the general public, they may be selected to represent

[24] These focus groups were held in San Diego, Walnut Creek, Riverside, Sacramento, and Irvine. Later, two additional focus groups were held in San Diego and San Mateo.

a range of demographic categories including age, sex, and education. Insights discovered in one group can be checked in subsequent groups, and design decisions based on group discussions can be tested in pretest and pilot studies.

For this study, participants were called randomly from the telephone directory and recruited to represent a range of demographic categories including age and education levels. Participants came to a local focus group facility for a group discussion which lasted about two hours. To avoid self-selection bias (*i.e.*, people choosing to attend or not to attend based on their level of interest in the topic), respondents were told only that the group was being conducted to gather opinions on a current state public policy issue. Following standard focus group practice, each participant received an incentive payment for attending, the amount depending on where the group was held and the time and day of the week. A member of the research team, acting as moderator, introduced the topics and guided the discussion.

In the first focus group, we explored participants' assumptions, knowledge, and attitudes about a number of topics, including whether they were aware of past California spills and the spills' effects, whether they connected the extent and type of harm with the spill's location, whether they were more concerned about some types of spills than others, and whether they felt it was credible for the State to initiate a program to prevent the harm from spills in different coastal locations. In the course of the discussion, the participants raised a number of questions:

- How can scientists predict the number of future spills and where they will occur?
- Why wouldn't the program be state-wide?
- How would the program be paid for?
- Why should citizens (rather than oil companies) pay to prevent oil spills?
- Would the money collected actually be used for the stated purpose?

In the next four focus groups, held at different locations throughout California, we checked whether the concerns raised in the first group were representative and explored different scenario features (*e.g.*, preventing varying numbers of future oil spills, targeting specific coastal

areas and types of shoreline). Beginning with the second group, we tested draft versions of the scenario by having the group either observe or take part in a simulated interview. During each simulated interview, group members were encouraged to comment and ask questions. We increased the amount of time devoted to the interview simulations in the third and fourth groups; and the fifth group was entirely given over to this activity. The simulated interview technique was very helpful for discovering potential problems in the content and wording of the draft scenarios. By the fifth group, we had developed a complete working draft.

Several findings emerged from the focus groups. Most participants were familiar with oil spills and interested in learning more about them. Preventing future damage from oil spills along the California coast was regarded as a plausible program which participants felt comfortable favoring or opposing; and, they were reluctant to spend money on this type of program unless they were convinced that the benefits would be worthwhile. Focus group participants, not surprisingly, also seemed to have preferences for spending money on programs that would prevent spills in their own geographic area. For example, San Francisco Bay area participants were more concerned about oil spills in the San Francisco Bay whereas Southern California participants were more concerned about oil spills along Southern California's coastline.

We also learned a great deal during this part of our research about how to make the prevention program as credible as possible. The group participants raised helpful questions such as who would operate the escort ships, whether the escort ships would be needed if tankers had double-hulls, and how escort ships would prevent harm once an oil spill occurred. We used this information throughout survey development to increase the program's credibility.

The first stage of our development work helped us to assess what information was important to present during the interview and which potential sources of misunderstanding

required addressing in the instrument's wording to avoid biasing the findings. On the basis of this information and the basic set of design features summarized above, the initial working draft of the survey instrument was developed.

§ 2.4.2 Survey Design Issues

In the next two stages of our development work, we turned our attention to improving the draft survey instrument and resolving several design issues: the nature of the payment vehicle, the number of scenarios to include, the quantity and type of information to provide, and the types of visual aids to use.

§ 2.4.2.1 Payment Vehicle

The payment vehicle is the mechanism by which the program, *i.e.*, the object of choice, will be funded. The link between the payment vehicle and the program must be plausible, and it should implicitly bring the relevant budget constraints to mind. In the focus groups, we explored several possible payment vehicles, for example, a special surcharge at the gas pump and higher gas and oil prices. We ultimately settled on a one-time increase in California income taxes. The lump-sum nature of the one-time payment has the advantage of clearly focusing respondent attention on the total magnitude of the cost of the program. It has the drawback of imposing a much tighter budget constraint on some households than a program which could be paid for over multiple years. The income tax nature of the payment mechanism has the advantage of being unavoidable and highly visible for most respondents; but for those respondents who do not pay state income taxes, this payment mechanism may not have appropriate incentives for revealing information about their true willingness to pay for the program. The rationale underlying our choice of the income tax payment vehicle and the

manner in which we dealt with respondents who do not pay state income taxes is discussed at more length below and in section 6.8.

§ 2.4.2.2 Number of Scenarios

The purpose of the COS CV Study was "to execute and document a contingent valuation study of natural resource damages associated with offshore, coastal, or estuarine oil spills in California".[25] In order to effect this general goal, we first had to decide how many scenarios could be successfully incorporated into a single survey instrument without compromising the plausibility of any one scenario and, related to that, what type of spill(s) each scenario would describe. This decision had to take into account a budget that provided for a sample size of approximately 1,000 completed interviews for the final survey, the interviews having an average administration time of approximately 30 minutes. Furthermore, this decision could not compromise any of our five design objectives.

On the basis of a comprehensive review of California's past oil spills and their effects conducted at the beginning of this study, we determined that in order for the scenario(s) to be realistic (*i.e.*, to reflect the appropriate level of specificity) the description of the spill(s) should include information about the following: (1) the type of shoreline affected (sandy beach, rocky shore, saltwater marsh), (2) the geographic location of the spill (far north, San Francisco Bay, Central Coast, Greater Los Angeles, San Diego), and, (3) the amount and types of harm (primarily, the harm to birds and shoreline habitat). While an instrument with multiple scenarios was explored, for the type of good to be valued in this study, the multiple scenarios made the choices appear too artificial for most respondents. While multiple instruments with different scenarios was another possible option, the results of our development work suggested that it was

[25] State of California, Department of Justice, Contract Number 89-2126.

not feasible given our sample size.

Due to the above considerations, we decided to include a single scenario which describes injuries to a type of shoreline that is fairly representative of a vast portion of California's coastline (*i.e.*, mostly rocky shoreline with some scattered sandy beaches). This scenario was successfully tested in Pilot 2 and used in the final survey. Some of the results which influenced this decision are elaborated on below.

§ 2.4.2.3 Quantity and Type of Information

What information is necessary for respondents to perceive the scenario as plausible? The amount of information respondents can meaningfully process is limited. Presenting information beyond that threshold can lead to information overload and inattentiveness. Our goal was to include only the information that most, if not all, respondents need to make a meaningful choice. A key reason for the use of focus groups and in-depth pretest interviews during the development process is to identify this information and how best to present it. Respondents' answers to open-ended questions and their spontaneous comments help us evaluate whether we have inadvertently omitted information from the survey instrument that some respondents consider important.

In making decisions about what kinds of information to present, we were motivated by four principles. First, the scenario should describe the harm in sufficient relevant detail so that respondents understand what the program would and would not protect. For example, we describe the harm as affecting shoreline ecosystems, including saltwater plants and small animals, and as killing or injuring several types of sea and shore birds. Some focus group participants and pretest respondents thought that oil spills also harm marine mammals and fish. As any particular spill along the Central Coast may not affect these resources in large numbers, the scenario informs respondents that marine mammals and fish are usually not affected by

California oil spills.

Second, the scenario should provide information about relevant substitutes and recovery times so that respondents can place the injury in context. The survey designer may have several options by which to establish the proper context. For example, depending on the available information, when evaluating the seriousness of an injury such as 12,000 bird deaths, respondents might be told about the effect such deaths would have on the species as a whole, either in terms of extinction or endangerment or in terms of the relevant extant populations of the birds in question.

Third, other scenario elements must be described in such a way that respondents accept them as reasonable for their intended purpose. In addition to information about the good to be valued, for example, the scenario must clearly explain how the program would accomplish its purpose.

Fourth, the material presented should be based on the best available information. The Study Managers provided the set of injuries to be valued in the main study survey instrument. Although endangered bird species, marine mammals, and fish have been injured in past California oil spills, the Study Managers excluded them to value a more typical set of injuries. We also relied on OSPR—in consultation with NOAA and the United States Fish and Wildlife Service (USFWS)—to identify for us the bird species most affected by particular types of spills and to provide us with estimates of the California populations of these bird species.

§ 2.4.2.4 Types of Visual Aids

In-person interviews commonly use visual displays to provide respondents with a graphic representation of some of the material which the interviewer presents verbally. These are particularly indispensable in a survey such as this one which must describe bundles of attributes:

geographical relationships, spill effects, different species of birds, and shoreline habitats.[26] In the course of revising the instrument, we developed a set of showcards which were continually refined on the basis of respondent, interviewer, and peer reviewer feedback. The showcards are used to display lengthy lists of closed-ended answer categories and drawings and tables which illustrate various features of the scenario. The development of the display format for the key scenario information (*i.e.*, depiction of affected shoreline habitats and bird species as well as the cumulative harm) is elaborated on below.

§ 2.4.3 Phase II—In-Depth Pretest Interviews

The goal of the second phase of our work was to take the focus group insights and the working draft and ready a working survey instrument for pilot testing. Through an iterative process based on eight sets of in-depth pretest interviews, we revised and expanded the initial working draft. We then began testing the survey instrument in a field setting. Several pretests helped identify potential problems with the different survey design dimensions, for example, question wording, flow of the narrative presentation, and interviewer skip instructions.

Each set of in-depth pretest interviews consisted of five to ten interviews conducted at one or two sites by a member of the research team. The respondents were recruited by market research firms who paid them to come to their facility for this purpose. This type of interview gave us the opportunity to observe the overall flow of the interview and the way respondents reacted to its various parts. Each respondent was debriefed at the end of the interview. After each interview and before conducting the next interview in the set, we revised the instrument to address any problems with wording and flow. After the set was complete, we reviewed the

[26] The use of extensive visual aids has a long history in CV surveys valuing changes in air and water pollution (*see, e.g.*, Randall, Ives, and Eastman, 1974).

results for that set of interviews; and the instrument was revised further.

In the first field pretest, NRDA staff administered the survey to a convenience sample of 11 respondents in three San Diego neighborhoods. Based on the interviewers' debriefing comments and our analysis of the results, we made a number of changes after the first pretest to improve the communication of information (both verbally and visually) and the plausibility of the scenario and of the choice.

In our second field pretest, with a similar sample, NRDA staff interviewed 14 respondents. The interviewers reported the survey revisions improved the communication of information and ease of administration. All of the respondents seemed to understand the material; most of the respondents were either extremely or very attentive throughout the interview; and all of the respondents either gave extremely or very serious consideration to the vote questions. However, three respondents found the survey to be too long.[27] The interviewers pointed out several places in the narrative that seemed wordy or that had lost the respondent's attention. The interviewers also mentioned that some respondents were skeptical (a recurring theme in the focus groups) about some of the information presented to them, particularly the certainty expressed regarding the number of future spills (with and without the program) as well as the characterization of expected harm.

§ 2.4.4 Phase III—Pilot Studies

In light of our pretest findings, we further revised the instrument prior to field testing in pilot studies.[28] This pilot testing simulated more of the procedures that would be used in the main survey by interviewing a larger and more diverse set of respondents. Professional Westat

---

[27] The length of the survey ranged from 27 to 50 minutes with an average of 38 minutes.

[28] We occasionally conducted a few in-depth pretest interviews during the pilot research phase to help further assess proposed changes.

interviewers[29] conducted the pilot interviews in primary sampling units (PSU's)[30] selected to represent the California population. Respondents meeting specified criteria were chosen at the household level on the basis of information collected with the Westat screener interview.

Table 2.2 describes the basic features of the pilot studies. The first pilot was administered in two stages (referred to as Stages 1-A and 1-B) to give greater flexibility in testing different wording and design features. In addition to the changes noted in the table, after each field effort, we made extensive wording changes based on interviewer comments and our review of the results. We now turn to a discussion of the design features tested in each pilot as well as the design points used and the substantive revisions made between each field test.

Table 2.2 Phase III—Pilot Studies

Pilot	Sample Size	Design Points	Scenario Design Features
I-A	56	$15, $60	First scenario described expected harm from oil spills in Greater Los Angeles area and program to prevent this harm and was followed by a vote question; second scenario described expected harm from spills along Central Coast and was followed by a preference question.
I-B	112	$15, $60	First scenario same as Pilot 1-A; depending upon treatment administered, second scenario described expected harm from oil spills along Central Coast *or* the San Francisco Bay area and was followed by a preference question if voted *for* in first scenario or by a second vote question if voted *not-for*.
II	154	$10, $120	Single scenario described expected harm from oil spills along Central Coast and program to prevent this harm.

[29] Westat, Inc. is one of the country's most respected survey research firms and is often retained by government agencies to conduct their most exacting surveys (*e.g.*, National Medical Expenditure Survey, National Household Education Survey).

[30] The pilot PSU's were Los Angeles City, Los Angeles County, Alameda/Contra Costa/Marin/San Francisco/San Mateo, San Diego, Sacramento/Placer/El Dorado/Yolo, Sonoma, and Del Norte/Humboldt.

§ 2.4.4.1 Design Points

In each pilot study, respondents were randomly assigned to two equivalent subsamples which received questionnaires differing only in the dollar amount (referred to here as design point) they were told the program would cost their household. We used the information obtained for these four design points ($10, $15, $60, and $120) to help select the design points for the main survey.

§ 2.4.4.2 Pilot Scenarios and Revisions

One of the purposes of the first pilot was to further evaluate the feasibility of valuing two scenarios, more specifically, whether respondents would be able to put the first choice out of their minds and treat the second choice as totally new, whether sufficiently-detailed scenarios could be presented within the necessary time frame, and whether asking about a second scenario would compromise in any way the credibility of the choices. A new payment vehicle—a one-time increase in California income taxes—and a single show card which consolidated harm information previously shown on three different cards were also tested.

The first scenario in the Stage 1-A survey instrument described a program which would prevent the expected harm to sandy beach habitat in the Greater Los Angeles area from large oil spills.[31] Respondents were then asked how they would vote if the program cost their household a specified, one-time increase in California income taxes. Because much of the material presented in the first scenario was relevant to both scenarios (*e.g.*, the description of the program), the focus in the second scenario was primarily on the information that was unique to it. The second scenario described a program which would prevent the expected harm from

[31] Prior to the presentation of this second scenario, a series of questions about respondent assumptions regarding various aspects of the first scenario was administered.

large oil spills along the Central coast, one of the two other heavy traffic areas: the effects of a typical, large spill on a rocky shore habitat are described as well as the cumulative effects from three spills expected to occur in this area over the next ten years. The respondent is then asked the following preference question:

> If the cost to your household was the same, which program would you want the State to set up—the program to prevent large spills on rocky shorelines along the Central Coast or the program to prevent large spills on sandy beaches in the Greater Los Angeles area?

In the second stage of Pilot 1, respondents were randomly assigned to one of two treatments which differed only in the second scenario. (The first scenario in both treatments was similar to that in Stage 1-A, *i.e.*, the object of choice was characterized as a program that would prevent the expected harm from oil spills in the Greater Los Angeles area.) The second scenario in treatment 1 described a program to prevent the expected harm from oil spills along the Central Coast and the second scenario in treatment 2 described a program to prevent the expected harm from oil spills in the San Francisco Bay area. To improve our ability to rank-order by preference the different scenarios,[32] in the second scenario, the preference question for respondents who voted *against* the program in the first scenario was replaced with a choice question.

In both stages of Pilot 1, respondent reactions to the choice question following the second scenario suggested they found the second less plausible and more artificial (*i.e.*, respondents raised questions about whether the state government had carefully thought out how the program would be undertaken) than the first choice question. Though we had attempted to present the second scenario in an abbreviated format, the average interview time still exceeded our limit of 30 minutes. Furthermore, the results suggested that we were perhaps trying to convey too much

[32] Further, in the first stage of Pilot 1, those respondents who had just voted *against* the program to protect sandy beaches in the Greater Los Angeles area appeared to have some difficulties with the preference question asked in the latter part of the interview.

information too quickly.[33]

The split-sample test in the second stage was inconclusive but suggested that, overall, respondents valued preventing spills that injured rocky shorelines along the Central Coast and saltwater marshes in the San Francisco Bay area more than they valued preventing those that injured beaches in the Greater Los Angeles area. Combining the results from Stages 1-A and 1-B, overall, respondents opted for programs which protected rocky shorelines in the Central Coast or saltwater marshes in the San Francisco Bay area almost two to one over a program which would protect beaches in the Greater Los Angeles area. Further, as expected, respondents' preference with respect to protecting the different coastal areas were strongly influenced by where they lived. For example, in Stages 1-A and 1-B, respondents living in the Greater Los Angeles area were significantly more likely to prefer a program which would prevent injuries from oil spills in that area ($p=0.023$[34] and $p < 0.001$, respectively). For treatment 2 in Stage 1-B, respondents living in the San Francisco Bay area were significantly more likely to prefer a program which would prevent injuries from oil spills in that area ($p=0.020$).

On the basis of these results and discussions with the peer reviewers, we decided to focus only on a single scenario in Pilot II. The Pilot II scenario described a type of shoreline that is representative of a large portion of California's coastline: an intermingling of predominantly rocky intertidal habitat with some sandy beach habitat. In a related change, beginning with Pilot II, we dropped references to a specific number of future spills and no longer characterized spills

---

[33]Almost 70 percent of those who voted against the program described in the first scenario said they did so because they had concerns about the program and/or the payment vehicle.

[34] A p-value is a simple value used to describe test results. As a rule, hypotheses tests adopt a specific significance level (often 5 or 10 percent). This significance level specifies the probability of incorrectly rejecting a "true" null hypothesis. Thus with the selection of a 5 percent significance level, one would be accepting the chance of making mistakes 5 percent of the time this test was repeated with exactly the same hypotheses and type of information. The p-value computes what level of significance would have to be adopted to reject the hypothesis.

with adjectives like "large" or "medium". Previously our practice was to describe how many spills occur presently and how many of these the program would prevent; for example, the Stage I-B instrument said:

> During the next ten years, however, medium size spills will continue to occur on sandy beaches about as often as they have in the past, which is about once every five years.

Although some respondents consistently questioned the credibility of this type of forecast—"How can you know that two spills will be prevented?"—this type of language was needed to construct a meaningful scenario for preventing a specific number of spills.

Shifting to a single location left us free to address the respondents' skepticism about the precision of the forecast. In Pilot II, respondents were not given a number of spills but instead were told that "spills" occur causing harm and that "spills are expected to happen every few years along the Central Coast . . ." and that if the program is implemented a specific amount of harm from these spills would be prevented. Undistracted by their skepticism regarding the prediction of the number of spills, respondents' attention was focused more firmly on the harm prevented. This approach did not seem to diminish the scenario's plausibility.

Another change in the scenario was the different handling of the possibility that respondents would mistakenly believe the program would prevent the effects of a huge spill. Originally we informed respondents that an Exxon-Valdez size spill is unlikely to happen in California and that, in the event that it did, the program would not be able to handle a spill of this size. A review of respondents' reasons for voting *for* the program in Stage I-A suggested that some respondents believed they were valuing the prevention of spills larger than the program would prevent. We concluded that mentioning Exxon-Valdez type spills in the interview instigated the belief we were trying to avoid. The main study B-2 responses suggest that eliminating references to Exxon-Valdez type spills, in conjunction with other wording changes, was largely effective in avoiding this problem (see section 5.2.3.2).

Stage I-A marked a shift in payment vehicle from higher prices for oil products to a one-time increase in state income taxes. Having the oil companies pay the program's operating costs helped defuse protests that the cost of preventing oil spill damage should be borne entirely by the oil companies. To establish the propriety of the State asking the tax-payers to assume part of the burden, Pilot II respondents were told that "individual oil companies cannot legally be required to pay the cost of setting up the program." This rationale also helped make credible the one-time only nature of the tax:

> ..., all California households would pay a special one time tax for this purpose. ... Once the prevention program is set up, all the expenses of running the program for the next ten years would be paid by the oil companies.

Respondents' aversion to taxes and their desire to have the oil companies held completely responsible for the costs of the program tend to make this tax vehicle a conservative choice.[35]

Prior to the pilot studies, we used multiple showcards to communicate information on a particular topic in a step-by-step process.[36] This approach has the advantage of helping to break-up long portions of text in the survey. Influenced by Edward Tufte's work on the visual display of information,[37] we consolidated the showcards in the pilots and presented all of harm information on a single card.[38] Over the course of the pilot surveys, we developed a visual aid (two 11 by 8 ½ showcards displayed side by side)[39] for the main survey which consolidated injury information that previously appeared on three separate cards. The final result is a black

---

[35] Moreover, some respondents in the main study believed that the oil companies would pass their share of the costs on to consumers in the form of higher gas and oil prices; hence, these respondents are less likely to vote *for* the program for this reason (*see* section 6.4.1).

[36] For example, one card would show the bird species most commonly affected by a certain type of spill. Another would give a picture of the shoreline ecosystem showing the types of shoreline habitat the spill would harm. A third card would present a table summarizing the number of birds harmed by a spill and the miles of shoreline affected.

[37] *See, e.g.,* Tufte, 1983.

[38] Another impetus for this presentation revision was the need to further shorten the length of the interview.

[39] *See* Card D in Appendix A.

and white drawing[40] displaying a rocky intertidal and sandy beach ecosystem with resident bird, small animal, and saltwater plant species and a table summarizing the cumulative harm that the program would prevent over the next ten years. On this single visual aid, the respondent can see most of the important scenario information. During this same time, we also merged two maps into a single visual showing photographs of the three shoreline types, the location of the shoreline types along the California coast, and the two types of tanker routes (*i.e.*, super-tanker and barge and small tanker).[41]

Changes to the visual aids required some rewording and sequence changes in the questionnaire. The interviewer's comments during the debriefings held after each pilot survey were extremely helpful in this regard. A large number of additional changes to the instrument were made during this phase; particular attention was paid to enhancing the questionnaire's clarity, making it shorter, and improving the flow. Between and after the pilots, in-depth pretest interviews, including some that used a cognitive interview "think-aloud" approach, were used to test revisions.[42] The rationales for our final wording choices are discussed in the next chapter.

§ 2.4.5 Main Study Survey Instrument

We completed the main survey instrument in December of 1994. After additional review by the study's peer reviewers, some minor revisions were made to the instrument and the final survey instrument was tested in several in-depth pretest interviews. We then delivered the

---

[40] We used black and white rather than color because it is more conservative. The use of black and white also lowered the reproduction cost.

[41] *See* Card B in Appendix A.

[42] In addition, respondent concerns expressed in the in-depth pretest interviews about the effect of oil spills on human health led us to add a vote reconsideration question to the Pilot II instrument which asked respondents how they would vote if human health was definitely not affected.

completed survey instrument with interviewer training instructions to Westat.

§ 3 Structure of the Main Study Survey Instrument

§ 3.1 Introduction

This chapter describes section-by-section the wording, format, and sequence used in the main study survey instrument as well as the rationale underlying the key features of the final design. Unless otherwise indicated, all quoted text in this chapter is from the survey questionnaire itself. Any questionnaire text in uppercase is an interviewer instruction not read to the respondent. The complete survey instrument including a copy of the graphics booklet is provided in Appendix A.

To avoid self-selection bias from people deciding to be interviewed because of their interest in the survey's specific subject matter, prospective respondents were told that the State of California was conducting the study to "collect valuable information about how you feel the state should spend tax dollars."[43] If potential respondents asked for more information about why the survey was being conducted or what it was about, the interviewers were instructed to use only the replies provided on a laminated Q & A card.[44] For example, if a respondent asked "Why are you doing this survey?", the interviewer was to reply, "The study will provide information so State policy makers can understand how people like yourself feel the State should be spending tax dollars." If the respondent asked a question like "What is this survey about?", the interviewer was to reply "This study is about the priorities of Californians and how Californians feel the State should spend tax dollars."

§ 3.2 Section A—Initial Questions

The main interview begins with a series of questions (A-1A to A-1F) that ask the

---

[43] *See* Westat, 1995, p. 2-23.

[44] *See* Appendix B.2.

respondent how important six state-wide issues are to him or her personally.

> A-1. Let's start by talking for a moment about some issues in California. Some may not[45] be important to you, others may be.
>
> SHOW CARD A[46]
>
> First, (READ X'd ITEM). Is this issue not important at all to you personally, not too important, somewhat important, very important, or extremely important? (READ EACH ITEM, BEGINNING WITH X'd ITEM; CIRCLE ONE CODE FOR EACH; REPEAT ANSWER CATEGORIES AS NECESSARY.)[47]

The primary purpose of this series of questions and the following similar series (A-2A to A-2F described below) was to encourage respondents to think about a broad range of current policy issues as a reminder that the program described later in the interview is just one of many government-provided goods. These questions also validate the pre-interview description of the survey contained in the advance letter: "This study will collect valuable information about how you feel the State should spend tax dollars."[48] The six issues in A-1 that establish this context are improving education, reducing air pollution, maintaining local library services, reducing crime, protecting coastal areas from oil spills, and finding ways to reduce state taxes.

Question A-2 draws the respondent's attention to the fact that the State already spends money on a wide variety of programs by asking how important it is to the respondent that the State continue to spend money on six current programs.

> A-2. The State of California spends tax money on many programs for many

---

[45] Underlining was used throughout the questionnaire to indicate to the interviewer the need to emphasize certain words to help convey the passage's meaning and to hold the respondent's interest.

[46] These boxed instructions cue the interviewer to show a particular card or map, in this case, Card A. This card lists the five answer categories, from "not important at all" to "extremely important", for questions A-1 and A-2; *see* Appendix A.

[47] Following standard survey practice, the order in which the six items were asked was randomized. The interviewer was instructed to begin with the item marked "X" to minimize response order effects. Each item thus had an approximately equal chance of being asked first.

[48] *See* Appendix B-5.

different purposes. I'm going to read a list of some of these programs. For each one, I would like you to tell me how important it is to <u>you</u> that the State continue to spend money on it.

SHOW CARD A AGAIN

First, (READ X'd ITEM). (READ EACH ITEM, BEGINNING WITH X'd ITEM; CIRCLE ONE CODE FOR EACH; REPEAT ANSWER CATEGORIES AS NECESSARY.)

The six programs are providing job training for the unemployed, providing shelters for the homeless, protecting wildlife, providing lifeguards at state beaches, providing public transportation for Los Angeles, and building new state prisons.

§ 3.3 Section A—Description of the Scenario

The presentation of the scenario, which begins at this point, provides the circumstances of the choice which may be relevant to the decision the respondent is later asked to make, *i.e.*, to vote for or against the described program. Among the material presented here is a description of the three general types of shoreline found along California's coastline, the main routes oil tankers and barges take along the coastline, the types of wildlife that have been harmed in past oil spills along the Central Coast, the wildlife that are expected to be harmed by future spills in this area, and the proposed prevention program, in particular, how the program would work and how it would be paid for.

The interviewer training for this study emphasized the importance of presenting this material in a way that would maintain respondents' interest and enhance comprehension. For example, the interviewer's manual stated:

> This study may differ from most that you have conducted because the central portion of the questionnaire is a narrative you read to the respondent. In our pretests of earlier versions of this questionnaire, we found that the text reads smoothly and that most respondents find the material very interesting. Our pretests also show that reading this type of material requires a somewhat

different approach than reading regular question material.

> The narrative material about oil spills and the harm they cause is intended to provide respondents with important background information about the choice they will be asked to make in B-1. It is important that the respondent understand what you are reading so that he/she can take this information into account when answering the voting question.
>
> Because of the amount of material you will be reading, there is a risk that some respondents may become bored or disinterested. We have found that the show cards interest most respondents a great deal and help involve them in the interview. Another lesson from the pretests is that it helps to read the material in a manner that is conversational and interesting. To do this, you need to make use of effective "body language" and use a tone of voice and manner that is interesting. (Westat, 1995, pp. 4-4 and 4-5).

The interviewers were instructed via instruction boxes placed strategically in the text of the questionnaire to show the respondent at the indicated time visual aids (*e.g.*, maps, drawings, answer categories). These aids were designed and pretested to help respondents visualize important aspects of the scenario and to help them understand the material that was being read to them. For ease of administration, the 16 visual aids (11 by 8 ½) were individually enclosed in clear, plastic sheet protectors and bound together in a loose-leaf binder.

The scenario description begins with an introduction of the survey's subject matter. Respondents are first given a rationale for why they would be asked if they would be willing to pay for a new program:[49]

> These are just a <u>few</u> of the programs the State of California currently spends tax money on. Proposals are sometimes made to the State for <u>new</u> programs; but the State does not want to start any new programs <u>unless</u> taxpayers are willing to pay the additional cost for them. <u>One</u> way for the State to find out about this is to give people like you information about a program so that <u>you</u> can make up your <u>own</u> mind about it. STOP [50] Your views are useful to State decision makers in deciding what, if anything, to do about a particular situation.

[49] The text in the questionnaire (*see* Appendix A) is presented in very short paragraphs to help interviewers keep their place. In the interest of conserving space, that convention has not been maintained here.

[50] The stop sign symbol is an instruction to the interviewer to pause before continuing.

In order to avoid bias due to respondents perceiving an impression, either from the wording of the survey instrument or the interviewer's demeanor, that one response is favored over another, the respondent is then told that people responding to this type of interview have different views about the proposed program:

> In interviews of this kind, some people think that the program they are asked about is not needed; STOP others think that it is. We want to know what you think. STOP

Next, a question is asked to involve the respondent in the interview:

> A-3. Have you ever been interviewed before about whether the State should start a new program?

The final part of the introduction introduces the specific program the respondent will be asked about later in the interview. The wording emphasizes the routine nature of this type of inquiry:

> In the past, people have been asked about various types of programs. In this interview, I am going to ask you about a program that would prevent harm from oil spills off one part of the California coast. STOP

This section also serves to legitimize and normalize this survey as business as usual on the part of the State. The atmosphere of normality will minimize the impression that there is anything special about this particular good or the action contemplated by the State.

Respondents are also informed that they will be asked to decide whether the program should be implemented and asked for the reasons behind that decision.

> I will begin with important background information. Then I will ask you whether you think this particular program is worthwhile and why you feel the way you do. STOP

This paragraph is intended to encourage respondents to pay attention to the scenario and the

choice.[51]

The interview then presents the background material which begins with a description of the three basic types of shoreline found along the California coastline:

SHOW CARD B

Along the California coast, there are three different types of shoreline. [52] The areas shown here in green are mostly saltwater marshes. The areas shown in brown are mostly rocky shoreline. And, the areas in yellow are mostly sandy beaches. STOP

Card B consists of two 11 by 8 ½ cards presented concurrently on opposing pages. The top page of Card B shows a small color picture of each type of shoreline and the bottom page displays a map of the State showing the approximate locations of the three different types of shoreline along the California coastline.[53] Stretches of rocky shoreline are displayed in green, saltwater marsh in brown, and sandy beaches in yellow.[54] Also depicted on the bottom of Card B are "Super-tanker routes" which run from the top of the State down to the San Francisco Bay and the Greater Los Angeles area and a "Central Coast small tanker and barge route" joining the San Francisco Bay with the Greater Los Angeles area.[55]

Next, respondents are asked if they had visited California's shoreline in the last 12

---

[51] This technique of inducing accountability at the start an interview has been shown to promote optimal respondent effort. *See* Tetlock, 1983.

[52] The hand symbol instructed the interviewer to point to the relevant feature on the show card.

[53] Because of the size of the map and in particular the minimum width of the color band needed for the colors to be visible to the respondent, small sections of shoreline that intermingle with another type (*e.g.*, the small stretches of beaches between rocky coves) were not displayed on the card; *see* Appendix A.

[54] Each of the shoreline pictures shown in the top of Card B has a color border which matches the color used to depict that particular type of shoreline along the coast; *see* Appendix A.

[55] So as to not overburden the respondent with extraneous information, only the major coastal oil tanker and barge routes are included in this graphic.

months, and if so, which of the three types had they visited:

> A-4. Have you visited any of these three types of California shoreline in the last 12 months?
>
> A-5. And, which ones are those?

Information about the tanker and barge routes shown on Card B is then provided along with a brief description of how oil spills usually occur and what happens when they occur:

> Each year, tankers and barges carrying oil make about 3,000 trips in and out of California harbors and along the Central Coast. Large oil tankers called super-tankers deliver their cargo to storage tanks and oil refineries in the San Francisco Bay and in the Greater Los Angeles area. Small tankers and barges transport various types of refined oil back and forth along the 500 miles of coastline between San Francisco and the L.A. area.
>
> Tankers and barges occasionally run into things like underwater rocks, other ships, or pipelines, and spill some of their oil into the water. Unless the spill is very small, the oil can harm wildlife. After an oil spill, the company that caused it must pay to clean up as much oil as possible from both the water and the shoreline. STOP

The focus of the interview then narrows from California's coastline to the one area of particular interest in this survey, the Central Coast. In order to legitimate the attention given to just this one area of the coast, respondents are told that measures had already been taken to set up programs in both the San Francisco Bay and the Greater Los Angeles area.[56] Very general information about how frequently spills occur in the Central Coast is provided to help define the context for the policy choice.[57]

> Over the years, the State has taken various steps to prevent harm from oil spills. Recently, steps have been taken to set up programs to prevent harm

---

[56] In fact, a tug-escort program, which some pretest respondents were aware of, had been initiated in San Francisco Bay.

[57] Because we wanted to focus the respondent on the cumulative harm from oil spills rather than the actual number of oil spills, we were intentionally vague about the past number of spills.

from spills in the San Francisco Bay and in the L.A. area. The State wants to know whether people think this would be worth doing for the Central Coast.

STOP

As you can see here, most of the Central Coast is rocky shoreline with some scattered sandy beaches. Oil spills that harmed wildlife have happened here every few years.

The next portion of the interview describes the types of wildlife that have been harmed in past California oil spills. To enhance the credibility of this information, the respondent is told the data were provided by state and university scientists.[58] A black and white drawing (Card C) of a typical Central Coast ecosystem (*i.e.*, intertidal rocky shoreline and sandy beach habitat) is used as a conservative way to portray the wildlife affected by a spill in this area. Respondents are told that none of the wildlife shown are endangered. Furthermore, respondents are also provided with information on the size of the different bird populations in California and told that these birds also live in other States.

About mid-way through this description, the respondent is asked whether he or she is familiar with any of the five bird species (that are presented as those harmed the most by past spills) shown on the card. Question A-7 and follow-up question A-8 are asked at this particular point to enhance respondent's attention to the material.

State and university scientists were asked to provide information about the effects of these past spills.

SHOW CARD C

This drawing shows the types of wildlife that Central Coast spills have harmed. It shows five types of birds and other types of small animals that live in or near the water. Take your time to look it over.

[58] As noted in Chapter 2, the set of injuries was provided to the study team by OSPR scientists in consultation with NOAA and USFWS. The survey's description of the injury and recovery time information was reviewed and approved by Study Managers prior to the fielding of the main survey.

STOP UNTIL R IS FINISHED LOOKING AT CARD

The five birds shown here are the types of birds that past spills have harmed the most.

A-7. Do you happen to be familiar with any of these birds?

A-8. Which ones? **(CIRCLE THOSE MENTIONED)**

According to scientists, none of these birds are in any danger of becoming extinct. The number next to each bird shows how many of them live in California.[59] For example, there are about 290,000 Pacific Loons and 130,000 Western Gulls.[60] All five types of birds also live in other States.[61] STOP

Whenever oil washes up on the shoreline along the Central Coast, it harms many small animals and saltwater plants. Some are shown here. They include clams, sea stars, crabs, mussels, kelp, and other seaweed. None of these are in any danger of becoming extinct. STOP

To dispel possible misconceptions about what would be prevented, information is also provided about what resources are not consistently affected by a coastal spill:

Marine mammals—such as whales, seals, and dolphins—are not usually affected by the oil because they generally leave the area when a spill occurs.[62] Fish also leave the area and are not affected. STOP

---

[59] The number shown was either the breeding population or the peak migratory population, whichever was relevant given the species. OSPR, in consultation with NOAA and USFWS, provided the population figures as well as the five particular species that were described.

[60] Box 1 provides the interviewer with the following scripted response if the respondent asked if Western Gulls are the same as sea gulls: "Western gulls are one of a dozen types of sea gulls." The other three species shown on the Card are the Rhinoceros Auklet, Common Murre, and Brandt's Cormorant. *See* Appendix A.

[61] This statement reinforces the notion that these animals are not in danger of extinction.

[62] Box 2 provided the interviewers with the following scripted response if the respondent asked about what happens to sea otters or mentioned that he/she thought sea otters were also affected in oil spills: "Like other marine mammals, sea otters usually leave the area when a spill occurs. They have not usually been affected by past Central Coast spills." Some sea otters have been injured by oil spillage; see, *e.g.*, Mercer Management Consulting (1993). The omission of sea otters from the injury scenario results in a more conservative valuation.

Next, the interviewer describes a context that pretesting suggested was plausible for the prevention program:

> Recently, the federal government passed a new law to help reduce the number of oil spills. Ten years from now, all oil tankers and barges will be required to have two outer hulls instead of the single-hull most of them have now. Double-hulls provide much more protection against oil leaking after an accident. However, it will take ten years before all single-hulled tankers and barges can be replaced. Until then, spills are expected to happen every few years along the Central Coast, just as they have in the past, unless something is done. STOP

As some respondents may be aware of the coming change in tanker technology, this information provides a rationale for why the program is worthwhile during the next ten years.

Respondents are then told that, based on studies scientists have made of past spills along the Central Coast, a certain amount of harm to wildlife is expected in this area over the next ten years. Because some pretest respondents objected to an earlier version of the scenario which specified the number of spills that would be prevented on the grounds that the number of spills did not seem credible for one reason or another, the text focuses on the expected, cumulative harm over the next ten years if the present policy is continued.

SHOW CARD D

> This shows the total amount of harm to wildlife that state and university scientists expect will happen in the Central Coast area over the next ten years. It is based on studies scientists have made of past spills in this area. STOP In the next ten years: scientists expect that a total of about 12,000 birds of various types will be killed by oil spills off the Central Coast. In addition, about 1,000 more birds are expected to be injured but survive. Also, many small animals and saltwater plants are likely to be killed along a total of about ten miles of shoreline.[63] STOP

[63] Information about the number of miles of Central Coast shoreline was provided earlier in the interview to help place the injured area in the larger context of the uninjured area.

The harm from an oil spill is not permanent. Over time, waves and other natural processes break down the oil in the water and on the shoreline. Typically, within ten years or less after a spill, there will be as many of the affected birds as before the spill. The small animals and saltwater plants in the affected area recover somewhat faster, in about five years or less. STOP

Following the NOAA Panel's recommendation (Arrow, *et al.*, 1993, p. 4609), several different checks on respondent understanding and acceptance of the scenario are used in this survey. One such check was the following question which gives respondents the opportunity to say whether they would like to know anything more about the harm spills are expected to cause:

A-10. Is there anything more that you would like to know about the harm oil spills are expected to cause off the Central Coast over the next ten years?

Those who answered "yes" were asked an open-ended question:

A-10A. What is that?[64]

This was the first of several questions in the survey which required the interviewers to record verbatim the respondent's answer. The interviewers were instructed to record on the questionnaire what the respondent said as closely as possible, asking the respondent to pause, if necessary, so an answer or comment could be completely transcribed. The importance of accurately recording the comments in the interview, both the answers given in response to specific questions like A-10A above and spontaneous remarks made by the respondent at any other place during the interview, was emphasized in training and in the interviewer's manual (Westat, 1995, p. 4-8.)

The interviewers were also trained to use non-directive probing techniques to clarify respondents' answers to open-ended questions if the answers were vague or did not adequately

[64] The interviewers were trained in answering respondents' questions whether asked in response to specific questions (such as A-10A) or asked spontaneously (*see* Westat, 1995, p. 4-7).

answer the question.[65] Such probing is a standard survey procedure used to refocus respondent's attention on the question. It requires the interviewer to find a way to get the respondent to elaborate or think about an incomplete or irrelevant answer without influencing the content of the subsequent answer.

The next portion of the interview describes the prevention program.

> If taxpayers think it is worthwhile, the State could prevent this harm by setting up a prevention program for this part of the coast. This program would be similar to those successfully used by other states, such as the State of Washington. It would last for ten years, until all tankers and barges have double-hulls. STOP This program would do two things. First, it would help prevent oil spills from occurring. Second, if an oil spill does occur, it would prevent the oil from spreading and causing harm. STOP Here is how a Central Coast program would prevent spills from occurring.
>
> SHOW CARD E
>
> Oil spill prevention and response centers would be set up in three different locations along this part of the coast. Specially-designed ships, called escort ships, would be based at each center. An escort ship would travel alongside every tanker and barge as it sails along the Central Coast. This would help prevent spills in this area by keeping the tankers and barges from straying off-course and running into underwater rocks, other ships, or pipelines. STOP
>
> SHOW CARD F
>
> If any oil were spilled, here's how the program would keep it from spreading and causing harm. The crew of the escort ship would quickly put a large floating sea fence into the water to surround the oil.[66] To keep it from spreading in rough seas, this fence would extend 6 feet above and 8 feet below the surface of the water. Then skimmers, like the one shown here, would suck the oil from the surface of the water into storage tanks on the escort ship. Other ships would be sent from the nearest prevention and response center to

[65] Chapter 5 of the interviewer's manual is devoted entirely to probing; *see* Westat, 1995.

[66] If the respondent asked if a sea fence was the same thing as a boom, the interviewer was instructed to answer "yes" (Westat, 1995, p. 4-35).

aid in the oil recovery and clean-up. STOP

Card E shows the location of the three proposed oil spill prevention and response centers along the Central Coast; and Card F illustrates with a black and white drawing how the escort ship would keep the oil from spreading and causing harm if an oil spill were to occur.[67] A two-part, open-ended question is asked next to ascertain if the respondent wants any more information about how the program would work:

> A-12. Is there anything more that you would like to know about how this prevention program would work?
>
> A-12A. What is that?[68]

The payment vehicle used in this study is the California income tax. The payment is described as a one-time payment that would be in addition to what the respondent would normally pay in state income taxes.[69] The one-time household payment emphasizes the respondent's monetary obligation and is conservative relative to any payment plan that would allow the household to pay over the course of several years. Further, to avoid protest votes by respondents who felt that individual oil companies should pay for all of the program's costs, a feeling that was expressed frequently in our early pretesting, respondents are told that the oil companies could not legally be required to pay for setting up the program but could be required

---

[67] *See* Appendix A.

[68] In case a respondent asked about what happened to the oil, interviewers were provided with the following scripted response: "Within hours, an emergency rescue tanker would come to the scene and take the oil to storage tanks on shore" (*see* Westat, 1995, p. 4-35). In addition, if a respondent asked about how the program would be paid for or about the program's cost, the interviewer was instructed to check Box 3 and say "I will come to that in just a moment" (*see* Appendix A).

[69] If a respondent who has had taxes withheld from a paycheck asks whether this additional tax would be withheld from his/her paycheck, the interviewers were instructed to say "yes" (*see* Westat, 1995, p. 4-43).

to pay for all the expenses of running the program once it was set up.[70] This cost-sharing approach was perceived as fair by respondents in our later pretests.

> The money to pay for this program would come from both the tax-payers and the oil companies. Because individual oil companies cannot legally be required to pay the cost of setting up the program, all California households would pay a special one time tax for this purpose. This tax money would pay for providing the escort ships and setting up the three oil spill prevention and response centers along the Central Coast.
>
> Once the prevention program is set up, all the expenses of running the program for the next ten years would be paid by the oil companies. This money would come from a special fee the oil companies would be required to pay each time their tankers and barges were escorted along the Central Coast. Once the federal law goes into effect ten years from now, all tankers and barges will have double-hulls and this program would be closed down. STOP

The respondent is then told we want to know how he or she would vote if the program were on the ballot in a California referendum:

> We are interviewing people to ask how they would vote on this Central Coast prevention program if it were put on the ballot in a California election. STOP

The referendum format is the elicitation framework recommended by the NOAA Panel (Arrow, *et al.*, 1993, p. 4608).

To provide a balanced view of the voting choice and to focus the respondent on the choice after the necessarily long and detailed description of the program, possible reasons to vote for and against the program are displayed on Card G and read aloud by the interviewer:

---

[70] If, after hearing the description of the payment vehicle, the respondent expressed the view that the oil companies should pay all costs, the interviewers were instructed to check the appropriate box in Box 4 and give the following response:

> The State cannot legally force individual oil companies to pay for setting up the program. However, the oil companies can be required to pay a special fee each time of their ships is escorted along the Central Coast. These fees will pay to keep the program operating over the next ten years.

If the respondent asked about program costs, the interviewers were instructed to check the appropriate box in Box 4, and say "I will come to that in just a moment"; *see* Appendix A.

There are reasons why you might vote for setting up this program and reasons why you might vote against it.

SHOW CARD G

The reason offered to vote for the program summarizes the change in harm to the Central Coast that the respondent would receive for paying the specified amount of additional tax:

> The program would prevent harm from oil spills in the Central Coast area during the next ten years. Specifically, the program would:
>
> - prevent the deaths of about 12,000 birds as well as the deaths of many small animals and saltwater plants along about 10 miles of shoreline, and prevent 1,000 more birds from being injured. STOP

Next, the questionnaire offers three reasons that the respondent might want to vote against the program. We drew on our experience in focus groups, other pretesting activities, and past experience to select reasons that would be perceived as compelling. The first reason offered for voting against the program was that the number of birds and other wildlife that would be protected is small relative to their total numbers and that none of the species potentially affected by Central Coast oil spills are endangered:

> - the number of birds and other wildlife it would protect is small in comparison to their total numbers, and none are endangered. STOP

The second reason explicitly reminds respondents that there may be other issues that are more important to them than this one:

> Your household might prefer to spend the money to solve other social or environmental problems instead. STOP

The third reason to vote against the program is that the cost of the additional tax payment

may be more than the household wants to spend for what the program would accomplish. This wording was chosen to make the respondent feel comfortable choosing to vote either for or against the program.

> Or, the program might cost more than your household wants to spend for this.
>
> REMOVE CARD G

§ 3.4 Section B—Choice Questions

At the beginning of the next section, respondents are told how much the program would cost their household. Respondents were randomly assigned to one of five versions of the questionnaire which differed only by the tax amount—either $5, $25, $65, $120, or $220—which their household would pay if the program were to be approved.

> If the Central Coast prevention program were put into place, it would cost your household a total of $(ONE OF FIVE AMOUNTS). You would pay this as a special one time tax added to your next year's California income tax.

During the training, the interviewers were told that "household" has the same meaning as it had on the Household Screener and that if the household had more than one person who paid California income taxes, the amount would be split among the taxpayers in the household. Since some pretest respondents expressed confusion about this, the interviewers were instructed to say to any respondents that asked: "Think of this amount as the total amount for your household" (Westat, 1995, p. 4-41).

The choice question, B-1, asks the respondent to make a decision about the object of choice, *i.e.*, to vote *for* or *against* the prevention program given the specified cost.[71] To make

[71] Due primarily to time constraints and a decision to use the Turnbull lower-bound mean as the summary statistic (*see* Appendix E-1), we elected not to ask a follow-up choice question concerning the respondent's willingness to pay for the program.

the decision as realistic and immediate as possible, the choice is posed in terms of an election being held *today*:

> B-1. If an election were being held today, and the total cost to your household for this program would be $(ONE OF FIVE TAX AMOUNTS), would you vote for the program or would you vote against it?[72]

The respondent is offered two explicit options: *for* and *against*.[73] Some respondents do not respond to either of those two explicit options. In order to avoid pressuring respondents who give some other response, the interviewers were trained to accept other responses, such as "don't know" or "not sure," as valid answers for this question and to record them as "not sure" without additional probing (Westat, 1995, p. 4-43). If the respondent said something like "I don't vote," "I'm not registered," or "I'm not a citizen," the interviewers were instructed to say "If you did vote, would you vote for the program or against it?" (Westat, 1995, p. 4-43). The interviewers were also trained to handle any attempts by the respondent to ask them what they (the interviewer) think about the question by saying:

> 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; which way would you vote if the program cost your household $____? (Westat, 1995, p. 4-43)

Depending on the response to B-1, respondents are asked one of three follow-up, vote-motivation questions to ascertain the motivation for their votes, a procedure specifically recommended by the NOAA Panel (Arrow, *et al.*, 1993, p. 4609). After the motivation

---

[72] The interviewer's manual warned that a few respondents may look to them for cues as to how they should vote at this point and that,

> in fact, it doesn't matter at all how people vote; what does matter is that their answers represent their own best judgement about their actual willingness to pay based on the information provided to them in the interview and their preferences about how they should spend their money. This is why you should use a neutral tone and an unhurried manner. (*See* Westat, 1995, p. 4-41).

[73] This dichotomous choice (*for* or *against*) for a particular level of taxation is recommended by the NOAA Panel (Arrow, *et al.*, 1993, p. 4612). *See* also Chapter 4 of Carson, *et al.* (1994) for a discussion of the NOAA Panel's recommendation to also include an explicit "would not vote" answer category.

question, those who voted *for* are also asked a question which allows them to reconsider their vote *for* the program. Those who do not vote *for* the program are not given an opportunity to reconsider their votes at this point and are only asked the motivation question.

The motivation question for respondents who voted *for* (B-2) is worded to assess as specifically as possible, without leading the respondent to give any particular answer, the reason that the respondent's household would be willing to pay the stated tax amount.

> B-2. People have different reasons for voting for the Central Coast prevention program. What would the program do that made you willing to pay for it? **(PROBE: Was there something specific that the program would do that made you wiling to pay for it?)**

If the respondent's answer was vague or non-responsive, the interviewer was directed to probe for a more specific answer. As noted earlier, in order to clarify vague answers, the interviewers were trained to use neutral and nondirective probes whenever respondents gave answers that were not responsive to the particular question or were vague or unspecific. The interviewers were provided with a card listing a set of standard probes and were trained when and how to probe.[74] In addition, in question B-2, a standard non-leading probe was furnished for vague responses.

After the motivation question (B-2), respondents who voted *for* the program were offered a chance to change their vote from *for* to *against*.[75] Question B-3 explicitly deals with the concern of some respondents that oil spills may affect human health. In addition, respondents who wanted to change their *for* votes for any other reason could also take advantage of this opportunity to reconsider.

> B-3. Occasionally, people vote for the program because they are concerned

---

[74] *See* Chapter 5 of the interviewer's training manual (Westat, 1995).

[75] This is the first of two reconsideration opportunities which are offered to every respondent who voted *for* the program.

> that oil spills may somehow harm human health. Suppose human health was definitely not affected and the program would only prevent harm to birds, small animals, and saltwater plants. Would you vote for or against the program if it cost your household $(B-1 TAX AMOUNT)?

For respondents who voted *against* the program at question B-1, the motivation is asked in Question B-4:

> B-4. Did you vote against the program because it isn't worth that much money to you, or because it would be somewhat difficult for your household to pay that amount, or because of some other reason?

This way of asking about the respondent's motivation to vote *against* the program alleviates the discomfort some respondents might feel at revealing motivations they find unpleasant or too personal (*e.g.*, they couldn't afford to pay for the program). The interviewers had specific instructions to record verbatim all "other" answers to B-4. As a conservative measure, the option to reconsider was not offered at this point to those who voted against the program to avoid the possibility that they would feel pressured to vote *for* as a result. They are given a chance to reconsider later at Question D-15.

Finally, respondents who answered "don't know" or "not sure" at B-1 were asked B-5:

> Could you tell me why you aren't sure about how you would vote? **(PROBE)**

If the respondent's answer was vague, the interviewer was prompted to use a probe such as the following: "Can you tell what it is about the program that made you unsure" (Westat, 1995, p 4-47)?

§ 3.5 Section C—Perception of Expected Harm and Program

The questions in Section C ask the respondents what they had in mind or had assumed about various aspects of the scenario when they voted on the program. If the respondents asked

why they were being asked these types of questions, the interviewers were instructed to say the following:

> We find that some people have different ideas about this. It is important for us to know what you had in mind (Westat, 1995, p. 4-49).

Despite the potential difficulty some respondents have with this type of question, the answers nevertheless help assess whether scenario features are accepted by respondents when they voted.[76]

Question C-1 asks about respondents' perception of the described extent of the harm from Central Coast oil spills over the next 10 years:

> Please think back to a few moments ago when I asked you whether you would vote for or against the program.
>
> SHOW CARD H[77]
>
> C-1. At that time, did you think the harm from oil spills in the Central Coast over the next ten years would be about the same as that shown here, or a lot more or a lot less?

The next question in this sequence asks how seriously respondents' considered the amount of harm shown on Card H:

> SHOW CARD I[78]
>
> C-2. How serious did you consider this amount of harm to be . . .[79] not serious at all, not too serious, somewhat serious, very serious, or extremely serious?

---

[76] The answers to these questions are another type of check on respondent understanding and acceptance of the scenario (Arrow, *et al.*, 1993, p. 4609).

[77] Card H repeated the injury information shown on Card D; *see* Appendix A.

[78] Card I listed the answer categories for C-2 and was shown concurrently with Card H; *see* Appendix A.

[79] Text in the questionnaire that follows ". . ." was presented as lower case answer categories (recall that interviewers were instructed not to read anything that appeared in upper case) rather than as part of the actual question text; a NOT SURE answer category was also included but was not read aloud; *see* Appendix A.

Questions C-3 and C-4 explore respondents' assumptions about the program's effectiveness and the payment period.

SHOW CARD J[80]

C-3. Did it seem to you that the prevention program I told you about would be completely effective at preventing harm from Central Coast oil spills, mostly effective, somewhat effective, not too effective, or not effective at all?

C-4. When you decided how to vote, did you think your household would have to pay the special tax for the program for one year or for more than one year?

In order to learn about how the respondent perceived the interview, the final sequence of questions in Section C asks whether the respondent felt pushed to vote one way or the other by the interview. For those who felt they had been pushed one way or the other, two follow-up questions, C-6 and C-6A, asked which direction they felt pushed and why they thought they had been pushed.

C-5. Thinking about <u>everything</u> I have told you during this interview, <u>overall</u> did it try to push you to vote one way or another, <u>or</u> did it let you make up your <u>own</u> mind about which way to vote?

C-6. <u>Which</u> way did you think it pushed you?

C-6A. <u>What</u> was it that made you think that? (PROBE: "Can you be more specific about what you have in mind?" "Anything else?")

§ 3.6 Section D—Respondent Household and Demographic Questions

The next part of the interview shifts from retrospection about the choice decision to the collection of demographic and other information about the respondent and the respondent's household. The first seven questions in Section D ask about various types of household recreational activities:

Now I would like to ask you a few questions about your household's

[80] Card J lists the answer categories for question C-3; *see* Appendix A.

recreational activities.

D-1. Has anyone in your household ever driven along the Central Coast on Highway 1, the coast highway?

Those who said *yes* to D-1, were asked the next question:

D-2. And, was this in the last five years?

Everyone was asked D-3 through D-7.

D-3. In the past five years, has anyone in your household gone saltwater boating or saltwater fishing?

D-4. Does anyone in your household like to identify different species of birds?[81]

D-5. During this past summer, about how many times did people in your household go to beaches anywhere along the California coast . . . never, once or twice, three to ten times, or more than ten times?

SHOW CARD K[82]

D-6. How often do you personally watch television programs about animals and birds in the wild . . . very often, often, sometimes, rarely, or never?[83]

The next question asks whether the respondent perceives himself or herself as an environmentalist.[84]

SHOW CARD L[85]

[81] If the respondent asked the interviewer what was meant by "identify different species of birds", the interviewer was instructed to provide a standard survey reply: "Whatever it means to you" (Westat, 1995, pp. 4-55 and 4-57).

[82] Card K listed the answer categories for D-6; *see* Appendix A.

[83] If the respondent asked the interviewer what was meant by "animals and birds in the wild", the interviewer was instructed to provide a standard survey reply: "Whatever it means to you" (Westat, 1995, pp. 4-55 and 4-57).

[84] Here, as elsewhere, if the respondent asked what was meant by "environmentalist", the interviewer was instructed to respond, "whatever it means to you" (Westat, 1995, pp. 4-55 and 4-57).

[85] Card L lists the answer categories for D-7; *see* Appendix A.

D-7. Do you think of yourself as an . . . environmental activist, strong environmentalist, somewhat strong environmentalist, not particularly strong environmentalist, or not an environmentalist at all?

The respondent's household and personal characteristics are the subject of the next sequence of questions.

Now, just a few questions about your background.

D-8. First, in total, how many years have you lived in California?

D-9. In what month and year were you born?

D-10. What is the highest year of school you completed or the highest degree you received?[86]

The next several questions deal with household finances.

D-11. Currently, how many adults in your household, including yourself, work for pay?

The interviewer then asks the respondent to select the correct category of household income from a list of income ranges on a card.[87]

SHOW CARD M[88]

D-12. I'd like you to think about the income received last year by everyone in your household. Adding together all income for everyone in your household, which letter on this card best describes your household's total income for last year—1994—before taxes? Please include wages or salaries, social security or other retirement income, child support, public assistance, business income, and all other income.

Those respondents who reported incomes in the lowest two categories in D-12 are asked in

[86] The interviewer coded the respondent's answer to this question into one of eleven categories ranging from "through 8th grade" to "doctorate degree"; *see* Appendix A.

[87] This is a standard survey research device.

[88] Card M lists 11 income categories ranging from "under $10,000" to $100,000 or more"; *see* Appendix A.

Question D-13 if they paid any California income taxes in 1994.

> D-13. Did anyone in your household pay any California income taxes for last year, 1994, by having taxes withheld from wages, retirement income, or other money received, or has anyone in your household sent, or intend to send, tax money for last year to the State with a tax form?

Question D-14 asks all respondents about their perception of their financial situation in the near future.

> SHOW CARD N[89]

> D-14. When you look ahead to the next few years, do you see your personal financial situation getting . . . much better, a little better, staying about the same, getting a little worse, or much worse?

§ 3.7 Section D—Reconsideration and Miscellaneous Questions

At question D-15, the respondent was offered a final opportunity to change his or her vote:

> Now that we're almost at the end of the interview and you have been able to think a bit more about the situation, I'd like to give you a chance to review your answer to the voting question. You were asked if you would vote for or against a program that would prevent the harm that I showed you earlier on this card.

> SHOW CARD O[90]

> D-15. If an election were being held today, would you vote for the program or against the program if it cost your household a one-time tax payment of $(B-1 TAX AMOUNT)?

For those who had earlier voted *for* the program at Question B-1, Question D-15 is a second reconsideration question. For those who had voted *against* the program at B-1, D-15 is the first reconsideration question.

---

[89] Card N lists the answer categories for D-14; *see* Appendix A.

[90] Card O repeats the same injury information shown on Cards D and H; *see* Appendix A.

The last two questions in this section, D-16 and D-17, inquire which method of paying for environmental programs the respondent would prefer and, given that various State agencies were identified as the survey's sponsor, how much trust the respondent places in state government.

> D-16. There are different ways for people to pay for new programs to protect the environment. One way is for the government to pay the cost. This will raise everyone's taxes. STOP Another way is for businesses to pay the cost. This will make prices go up for everyone. STOP If you had to choose, would you prefer to pay for new environmental programs . . . through higher taxes, or through higher prices?
>
> SHOW CARD P[91]
>
> D-17. Generally speaking, how much confidence do you have in the California state government? Would you say . . . a great deal, some, hardly any, or none?

Question D-17 is the last question that is administered to the respondent. At this point, the interviewer is asked to thank the respondent for his or her cooperation.[92]

§ 3.8 Section E—Interviewer-Evaluation Questions

The interviewers were asked to give their impressions about certain aspects of the interview by filling out the questions in Section E. All questions in this section were answered by the interviewers after they left the respondents' homes. The interviewers were told the following:

> Section E of the questionnaire is designed to provide us with your feedback.

---

[91] Card P lists the answer categories to D-17; *see* Appendix A.

[92] In those cases in which the interviewer had elected to administer the abbreviated version of the Household Screener, the interviewer is instructed to say: "I have just a few more questions I need to ask about the other adults in your household. Let me verify that there are (number from AS-1) people 18 or older living in this household." (*see* Appendices A and B.1). The interviewer then asked questions S-3 through S-5, the questions in the enumeration table, and S-12.

It is important that you complete this section as soon as possible after you have conducted the interview so that it is still fresh in your mind. It is crucial to the evaluation effort that you answer all applicable questions as fully as possible. We are very interested in hearing about your experiences with the materials, procedures, and questions. You, as an interviewer, are our most important source of information for evaluating how well these worked. (Westat, 1995, p. 4-69).

Questions E-1 to E-3 ask the interviewer to record (by observation) the respondent's sex and race and to transfer the respondent's zip code to the boxes provided:

PLEASE NOTE THE FOLLOWING ABOUT THE RESPONDENT BY CIRCLING THE NUMBER OF THE CORRECT RESPONSE:

E-1. SEX

E-2. RACE

E-3. TRANSFER THE RESPONDENT'S ZIP CODE FROM THE ADDRESS LABEL ON THE CALL RECORD FOLDER: ☐☐☐☐☐

Questions E-4 to E-9A ask the interviewer to give his or her impression about how attentive to the interview the respondent had been and what difficulties the respondent appeared to have had.

E-4. What was the reaction of the respondent as you read **A-3 through A-13?** (This is the descriptive material including the maps and drawings).[93]

a. How distracted was the respondent?
b. How attentive was the respondent?
c. How interested was the respondent?

E-5. Did the respondent say anything suggesting that he or she had any difficulty understanding either the harm caused by Central Coast oil spills or the prevention program?

E-5A. Please describe the difficulties. [OPEN-ENDED]

[93] The scale includes the following categories: extremely, very, somewhat, slightly, not at all, and not sure; *see* Appendix A.

E-6. Did the respondent have any difficulty understanding the voting question, B-1?

E-6A. Please describe the difficulties. [OPEN-ENDED]

E-7. When you asked B-1, did you feel the respondent was impatient to finish the interview?

E-7A. How impatient was the respondent?

E-8. How serious was the consideration the respondent gave to the decision about how to vote?

E-9. Not counting you and the respondent, was anyone age 13 or older present when the respondent voted?

E-9A. Do you think the other person(s) affected how the respondent voted or don't you know?

The final question is an open-ended invitation to the interviewer to make any other comments about the interview. No specific instructions were provided other than "record any other comments you think would be useful here about the interview" (Westat, 1995, p. 4-75).

E-10. Do you have any other comments about this interview?

§ 4 Implementation of the Main Study Survey

§ 4.1 Introduction

Westat's implementation of the main study survey consisted of several distinct steps. A random sample of dwelling units (DU's) was drawn. An interviewer's training manual was prepared, and Westat's interviewers attended a two-day training session. During the 14 weeks of main survey data collection, the interviewers were supervised by regional field supervisors and a project director. As interviews were completed, Westat conducted quality control edits and validation interviews. At the end of the data collection, sample weights were constructed by the Survey Research Center at the University of Maryland. Finally, data sets containing the responses to both the close-ended and open-ended questions were prepared. This chapter provides a detailed discussion of each of these steps.

§ 4.2 Sample Design

The multi-stage probability sample drawn for the COS main study represents the population of English-speaking Californians, age 18 or older, living in private residences they own or rent (or to whose rent or mortgage they contribute). A sample of dwelling units was drawn from areas randomly selected in the fall of 1991 for inclusion in the National Adult Literacy Survey (NALS).[94] There were three stages of sample selection: primary sampling units (PSU's; one or several counties); segments (Census blocks or block groups) within PSU's; and dwelling units within segments.

COS used the same sample PSU's selected for NALS.[95] Within these PSU's, 167

---

[94] NALS was conducted by the Educational Testing Service and Westat for the U.S. Department of Education. We would like to thank the Adult Education Unit of the California Department of Education and the California State Library for granting Westat permission to use the NALS listing for use in this study.

[95] The ten PSU's were Sacramento, San Francisco/Oakland, Riverside/San Bernadino, Los Angeles City, Los Angeles County/Long Beach, Anaheim/Santa Ana, San Diego, Del Norte/Humboldt, Sonoma, and Bakersfield.

segments were chosen with probabilities proportionate to size. A random selection of dwelling units (DU's) was then drawn in the third and final stage. The number of DU's selected (1,747) took into account expected rates of occupancy (some DU's would be vacant), eligibility (some would not contain any English-speaking adults), and response (some would not agree to the request for an interview) so as to yield approximately 1,000 interviews. The selected DU's were randomly assigned to one of the five tax amount treatments described in Chapter 3.[96] Within each DU, one respondent was chosen from the eligible members of the household using a random sampling table generated prior to the respondent selection.

Interviewing took place over a 14 week period from January 30 to May 3, 1995. At the beginning of this period, interviewers followed a standard probability procedure[97] to sample DU's not included on the original listing of DU's.[98] The procedure corrected, in an unbiased manner, for DU's missed by the NALS listers as well as for any units constructed after the listing was conducted. Fifty additional DU's were added to the sample as a result of implementing this quality control procedure; thus, the total sample consisted of 1,797 dwelling units.

§ 4.3 Interviewer Training

The 33 professional interviewers participating in the study attended a two day in-person training session on January 28-29, 1995, in San Diego, California. All of the interviewers had prior household interviewing experience. The training session was conducted by the study's

---

[96] *See* section 3.4.

[97] A copy of the Missed DU Procedure form can be found in Appendix B.1.

[98] The Census Bureau's definition of a dwelling unit was used: a house, an apartment, or group of rooms or a single room occupied as separate living quarters, that is, the occupants do not live and eat with any other person in the structure, and there is direct access from the outside or through a common hall or area (Westat, 1995, pp. 2-10, 2-11).

project director, Naomi Everett, assisted by the two Regional Field Supervisors. Martha Berlin, Westat's vice-president of survey operations, also attended. The study was referred to as the State Policy Study (SPS); and the interviewers were told the study was being sponsored by various California state agencies.[99]

The interviewers had been given an initial set of study materials to read before attending training. The training consisted of scripted lectures, exercises, interactive small group sessions, and role-playing sessions (using prepared scripts) in which one trainee took the role of the interviewer and another, the role of the respondent.

After introductory remarks, the first morning of the training began with an overview of the study and the role of the interviewer. Next, a demonstration interview was conducted to illustrate how to administer the main interview. That was followed by a lecture on locating selected DU's and screening procedures. The rest of the first day was spent reviewing techniques for administering the main interview, practicing these techniques in an interactive session, and reviewing probing techniques.

The second day of training began with a further review of screening procedures and an interactive session on administering the screener. That was followed by a lecture on how to avoid refusals in the field, a review of the probing exercise (a self-administered test the interviewers were given to complete the night before), and a discussion of administrative procedures. The remainder of day two was devoted to role-playing sessions. Finally, the interviewers were told to practice administering the survey instrument at home before attempting interviews at sampled DU's.

---

[99] *See* Appendix B.2.

§ 4.4 Interviewer Supervision

All interviewers reported to one of two regional field supervisors, who in turn reported to the project director. Supervisors were responsible for conferring with the interviewers regularly, reporting on and managing progress in the field, performing quality control edits, and validating interviews.

Interviewers reported to their supervisor by telephone at least once a week. The discussions included a case-by-case review, feedback on quality and production, and strategy for the remaining assignment. In addition, interviewers participated in conference calls with other interviewers and supervisors to share strategies on obtaining agreement to be interviewed.

Supervisors entered data on interviewing production, time, and expenses into a machine-readable file that was designed to generate weekly field status reports. Supervisors also reported weekly by telephone to the project director on survey progress, case assignments, and refusal conversion strategies.

§ 4.5 Quality Control

Completed interviews were sent by field interviewers directly to their respective supervisor. Upon receipt, the supervisors were responsible for quality control, including an edit of a percentage of each interviewer's questionnaires for completeness and accuracy in following procedures and skip patterns.[100] The questionnaires were then sent to Westat's home office for further editing. Results of the edits were discussed with the interviewers.

The edits uncovered 4 cases in which the respondent selection within the household was carried out improperly. None of these cases were included in the final data set; they were counted as "other non-response" to the main interview.

---

[100] The form used for editing is shown in Appendix B.3.

§ 4.6 Validation of Interviews

Validation of ten percent of each interviewer's work (with the one exception described below) was conducted either by telephone by the supervisors or, for non-response cases and non-telephone households, by an in-person visit by another interviewer working in the same PSU. The cases to be validated were randomly pre-selected in advance of the field work and were performed using the form shown in Appendix B.4.

All the selected cases were successfully validated except one. Age discrepancies were noted on one non-response case; and as a result, all of the cases of the interviewer who handled that case were reviewed. This review revealed a total of two interviews where the interviewer altered the ages of the household members listed on the screener so that the selected respondent would be the person who was home at the time and available to be interviewed. Since information about the interview topic was conveyed to the household in the course of confirming that a problem existed with these two cases, the possibility of self-selection bias contaminated these two cases. In order to avoid self-selection bias arising from knowledge of the survey topic, no attempt was made to interview the correct respondent in these households. Instead, the two cases were treated as "other non-response" to the main interview. The remaining cases of that interviewer's assignment were validated successfully.

§ 4.7 Sample Completion

The household screener was designed to collect information on household composition and to select a main interview respondent randomly from the eligible members of the household.[101] The disposition of the total sample of 1,797 cases follows:

---

[101] A copy of the SPS Household Screener can be found in Appendix B.1. The other field materials (*e.g.*, advance letter, refusal conversion letters, "Sorry I Missed You" card, "No Hablo Español" card) used by the interviewers and, when appropriate, mailed to the selected dwelling units can be found in Appendices B.5 to B.7.

Screeners Completed	1,311	
Not an Occupied Dwelling Unit	219	
Language Barriers	31	
Refusals	175	*233 eligibility unknown*
Physical/Mental Handicaps	12	
Never Reached	39	
Other Non-response[102]	7	
Other Ineligibles[103]	3	
TOTAL	1,797	

The final disposition of the 1,311 cases in which a screener was completed and a main survey respondent selected follows:

Main Interviews Completed	1,085
Refusals	114
Language Barriers	65
Physical/Mental Handicaps	9
Never Reached	28
Other Ineligibles[104]	3
Other Non-response[105]	7
TOTAL	1,311

The response rate is the number of completed main interviews divided by the number of eligible households. Since the sample was intended to represent households of English-speaking Californians, age 18 or older, living in private residences they own or rent, the ineligible cases are not included in the response rate calculations. Computing the response rate involves making an assumption about the eligibility of the 233 occupied dwelling units that were non-responses to the Screener for other than language reasons. The standard survey practice is to assume the

[102] This category includes 5 cases where the household moved before the Screener could be administered; 1 case where the only resident of the DU was deceased; and 1 case where the renters were temporarily residing in the DU while their permanent home in California was under-going repairs.

[103] This category consists of 2 DU's that were occupied on a temporary basis by visitors who resided outside of California; and 1 case where the only resident of the DU was living away from home in a substance abuse treatment facility.

[104] This category consists of 3 cases where the residences were being occupied on a temporary basis by out-of-state residents.

[105] This category consists of four cases in which respondent selection within the household was carried out improperly, two failed validation cases, and one case where the selected respondent's whereabouts were unknown to the other residents at the DU.

same proportion of these cases was eligible as for those cases whose eligibility was determined during screener administration (Council of American Survey Research Organizations, 1982), which in this instance is 92.4 percent.[106] Using this approach, the response rate is 74.4 percent: 1,085 divided by [1,797 - (219 + 31 + 18 + 3 + 65 + 3)].[107] That is, in calculating the response rate, we removed from the denominator all the ineligible cases: the 219 addresses that were not occupied DU's, the 31 language barriers on the screener, the 18 cases representing our best estimate of the ineligibles among the screener non-responses,[108] 3 other ineligibles on the screener, 65 language barriers on the main interview, and 3 other ineligibles on the main interview.[109]

§ 4.8 Sample Weights

As information about the survey topic was not provided to individuals until after the main interview began, willingness to pay for the prevention program could not have directly affected whether or not a household responded. It is possible, however, that other characteristics, *e.g.*, household location, were related to responding/non-responding status. Thus the composition of the interviewed sample could differ from that of the total sample initially chosen. In addition, the composition of the sample initially chosen might differ from the total population by chance. These problems were addressed through weights constructed by the University of Maryland Survey Research Center.

---

[106] Of the 1,345 occupied DU's whose status was determined (completed screeners, screener language barriers, and screener other ineligibles), 1,243 (or 92.42 percent) were members of the eligible population (1,345 less the screener language barriers, main interview language barriers, and screener/main interview other ineligibles).

[107] Response rates by PSU are provided in Appendix B.8.

[108] This estimate is obtained by multiplying (1 - 0.9242) by 233.

[109] The lower-bound estimate of the response rate, assuming that all of the 233 unknown eligibility cases were in fact eligible, is 73.5 percent.

In order to correct for nonresponse, each interview was multiplied by the ratio of the number of interviews that would have been completed in the segment had there been no nonresponse divided by the actual number of completed interviews in the segment.[110] For example, if all sampled households in suburban San Francisco segments responded to the interview, but only half of those in the city of San Francisco did, each of the city cases would be weighted at 2.0; and each of the suburban cases at 1.0.

In order to correct for possible chance variations that remained after the nonresponse adjustment was applied, the weights were poststratified with respect to type of dwelling unit, one household-level variable that was measured in both our survey and in the 1990 Census. Disproportionately more interviews were conducted in single-family detached units, therefore these interviews were weighted down; other types of housing structures were correspondingly weighted up.

§ 4.9 Data Entry

After review at the home office, Westat sent the completed questionnaires to NRDA for data entry. Upon receipt at NRDA, staff logged the questionnaires and entered the numeric and verbatim responses into a machine-readable file. NRDA staff corrected skip pattern violations and recording errors. A computer program was designed to assign a value of "9" (categorized as *not ascertained* in the Appendix C.2 tables) to those questions that the respondent was not asked but should have been asked. A value of "." (*i.e.*, missing) was assigned to those questions that the respondent was asked but should not have been asked. The cleaned data set was used in the analysis reported elsewhere in this report. Tabulations, both unweighted and weighted, of the cleaned data set are found in Appendix C.2 and the verbatim responses in Appendix D.

[110] This ratio was capped at 2.6 to reduce the impact on the sampling error due to variance in the weights.

§ 5 Evaluation of Open-Ended, Vote-Assumption, Reconsideration, and Interviewer-Evaluation Questions

§ 5.1 Introduction

This chapter examines the qualitative measures related to the reliability of the choice data.[111] In section 5.2, open-ended responses to certain open-ended questions are examined. The primary focus is on the open-ended, follow-up questions recommended by the NOAA Panel that ask respondents to explain their reason(s) for voting *for* or *against* the Central Coast prevention program or why they were not sure about how they would vote. Section 5.2 also examines responses to questions embedded in the presentation of the scenario. In section 5.3, responses to the vote-assumption questions, which suggest how some respondents perceived various aspects of the scenario and whether they felt pressured to vote a particular way, are examined. Section 5.4 explores the characteristics of those respondents who changed their initial vote when they were given opportunities to reconsider. In section 5.5, interviewer assessments of various aspects of the interview are examined; and finally, section 5.6 presents a summary of our qualitative analysis.

§ 5.2 Open-Ended Questions

Several questions are examined in this and the following sections; for example, whether respondents paid attention and took the choice opportunity seriously, whether respondents' decisions reflect their perceptions of the object of choice and their preferences for it, and whether extraneous factors influenced respondents' choices. The concern for the meaningfulness of the respondents' voting choices underlies these questions, a concern that motivated both the NOAA Panel's methodological recommendations and the design and administration of this study.

[111] The quantitative measures are presented in Chapter 6.

§ 5.2.1 Coding of Open-Ended Questions

Verbatim responses to the open-ended questions A-10A, A-12A, B-2, B-4, and B-5 were coded into discrete categories.[112] We devised coding categories for each question after examining the typical comments made in response to these questions in the first batch of questionnaires.[113] Next, each individual, discrete thought was assigned to one of the categories developed for that question.[114] Because a verbatim response may consist of several different thoughts and the percentaging base used is the number of respondents who responded to the question, the sum of the percentages for the coded answers totals more than 100 percent in the tables below.

§ 5.2.2 Questions During Presentation of the Scenario

The first section of the interview contained two sets of paired questions, A-10/A-10A and A-12/A-12A, that asked respondents whether they would like to know anything more about the material that had just been presented. Those who said *yes* to the first questions of the pairs were asked the follow-up question: "What it that?" The responses to these questions can provide insight into how respondents' reacted as the information on the expected harm and the prevention program was presented and into whether respondents had difficulty understanding this part of the presentation.

Question A-10 was asked after the description of the expected harm. When asked if there was "anything more that you would like to know about the harm oil spills are expected to cause

---

[112] The only other open-ended question in the survey was C-6A which asked the respondent why he or she felt pushed to vote one way or another. Because only the 5.5% of the sample who felt pushed were asked this question, the responses to C-6A were not coded into explicit categories. *See* section 5.3.4 below.

[113] *See* Appendix C.1.

[114] In many cases, the open-ended responses were recorded in short sentences; hence the task of identifying individual, separate thoughts was relatively straightforward. *See* Appendix D.

off the Central Coast over the next ten years?," 85.5 percent said *no*. The A-10A verbatim responses of the 142 respondents who wanted more information are summarized in Table 5.1.[115] The percentaging base used in Table 5.1 is the total number of respondents who answered the preceding filter question, A-10.

Table 5.1 Respondents Who Asked Questions at A-10A as a Percentage of Respondents Who Answered A-10 [N=1085][a]

WANTED MORE INFORMATION ABOUT . . .		
HARM OR POSSIBLE HARM		8.30%
Some aspect of harm not described.	4.70%	
Possible impacts on humans/human health/drinking water.	2.86%	
Possible impacts on economy/recreational use/tourism.	0.74%	
LEGITIMACY		2.12%
Validity/source of information.	1.75%	
Sponsor of survey.	0.37%	
PREVIOUS/SUBSEQUENT ITEMS		3.41%
Some aspect of harm already described.	1.84%	
Cost of program/who pays for program?	0.74%	
Can anything be done to prevent harm?	0.83%	
OTHER[b]		1.84%
DID NOT REQUEST MORE INFORMATION		85.53%

[a] Percentaging base is 1,085 (*i.e.*, the number of respondents who answered the preceding filter question, A-10). Percentages total more 100 as multiple responses were allowed.

[b] Only includes those for whom no other category was coded.

A little over 8 percent of the sample requested more detailed information on a specific aspect of the harm (*e.g.*, affected wildlife, recovery times) or questioned the possible harm to humans or the economy, including recreational use of the affected shoreline.[116] About 2.1

[115] The verbatim responses given at A-10A that were comments rather than queries about other types of information are excluded from Table 5.1. Fifteen A-10A responses were in this "comments" category.

[116] The issue of possible effects on human health is addressed later in the main interview at question B-3; *see* section 5.4.

percent of the sample asked about the validity and source of the information presented and, related to that, the sponsor of the survey. Another 3.4 percent inquired about an aspect of the harm that had already been described or an element of the scenario that was to be presented in a later section of the questionnaire (*i.e.*, questions about cost or what could be done to prevent the harm).[117]

Question A-12 followed the description of the prevention program. When asked, "Is there anything more you would like to know about how this prevention program would work?", 78.6 percent said *no*. The verbatim responses of the 218 respondents who wanted more information are summarized in Table 5.2.[118] As in the preceding table, the percentaging base is the total number of respondents who answered the preceding filter question, A-12.

As one might expect, the most commonly asked question concerned the cost of the program or who would pay for it; based on the number of respondents who answered A-12, 8.3 percent asked about this at A-12A. This information was presented in the very next section of the interview. The second type of query, asked by 8.3 percent of the sample, was either about various aspects of how the program would work (*i.e.*, a feature already described or additional information about a specific feature of the program that was not discussed) or about other alternative programs. The third type, asked by 1.7 percent, related to the program's effectiveness and often involved expressions of skepticism about whether the program would actually work. Last, 1.7 percent asked about what happens to oil after it is spilled or about what happens to oil after it is recovered by the escort ships.[119]

---

[117] If the interviewer had already covered the answer to the question in material presented earlier, they were instructed to go back to that material and re-read the pertinent information. If the question would be covered later in the questionnaire, the interviewers were instructed to say: *"I will come to that shortly"* (Westat, 1995, p. 4-7).

[118] The verbatim responses given at A-12A that were comments rather than queries are excluded from Table 5.2. Thirteen A-12A responses were in this "comments" category.

[119] If the respondent asked about what happens to the oil, the interviewers were instructed to answer: *"Within hours, an emergency tanker would come to the scene and take the oil to storage tanks on the shore"* (Westat, 1995, p. 4-35).

Table 5.2 Respondents Who Asked Questions at A-12A as a Percentage of Respondents Who Answered A-12 [N=1,085][a]

WANTED MORE INFORMATION ABOUT . . .		
COST OF PROGRAM/WHO PAYS?		8.29%
PREVENTION PROGRAM		8.29%
Additional information about specific feature of program (*e.g.*, escort ship, sea fence).	5.16%	
Feature of program already described.	1.57%	
Possible alternative solutions/programs.	1.57%	
PROGRAM EFFECTIVENESS		1.66%
WHAT HAPPENS TO SPILLED/RECOVERED OIL?		1.66%
OTHER[b]		2.67%
DID NOT REQUEST MORE INFORMATION		78.71%

[a] Percentaging base is 1,085 (*i.e.*, the number of respondents who answered the preceding filter question, A-12). Percentages total more than 100 percent as multiple responses were allowed.

[b] Only includes those for whom no other category was coded.

§ 5.2.3 Vote-Motivation Questions

Voting choices should usually relate to (1) what respondents perceive the program to offer, (2) the cost of the program to the respondent's household, and (3) the respondent's preferences for environmental amenities of this sort.[120] Particularly relevant to the first two items is the set of respondents' answers to the open-ended introspection questions immediately following the choice questions. These follow-up questions in Section B asked respondents to explain why they voted as they did.[121] B-2 was asked of those who said they would vote for the program; B-4 was asked of those who said they would not vote for the program; and B-5 of those who said they were not sure about how they would vote. Before examining the responses to these introspective questions, a framework for interpreting the responses is needed.

[120] Important sources of evidence for these relationships are presented in Chapter 6, including the construct validity equation and sensitivity of respondents' choice to the dollar amounts they would pay.

[121] The NOAA Panel recommended the use of such questions and the careful coding of responses to show the types of responses (Arrow, *et al.*, 1993, p. 4609).

§ 5.2.3.1 Interpreting Introspective Questions

Introspective questions can provide useful information; however, such questions and their responses have inherent limitations that require caution. For reasons partially discussed below, the recorded responses to such questions need not be a complete and fully accurate accounting of all factors that shaped people's judgments. This limitation imposes practical constraints on some uses of such questions.

First, a number of psychological studies indicate that although people generally have good insights into their likes and dislikes and can report their attitudes well, the process *underlying* their thinking is more difficult to elicit (Nisbett and Wilson, 1977). Further, studies suggest that people are sometimes unaware of factors that shape their judgments (*e.g.*, Nisbett and Wilson, 1977) or sometimes forget what factors influenced judgments made previously (Lodge, McGraw, and Stroh, 1989). Therefore, when asked why they voted as they did in this survey, some respondents may fail to mention considerations that shaped their voting decisions and may mention factors that were not significant causes.

Second, in typical everyday conversations, speakers tend to conform to certain conversational norms or conventions (Grice, 1975). These conventions seem to affect respondent answers in survey questionnaires. For example, one such conversational convention is the implicit understanding that one should not waste time saying things that one's conversational partner already knows.[122] In this survey, when explaining their decisions to vote in favor of the program, respondents sometimes made general statements such as "the program will help the environment" which were probably intended to be understood in the context of the information that was shared with the interviewer. Most respondents would have

[122]Grice provides for this convention in his discussion of two maxims: "[d]o not make your contribution [to the conversation] more informative than is required" and "[b]e relevant".

considered it inappropriate to repeat to the interviewer all the details of the scenario since they would have presumed that the interviewers were well aware of all the details of the prevention program.

To explore whether this convention of conversational parsimony affects responses to our introspective questions, we tested whether respondents would explain the reasons for their vote in greater detail if they believed their conversational partner, the interviewer, did not necessarily share all their knowledge about the program. One way to test for such an effect is to have a different interviewer ask the *vote-motivation* questions, rather than the interviewer who had originally explained the prevention program to the respondent.

We retained a marketing research facility to recruit adult respondents (from their database of San Diego area residents) to be interviewed at their facility. The respondents were not told the specific topic of the interview in advance, only that the discussion would focus on a current state issue. During these sessions, respondents were randomly assigned to one of two conditions: (1) a slightly reduced version of the main study questionnaire[123] administered by a single interviewer (the "control" condition) or (2) this same version administered by two interviewers (the "treatment" condition). At the beginning of the treatment interview, the respondent was told:

> This session will be in two parts. First, I'll give you some information about a specific situation. When I've finished, someone else will come in to ask you a few additional questions.[124]

Then the first interviewer administered the questionnaire up to and including B-1, the vote

---

[123] This version used the same questionnaire up to and including the vote question as that used in the main study. Some of the questions asked after the vote-motivation questions in the main study questionnaire, including the reconsideration questions and household recreational questions, were not asked to shorten the interview session.

[124] In order to keep the treatment protocol as close to the control as possible, for purposes of comparability we did not offer any additional explanation about why two interviewers were involved or about how knowledgeable the second interviewer was about the program.

question (the tax amount was $25 for all respondents). After the first interviewer left the room, a second interviewer immediately came in to complete the questionnaire, beginning with the appropriate *vote-motivation* question.

The interviewers were all experienced in-person interviewers familiar with administering the main study instrument. They were randomly assigned to roles in the individual sessions, and none were informed about the particular hypothesis being tested. The interviews were conducted in a research facility that permitted continuous quality monitoring of the data collection process via one-way mirrors and sound systems.

As expected, given that the control and treatment interviews were administered identically up to and including the vote question, there was no significant difference between the percentage voting *for* the program in the two conditions ($t=0.86$; $p=0.39$). Sixty-nine percent of 94 respondents in the control condition voted *for* the program, compared to 61 percent of the 84 respondents in the treatment condition.[125]

Our analysis focused on respondents voting *for* the program. In both conditions, respondents voting *for* the program were asked first:

> B-2. People have <u>different</u> reasons for voting for the program that was described to you. What would the program <u>do</u> that made <u>you</u> willing to pay for it?

After respondents appeared to be completely finished answering B-2, in both conditions every respondent was asked:

> B-2A. Could you be any more specific about what the program would accomplish that made you decide to vote for it?

To obtain an accurate record of everything the respondent said to these two open-ended

[125]Furthermore, the basic demographics of the control and treatment conditions were statistically equivalent: gender ($t=-0.53$; $p=0.60$), education ($t=0.49$; $p=0.62$), age ($t=-0.49$; $p=0.62$), and income ($t=1.10$; $p=0.27$).

questions, the interviews were tape-recorded; and the answers to questions B-2 and B-2A were transcribed.

The simplest approach to testing our hypothesis is to examine the numbers of words respondents used to explain their voting decisions. If changing the interviewers led respondents to believe that the interviewer asking the *explanation* questions might not share all of their knowledge about the prevention program, then respondents in the treatment condition should have provided more extensive explanations. And indeed, respondents in the treatment condition used 33 percent more words than the control group. Respondents in the control condition explained their votes with an average of 149 words ($n=61$), whereas the treatment condition respondents did so using an average of 198 words ($n=50$). These means are significantly different from one another: $t=2.35$, which has a p-value of 0.021 for a two-sided test and a p-value of 0.010 for the one-sided test suggested by the conversational parsimony hypothesis.[126]

This result supports the notion that respondents in CV studies such as this one are less complete in explaining their vote decisions to the interviewer than they would to someone else because of the conversational context in which those explanations are solicited. Because respondents believe they share information about the prevention program with the interviewer, they are likely to feel it is inappropriate or unnecessary to provide as much information in explaining their vote decisions to the interviewer as they would supply to another person.

Because of the constraints on their interpretation, particularly that from introspective inaccessibility and that of conversational conventions, answers to the vote-motivation questions in our main study should be viewed as providing insight into, though not necessarily a complete accounting of, the factors influencing respondents' choices. That is, considerations mentioned

[126] A Wilcoxon test on the ranks of the word counts also rejects the equivalence of the two distributions with a z-test statistic of 2.18 ($p=0.029$).

by respondents should be interpreted as "traces" of the judgmental process underlying their choices rather than as comprehensive accounts of that process. The interpretation and use of respondents' responses to introspective questions such as B-2 must take their limitations into account. Despite their constraints, such questions and their answers are essential to survey design and, in the aggregate, they are indicative of the overall validity of the study.

§ 5.2.3.2 Reasons for Choosing to Vote *For* the Program

Question B-2 was administered only to those who voted *for* the program at B-1:

> People have different reasons for voting for the Central Coast program. What would the program do that made you willing to pay for it?

This particular wording was designed, in light of the conversational convention that one should not tell someone what they already know, to focus the respondent on the outcome of the program. The interviewers were trained to use neutral and nondirective probes when respondents gave answers that seemed vague or non-responsive to the question to determine whether the respondent had anything more specific in mind.[127]

Every distinct idea in the verbatim responses to B-2 was coded into the categories listed in Table 5.3. The percentage distribution across the categories for the 552 respondents who answered this question shows that a large majority, 76.1 percent, said they voted for the program to protect the wildlife and/or environment that the program would protect from injuries caused by oil spills along the Central Coast over the next ten years. About eighteen percent responded that they felt the cost of the program was reasonable given what the program would accomplish. Those voting for the program who received the two lowest tax amounts, $5 or $25, were significantly more likely to give this response compared to those voting for the program

[127] An example probe was included in the questionnaire after B-2: *"Was there something specific that the program would do that made you willing to pay for it?"*

who received one of the other tax amounts (p=0.035). The third most common category (14%) of respondents expressed personal interest in the program because it would accomplish goals that were important to the respondent. These reasons were prefaced by "I" or "we" (*e.g.*, "I'm a bird lover and I don't want that harm to birds" or "we are always going to the coast").[128] Other reasons in this category reflected the respondents' personal interest in water-based recreation activities such as fishing or swimming.

Table 5.3 Reasons for Choosing to Vote *For* the Program

B-2. What would the program do that made you willing to pay for it?

CODING CATEGORY	PERCENTAGE [N=552][a]
Protect the wildlife and/or affected environment.	76.09%
Cost affordable/reasonable.	17.75%
Respondent personally concerned about wildlife/environment or perceives household would benefit in some way.	14.31%
Program would work.	10.69%
Others such as grandchildren or people living in the area would benefit.	9.42%
Prevent possible physical harm to respondent or others.	7.97%
Feel responsible to help prevent harm.	6.16%
Would make oil companies more responsible.	4.53%
Might help other specific animals not mentioned in survey.	4.35%
Prevent possible permanent harm.	2.54%
Protect environment in general.[b]	1.99%
Other.[c]	2.17%

(a) Percentaging base is the number of respondents who gave a response to this question. Percentages total more than 100 percent as multiple responses were allowed.

(b) Only includes those for whom no other verbatim response category was coded or for whom the only other response was coded "other".

(c) Only includes those for whom no other verbatim response category was coded.

[128] Compared to others who also voted *for* the program, those respondents who gave this response were significantly more likely to report at D-4 that their households liked to identify different species of birds (p=0.009) and to report at D-5 that their households went to the beach more than 3 times over the past summer (p=0.072).

The reasons coded in the "program would work" category (10.7%) involved something intrinsic to the program itself, *e.g.*, the response time would be shorter, it was a good plan and would work, or the program was proactive and preventative. Next were those (9.4%) who expressed satisfaction that others, such as grandchildren or people living in the area, would benefit from the program.

The reasons coded in the category "prevent possible physical harm to respondent or others" (8.0%) usually involved a desire to avoid the possibility that they or others would need to worry about eating contaminated food or swimming in contaminated water. We had anticipated that some respondents would be concerned about human health and hence the first reconsideration question, B-3, addressed this concern and emphasized that the only outcome of the program would be the prevention of harm to birds, small animals, and saltwater plants. Respondents who had expressed a concern about possible physical harm to humans at B-2 were significantly more likely to change their votes *for* to votes *not-for* at B-3 compared to other respondents who voted *for* ($p < 0.001$).[129] This result suggests that respondents paid attention to the information conveyed in B-3. Furthermore, at B-2, only six respondents gave a reason to vote *for* the program that related to possible physical harm and did not also give another reason.

Among the other types of reasons given at B-2 were expressions of personal or collective responsibility to do something about the harm from oil spills because it is caused by humans (6.2%), expectations that the program would make the oil companies more responsible (4.5%), that the program might help other animals (4.4%), and that it may prevent possible permanent harm (2.5%). Responses coded as "protect environment" (2.0%) are those for which the respondent did *not* give any other type of reason (that was not coded as "other") in their answer

[129] Eighteen respondents changed their vote at B-3; *see* section 5.4 for a more detailed discussion.

to B-2. Other respondents giving answers coded in this category clarified their initial response with more specific reasons in response to the non-directive probes.[130]

§ 5.2.3.3 Reasons for Choosing to Vote *Against* the Program

B-4 was administered to respondents who voted *against* the program:

> Did you vote against the program because it isn't worth that much money to you, or because it would be somewhat difficult for your household to pay that much, or because of some other reason?

As noted in Chapter 3, this way of asking about the respondent's motivation to vote *against* the program can alleviate the discomfort some respondents might feel at revealing motivations they find unpleasant or too personal (*e.g.*, they couldn't afford to pay for the program). Overall, 19.3 percent of those who answered this question chose the "somewhat difficult to pay" response. As one would expect, the likelihood of this response was strongly related ($p < 0.001$) to the tax amount the respondents were asked about in B-1; almost five times as many respondents at the two higher tax amounts (*i.e.*, $120 and $220) said they could not afford it than at the two lower amounts (*i.e.*, $5 and $25).

Eleven percent gave the first-offered response category, "isn't worth that much money", and 73.9 percent said they were against the program for some other reason (albeit, as noted below, the majority of these "other" reasons were different ways of stating the two pre-coded responses).[131] If the respondent said he or she had another reason, the interviewer was instructed to probe as to what that reason was. Each of the reasons expressed in these "other"

---

[130] For example, in the following verbatim, the respondent elaborated on what he or she meant by "help the environment" in response to interviewer probing: *"Because it would help the environment."* (PROBE) *"The program will keep the shorelines clean and save the birds at the coast."* (PROBE) *"I want my children to be able to enjoy the shoreline as I did when I was young."* *See* Appendix D.

[131] In 13 cases, the interviewer circled more than one B-4 answer category, hence the percentages total more than 100.

open-ended responses were coded into the categories shown in Table 5.4. In order to give a complete picture of the responses to question B-4, also included in this table (shown in italics) are the answers to the two pre-coded categories, "somewhat difficult to pay" and "isn't worth that much money".

Table 5.4 Reasons for Choosing to Vote *Against* the Program

B-4. Did you vote against the program because it isn't worth that much money to you, or because it would be somewhat difficult for your household to pay that much, or because of some other reason?

CODING CATEGORY	PERCENTAGE [N=486][a]
Concerns about program or payment vehicle.	38.68%
Problem not that important (or *isn't worth that much money*)/Others problems more important.	37.65%
Cost too high/*somewhat difficult to pay*.	31.69%
Wants more information.	1.85%
Other.[b]	2.47%

[a] Percentaging base is the number of respondents who answered B-4. Categories in italics were the explicit answer categories offered to respondents. Percentages total more than 100 percent as multiple responses were allowed.

[b] Only those for whom no other category was coded.

Overall, 38.7 percent of those who voted *against* voiced a concern they had about the program, such as skepticism about whether it would work, or about the payment vehicle.[132] An almost equal percentage (37.7%) mentioned that the problem was not that important and/or other programs were more important to them.[133] Thirty-two percent said that the cost was too high or that the tax amount was somewhat difficult to pay.[134] Only two percent wanted more information, a percentage which suggests that almost all respondents who voted against the

[132] Respondents who expressed concern with the program or payment vehicle at B-4 were at C-3 more likely to perceive the program as only somewhat effective ($p=0.008$) or not too effective/not effective at all ($p < 0.001$) and at C-4 to think they would have to pay the tax for more than one year ($p=0.003$).

[133] This includes those responses that were pre-coded into "isn't worth that much money."

[134] This includes those responses that were pre-coded into "somewhat difficult to pay."

program thought the information provided in the interview was sufficient. Moreover, the types of B-4 responses and our further analysis of these responses (described in section 6.6) strongly suggests that respondents who voted *against* the program were attentive to the object of choice and to the financial implications of voting for it and that they weighed the object of choice against other concerns when making their decision.[135]

§ 5.2.3.4 Reasons for Uncertainty about Program Vote

Respondents who said at B-1 that they were not sure about how they would vote were asked B-5: "Could you tell me why you aren't sure?" As shown in Table 5.5, the verbatim responses given by the 42 respondents who were asked B-5 are similar to the responses given at B-4 for voting against the program; however, as one might expect, these *not sure* respondents were more likely to express a desire for more information, 23.8 percent at B-5 versus 1.9 percent of the *against* at B-4. Thirty-six percent raised a concern about the program or, more commonly, the payment plan. Equal percentages gave the two most common B-4 responses: 16.7 percent commented either that the problem was not that important or that other problems were more important to them, and 16.7 percent commented that the cost was too high or difficult to pay. These verbatim responses are consistent with the evidence in Carson, *et al.*, 1995, that *not sure* voters tend to vote *against* when forced to make a choice between *for* and *not-for*.

[135] *See* discussion of construct validity equation in section 6.6.

Table 5.5 Reasons Why *Not Sure* About Program Vote

B-5 Could you tell me why you aren't sure?

CODING CATEGORY	PERCENTAGE [N=42][a]
Concerns about program or payment vehicle.	35.71%
Problem not that important/Other problems more important.	16.67%
Difficult to pay/cost too high.	16.67%
Wants more information.	23.81%
Other (including not sure).[b]	19.05%

[a] Percentaging base is the number of respondents who gave a response to B-5. Percentages total more than 100 percent as multiple responses were allowed.

[b] Only those for whom no other category was coded.

§ 5.3 Vote-Assumption Questions

Respondents were asked to choose between the status quo and a program to prevent the expected harm from Central Coast oil spills over the next ten years which would cost their household a specified amount in higher taxes. As the NOAA Panel pointed out, the reliability of respondents' choices depends on the degree to which they accepted or believed certain basic assumptions underlying the choice.[136] For example, to the extent that some respondents did not believe that the prevention program would be effective, their choices would tend to under-state their values for preventing the injuries presented in the scenario. Similarly, if respondents believed that the injuries prevented by the program over the next ten years would actually be more than that described, their choices would over-state their values for the injuries presented in the scenario.[137]

[136] The NOAA Panel states (with reference to what happens when respondents do not accept information of this type) "in effect they (the respondents) will be answering a different question from that being asked" (Arrow, *et al.*, 1993, p. 4605).

[137] In Chapter 6 we examine, in a quantitative sense, the likely impacts of respondent assumptions that diverged from those features described in the survey on the estimate of willingness to pay.

During this project, we devoted a great deal of effort to developing a program that as many respondents as possible would perceive as effective in preventing the specific set of injuries. Given the *ex ante* nature of the prevention program and the probabilistic nature of oil spills, some divergence between the scenario and respondent beliefs is inevitable. To monitor this, the questionnaire incorporated a set of checks on respondent acceptance of certain choice elements. Below, we examine the responses to those questions that asked respondents what they were assuming about various aspects of the expected harm and prevention program when they voted.

A series of questions at the beginning of Section C of the survey monitored respondent acceptance of several elements of the choice, including two key items, the extent of the expected harm and the effectiveness of the program to prevent this harm. These questions asked respondents to recall what they had in mind about certain elements of the scenario when they had voted. As discussed above, introspective questions present issues of design and interpretation; these questions were designed to avoid, as much as possible, respondents misunderstanding what information we were requesting.[138]

§ 5.3.1 Respondent Assumptions Regarding Expected Harm

The first debriefing question, C-1, asked respondents whether they thought that "the harm from oil spills in the Central Coast over the next ten years would be about the same as that shown here [in Card H],[139] or a lot more or a lot less." Ideally all respondents should have responded that the harm would be the same. As shown in Table 5.6, roughly equal percentages

[138] For example, respondents sometimes take this type of question as an invitation to speculate about the topic of the question instead of reporting what they had been thinking at the time they decided to vote. Some respondents may react to this type of question with annoyance because they believe the interviewer is giving them a quiz to see if they had paid attention to these features earlier in the interview.

[139] Card H summarized the harm described in the scenario; *see* Appendix A.

(about 35 percent) thought that either the harm would be about the *same* or *a lot more*; 15.7 percent thought it would be a lot less; 6.4 percent were not sure; and the remaining 8.7 percent gave a response coded as "other."[140] As expected, those who said *a lot more* were significantly more likely to vote *for* the program ($p < 0.001$), and those who said *a lot less*, significantly less likely to vote *for* ($p < 0.001$).[141]

Table 5.6 Respondents' Perceptions About Expected Harm[a]

C-1: At that time, did you think the harm from oil spills in the Central Coast over the next ten years would be about the same as that shown here, or a lot more or a lot less?

Answer categories:	Same	A lot more	A lot less	Other	Not Sure
	34.50%	34.78%	15.68%	8.67%	6.37%

[a] Percentaging base is the number of respondents who answered C-1.[142]

Also included in this debriefing section was a question that serves as a general check on acceptance. Question C-2 asked, "how serious did you consider this amount of harm to be?" A little more than 11% perceived the expected harm as *extremely* serious, 32.1% as *very* serious, 35.4% as *somewhat* serious, 16.6% as *not too serious*, and 3.9% as *not serious at all*. Those who felt that the harm was either *very* or *extremely* serious were significantly more likely to vote *for* the program ($p < 0.001$) and those who felt the harm was either *not too serious* or *not serious at all* were significantly less likely to vote *for* the program ($p < 0.001$).

[140] An examination of the "other" responses suggests that most of these respondents were not sure; *see* Appendix D.

[141] In Chapter 6, we examine the effect of respondent assumptions regarding expected harm on the estimate of willingness to pay.

[142] Unless stated otherwise, through out Chapters 5 and 6, we have used the number of respondents who answered the question as the percentaging base. *See* Appendix C.2 for tabulations using all question responses as the percentaging base.

§ 5.3.2 Respondent Assumptions Regarding Program Effectiveness

Another key respondent assumption examined in Section C was how effective the respondents believed the program would be in preventing the expected harm from Central Coast oil spills. Question C-3 asked: "Did it seem to you that the prevention program I told you about would be completely effective at preventing harm from Central Coast oil spills, mostly effective, somewhat effective, not too effective, or not effective at all?" As shown in Table 5.7, 6.0 percent thought it would be *completely* effective, and 44.7 percent thought that program would be *mostly* effective. Another 38.8 percent thought it would be *somewhat* effective. Only 8.4 percent held serious doubts about its effectiveness (answering either *not too effective* or *not effective at all*) and an additional 2.1 percent expressed uncertainty about its effectiveness. Given repeatedly expressed respondent concerns about the government's competence to run such a program effectively and the level of uncertainty associated with future events, incomplete acceptance is not surprising. Further, as should be expected, those respondents who did not think that the program would be *completely* or *mostly* effective were less likely to vote *for* the program ($p < 0.001$).[143]

Table 5.7 Respondents' Perceptions About Effectiveness of Program

C-3: Did it seem to you that the prevention program I told you about would be completely effective at preventing harm from Central Coast oil spills, mostly effective, somewhat effective, not too effective, or not effective at all?

Answer categories:	Completely effective	Mostly effective	Somewhat effective	Not too effective	Not effective at all	Not sure
	6.00%	44.69%	38.78%	5.54%	2.86%	2.12%

(a) Percentaging base is the number of respondents who answered C-3.

[143] In Chapter 6, we examine the effect of respondents' assumptions about the program's effectiveness on the estimate of willingness to pay.

§ 5.3.3 Respondent Assumptions Regarding Length of Payment

Question C-4 asked respondents whether they had thought when they voted that their households would have to pay the special tax for the program "for one year or for more than one year?" Fifty-six percent said one year, 38.9 percent said they doubted that it would be for just one year, and 5.3 percent were unsure. This level of skepticism about the representation that the special tax payment would be in place for only one year reflects the frequently cynical views about government promises expressed by participants in focus groups and pretesting (*e.g.*, one item mentioned was the one-year sales tax increase for San Francisco earthquake repairs approved by voters stayed in place for several years beyond its voter-approved term). This lack of acceptance is associated with a lower willingness to pay for the program that one should expect if respondents believed the object of choice actually entailed a higher cost than was described to them.[144]

§ 5.3.4 Did Respondents Feel Pressured to Vote One Way or Another?

Question C-5 asked respondents whether they perceived that the interview, overall, tried to push them to vote one way or another or let them make up their own mind. Only 5.5 percent of the total sample, or 60 respondents, said that they thought the interview had tried to push them to vote one way or another; 5 respondents were not sure. Question C-6 asked the 60 respondents who felt pushed to vote one way or another,[145] "Which way did you think it pushed you?": 32 (2.9% of the total sample) said they felt pushed to vote *for* the program, 23 (or 2.1%) felt pushed to vote *against*, and 5 respondents were either not sure about the direction

---

[144] *See* Chapter 6.

[145] The five respondents who said "not sure" at C-5 were not asked this follow-up question but rather skipped to the next question.

or gave a response coded as "other".[146] These respondents were asked to explain in C-6A: "What was it that made you think that?"[147] Those who said they felt pushed to vote *for* mentioned the presentation of all the information about the expected harm, particularly to the birds, the fact that only the one program was described, or that the perspective of the oil companies was not presented. Those who said they felt pushed to vote *against* felt that the harm was down-played (*e.g.*, "none are endangered," "only 12,000 [birds] involved") or that the harm was made to seem minor in contrast to the reasons to vote against the program. The small percentage of respondents that said they felt pushed suggests that most respondents perceived the survey as neutral, and the split among those respondents in regard to the direction they felt pushed suggests that the survey design achieved a reasonable degree of neutrality.

Also reassuring is the relationship between the direction these respondents perceived they were pushed and their votes at B-1.[148] Table 5.8 shows that those who felt pushed to vote *against* actually voted *for* the program with virtually the same frequency ($p=0.642$) as the 94 percent of the sample who said they felt the interview let them make up their own mind. This result is consistent with the interpretation that although they perceived pressure, they were not influenced to vote for or against. In contrast, there is a significant difference ($p=0.023$) in the voting behavior of those who said they felt pushed to vote *for* the program and those who said they did not feel pushed: those who felt pushed to vote *for* voted *against* the program more often than those who felt the interview let them make up their own mind. This is the opposite result from that one might have thought *a priori*. Using the more appropriate one-sided Fisher's exact test, which takes into account the direction of the difference, to compare the voting

---

[146] *See* Appendix D.

[147] *See* Appendix D.

[148] Two respondents who reported they felt pushed to vote against later changed their vote when given an opportunity to reconsider at D-15. The reconsideration questions, B-3 and D-15, are discussed below.

behavior for those who felt pushed *for* to those who did not feel pushed results in a p-value of 0.993.[149] Similarly, comparing those who felt pushed *against* to those who did not feel pushed results in a p-value of 0.401.

Table 5.8 Voting Patterns by Direction Felt Pushed

Direction Felt Pushed	Voted *For*	Voted *not-for*
Pushed *For* [N=32]	31.25%	68.75%
Not Pushed [N=1019]	51.62%	48.38%
	$\chi^2_{(1)} = 5.15$; p=0.023	
Pushed *Against* [N=23]	56.52%	43.48%
Not Pushed [N=1019]	51.62%	48.38%
	$\chi^2_{(1)} = 0.22$; p=0.642	

§ 5.4 Reconsideration Questions

Two questions included in the survey instrument gave respondents an opportunity to change their votes. The first reconsideration opportunity, B-3 was only asked if the respondent had voted *for* the program at B-1. B-3 directs respondents to "suppose human health was definitely not affected" and that "the program would <u>only</u> prevent harm to birds, small animals, and saltwater plants." B-3 then asks: "Would you vote for or against the program if it cost your household [B-1 tax amount]?" Although this question was directed toward cleansing the health concerns from their vote, it also offered respondents who voted *for* a chance to change their vote for *any* other reason. To be conservative, we counted those who said *not sure* to B-3

[149] This test compares the null hypothesis of no difference to the alternative hypothesis that respondents who felt pushed to vote *for* were more likely to vote *for*. *See* Lehmann, 1986 for a discussion of the test.

as having changed their vote to *against*.[150] These combined categories (*i.e.*, *against* or *not sure*) are referred to below as *not-for* the program. The other reconsideration question, D-15, was asked much later in the interview and was asked of all respondents. Here again, those who expressed uncertainty were treated as *not for* votes.

Four percent of those who originally voted *for* the program, a total of 22 respondents, changed their votes from *for* to *not-for*, with most (N=18) switching their vote at the first opportunity offered, B-3. The respondents who changed their votes from *for* to *not-for* have certain distinguishing characteristics when compared to those who did not change. Respondents who did not pay taxes (and for whom the tax payment may not have been incentive-compatible) were much more likely to change from *for* to *not for* ($p=0.034$); and those in the lowest three income categories were almost twice as likely to change to *not-for* ($p=0.015$). Another category of respondents who were more likely to switch from *for* to *not for* are those who the interviewer identified at E-5 or E-6 as having some difficulties understanding either the harm, the program, or the voting question ($p=0.006$).

In past studies, we have usually asked vote reconsideration questions only of respondents who voted *for* the program to avoid the possibility of over-estimating support for the program. In this study, for research purposes, the second reconsideration question, D-15, was asked of everyone; but we continued our usual practice of using for estimation purposes only the votes of those who changed their votes from *for* to *not-for*.

At D-15, 42 respondents, 7.9 percent of those who originally voted *not-for* or 3.9 percent of the total sample, changed their votes to *for*. As one should expect, respondents who changed from *not-for* to *for* were significantly more likely to have said *not sure* to the first choice

[150] This treatment is consistent with the treatment in the next chapter of those who responded *not sure* to the initial voting question.

question, B-1 ($p < 0.001$). Twelve, or 28.6 percent of those who changed from *not-for* to *for*, were unsure at B-1. The other notable characteristic of this group was that those who changed from *not-for* to *for* were significantly more likely ($p < 0.001$) to have lower levels of schooling (high school diploma or less) than the rest of the sample. While it may be the case that some respondents who changed from *not for* to *for*, particularly those who were initially *not sure* at B-1, were truly willing to pay the specified tax amount, the significant relationship of this group with low education suggests however that some respondents may have felt prompted to change their vote.

§ 5.5 Interviewer-Evaluation Questions

Another source of information about whether respondents understood the voting choice is the series of questions in Section E which the interviewers answered after leaving the respondent's home. Although interviewer evaluations are necessarily subjective, experienced interviewers have the advantage of being able to closely observe the respondent. Our training emphasized that interviewers should be frank in making these appraisals.[151]

E-4 asked the interviewers to assess the respondent's reactions "as you read through the material beginning with A-3 through A-13" to the portion of the interview that presented the elements of the scenario, such as the extent of the expected harm and the prevention program.[152] Table 5.9 shows the interviewer ratings from E-4 for how distracted, attentive, and interested the respondent appeared during this part of the presentation. As shown in the table, very low percentages of respondents were said to be *extremely* or *very* distracted (1.6%) during the presentation, and a large majority, 72.5 percent, were rated as *not at all* distracted.

---

[151] Westat, 1995, pp. 4-3 and 4-69.

[152] The first three questions in Section E asked the interviewer to record the respondent's sex, race, and zip code.

Less than one percent were rated by the interviewers as only *slightly* or *not at all* attentive and 86.2 percent were rated as either *very* or *extremely* attentive. A very low percentage was rated as only *not at all* or *slightly* (3.1%) interested, with most rated as either *extremely* or *very* interested (74%). Those who were rated as either *very/extremely* distracted or *slightly/not at all* attentive or interested, a total of 47 respondents, were significantly less likely to vote *for* the program ($p < 0.001$).

Table 5.9 Interviewer Evaluation of Respondent Reaction to Choice Elements[a]

E-4 Questions	Extremely	Very	Somewhat	Slightly	Not at all	Not Sure
How distracted was the respondent?	0.28%	1.29%	6.93%	18.93%	72.48%	0.09%
How attentive?	24.65%	61.59%	13.30%	0.37%	0.09%	0.00%
How interested?	22.85%	51.16%	22.66%	2.13%	0.93%	0.28%

(a) Percentaging base is the number of respondents who answered each question.

Question E-5 asked the interviewer if the respondent had said anything that suggested difficulty understanding either the harm caused by oil spills or the prevention program. A total of 48 respondents (or 4.42% of the total sample) were identified as having had a difficulty of some sort. In the open-ended question E-5A, the interviewers were asked to "describe the difficulties".[153] As described by the interviewers, these respondents had either difficulty seeing the visual aids or misunderstood an aspect of the expected harm or program (which was sometimes subsequently clarified). Some interviewers reiterated respondents' questions about aspects of the harm or the program that were asked and recorded earlier in the interview. With respect to the vote questions, respondents who the interviewers identified as having a problem understanding this aspect of the interview were not significantly different from other respondents

[153] *See* Appendix D.

($p=0.451$); these respondents were however more likely to change their vote from *for* to *not-for* the program ($p=0.002$).

The following two questions asked about the respondent's reaction to the choice question, B-1. E-6 asked if the respondent had any difficulty understanding B-1, if so, to describe the difficulties at E-6A.[154] A total of 19 respondents (1.8% of the total sample) fell into this category. Their difficulties, as described by the interviewers, included respondent skepticism about the length of time of the tax payment or about another aspect of the payment plan. Some of these difficulties appear to have been overcome after the interviewer re-read the pertinent material or question.

Another factor that might affect a respondent's understanding of the choice is impatience to get through the interview. Questions E-7 and E-7A asked the interviewer to rate the degree of impatience the respondent displayed when asked the voting questions. The vast majority of the respondents (84.9%) were not thought to be impatient, and another 6.2 percent were rated as *not very* impatient or only *a little* impatient. Four percent were said to be *somewhat* impatient, and 1.7 percent were said to be *very* impatient. Those who were considered by the interviewer to be *very* or *somewhat* impatient were significantly less likely to vote for the program ($p < 0.001$).

Interviewer ratings in E-8 of "how serious was the consideration the respondent gave to the decision about how to vote" can be used to gauge whether the choice mechanism was plausible and taken seriously by respondents. As shown in Table 5.10, 80.5 percent of the total sample were thought to have given the matter *extremely* or *very* serious consideration. Only about 1.2 percent (or 13 cases) were rated as giving it only *not at all* or *slightly* serious consideration, and these respondents were significantly less likely to vote for the program

[154] *See* Appendix D.

(p=0.015).

Table 5.10 Interviewer Evaluation of the Seriousness of Respondent Consideration of the Voting Decisions[a]

Question E-8	Extremely	Very	Somewhat	Slightly	Not at all	Not Sure
How serious was the consideration the respondent gave to the decision about how to vote?	24.95%	55.57%	18.00%	0.93%	0.28%	0.28%

(a) Percentaging base is the number of respondents who answered E-8.

§ 5.5.1 Were Respondents' Choices Influenced by Others?

In order to avoid distractions, interviewers were instructed to try to conduct interviews without other persons present. Frequently, when other people were present, they were young children in the respondent's care. In order to differentiate these cases from those where teenagers or adults were present, the interviewers were asked in E-9 to report whether anyone age 13 or older was present when the respondent voted. In 21 percent of the interviews, someone age 13 or older was present.

In question E-9A, the interviewer was asked whether the other person(s) seemed to affect how the respondent voted. In almost 90 percent of the cases where someone age 13 or older was present while the respondent voted, the interviewers judged that there was no effect. In six cases (less than 1% of the sample), the interviewer believed that the other person present did have an effect; and, in 18 cases the interviewer indicated that he or she did not know. We examined the E-9A and E-10 responses of these 24 cases; whenever the interviewer mentioned

possible influence, it was almost always by another household member.[155] Since the goal of the survey was to measure *household* preferences, if other household members did influence the respondent in this small number of cases, such influence would not be inconsistent with that goal.

§ 5.6 Summary

The pattern of verbatim responses to the open-ended questions considered in this chapter was consistent with respondents paying attention to the survey and taking the choice opportunity seriously. Answers to questions about the reasons for their voting choices (B-2, B-4, and B-5) generally referred to relevant features of the prevention program such as what the program would accomplish and its cost. Overall, the types of queries raised during the presentation of the expected harm and the prevention program were usually related to the material in a meaningful way and provide assurance that the respondents were paying attention to this part of the interview. Further, coupled with the responses to Section C vote-assumption questions and the Section E interviewer evaluation questions, these patterns suggest that respondents' decisions reflected their perceptions of the object of choice and their preferences for it.

An important feature of our design was to offer respondents opportunities to reconsider their choices. Question B-3 allowed those who voted *for* the program to reconsider their votes shortly after they voted; and another reconsideration question, D-15, was offered to all respondents near the end of the interview. Those who gave a B-2 verbatim response related to "possible effects on humans" were more likely to reconsider and change their *for* vote to an *against* vote compared to the rest of the sample.

Analysis of the interviewer-evaluation questions in Section E of the survey found that in

[155] *See* Appendix D.

only a few cases the interviewers identified possible problems with respondents' attentiveness, interest, and impatience. Further, those who the interviewers identified as less focused were generally less likely to vote for the program compared to the rest of the sample.

§ 6 Analysis of Choice Questions

§ 6.1 Introduction

In this chapter the choices made by respondents in the main survey are used to construct a lower-bound estimate of the *ex ante* total value for preventing the expected harm from oil spills along the California Central Coast over the next decade. The relationships between the choice measure and other respondent characteristics measured by the survey are also examined. Section 6.2 presents two versions of the choice measure. Section 6.3 discusses the non-parametric (Turnbull, 1976) statistical framework used in much of our analysis of the estimate of *ex ante* total value. Section 6.4 provides the Turnbull lower bound estimate on the sample mean[156] and examines the sensitivity of this estimate to various assumptions regarding the treatment of the data. Using the categories suggested by the NOAA Panel[157] as a framework, section 6.5 examines the bivariate relationships between choice measures and respondent characteristics. Section 6.6 examines construct validity using a multivariate counterpart to the evaluations of individual variables reported in the prior section. Section 6.7 provides a sensitivity analysis which looks at possible shifts in value related to respondent assumptions at variance with key scenario features. Section 6.8 presents a more conservative treatment of the respondents who said that they did not pay California income taxes and their impact on the total value estimate.

§ 6.2 Definition of Choice Measures

B-1, the principal choice question, asked respondents if they would vote for or against a Central Coast prevention program if it cost their household a certain tax amount. Before survey administration, the survey research company had randomly assigned each respondent to

---

[156] *See* Appendix E.1 for a more detailed discussion of the Turnbull estimator and Appendix E.2 for a comparative analysis of COS and the Exxon Valdez survey results.

[157] Arrow, *et al.*, 1993, p. 4609.

one of five different tax amounts (referred to as B1AMT below): $5, $25, $65, $120, or $220. Table 6.1a summarizes the responses to the B-1 choice question by B1AMT. In the following analysis, the *against* and *not-sure* categories of question B-1 (displayed in the last two columns of Table 6.1a) are combined into a single *not-for* category; the choice measure incorporating this coding will be referred to as **B1**.[158]

Table 6.1a B-1 Response by B1AMT

B1AMT	For	Against	Not Sure[159]
$5	69.86%	27.40%	2.74%
$25	58.33%	39.35%	2.31%
$65	51.04%	45.23%	3.73%
$120	45.30%	49.17%	5.52%
$220	29.82%	64.47%	5.70%

The key prediction of economic theory is that the percent of the sample voting *for* should decrease as the tax amount increases. Table 6.1b displays the percentages of *for* and *not-for* responses to **B1** by B1AMT. Based on these data, a chi-squared test[160] ($\chi^2_{(4)}=79.08$; $p < 0.001$) clearly rejects the null hypothesis that the percentage *voting-for* does not systematically vary with B1AMT. A one-sided Fisher's exact test which takes into account the theoretically predicted direction of the variation of **B1** with B1AMT provides an even stronger rejection of the null hypothesis.

[158] All choice measure variables are denoted in bold capital letters.

[159] In a $65 version questionnaire, the interviewer did not circle an answer category at B-1. Given the nature of the verbatim comments recorded by the interviewer at B-1 (*"I'm not going to answer that ... can't say for or against"*), the B-1 response for this case was coded as *not sure*.

[160] The null hypothesis tested in a chi-squared (χ^2) test is that the rows and columns in a two-way table are independent.

Table 6.1b B1 Choice Measure by B1AMT

B1AMT	For	Not For
$5	69.86%	30.14%
$25	58.33%	41.67%
$65	51.04%	48.96%
$120	45.30%	54.70%
$220	29.82%	70.18%
	$\chi^2_{(4)} = 79.08$; $p < 0.001$	

A choice measure defined only by the B-1 responses (*e.g.*, the **B1** choice measure defined above) results in *single-bounded* interval data.[161] That is, if a respondent votes *for*, the respondent's willingness to pay (WTP) for the program is bounded from below by B1AMT (*i.e.*, the respondent is willing to pay at least B1AMT). How much more the respondent might be willing to pay is not revealed; we know only that the respondent's WTP is not less than B1AMT ($WTP_R \geq$ B1AMT). If the respondent gives a *not-for* answer, the respondent's willingness to pay is bounded from above by B1AMT (*i.e.*, the respondent may be willing to pay some tax amount below B1AMT or may not be willing to pay anything at all). Thus, the respondent's WTP is less than B1AMT ($0 \leq WTP_R <$ B1AMT).[162]

§ 6.3 Statistical Framework for Analysis

In developing an estimate of the *ex ante* total value for preventing the expected harm from oil spills off California's Central Coast over the next ten years, we have consistently

[161] The seminal contingent valuation paper on the use of binary discrete choice data is Bishop and Heberlein (1979). Hanemann (1984) developed the utility-theoretic approach to such models. Cameron and James (1986) look at choice data using an approach based on the willingness-to-pay function. McConnell (1990) compares the two approaches.

[162] We assume that no respondent would demand compensation for implementing a program to prevent oil spills along the Central Coast, *i.e.*, no respondent has a negative WTP.

chosen conservative design features and statistical assumptions.[163] Respondents who voted *for* the program at B-1 were allowed to reconsider that vote in question B-3 and again in question D-15.[164] Revising the **B1** choice measure (defined above) to take into account those respondents who changed their votes *for* to votes *not-for* results in a second choice measure—**B1CH**. **B1CH** treats as votes *for* only those respondents who voted *for* the program at B-1 and who did not change their *for* vote at either of the two opportunities (*i.e.*, at B-3 and D-15); by construction, **B1CH** is a more conservative choice measure than **B1**.

Table 6.2 displays the **B1CH** choice measure by B1AMT; the null hypothesis that the **B1CH** choice measure does not systematically vary with B1AMT is rejected ($\chi^2_{(4)}=82.48$, $p < 0.001$). Unless otherwise indicated, we use the **B1CH** choice measure for the remaining analysis presented in this chapter.

Table 6.2 B1CH Choice Measure by B1AMT

B1AMT	For	Not For
$5	68.95%	31.05%
$25	56.94%	43.06%
$65	48.55%	51.45%
$120	40.33%	59.67%
$220	28.95%	71.05%
	$\chi^2_{(4)} = 82.48$; $p < 0.001$	

To estimate *ex ante* total value, we have chosen as our summary statistic the Turnbull

[163] This is in keeping with the NOAA Panel's recommendation: "Generally, when aspects of the survey design and the analysis of the responses are ambiguous, the option that tends to underestimate willingness to pay is preferred" (Arrow, *et al.*, 1993, p. 4612).

[164] Respondents who voted *against* the program at B-1 were also given an opportunity to reconsider their vote at D-15.

(1976) non-parametric, maximum likelihood (ML) estimator for interval-censored data.[165] The Turnbull estimator uses respondents' choices to construct an interval estimate for the latent willingness to pay implied by each respondent's choice. As described above, an individual's choice will distinguish either a lower or an upper bound for his or her WTP. By combining respondents' choices, we obtain estimates for the relative frequency of responses at different WTP intervals, (0, $B1AMT_i$) and ($B1AMT_i$, ∞), where $B1AMT_i$ is one of the five B-1 tax amounts administered to independent sub-samples. The first pair, (0, $B1AMT_i$), defines the interval with $B1AMT_i$ as an upper bound, and the second pair, ($B1AMT_i$, ∞), the interval with $B1AMT_i$ as a lower bound. The six intervals or "steps" defined by B1AMT are: (1) \$0–\$5, (2) \$5–\$25, (3) \$25–\$65, (4) \$65–\$120, (5) \$120–\$220, and (6) \$220–∞.

A range of summary statistics, related to the sample mean, can be defined based on the Turnbull estimates of the fraction of the sample in each of the six intervals. These differ in the assumed distribution of respondents within the intervals. The lowest of these we will refer to as the lower bound on the sample mean. The fraction of the sample estimated to be in each interval is treated as having a willingness-to-pay value equal to the lower end-point of the interval, and the ordinary sample mean is calculated.[166] The highest of these summary statistics is the upper-bound mean. The fraction of the respondents estimated to be in an interval is treated as having a WTP at the high end-point of the interval, and the ordinary sample mean is calculated.[167]

---

[165] The initial uses of this framework in the CV literature are found in Carson and Steinberg (1990) and Kristrom (1990). See Carson, Willis, and Imber (1994) for a recent large scale application. Appendix E.1 provides a detailed discussion of the Turnbull estimator.

[166] For example, if 20% of the sample is estimated to be in the interval \$25 to \$65, the lower-bound mean is calculated by assuming that this 20% of the sample is willing to pay exactly \$25.

[167] For example, if 20% of the sample is estimated to be in the interval \$25 to \$65, the upper-bound mean is calculated by assuming that this 20% of the sample is willing to pay \$65. As the high end-point in the last interval, \$220-∞, is infinity, the upper-bound mean is infinite unless reasonable additional assumptions are imposed. As we are asking about WTP, it would be possible to substitute for infinity an upper bound based on either income or wealth. *See*

Irrespective of the particular tax amounts used to define the intervals, the unobserved sample mean is always bounded below by the lower bound on the sample mean and above by the upper bound on the sample mean if there are equivalent sub-samples at each of the tax amounts.[168] However, the particular tax amounts respondents are asked about influence how much *less* the Turnbull estimate of the lower bound on the sample mean is than the sample mean and how much *greater* the Turnbull estimate of the upper bound on the sample mean is than the sample mean. Any estimate of the sample mean which is lower than the Turnbull estimate of the lower bound on the sample mean or higher than the Turnbull estimate of the upper bound on the sample mean is inconsistent with the observed choices made by respondents. Without additional statistical assumptions about the latent willingness-to-pay distribution, any other observed choice measure is uninformative about where, within the two Turnbull bounds, the sample mean lies. The most conservative assumption consistent with the observed choices is that the sample mean is equal to the Turnbull estimate of the lower bound on the sample mean.

§ 6.4 Turnbull Estimate of the Lower Bound on the Sample Mean

Table 6.3 reports the Turnbull estimate of the lower bound on the sample mean for the WTP distribution using the **B1CH** choice measure. Note that the third column in the table (labeled "Probability of Voting *For* at Upper-Bound") is simply the estimated fraction of those in Table 6.2 who would vote *for* the program at each B1AMT. The elements in the table describe the interval defined by B1AMT and respondents' choices. For example, we know a respondent's willingness to pay for the Central Coast prevention program is greater than or equal to $5 if the respondent voted *for* the program at $5. If, on the other hand, a respondent voted

section 2 in Appendix E.1.

[168] In large but finite random samples, such as the one used for this study, the number of respondents receiving each tax amount is approximately equivalent. The standard error of the estimate reflects the sampling variation.

against the program at $5, we know that the respondent's willingness to pay is less than $5 and possibly $0. Likewise, for a respondent who was asked about $25, a vote *against* the program implies that the respondent's willingness to pay for the prevention program lies somewhere in an interval from $0 to $25, while a vote *for* implies a minimum willingness to pay of at least $25. In this way, we can classify each respondent's willingness to pay into an interval depending on the B1AMT the respondent received. The Turnbull estimate of the lower bound is then calculated by multiplying the lower bound of the interval column by the change in density column and then summing the products.[169]

Table 6.3 Turnbull Estimate of WTP Distribution and Lower Bound on the Sample Mean: B1CH Choice Measure [N=1,085]

Lower Bound of Interval	Upper Bound of Interval	Probability of Voting *For* at Upper Bound	Change in Density[170]
$0	$5	0.6895	0.3105
$5	$25	0.5694	0.1201
$25	$65	0.4855	0.0840
$65	$120	0.4033	0.0822
$120	$220	0.2895	0.1138
$220	∞	0.0000	0.2895
	Log-Likelihood	-709.48	
	Estimate of lower bound on sample mean	$85.39	
	Standard error of the estimate	$3.90	

[169] *See* Appendix E.1.

[170] The values shown in the change in density column are the percentage of respondents who fall into each interval; for example, 12.0 percent of respondents fall into the $5–$25 interval and, hence, the Turnbull assumes 12.0 percent are willing to pay $5. The z-statistics for the five change in density parameters estimated by the model are 9.93, 2.61, 1.80, 1.69, and 2.41, respectively. The significance of each individual parameter value is of little importance; the set of parameters taken together, however, is reflected in the standard error of the estimate. Here, the standard error of $3.90 suggests reasonable precision in the estimate.

The Turnbull estimate of the lower bound on the sample mean, $85.39,[171] is obtained by assuming that all of the fraction of the sample estimated to be in a particular interval falls at the lower end of that interval. For example, respondents who voted *against* at $5 fall into the $0–$5 interval and are assumed to have a willingness to pay of $0. Respondents who voted *for* at $220 fall into the $220–∞ interval and are assumed to have a willingness to pay of $220. The median (50th percentile) respondent falls in the $25-$65 interval.

§ 6.4.1 Sensitivity of the Turnbull Estimate of the Lower Bound

In this section, we examine the sensitivity of the Turnbull estimate of the lower bound on the sample mean if the choices of one of ten categories of respondents (summarized in Table 6.4 below) are not included in the estimation.

The first category of respondents are those who were unsure about how they would vote on the program. In the prior section, these 43 respondents were treated as *not-for* the program. A less conservative treatment is to not include them in the estimation; not including these observations raises the Turnbull estimate of the lower bound on the sample mean by $4.16, from $85.39 to $89.55.

Another group of respondents are those who "protest" the payment mechanism, *i.e.,* they object that oil companies, not taxpayers, should pay for the prevention program. The common locations at which this sentiment was spontaneously expressed by respondents are, first, before the choice question (*e.g.,* during the presentation of the payment vehicle) and, second, in response to the choice question. Some respondents also protested in their responses to the open-ended vote-motivation questions (B-2, B-4, and B-5) or spontaneously at later points in the

[171] The Turnbull estimate of the lower bound on the sample mean using the sample weights is $85.50, only $0.11 higher than the unweighted sample estimate. The standard error of the weighted estimate is $3.84.

Table 6.4 Summary of Sensitivity Tests

Part of Sample Not Included	Percentage of Sample Not Included In Estimation	Change in $85 **B1CH** Lower-Bound Estimate
not sure responses to choice question [B-1]	3.96%	+$4.16
protested that oil companies should pay before B-1	7.10%	+$2.32
protested that oil companies should pay before or during B-1	10.32%	+$3.87
protested that oil companies should pay at any point during interview	15.58%	+$8.74
protested that oil companies would pass cost on to consumers at any point during interview	5.35%	+$3.24
felt pushed to vote one way or another [C-5]	5.07%	+$0.43
not at all/slightly serious consideration of B-1 [E-8]	1.20%	+$0.74
negative evaluations by interviewer on one of six indices (includes R's in previous category)	9.95%	-$0.33
WTP more than 5% of income	0.37%	-$1.27

interview. Not including the 7.1 percent who protested before the choice question results in a $2.32 increase in the lower bound on the sample mean. Not including, in addition, those who protested at the choice question results in a slightly higher increase of $3.87. Using the most inclusive definition—not including the 15.6 percent who protested that the oil companies should pay for the program at any point in the interview—results in an increase of $8.74.

Some respondents also protested that oil companies would pass their share of the program costs (*i.e.*, operating costs over the next 10 years) to the consumer in the form of higher prices. Not including the 5.4 percent of the sample who expressed this at any point in the interview results in an increase of $3.24 in the estimate. In sum, not including in the estimation those respondents who protested the payment mechanism either because they believed oil companies should be completely responsible for the program's costs or because they believed they would in fact have to pay for the oil companies share of the program costs results in a $2.32 to $8.74

increase in the estimate.

A third group of respondents are those who felt pushed to vote one way or another.[172] Not including the 5.1 percent who felt pushed to vote either *for* or *against* the program results in an increase of $0.43.

The interview evaluation questions in Section E can be used to identify respondents who may have had problems understanding an element of the survey. The most obvious group is that of respondents who the interviewers said gave the choice question "slightly/not at all" serious consideration (E-8). Not including this 1.2 percent of the sample raises the estimate of the lower bound on the sample mean by $0.74. A more expansive definition encompasses, in addition, respondents the interviewers identified as having some difficulty understanding the harm caused by Central Coast oil spills or the prevention program [E-5], as having some difficulty understanding the choice question [E-6], or as being "very/extremely" distracted [E-4a], "slightly/not at all" attentive [E-4b], or "slightly/not at all" interested [E-4c] during the presentation of the descriptive material (*i.e.*, A-3 through A-13; see Appendix A). Not including this 10 percent in the sample results in a $0.33 decrease in the lower bound on the sample mean. This analysis suggests that those who did not take the choice seriously, who had any difficulties understanding, or who were distracted, inattentive, or uninterested were willing to pay slightly less than other respondents in the sample.

Respondents voting to pay a large percentage of their income is an indication that they may not have taken their budget constraints seriously. Taking the ratio of the respondent's income to the lower bound of the interval where the respondent's willingness-to-pay amount lies, we find that no respondents are willing to pay more than 10 percent of their household income and only four respondents, all of whom reported that they did not pay state income taxes, are

[172] *See* section 5.3.4.

willing to pay more than 5 percent. Not including these four respondents in the estimation results in a decrease of $1.27 in the estimate.[173] We consider a possible adjustment for respondents not paying state income taxes in section 6.8 below.

§ 6.4.2 Section 6.4.2 Tests for Nay-sayers and Yea-sayers

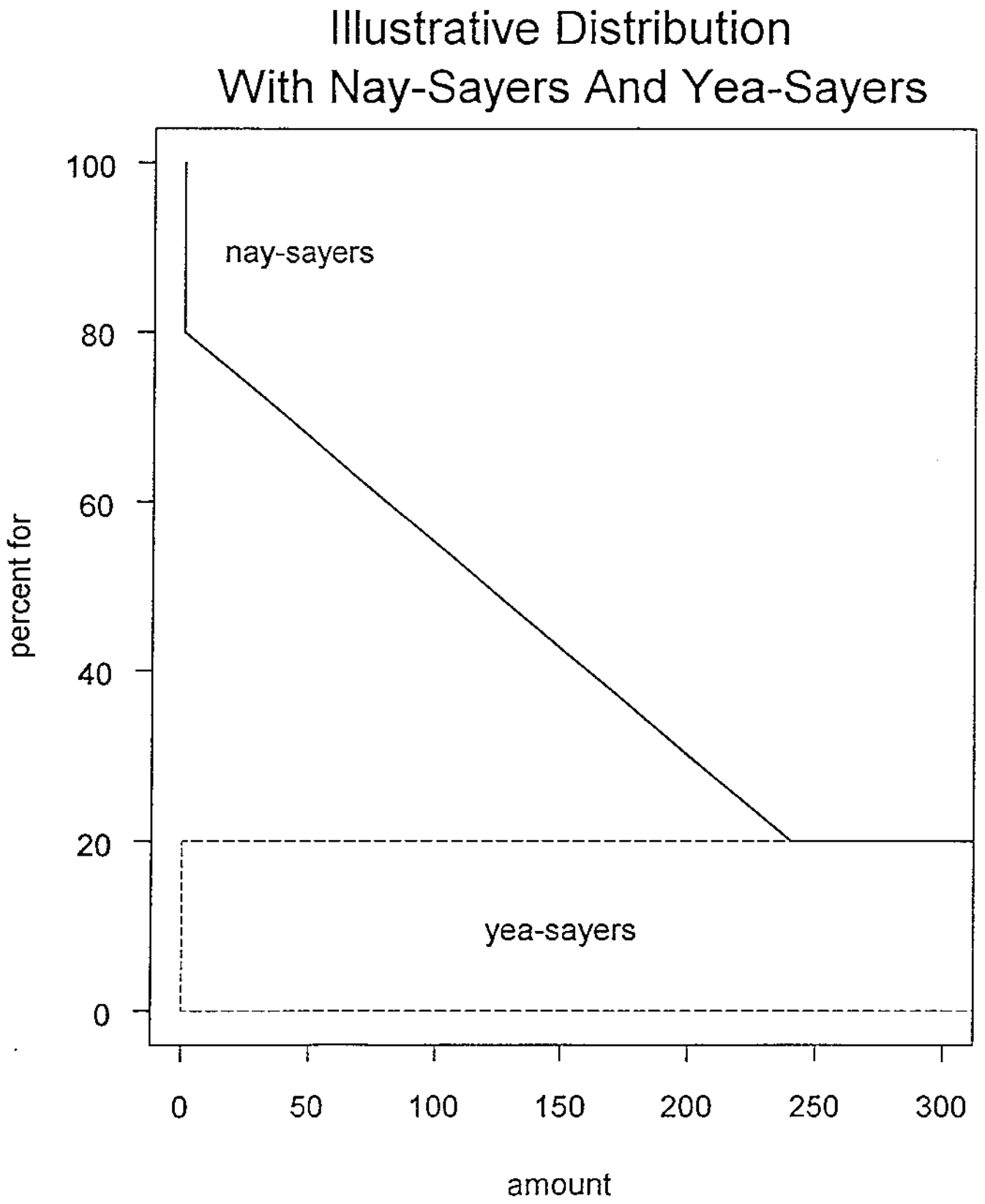
The Turnbull non-parametric estimator fit to the B1CH choice measure allows for the possibility that some fraction of the respondents might vote *not-for* even if the cost to the respondent was $0, and it also allows for the possibility that some fraction of the respondents might vote *for* the program irrespective of the tax amount. Such respondents are often referred to as nay-sayers and yea-sayers in the social science and contingent valuation literature (Mitchell and Carson, 1989; Hanemann and Kanninen, forthcoming) and as *natural mortality* and *immunes* in the biometrics literature where statistical methods of dealing with these phenomena were first developed (Finney, 1949).

Figure 6.1 depicts an illustrative distribution of which 60% of the respondents react to the amount they are asked to pay in a linear fashion; 20% of the respondents are nay-sayers; and 20% of the respondents are yea-sayers. The nay-sayers are represented as a *vertical spike* at zero while the yea-sayers are represented as a *horizontally-oriented slice* at the bottom of the graph. The regression line now effectively starts at 80% *for* at $0.01 and ends with 20% being willing to pay an arbitrarily large amount.

The presence of nay-sayers or yea-sayers respondents can influence the interpretation of

[173] Eighteen respondents violate a strict two percent of income criterion and 30 violate a very strict one percent criterion. These respondents are significantly more likely to have reported they did not pay income taxes ($p < 0.001$).

Figure 6.1


Illustrative Distribution
With Nay-Sayers And Yea-Sayers
100
80
60
40
20
0
percent for
nay-sayers
yea-sayers
0
50
100
150
200
250
300
amount

the Turnbull lower bound on the sample mean.[174] Due to its non-parametric nature, the Turnbull approach is incapable of providing estimates of either the fraction of nay-sayers or the fraction of yea-sayers. To obtain such estimates, it is necessary to make a parametric assumption concerning the distribution of the underlying WTP distribution. The standard parametric distributions for survival analysis (Nelson, 1982) do not allow for the possibility of either nay-sayers or yea-sayers. One may include parameters allowing for the possibility of nay-sayers and/or yea-sayers and then test whether the fit of the parametric model is improved.

To perform such a test for nay-sayers, we consider three commonly used parametric models: the log-normal, the Weibull and the Box-Cox. The log-normal and the Weibull are two-parameter distributions. Allowing for the possibility of a spike at zero in the log-normal model (i.e., nay-sayers), results in a an estimate that 28.7% of the respondents have a zero WTP; and in the Weibull model, an estimated 24.2% of the respondents have a zero WTP. The individual z-statistics on the spike parameters (4.75 and 2.61, respectively) are both significant at the $p < .01$ level, but the likelihood ratio test for the log-normal model with and without the naysayer-parameter ($\chi^2_{(1)}=2.99$, $p=.08$) and that for the Weibull model with and without the naysayer-parameter ($\chi^2_{(1)}=1.73$, $p=.19$) suggest that these two distributions do not fit dramatically better as result of adding the naysayer-parameter.[175]

The Box-Cox is a three-parameter distribution which nests the log-normal and normal as special cases. The three-parameter Box-Cox model allows more flexible curvature with respect to the shape of the underlying WTP distribution than either of the two-parameter models above,

---

[174]True zeros are correctly taken into account in the calculation of the Turnbull lower bound on the sample mean, while nay-sayers who do have a positive WTP for the good under the scenario depicted will bias the Turnbull lower-bound on the sample mean downward. The presence of yea-sayers will bias the Turnbull lower bound on the sample mean upward.

[175]The log-likelihoods for the log-normal model and the log-normal model with a spike at zero are -711.31 and -709.81, respectively. For the Weibull and Weibull spike models, the log-likelihoods are -710.54 and -709.67, respectively.

the log-normal and the Weibull. The estimate of the Box-Cox λ parameter of .3665 suggests a distributional shape between that of the normal ($\lambda=1$) and that of the log-normal ($\lambda=0$). Likelihood ratio tests reject the normal in favor of the Box-Cox at $p=.01$ and the log-normal in favor of the Box-Cox at $p=.07$. An examination of the Box-Cox fit in this case shows the model predicting a steep drop in the percentage willing to pay as one moves away from zero. As a result the inclusion of a zero spike parameter does not improve the statistical fit of the Box-Cox model to the data. Allowing for the possibility of a spike at zero in the Box-Cox model results in an estimate of the fraction of respondents at zero of less than 0.1%.

Thus, while the nay-sayer tests based on the two-parameter distributions support the possibility of a downward bias in the Turnbull lower bound on the sample mean, the Box-Cox model suggests allowing more flexibility in the percentage of the respondents who are estimated to have very small, but still positive, WTP values also provides an equally good fit to the data. Comparison of the log-normal with a spike, the Weibull with a spike, and the Box-Cox, all three-parameter distributions, to the Turnbull (which is a perfect fit to the data) suggests that all three of the three parametric distributions fit the data quite well.[176]

Turning to the possibility of yea-sayers, all three parametric models suggest the same answer: the fraction of respondents that are yea-sayers is less than 0.1%. The yea-sayer parameter never approaches significance in any of the three models. Further, the log-likelihood of each of the three models with the yea-sayer parameter is almost identical to that of the corresponding model without the yea-sayer parameter, a result which suggests that allowing for the possibility that some respondents are willing to pay any amount provides almost no improvement in the distributional fit.

[176]The log-likelihoods for the log-normal with a spike at zero, the Weibull spike model, the Box-Cox model, and the Turnbull are -709.81, -709.67, -709.63, and -709.48, respectively.

Allowing for the possibility of both nay-sayers and yea-sayers at the same time does not change this conclusion. If the possibility of both nay-sayers and yea-sayers is allowed in the log-normal and Weibull distributions, the nay-sayer parameter again suggests 20–30% of the observations should be placed in a spike at zero; and the yea-sayer parameter estimates are again very close to zero. In the Box-Cox model with nay-sayer and yea-sayer parameters, both of those parameter estimates are essentially zero and insignificant. In all three cases, the resulting distributions look equivalent to the models which only allowed for nay-sayers. Thus, none of the tests performed with both parameters provide any support for the presence of yea-sayers who might bias the Turnbull lower bound on the sample mean upward; nor do they provide any ground for choosing as the correct model either a spike of nay-sayers at zero or a substantial number of small but non-zero WTP values.

§ 6.5 Bivariate Relationships Including Cross-Tabulations Recommended by NOAA Panel

The NOAA Panel recommends categorizing the responses to the "primary valuation" question (*e.g.*, in this survey, the **B1** and **B1CH** choice measures) to facilitate interpretation of the responses to this question. The recommended categories include income, prior knowledge of the site, variables related to prior interest in the site, distance to the site, attitudes toward the environment, attitudes toward big business, understanding of the task, beliefs about the scenario, and ability/willingness to perform the task (Arrow, *et al.*, 1993, p. 4609). First, we report the cross-tabulation results for variables reflecting each of the recommended categories with the choice measures, **B1** and **B1CH**. Second, for illustrative purposes, we consider how a subset of these recommended categories would influence the Turnbull estimate of lower bound on the sample mean. Third, in section 6.6, we present a multivariate analysis addressing a subset of the categories recommended by the Panel and others that economic theory suggests should

influence respondents' choices.

Table 6.5 describes the specific source of the information used in each of the cross-tabulations. In most cases, the variables are constructed from single questions in the main study survey instrument. In a few cases, we constructed the measure using two or more questions. The table also includes a short descriptive summary of the information and an indication of whether the variable directly ("D") or indirectly ("I") measures the category identified by the Panel. Furthermore, as shown by the table, the survey instrument contains multiple variables for some of the Panel's recommended categories.

Table 6.5 Description of Sources of Information for Cross-Tabulations

Recommended Category	Source	Relationship to Category	Description
1. Income	D-12	D	Total household income before taxes in 1994
2. Prior Knowledge of Site	A-7 D-1	D	Familiar with 5 types of birds affected in spills; Driven along the Central Coast on Highway 1
3. Prior Interest in the Site	A-4 D-5	I	Visited any of three types (beach, marsh, rocky shore) of California shoreline in last twelve months; Visited beach at least three times last summer
	D-3 D-4	I	Saltwater boating or fishing in the last 5 years; Bird watcher
4. Attitudes Toward the Environment	A-1b A-1e A-2c	D	Reducing air pollution in California cities; Protecting coastal areas from oil spills; Protecting wildlife
	D-7	I	Respondent's self-evaluation on environmentalist scale
5. Attitudes Toward Big Business	Oil companies should pay	I	Oil companies should pay for all of program costs (*i.e.*, Box 4 or Box 5 checked by interviewer)
6. Distance to the Site	Central Coast PSU's	D	Location of respondent's residence in Central Coast PSU's (*i.e.*, San Francisco Bay down to greater Los Angeles area)
7. Understanding of the Task	E-5	I	Interviewer evaluation of respondent comments indicating any difficulty in understanding harm caused by spills or the prevention program

Recommended Category	Source	Relationship to Category	Description
	E-6	I	Interviewer evaluation of respondent having any difficulty understanding the voting question
8. Beliefs about the Scenario	C-1	D	Respondent judgment about oil spill effects
	C-3	D	Respondent judgment about effectiveness of prevention program
	C-4	D	Respondent judgment about limit of special tax to single year
9. Ability/Willingness to Perform Task	E-7	I	Interviewer evaluation of whether respondent impatient to complete interview

Table 6.6 summarizes the cross-tabulation results for the **B1** and **B1CH** choice measures.[177] These cross-tabulations test the null hypothesis that the *for* and *not-for* choices are not influenced by each of the variables. The reported p-value is the probability that the test result would call for incorrectly rejecting a "true" null hypothesis of no effect of the source variables on respondents' choices. The last column in the table reports the decision—assuming as a threshold the commonly used p-value of 0.05—that would be made about differences in the distribution of responses between *for* and *not-for* choices (using both the **B1** and **B1CH** choice measures) and the categories in each of the source variables. The label "R" indicates that the null hypothesis of independence was rejected in favor of the alternative hypothesis of association between the choice measure and the variable; and "N" indicates that the null hypothesis was not rejected which suggests no association. For example, in the case of income, we reject the null hypothesis at the 95 percent confidence level that income does not affect the distribution of votes *for* and *not-for* since the calculated p-value of 0.04 is smaller than 0.05.

[177] Appendix E.3 presents cross-tabulation tables for each of the variables in Table 6.6 by the **B1** and **B1CH** choice measures.

Table 6.6 Cross-Tabulation Summary

Recommended Category	Source[a]	Choice Measure	Statistic of Association[b]	p-value[c]	Reject/Not Reject Hypothesis of No Association
1. Income	D-12	**B1**	γ=-0.006	0.04	R
		B1CH	γ= 0.024	0.04	R
2. Prior Knowledge of Site	A-7	**B1**	$\chi^2_{(1)}$= 8.366	0.00	R
		B1CH	$\chi^2_{(1)}$=10.728	0.00	R
	D-1	**B1**	$\chi^2_{(1)}$=4.811	0.03	R
		B1CH	$\chi^2_{(1)}$=7.032	0.01	R
3. Prior Interest in the Site	A-4	**B1**	$\chi^2_{(1)}$=1.942	0.16	N
		B1CH	$\chi^2_{(1)}$=3.642	0.06	N
	D-5	**B1**	$\chi^2_{(1)}$=5.746	0.02	R
		B1CH	$\chi^2_{(1)}$=7.032	0.01	R
	D-3	**B1**	$\chi^2_{(1)}$=7.573	0.01	R
		B1CH	$\chi^2_{(1)}$=5.243	0.02	R
	D-4	**B1**	$\chi^2_{(1)}$= 9.878	0.00	R
		B1CH	$\chi^2_{(1)}$=10.975	0.00	R
4. Attitudes Toward Environment	A-1b	**B1**	$\chi^2_{(4)}$=47.561	0.00	R
		B1CH	$\chi^2_{(4)}$=45.007	0.00	R
	A-1e	**B1**	$\chi^2_{(4)}$= 98.372	0.00	R
		B1CH	$\chi^2_{(4)}$=100.051	0.00	R
	A-2c	**B1**	$\chi^2_{(4)}$=123.309	0.00	R
		B1CH	$\chi^2_{(4)}$=118.025	0.00	R
	D-7	**B1**	$\chi^2_{(4)}$=44.130	0.00	R
		B1CH	$\chi^2_{(4)}$=40.851	0.00	R
5. Attitudes Toward Big Business	Oil companies should pay	**B1**	$\chi^2_{(1)}$=5.089	0.02	R
		B1CH	$\chi^2_{(1)}$=6.295	0.01	R
6. Distance to Site	Central Coast PSU's	**B1**	$\chi^2_{(1)}$=5.502	0.02	R
		B1CH	$\chi^2_{(1)}$=6.655	0.01	R
7. Understanding of Task	E-5	**B1**	$\chi^2_{(1)}$=2.780	0.10	N
		B1CH	$\chi^2_{(1)}$=0.600	0.44	N
	E-6	**B1**	$\chi^2_{(1)}$=4.009	0.05	R
		B1CH	$\chi^2_{(1)}$=4.757	0.03	R
8. Beliefs about Scenario	Oil Spill Effects: More Harm[d]	**B1**	$\chi^2_{(1)}$=46.032	0.00	R
		B1CH	$\chi^2_{(1)}$=45.428	0.00	R

Recommended Category	Source[a]	Choice Measure	Statistic of Association[b]	p-value[c]	Reject/Not Reject Hypothesis of No Association
	Oil Spill Effects: Less Harm[d]	**B1**	$\chi^2_{(1)}$=35.150	0.00	R
		B1CH	$\chi^2_{(1)}$=30.477	0.00	R
	Prevention Program: Might Work[e]	**B1**	$\chi^2_{(1)}$=35.333	0.00	R
		B1CH	$\chi^2_{(1)}$=37.575	0.00	R
	Prevention Program: Not Work[e]	**B1**	$\chi^2_{(1)}$=70.391	0.00	R
		B1CH	$\chi^2_{(1)}$=67.334	0.00	R
	C-4	**B1**	$\chi^2_{(1)}$=13.183	0.00	R
		B1CH	$\chi^2_{(1)}$=16.310	0.00	R
9. Ability/Willingness to Perform Task	E-7	**B1**	$\chi^2_{(1)}$=4.173	0.04	R
		B1CH	$\chi^2_{(1)}$=5.117	0.02	R

(a) The source is the question number in the main survey unless otherwise indicated; see Table 6.5. Refused/not sure/not ascertained categories have been set to missing for the source variables and excluded from the cross-tabulations.

(b) When there are many categories, as with income, it is appropriate to report statistics such as the gamma and Kendall's tau-b. Here, we report the gamma statistic, γ. For the other variables, we report the Pearson chi-squared statistic.

(c) The p-value is the probability level estimated for a Type-I error for a χ^2 statistic using a cross-tabulation of the choice measure and the recommended variable.

(d) Question C-1 was used to construct a (0,1) indicator/dummy variable for whether respondents felt oil spills off the Central Coast would cause "a lot more" (MOREHARM) or "a lot less" harm (LESSHARM) than that described in the survey.

(e) Question C-3 was used to construct a (0,1) indicator/dummy variable for whether respondents felt the prevention program would be "somewhat effective" (PMWORKS) or would be "not too effective" or "not effective at all" (PNOTWORK).

These cross-tabulations permit a simple test of association between respondents' choices and four different types of information.[178] The first type includes the characteristics and attitudes of respondents (*i.e.*, the first six categories in Table 6.6). Each of the variables in these six categories displays a significant association with the **B1** and **B1CH** choice measures. Variables with economic interpretations, such as income (D-12), as well as measures of activities

[178] The parameter estimates and z-statistics for the variables included in the construct validity model below (*see* Table 6.7) provide additional information about the extent to which each variable, controlling for the other variables in the equation, influences the percentage who voted *for* the program and in what direction this influence is exerted.

that might be related to the injured resources, such as participation in various forms of saltwater recreation (D-3 and D-5) and identifying bird species (D-4), all influence the choice measure. As expected, respondents with higher incomes and those who engage in activities related to the resources described in the survey are significantly more likely to vote *for* the program. The selection of the choice measure used in the cross-tabulation, **B1** or **B1CH**, does not influence this conclusion.

Respondents' environmental attitudes consistently relate to differences in the decisions about the program. Variables from survey questions (A-1b, A-1e, and A-2c) asked before the expected harm and prevention program is described and a later question (D-7) which asks for a general self-evaluation on an environmentalist scale all significantly influence respondents' choices in the expected directions. For example, those respondents who feel that protecting coastal areas from oil spills and those who identify themselves as environmentalists are both significantly more likely to vote *for* the program. We also indirectly measure respondents' attitudes towards big business by recording whether respondents volunteered that oil companies should pay all of the costs of the spill. The null hypothesis of no association is rejected; and, the cross-tabulation suggests that those who volunteered that oil companies should pay were significantly less likely to vote *for* the prevention program.

Variables in the *prior knowledge* and *distance to site* categories display the expected relationships with respondents' choices. Respondents indicating *prior knowledge*, *i.e.*, whether they were familiar with the five types of birds affected by past Central Coast spills and whether they had driven along the Central Coast on Highway 1, were more likely to vote *for* the program. *Distance to site*, measured here by a coastal proximity variable, also affected respondents' choices: respondents whose residences fall within the Primary Sampling Units (see Chapter 4) along the coast between the San Francisco Bay and the greater Los Angeles area

(CCOAST) were significantly more likely to vote *for*.

The second type of information summarized in the above table is respondents' *understanding of the task* as assessed by the interviewer. Those respondents having difficulty understanding either the harm caused by Central Coast oil spills or the prevention program (E-5) were not significantly more likely to vote *for* the program; however, those having difficulty understanding the voting question (E-6; less than 2% of the total sample) were more likely to vote *for*.

The third type of information summarized in Table 6.6 is *beliefs in the scenario*. As expected, respondents' perceptions of the effects of oil spills and of the effectiveness of the prevention program both affect the pattern of choices. For example, those respondents who thought that the effects of the oils spills were less severe than described and those who thought that the prevention program would not work were both significantly less likely to vote *for* the program. In addition, those who thought the tax would not be limited to just one year were less likely to vote *for*.

The fourth type of information summarized in the table is the respondent's *ability/willingness to perform the task*. The interviewers' evaluations of respondents' impatience to complete the interview offer an indirect gauge of respondent *willingness to perform the task*. The cross-tabulation suggests a significant association with impatient respondents tending to vote *not-for* the program.

The categories identified by the NOAA Panel may also be used as a basis for dividing the sample into subsamples; and separate WTP estimates may be computed for the various subsamples. For example, respondents who stated at D-1 that someone in their household had driven along the Central Coast on Highway 1 have a higher **B1CH** lower bound estimate than those who stated that someone had not ($88.93 versus $60.51; $t=4.44$, $p < 0.001$). Also, the

estimated mean derived from the choices of those respondents who indicated that oil spills cause less damage than described in the injury scenario should be significantly smaller than that estimated from the choices of those who did not indicate that oil spills cause less damage; the difference between the subsample means is highly significant ($41.58 versus $92.51; $t=7.40$; $p < 0.001$).

Very large differences between subsample means may be seen for questions relating to respondent preferences for the general class of resources protected by the program described in this study. For instance, the **B1CH** lower-bound estimate for those who said at A-1e that protecting coastal areas from oil spills was *extremely* important is $121.57 while the lower bound on the sample mean for those who said it was *not important at all* is $5.68 ($t=2.91$; $p < 0.001$). These relationships reinforce the test results derived from the examination of the cross-tabulations recommended by the NOAA Panel.

§ 6.6 Examination of Construct Validity Using a Multivariate Approach

It is possible to look at the simultaneous influence of multiple variables on respondent choices with respect to the oil spill prevention program. The standard way to do this is through the estimation of a multivariate choice function which relates respondent choices to various respondent characteristics. Such a function can also be used to demonstrate the *construct validity* of the CV results (Mitchell and Carson, 1989). Construct validity, one of the standard validity concepts widely accepted for use in evaluating models, refers to the degree to which a measure relates to other measures as predicted by economic theory; in this case, whether variation in the **B1CH** choice measure is systematically related to factors such as preferences for the object of

choice, the cost of program, and the ability to pay for it.[179] Other factors which may be economically relevant include measures of respondents' evaluations of the expected harm and the characteristics of the prevention program. For example, respondents who thought oil spills would cause more harm than that described in the survey should be more likely to vote *for* the program.

§ 6.6.1 Definition and Interpretation of Covariates in Choice Function

Economic theory suggests that B1AMT, the randomly assigned treatment variable, should have a negative coefficient and be a major determinant of a respondents' choices (**B1CH**) (Deaton and Muellbauer, 1980); that is the observed result here. By itself in a simple probit equation, the coefficient on B1AMT is negative and highly significant with a z-statistic of -8.67 ($p < 0.001$).[180]

Economic theory, however, only predicts that the percentage voting *for* should decline monotonically as the tax price increases. It does not suggest a specific functional form for this relationship. As a consequence, we allow for a flexible (*i.e.*, Box-Cox) transformation (Box-Cox, 1964; Greene, Greene and Seaks, 1995; Hanemann and Kanninen, forthcoming) on B1AMT in the multivariate choice function presented in Table 6.7.[181] In the formulation presented, the coefficient on B1AMT and the Box-Cox λ parameter are highly correlated masking their joint significance level.[182] A test of the joint significance of the two parameters

---

[179]There is a long history of estimating construct validity equations in CV studies, see, *e.g.*, Knetsch and Davis (1966). For an example involving oil spill prevention, see Carson *et al.* (1992).

[180] With **B1CH** as the dependent indicator variable, the simple probit model yields a constant term of 0.3483 (0.0578) and a slope coefficient on B1AMT of -0.0044 (0.0005), where the standard errors are in parentheses.

[181] A Weibull choice model rather than the Box-Cox probit model used in Table 6.7 yields the same basic results; *see* Appendix E.4, Table 1.

[182]This high correlation between λ and the B on B1AMT has long been noted in the biometrics literature on fitting dose-response models (Morgan, 1992).

can be performed using a likelihood ratio test. This test yields a $\chi^2_{(2)}=199.068$ which rejects the null hypothesis that β * B1AMT$^{\lambda}$ does not contribute to the model at $p < 0.001$.[183] Log-likelihood tests can also be used to compare the Box-Cox formulation against the $\lambda=1$ linear specification and the log specification: both of these simpler specifications are rejected in favor of the Box-Cox model ($\chi^2_{(1)}=10.678$ ($p=0.001$), and $\chi^2_{(1)}=4.462$ ($p=0.035$), respectively).

Table 6.7 Multivariate Analysis of Construct Validity: Probit Estimates for B1CH Choice Valuation Function

Variable	Coding	Parameter Estimate	Z-Statistic	p-value (two-sided)	Variable Mean
CONSTANT	Equals 1 for all respondents	-2.0265	-2.88	0.004	—
B1AMT	B1 tax amount	-0.0801	-1.48	0.138	86.67
λ(B1AMT)	Box-Cox parameter	0.3424	1.51	0.132	—
LINC1	Log of income if household income < $150,000; 0 otherwise	0.1945	3.19	0.001	9.00
LINC2	Log of income if household income ≥ $150,000; 0 otherwise	0.1573	2.65	0.008	0.44
NOTAX	Did not pay California taxes=1; 0 otherwise	0.2147	1.25	0.212	0.10
CCOAST	Resides in Central Coast PSU (807, 812, 813, and 814)=1; 0 otherwise	0.2187	2.31	0.021	0.45
COASTIP	A-1e protect coastal area very important or extremely important=1; 0 otherwise	0.5031	4.01	0.000	0.78
WILDIP	A-2c increase spending to protect wildlife very important or extremely important=1; 0 otherwise	0.5101	4.98	0.000	0.57
ENVIST	D-7 strong environmentalist or activist=1; 0 otherwise	0.3717	3.09	0.002	0.21

[183] Assuming λ was known *a priori* to be 0.3424, its estimated value in Table 6.7, the t-statistic on the transformed B1AMT would be -10.40 ($p < 0.001$).

Variable	Coding	Parameter Estimate	Z-Statistic	p-value (two-sided)	Variable Mean
LOWSPEND	Wants increased spending only on one or no programs (A-2a, A-2b, A-2d, and A-2e)=1; 0 otherwise	-0.4835	-2.13	0.033	0.07
PAYVEH	D-16 prefer tax vehicle over higher prices or indifferent=1; 0 otherwise	0.4993	5.00	0.000	0.40
HWY1	Traveled along the Central Coast on Highway 1=1; 0 otherwise	0.3328	2.35	0.019	0.88
FAMBIRD	Familiar with any of five types of birds often harmed in oil spills=1; 0 otherwise	0.2700	1.94	0.053	0.86
MOREHARM	C-1 oil spills more harmful than described=1; 0 otherwise	0.1867	1.74	0.083	0.35
LESSHARM	C-1 oil spills less harmful than described=1; 0 otherwise	-0.3298	-2.30	0.021	0.16
PMWORKS	C-3 expect program to be somewhat effective=1; 0 otherwise	-0.6403	-6.46	0.000	0.39
PNOTWORK	C-3 expect program to be not too effective or not effective at all=1; 0 otherwise	-1.5289	-7.18	0.000	0.08
PAYMORE	C-4 does not think will only have to pay special tax for one year=1; 0 otherwise	-0.2425	-2.56	0.011	0.44
PROTEST	Stated oil companies should pay for program or that oil companies would pass program costs on to consumers=1; 0 otherwise	-0.7283	-5.93	0.000	0.20
	N = 1085 Log(L) = -500.26 Pseudo R^2 = 0.335				

The other variables selected for inclusion in the choice model can be grouped into five broad categories: economic and demographic characteristics, preferences and attitudes, interest in and use of the affected natural resources, evaluations of the expected harm and prevention

program, and interpretations of the payment mechanism.[184] The model in Table 6.7 explains a substantial fraction of the variability in the choices made in the survey. Respondents with some sets of characteristics are predicted by the estimated choice model to be willing to pay less than one dollar while respondents with other sets of characteristics are predicted to be willing to pay as much as several hundred dollars.

The economic and demographic variables are LINC1, LINC2, NOTAX, and CCOAST.[185] The model in Table 6.7 allows income to have a different coefficient depending upon the level of household income.[186] Two income classes are identified, those below $150,000 (LINC1) and those of $150,000 and above (LINC2). The coefficients of both income terms are positive and statistically significant with p-values less than 0.01.

NOTAX is an indicator variable for households that reported they did not pay California income taxes in 1994. The coefficient on this variable is positive but not significant at conventional significance levels. However, the one-sided p-value, 0.106, is suggestive that those not paying taxes are willing to pay more than their characteristics would otherwise imply.

The demographic variable in the model is CCOAST, a qualitative variable identifying the

---

[184] Note that Table 6.7 reports p-values for two-sided hypothesis tests. In most instances, the hypothesis about the coefficient on a particular test is of the one-sided form (*e.g.*, a null hypothesis that respondents who do not think the program works are as likely to vote *for* the program as other respondents versus the alternative that they are less likely). For one-sided hypothesis tests, the reported (two-sided) p-values should be divided by 2.

[185] Missing values for income (n=86) have been replaced with an estimate based on the median income in the 1993 zip code, housing type, education, gender, race, age, and qualitative variables for the number of employed adults in the household; Tables 2 and 3 in Appendix E.4 present more detailed definitions of the variables included in the income prediction equation and the model for estimating income, respectively. Excluding from the sample the households who did not report income does not change the sign or significance of the income measure or the role of any other variables; *see* Table 4 in Appendix E.4. It does reduce the sample from 1085 to 999, so the p-values for some of the tests for relationships between these variables and respondents' choices necessarily decrease somewhat.

[186] Allowing the income coefficient to vary with the level of household income, the effect of the log of income on the probability of favoring the program is still positive and significant with a p-value of 0.006. This result holds regardless of the treatment of missing values for income. The one-variable income specification can be rejected in favor of the two-variable specification used here using a likelihood-ratio test ($\chi^2_{(1)}$=2.86, p=.09). If LINC1 is split into two categories, one consisting of those households with income greater than the median California household income and one consisting of those below, the estimated income effect in the second category is smaller but not significantly so.

respondent's location in relationship to the area of the natural resource injuries. Respondents living in CCOAST PSU's between San Francisco Bay and the greater Los Angeles area are significantly more likely to vote *for* the program than those in the rest of the state.

The preference and attitudes variables are COASTIP, WILDIP, ENVIST, LOWSPEND, and PAYVEH. The three preference variables, COASTIP, WILDIP, and ENVIST, are directly related to the environment. COASTIP is a qualitative variable identifying those respondents who, at A-1e, rated preventing oil spills in coastal area as "very important" or "extremely important." WILDIP, also a qualitative variable, identifies those respondents who, at A-2c, thought spending to protect wildlife is "very important" or "extremely important." ENVIST identifies individuals who, at D-7, considered themselves to be either strong environmentalists or environmental activists. The positive signs on the COASTIP, WILDIP, and ENVIST coefficients and the associated p-values suggest that those who support the relevant class of environmental programs and who identify themselves as environmentalists are significantly *more* likely to vote for the program.

The two attitude variables, LOWSPEND and PAYVEH, relate more generally to respondents' attitudes about government programs. LOWSPEND identifies those respondents who view increased spending as "not too important" or "not important at all" for at least three of four other programs asked about in question A-2 (*i.e.*, job training for the unemployed, shelters for the homeless, lifeguards at state beaches, and public transportation in Los Angeles).[187] PAYVEH is defined from respondents' evaluations of whether taxes are the appropriate way to pay for new programs to protect the environment. Both LOWSPEND and PAYVEH have the expected signs and are significant: those not favoring increased government

[187] Note that the other program asked about in the A-2 series, spending on prisons, was not included in LOWSPEND as its inclusion resulted in perfect failure (*i.e.*, all of the respondents who meet this more inclusive criterion voted *not for* the program).

spending are *less* likely to vote for the program and those favoring the use of government taxes to effect environmental improvements or who are indifferent between higher taxes and higher prices to effect environmental improvements are *more* likely to vote for the program.

The next two variables reflect interest and use of the affected natural resources: driving along the Central Coast on Highway 1 (HWY1) and familiarity with at least one of the five species of birds most often harmed by past oil spills (FAMBIRD). HWY1 and FAMBIRD are both positive and significant: individuals whose activities and knowledge are related to the Central Coast (HWY1) and the five species of birds (FAMBIRD) are *more* likely to vote for the program.

The next four variables are related to respondents' evaluations of the expected harm and the prevention program. Those respondents who thought that oil spills along the Central Coast over the next 10 years would cause more harm than that described in the survey (MOREHARM) should be and were more likely to vote *for* the program; and those who thought that oil spills would cause less harm (LESSHARM) should be and were more likely to vote *not-for* the program. The two coefficients in Table 6.7 have the expected sign and both effects are statistically significant. The effects of MOREHARM and LESSHARM on willingness to pay offset almost exactly.[188] Also, those who thought the program would be somewhat effective (PMWORKS) or not effective (PNOTWORK) should be and were less likely to vote *for* the program. Again, the effects are highly significant. Had all respondents thought that the program would be completely effective, the percentage of *for* votes would have been higher. These results on the perceived degree of harm prevented and the perceived program effectiveness provide strong within sample evidence that respondents are sensitive to the scope of the good

[188] The absolute value of the coefficient on LESSHARM is almost twice that of MOREHARM. However, the percentage of respondents giving a MOREHARM answer is more than double that of those giving a LESSHARM answer.

they were asked to value.

The final two variables relate to the respondent's interpretation of the payment mechanism. PAYMORE identifies those respondents who thought the tax payment might not be limited to just one year. PROTEST identifies those respondents who protested at any point during the interview that either the oil companies should pay for all of the program costs or that the oil companies would pass on their share of the costs to consumers in the form of higher gas and oil prices. Those who did not think they would have to pay the amount for only one year (PAYMORE) should be and were less likely to vote *for* the program as were those who protested the payment mechanism.

Other information on demographics, knowledge, and attitude/behavioral information were also considered in our evaluation of the construct validity model reported in Table 6.7. The two demographic variables age and gender were not significant determinants of choices when income, attitude, and program-evaluation variables were included in the model. However, in bivariate relationships with **B1CH**, age (negative relationship) and gender (positive relationship with female) were statistically significant. Other variables, positive and significant in bivariate relationships with **B1CH**, but not significant in the model in Table 6.7, include four variables related to the use of and interest in natural resources: D-3, saltwater boating or fishing; D-4, birdwatching; D-5, going to the beach; and D-6, watching television programs about animals and birds in the wild.

§ 6.6.2 A Cluster Analysis Interpretation

The multivariate regression results presented in Table 6.7 is the standard approach in the contingent valuation literature to demonstrating construct validity (Mitchell and Carson, 1989). That equation parsimoniously summarizes the dependence of the probability of a *vote for* the

program on the tax amount and various other covariates. An alternative approach, used particularly in marketing research, is cluster analysis, which partitions the sample into clusters of respondents based on the respondents' values for the covariates.[189] One can then look at the probability of a respondent voting *for* the program, given the particular cluster to which the respondent has been assigned.

A large number of clustering algorithms have been proposed (Hartigan, 1975; Kaufman and Rousseeuw, 1990). Perhaps the most commonly used clustering approach is a partitioning method known as k-means clustering. Given a predetermined number of clusters, k, the sample is divided into k distinct clusters which minimize the within cluster variation through the choice of the "multidimensional" centers of each of the k clusters and assigning particular observations to particular clusters. Each of the k cluster centers can be thought of as a representative respondent. The variables are typically normalized so that each variable plays an equal role in terms of the variance function being minimized.[190] The covariate values for each of the representative respondents are expressed as deviations from the average respondent in the sample as a whole.

The k-means cluster analysis is performed on the set of predictor variables used in the model presented in Table 6.7 with the exception of CONSTANT which does not vary across respondents and B1AMT which was randomly assigned to respondents.[191] The value of k

---

[189]The usual objective of cluster in marketing research is to "segment" potential consumers into groups which differ by one or more key characteristics (Lunn, 1986). These segments, often based on demographic variables, typically differ on the basis of their purchasing patterns or the marketing channels, *e.g.*, direct mail, which can most successfully be used to reach them.

[190]Normalization of a variable is accomplished by subtracting the variable's mean value from each observation in the data set and dividing by the variable's standard deviation. The normalized variable has a mean value of zero and a standard deviation of one. The value of a normalized variable is interpreted in terms of the number of standard deviations from the mean.

[191]LINC1 and LINC2 have been added together to form LINC, the log of the respondent's household income. There is no gain in the cluster approach to using two income variables as the clustering algorithm can perform its partitioning at any point along the income distribution.

chosen for this analysis is four.[192] With respect to the fit of the cluster solution, the sum of the within cluster variance is 88.5% of the sum of the variance of the sample; the pseudo F-statistic is 100.91; and the over-all R^2 measure is .21.

The four representative respondents, *i.e.*, the centers of the four clusters, are given in Table 6.8. Each cluster center takes on a value for each of the 16 predictor variables used in the cluster analysis. Positive values represent a cluster center for that variable which is larger than the average of that variable for the entire sample, and negative values represent the opposite. For instance, looking at the row for the COASTIP variable, the representative respondent for cluster A was more likely than the average respondent in the sample to think that protecting coastal areas was important. The representative respondent for cluster B was slightly more likely than the average respondent to think that protecting coastal areas was important; the representative respondent for cluster C was less likely than the average respondent; and the representative respondent for cluster D was much less likely. This comparison of the values of the normalized variables across the four clusters can be made for each variable. The k-means clustering algorithm assigns 521 respondents to cluster A, 105 respondents to cluster B, 389 respondents to cluster C, and 70 respondents to cluster D.

The first column of Table 6.8, which contains the cluster A's variable centers for each variable, presents a very clear impression of the type of respondent assigned to this cluster. They are more likely than the average respondent to believe that protecting coastal areas (COASTIP) and wildlife (WILDIP) is important and to identify themselves as strong environmentalists (ENVIST). They are somewhat more familiar with the specific resources

[192]The choice of k in a cluster analysis is largely dependent upon the purpose for which the analysis is intended and the nature of the data being clustered. Allowing too few clusters can suppress key detail in the data. Allowing too many clusters makes interpretation difficult and eventually will largely reproduce the regression results already provided in Table 6.7. We have chosen k equal to four as a compromise. Much of the same insight is gained *if* k is equal to three or k is equal to five.

Table 6.8 Cluster Centers

Variable	Cluster A	Cluster B	Cluster C	Cluster D
COASTIP	0.406	0.141	-0.436	-0.812
WILDIP	0.757	0.173	-0.931	-0.721
FAMBIRD	0.193	-0.210	-0.228	0.150
HWY1	0.077	-0.524	-0.012	0.275
ENVIST	0.241	0.141	-0.325	-0.198
CCOAST	0.051	0.109	-0.043	-0.303
MOREHARM	0.302	0.130	-0.347	-0.519
LESSHARM	-0.204	-0.169	0.269	0.276
PMWORKS	-0.097	-0.091	0.145	0.056
PNOTWORK	-0.212	-0.096	0.170	0.779
PAYMORE	-0.039	0.054	0.026	0.064
PAYVEH	0.220	0.223	-0.280	-0.417
LOWSPEND	-0.268	-0.154	-0.268	3.722
PROTEST	-0.083	-0.182	0.082	0.441
LINC	0.169	-1.783	0.206	0.271
NOTAX	-0.332	2.879	-0.332	0.002

(FAMBIRD, HWY1). They are also are more likely than the average respondent to believe that there would be more injury from oil spills (MOREHARM, LESSHARM) and more likely to think that the plan would be effective (PMWORKS, PNOTWORK). They are also more likely to prefer the tax payment vehicle (PAYVEH) and less likely to be unsupportive of other public programs (LOWSPEND) than the average respondent in the sample. They have higher incomes (LINC) than the average respondent in the sample and are less likely than the average respondent not to pay California income taxes (NOTAX). This is the general profile of a respondent who would be expected to be more likely to vote *for* the program than the average sample respondent. We can look at the implications of being assigned to this cluster in two ways. One way is to compare the percentage of respondents voting *for* in this cluster, 68.1%, compared to

the 48.8% percent voting *for* in the sample as a whole, or the 31.0% *for* in the sample excluding cluster A respondents.[193] The other way is to look at the Turnbull lower bound on the sample mean. Table 6.9 summarizes the information on a probability of a *for* vote and the Turnbull lower bounds for the sample as a whole and each of the four clusters. Cluster A respondents have a Turnbull lower bound mean of $121.53, over 40% higher than the sample as a whole ($85.39). Cluster A's respondents comprise half the sample and over two-thirds ($57.20) of the magnitude of the sample Turnbull lower bound mean ($85.39).

Table 6.9 Summary Statistics for Sample and Cluster

Cluster	Sample Size N	Percentage Voting For	Turnbull Lower Bound Mean	Standard Error of the Mean
Sample	1085	48.8	85.39	3.90
A	521	68.1	121.53	5.91
B	105	46.7	81.39	11.63
C	389	30.0	50.18	5.58
D	70	12.9	20.77	24.81

Looking down the column for cluster B in Table 6.8, a picture emerges of a set of respondents who have moderately stronger environmental preferences (COASTIP, WILDIP, ENVIST) than the average respondent in the sample but who have less familiarity with the resource depicted in the scenario (FAMBIRD, HWY1). The dominant characteristic of cluster B respondents, however, is their low income (LINC) and their much higher likelihood of not having paid California income taxes (NOTAX).[194] Cluster B respondents have a 46.7% probability of providing a *for* vote and have a Turnbull lower bound mean of $81.39. Both of

[193]Because the B1AMT's were assigned to respondents independently of their characteristics, and B1AMT was not used in determining the clusters, it is meaningful to look at the average probability of a *for* vote across the different clusters. Indeed, B1AMT is not significantly correlated with the cluster indicator variables and there is no significant difference in the average B1AMT for each cluster.

[194]The influence of respondents not paying California income taxes is considered in section 6.8.

these are close to the sample average probability.

Looking down the column for cluster C, the picture which emerges is that of a group of respondents who should have a lower probability of voting for the program. The representative respondent of this cluster has less preference for environmental amenities (i.e., COASTIP, WILDIP, ENVIST), less knowledge about the specific resources (FAMBIRD, HWY1), is more likely to think that the harm from a spill will be less than that described by the scenario (MOREHARM, LESSHARM), is more skeptical about the plan working (PMWORKS, PNOTWORK), and dislikes the payment vehicle (PAYVEH). The only positive factor with respect to a *for* vote is a higher income (LINC). Cluster C is the second largest cluster with 389 respondents. These respondents have a 30.0% probability of having a *for* vote and have a Turnbull lower bound mean of $50.18.

Looking down the column for cluster D, one sees a representative respondent who is fairly negative with respect to the environment (COASTIP, WILDIP, ENVIST). This respondent representative is more familiar than the average respondent with the specific resources (FAMBIRD, HWY1) but less likely to live anywhere along the coast between the greater Los Angeles area and the San Francisco Bay area (CCOAST). This respondent is more likely to be skeptical about the amount of harm done by oil spills (MOREHARM, LESSHARM) and believes that the proposed program is not likely to work (PNOTWORK). The representative respondent of this cluster dislikes the payment vehicle; but the most distinguishing characteristics of cluster D respondents is their dislike of spending on any government program (LOWSPEND) and their much greater proclivity to protest that the oil companies should be paying (PROTEST). Cluster D is the smallest cluster with 70 respondents. Given the above characteristics of the representative respondent of this cluster one would expect a low probability of a vote *for*; indeed, cluster D respondents have a 12.9% probability of voting *for* the program; and they have

a Turnbull lower bound mean of $20.77. As one might expect, none of the cluster D respondents are willing to pay the highest amount asked, $220, while over 65% of the *for* respondents at this dollar amount come from the cluster A respondents.

An equation similar to that in Table 6.7 may be created by substituting indicator variables of cluster membership for the predictor variables other than B1AMT. This probit model is reported in Table 6.10. The cluster A indicator variable has been absorbed into the CONSTANT term. Membership in cluster B, cluster C, or cluster D decreases the probability of a *for* vote. As one would expect from the information in Table 6.9, the parameter estimates become more negative as one goes from cluster B to C to D. Note that the parameter estimates on the cluster membership indicator variables are all highly significant.[195] The pseudo R-

Table 6.10 Box-Cox Model Using Cluster Indicators

Variable	Parameter Estimate	Z-Statistic	p-value (two-sided)	Variable Mean
CONSTANT	1.1716	6.85	.000	-----
B1AMT	-0.0526	-1.21	.226	86.67
λ(B1AMT)	0.4628	2.37	.021	-----
CLUSTER_B	-0.5245	-3.73	-000	0.10
CLUSTER_C	-1.0582	-11.67	.000	0.36
CLUSTER_D	-1.6827	-8.27	.000	0.06
		N=1085 Log(L)=-616.86 Pseudo R2=.180		

square of this model, is a little over half that of the much larger model reported in Table 6.7 reflecting the fact that this set of four clusters has substantial explanatory power but does not use all the relevant information contained in the model presented in Table 6.7.

[195]The basic set of Box-Cox parameters are, as in Table 6.7, highly correlated with each other; this correlation is responsible for their fairly small overall z-statistics. In a probit model of the cluster indicators with either B1AMT or log(B1AMT) as the stimulus variable, the z-statistic on the stimulus variable is over 9 ($p < .001$). The linear form of the model can be rejected using a likelihood ratio test at $p=.02$ and the log form of the model at $p=.01$.

§ 6.7 Sensitivity of WTP Estimate

In this section, using the construct validity equation reported in Table 6.7, we examine the effects of respondent assumptions which deviate in some way from the scenario presented in the survey instrument. The information about these assumptions comes from respondents' answers to the Section C questions about what they had in mind when they voted for or against the plan and from respondent answers to questions concerning the payment mechanism. We also looked at the sensitivity of the WTP estimate to possible protest responses identified by the PROTEST variable used in Table 6.7.

One may look at a number of different summary statistics in regard to shifts in the WTP distribution related to these deviations. In this section, we focus on the shift in the estimate of median WTP, $60.56,[196] because this statistic is fairly robust to statistical assumptions about the general shape of the underlying WTP distribution.[197]

The first possible shift we consider is related to those respondents who believed that the harm would be more or less than the harm described, one of two dummy variables—MOREHARM or LESSHARM—in the construct validity model has a value of one to represent the particular deviation from the desired perception. Setting the value of these two dummy variables to zero effectively forces the perceptions to the "same" harm category. This adjustment, however, does not change the estimate of median household willingness to pay by

---

[196] The median is the point in the WTP distribution above which 50 percent of the respondents are predicted to be willing to pay more and below which 50 percent are willing to pay less,

[197] In contrast, the estimate of mean WTP is quite sensitive to distributional assumptions and tends to be dominated by the assumption regarding a very small percentage of observations in the right tail of the distribution. A number of specific technical difficulties (Collins, 1991) are associated with inverting the Box-Cox model in Table 6.7 to get mean estimates, particularly with estimates close to zero. Setting predicted values which are close to zero equal to zero, the changes in the predicted Box-Cox means are generally similar in magnitude to the changes in the predicted medians reported in this section. The Turnbull estimate of the lower bound on the sample mean used in earlier sections of this chapter avoids all of these problems by using the highest amount asked as the maximum possible WTP estimate for a respondent. Unfortunately, while the Turnbull framework can be used to look at differences with respect to WTP based on a simple dichotomous split of respondents into two groups based their answer to one particular question, that framework does not generalize in a straightforward manner to the focus of this section—looking at the implications of changes in particular covariate values while holding other covariate values constant.

more than a few cents; as noted above, the combined MOREHARM and LESSHARM effect is almost exactly off-setting.

Another possible shift is that related to the perceived effectiveness of the program. Ideally, all respondents would have perceived the plan as being effective. One of two dummy variables—PMWORKS and PNOTWORK—in the construct validity model have a value of one if a respondent indicated that he or she believed that plan would not be either mostly or completely effective. Setting both of these dummy variables to zero effectively forces the perception that the plan is effective. This adjustment increases the estimated median willingness to pay by $58.86.

Some respondents believed that they would have to pay the special tax for more than one year. Ideally, all respondents would have perceived the tax as limited to a one-time payment. Setting the dummy variable, PAYMORE, to zero effectively forces the perception that the tax was a one-time payment. This adjustment increases the estimated median WTP by $18.25. A fourth possible shift is that related to protest responses. As noted in Table 6.7, PROTEST identifies those respondents who protested at any point during the interview that either the oil companies should pay for all of the program costs or that the oil companies would pass on their share of the costs to consumers in the form of higher gas and oil prices. Setting the PROTEST dummy variable to zero in the construct validity model forces out that consideration and increases the estimated median WTP by $23.77.

In sum, these shifts in the median WTP suggest that the overall effect of respondent assumptions that deviate from these four scenario features is a downward bias. Although the direction of the estimated changes in median WTP are likely to be a reliable indicator of the direction of the bias in the Turnbull estimate reported earlier, an extremely strong set of assumptions would be necessary to justify translating the absolute magnitude of these changes

into specific changes in the Turnbull estimate of the lower bound on the sample mean WTP.

§ 6.8 Correction for Non-Taxpayers

The payment vehicle used in this study—a one-time increase in California income taxes—presents a problem different in type from that of the potential divergences between respondent perceptions and the scenario presented: respondents not currently paying state income taxes do not necessarily treat the tax payment obligation in the same way as those respondents paying state income taxes. The sign and significance of NOTAX in the multivariate choice model suggests that this group of respondents is willing to pay more than other respondents with otherwise identical characteristics. Below, we treat the B1CH responses of those respondents who did not pay California income taxes in 1994 as *not-for* responses, the most conservative adjustment for this tendency, by re-coding the 108 respondents who did not pay California income taxes in 1994 and who voted *for* the program given the **B1CH** choice measure to *not-for* votes. This adjustment effectively sets the lower-bound estimate for this group of respondents to zero.[198] This choice measure is referred to as **B1CHNT** below.

Table 6.11 reports the distribution of **B1CHNT** by B1AMT and Table 6.12 reports the Turnbull estimate of the lower bound on the sample mean for this choice measure. As for the **B1** and **B1CH** choice measures, a $\chi^2_{(4)}$ test (71.98) for the **B1CHNT** measure rejects the hypothesis ($p < 0.001$) that responses are not sensitive to B1AMT. As shown in Table 6.12, the estimated lower bound on the sample mean for the **B1CHNT** choice measure is $76.45[199]

[198] *See* Appendix E.1.

[199] For the B1CHNT choice measure, the Turnbull estimate of the lower bound on the sample mean using the sample weights is $77.36, $0.91 higher than the unweighted sample estimate. The standard error of the weighted estimate is $3.73.

Table 6.11 B1CHNT Choice Measure by B1AMT

B1AMT	For	Not For
$5	62.10%	37.90%
$25	51.85%	48.15%
$65	45.23%	54.77%
$120	35.36%	64.64%
$220	25.44%	74.56%
	$\chi^2_{(4)} = 71.98$; $p < 0.001$	

with a standard error of $3.78.[200] This estimate, smaller than that from the **B1CH** choice measure ($85.39, with a standard error of $3.90), represents a conservative adjustment to the lower-bound estimate on the sample mean WTP.

Table 6.12 Turnbull Estimate of WTP Distribution and Lower Bound on the Sample Mean: B1CHNT Choice Measure [N=1,085]

Lower Bound of Interval	Upper Bound of Interval	Probability of Voting *For* at Upper Bound	Change in Density
$0	$5	0.6210	0.3790
$5	$25	0.5185	0.1025
$25	$65	0.4523	0.0662
$65	$120	0.3536	0.0987
$120	$220	0.2544	0.0992
$220	∞	0.0000	0.2544
	Log-Likelihood	-707.73	
	Estimate of lower bound on sample mean	$76.45	
	Standard error of the estimate	$3.78	

[200] The z-statistics for the five change in density parameters estimated by the model are 11.56, 2.17, 1.42, 2.06, and 2.17, respectively.

§ 6.9 Summary

In this chapter, after adjusting for non-taxpayers, we estimate a conservative lower bound on the average *ex ante* economic value of the oil spill prevention program to sample households of $76.45, with a standard error of $3.78. Two types of quantitative evaluations provide extensive quantitative evidence on the validity and reliability of the choice data collected in this survey, complementing the qualitative analysis presented in the previous chapter.

The first such evaluation consists principally of bivariate relationships in the form of cross-tabulations between the information variables recommended by the NOAA Panel and the **B1** and **B1CH** choice measures. Overall, the bivariate analysis provides support for the presence of relationships that economic theory suggests should influence respondents' choices regarding the prevention program.

The second evaluation considers the same issues in a more structured format using a conventional, multivariate choice model. A probit model is estimated to identify the determinants of **B1CH** choices. In each instance, the factors hypothesized to be associated with the choices are found to be consistent with prior expectations; and the relationships are statistically significant determinants of **B1CH**. Moreover, these effects are robust and generally do not change much with the specific coding of the variables involved. This construct validity equation is used to examine the effect that various adjustments would have on the WTP estimated by the model. In adjusting the WTP estimate to account for shifts due to perceptions of harm differing from those presented in the survey, the increase in WTP due to those perceiving more harm is almost exactly offset by the decrease in WTP due to those perceiving less harm. Adjustments for perceptions that the program would be less than effective, perceptions that the tax would not be one-time, and for protest responses, all result in substantially higher estimates.

REFERENCES

1. Arrow, Kenneth, Robert Solow, Paul R. Portney, Edward E. Leamer, Roy Radner, and Howard Schuman. "Report of the NOAA Panel on Contingent Valuation." *Federal Register* 58, 1993, 4601-4614.

2. Audience Research, Inc. *A Study of Outdoor Recreational Activities and Preferences of the Population Living in the Region of the Delaware River Basin*, Report prepared for the U.S. National Park Service, 1958.

3. Ayer, M., H.D. Brunk, G.M. Ewing, W.T. Reid, and E. Silverman. "An Empirical Distribution Function for Sampling with Incomplete Information." *Annals of Mathematical Statistics* 26, 1955, 641-647.

4. Bishop, Richard C. and Thomas A. Heberlein. "Measuring Values of Extra-Market Goods: Are Indirect Measures Biased?" *American Journal of Agricultural Economics* 61, no.5, 1979, 926-930.

5. Box, G.E.P. and D.R. Cox. "An Analysis of Transformation." *Journal of Royal Statistical Society Series B* 26, 1964, 211-52.

6. Cameron, Trudy Ann and Michelle D. James. "Efficient Estimation Methods for 'Closed-Ended' Contingent Valuation Surveys." *Review of Economics and Statistics* 69, no. 2, 1987, 269-276.

7. Carson, Richard T. "Contingent Valuation Surveys and Tests of Insensitivity to Scope," in *Determining the Value of Non-Marketed Goods: Economic, Psychological, and Policy Relevant Aspects of Contingent Valuation Methods*, edited by Raymond J. Kopp and Norbert Schwarz. Boston: Kluwer, forthcoming.

8. Carson, Richard T., Nicholas E. Flores and W. Michael Hanemann. "On the Creation and Destruction of Public Goods: The Matter of Sequencing." Discussion Paper 95-21. Department of Economics, University of California, San Diego, April, 1995.

9. Carson, Richard T., Nicholas E. Flores, Kerry M. Martin and Jennifer L. Wright. "Contingent Valuation and Revealed Preference Methodologies: Comparing the Estimates for Quasi-Public Goods." *Land Economics*, forthcoming.

10. Carson, Richard T., Nicholas E. Flores and Norman F. Meade. "Contingent Valuation: Controversies and Evidence." Paper presented at the Western Economic Association Meeting. San Diego, July, 1995.

11. Carson, Richard T., Nicholas E. Flores and Robert Cameron Mitchell. "The Theory and Measurement of Passive Use Value," in *Valuing Environmental Preferences: Theory and Practice of the Contingent Valuation Method in the US, EC, and Developing Countries*, edited by Ian J. Bateman and Ken G. Willis. New York: Oxford University Press, forthcoming.

12. Carson, Richard T., W. Michael Hanemann, Raymond J. Kopp, Jon A. Krosnick, Robert C. Mitchell, Stanley Presser, Paul A. Ruud, and V. Kerry Smith. *Prospective Interim Lost Use Value Due to DDT and PCB Contamination in the Southern California Bight: Volumes 1 & 2*, Report to the National Oceanic and Atmospheric Administration. La Jolla, CA: Natural Resource Damage Assessment, Inc., September, 1994.

13. Carson, Richard T., W. Michael Hanemann, Raymond J. Kopp, Jon A. Krosnick, Robert C. Mitchell, Stanley Presser, Paul A. Ruud, and V. Kerry Smith. "Temporal Reliability of Estimates from Contingent Valuation." Discussion Paper 95-37. Resources for the Future, August, 1995a.

14. Carson, Richard T., W. Michael Hanemann, Raymond J. Kopp, Jon A. Krosnick, Robert C. Mitchell, Stanley Presser, Paul A. Ruud, and V. Kerry Smith. "Referendum Design and Contingent Valuation: The NOAA Panel's No-Vote Recommendation." Discussion Paper 96-05. Resources for the Future, November, 1995b.

15. Carson, Richard T., Norman F. Meade and V. Kerry Smith. "Passive Use Values and Contingent Valuation: Introducing the Issues." *Choices* 8, no. 2, 1993, 4-8.

16. Carson, Richard T. and Robert Cameron Mitchell. "The Value of Clean Water: The Public's Willingness to Pay for Boatable, Fishable, and Swimmable Quality Water." *Water Resources Research* 29, no. 7, 1993a, 2445-2454.

17. Carson, Richard T. and Robert Cameron Mitchell. "The Issue of Scope in Contingent Valuation Studies." *American Journal of Agricultural Economics* 75, 1993b, 1263-1267.

18. Carson, Richard T. and Robert Cameron Mitchell. "Sequencing and Nesting in Contingent Valuation Surveys." *Journal of Environmental Economics and Management* 28, 1995, 155-173.

19. Carson, Richard T., Robert C. Mitchell, W. Michael Hanemann, Raymond J. Kopp, Stanley Presser, and Paul A. Ruud. *A Contingent Valuation Study of Lost Passive Use Values Resulting from the Exxon Valdez Oil Spill*, Report to the Attorney General of the State of Alaska, November 1992.

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

21. Carson, Richard T., Leanne Wilks and David Imber. "Valuing the Preservation of Australia's Kakadu Conservation Zone." *Oxford Economic Papers* 46, 1994, 727-749.

22. Carson, Richard T., Jennifer L. Wright, Nancy J. Carson, Anna Alberini, and Nicholas E. Flores. *A Bibliography of Contingent Valuation Studies and Papers*. Natural Resource Damage Assessment, Inc., January, 1995.

23. Ciriacy-Wantrup, S.V. "Capital Returns from Soil-Conservation Practices." *Journal of Farm Economics* 29, 1947, 1181-1196.

24. Collins, Sean. "Prediction Techniques for Box-Cox Regression Models." *Journal of Business and Economics Statistics* 9, 1991, 267-277.

25. Council of American Survey Research Organizations. "On the Definition of Response Rates." Report of the CASRO Task Force on Completion Rates. Port Jefferson, NY, June, 1982.

26. Davis, Robert K. "The Value of Big Game Hunting in a Private Forest," in *Transactions of the Twenty-Eighth North American Wildlife and Natural Resources Conference*. Washington, DC: Wildlife Management Institute, 1963.

27. Deaton, Angus and John Muellbauer. *Economics and Consumer Behavior*. New York: Cambridge University Press, 1980.

28. Desvousges, William H., Alicia R. Gable, Richard W. Dunford and Sara P. Hudson. "Contingent Valuation: The Wrong Tool for Damage Assessment." *Choices* 8, no. 2, 1993, 9-11.

29. Diamond, Peter A. and Jerry A. Hausman. "Contingent Valuation: Is Some Number Better than No Number?" *Journal of Economic Perspectives* 8, no. 4, 1994, 45-64.

30. Finney, D.J. "The Estimation of the Parameters of Tolerance Distributions." *Biometrika* 36, 1949, 239-256.

31. Flores, Nicholas E. and Richard T. Carson. "The Relationship Between the Income Elasticities of Demand and Willingness to Pay." Paper presented at the Association of Environmental and Resource Economists Meeting. Washington, DC, January, 1995.

32. Freeman, A. Myrick, III. *The Measurement of Environmental and Resource Values: Theory and Methods*. Washington, DC: Resources for the Future, 1993.

33. Greene, William H., Laura Knapp Greene, and Terry G. Seaks. "Estimation the Functional Form of the Independent Variables in Probit Models." *Applied Economics* 27, 1995, 193-196.

34. Grice, H.P. "Logic and Conversation," in *Syntax and Semantics 3: Speech Acts*. New York: Academic Press, 1975.

35. Haab, Timothy C. and Kenneth E. McConnell. "Referendum Models and Negative Willingness to Pay: Alternative Solutions." Unpublished paper. Department of Agricultural and Resource Economics, University of Maryland, April, 1995.

36. Hanemann, W. Michael. "Discrete/Continuous Models of Consumer Demand." *Econometrica* 52, 1984, 541-561.

37. Hanemann, W. Michael. "Willingness to Pay and Willingness to Accept: How Much Can They Differ?" *American Economic Review* 81, no. 3, 1991, 635-647.

38. Hanemann, W. Michael. "Preface," in *Pricing the European Environment*, edited by Ståle Navrud. New York: Oxford University Press, 1992.

39. Hanemann, W. Michael. "Valuing the Environment Through Contingent Valuation." *Journal of Economic Perspectives* 8, no. 4, 1994, 19-43.

40. Hanemann, W. Michael. "Contingent Valuation and Economics," in *Environmental Valuation: Some New Perspectives*, edited by Ken Willis and John Corkindale. Wallingford, Oxon, UK: CAB International, forthcoming.

41. Hanemann, W. Michael and Barbara J. Kanninen. "Statistical Considerations in CVM," in *Valuing Environmental Preferences: Theory and Practice of the Contingent Valuation Method in the US, EC, and Developing Countries*, edited by Ian J. Bateman and Ken G. Willis. Oxford: Oxford University Press, forthcoming.

42. Harrison, Glenn W. and Bengt Kriström. "On the Interpretation of Responses in Contingent Valuation Surveys," in *Current Issues in Environmental Economics*, edited by P.O. Johansson, B. Kriström and K.G. Mäler. Manchester: Manchester University Press, 1995.

43. Hartigan J.. *Clustering Algorithms*. New York: Wiley, 1995.

44. Hausman, Jerry A., ed. *Contingent Valuation: A Critical Assessment*. Contributions to Economic Analysis. Amsterdam: North-Holland Publishing Co., 1993.

45. Havemann, Joel. "Why Oil Spills are Increasing." *Los Angeles Times*, Part A:1, 26 March 1993.

46. Kaufman, Leonard and Peter J. Rousseeuw. *Finding Groups in Data: An Introduction to Cluster Analysis*. New York: Wiley, 1990.

47. Knetsch, Jack L. and Robert K. Davis. "Comparisons of Methods for Resource Evaluation," in *Water Research*, edited by Allen V. Kneese and Stephen C. Smith. Baltimore: Johns Hopkins, 1966.

48. Kriström, Bengt. "A Non-Parametric Approach to the Estimation of Welfare Measures in Discrete Response Valuation Studies." *Land Economics* 66, no. 2, 1990, 135-139.

49. Lehmann, E.L. *Testing Statistical Hypotheses*, Second Edition. New York: John Wiley and Sons, 1986.

50. Lodge, Milton, Kathleen M. McGraw and Patrick Stroh. "An Impression-Driven Model of Candidate Evaluation." *American Political Science Review* 83, no. 2, 1989, 397-419.

51. Lunn, Tony. "Segmenting and Constructing Markets," in *Consumer Market Research Handbook*, edited by Robert M. Worcester and John Downham. New York: North-Holland, 1986.

66. Westat, Inc. *State Policy Study — Main Study Interviewer's Manual*. Rockville, MD, January, 1995.

Appendix A
Main Study Survey Instrument

START INTERVIEW TIME: ☐☐:☐☐ A.M. P.M.

SECTION A

A-1. Let's start by talking for a moment about some issues in California. Some may not be important to you, others may be.

SHOW CARD A

First, (READ X'd ITEM). Is this issue not important at all to you personally, not too important, somewhat important, very important, or extremely important? (READ EACH ITEM, BEGINNING WITH X'd ITEM; CIRCLE ONE CODE FOR EACH; REPEAT ANSWER CATEGORIES AS NECESSARY.)

	NOT IMPORTANT AT ALL	NOT TOO IMPORTANT	SOMEWHAT IMPORTANT	VERY IMPORTANT	EXTREMELY IMPORTANT	N/S
( ) a. Improving education in California elementary and secondary schools	1	2	3	4	5	8
( ) b. Reducing air pollution in California cities	1	2	3	4	5	8
( ) c. Maintaining local library services	1	2	3	4	5	8
( ) d. Reducing crime	1	2	3	4	5	8
( ) e. Protecting coastal areas from oil spills	1	2	3	4	5	8
( ) f. Finding ways to reduce state taxes	1	2	3	4	5	8

A-2. The State of California spends tax money on many programs for many different purposes. I'm going to read a list of some of these programs. For each one, I would like you to tell me how important it is to you that the State continue to spend money on it.

SHOW CARD A AGAIN

First, (READ X'd ITEM). (READ EACH ITEM, BEGINNING WITH X'd ITEM; CIRCLE ONE CODE FOR EACH; REPEAT ANSWER CATEGORIES AS NECESSARY.)

	NOT IMPORTANT AT ALL	NOT TOO IMPORTANT	SOMEWHAT IMPORTANT	VERY IMPORTANT	EXTREMELY IMPORTANT	N/S
( ) a. Providing job-training for the unemployed	1	2	3	4	5	8
( ) b. Providing shelters for the homeless	1	2	3	4	5	8
( ) c. Protecting wildlife	1	2	3	4	5	8
( ) d. Providing lifeguards at state beaches	1	2	3	4	5	8
( ) e. Providing public transportation for Los Angeles	1	2	3	4	5	8
( ) f. Building new state prisons	1	2	3	4	5	8

These are just a few of the programs the State of California currently spends tax money on. Proposals are sometimes made to the State for new programs; but the State does not want to start any new programs unless taxpayers are willing to pay the additional cost for them.

One way for the State to find out about this is to give people like you information about a program so that you can make up your own mind about it. STOP

Your views are useful to State decision makers in deciding what, if anything, to do about a particular situation. STOP

In interviews of this kind, some people think that the program they are asked about is not needed; STOP others think that it is. We want to know what you think. STOP

A-3. Have you ever been interviewed before about whether the State should start a new program?

YES	1
NO	2
N/S	8

In the past, people have been asked about various types of programs. In this interview, I am going to ask you about a program that would prevent harm from oil spills off one part of the California coast. STOP

I will begin with important background information. Then I will ask you whether you think this particular program is worthwhile and why you feel the way you do. STOP

SHOW CARD B

Along the California coast, there are three different types of shoreline.

- The areas shown here in green are mostly saltwater marshes.
- The areas shown in brown are mostly rocky shoreline.
- And, the areas in yellow are mostly sandy beaches. STOP

A-4. Have you visited any of these three types of California shoreline in the last 12 months?

YES	1	
NO	2	**(SKIP TO A-6)**
N/S	8	

A-5. And, which ones are those? **(CIRCLE THOSE MENTIONED)**

SALTWATER MARSH	1
ROCKY SHORE	2
SANDY BEACHES	3

A-6. Each year, tankers and barges carrying oil make about 3,000 trips in and out of California harbors and along the Central Coast.

- Large oil tankers called super-tankers deliver their cargo to storage tanks and oil refineries in the San Francisco Bay and in the Greater Los Angeles area.
- Small tankers and barges transport various types of refined oil back and forth along the 500 miles of coastline between San Francisco and the L.A. area.

SHOW CARD B AGAIN

Tankers and barges occasionally run into things like underwater rocks, other ships, or pipelines, and spill some of their oil into the water.

Unless the spill is very small, the oil can harm wildlife.

After an oil spill, the company that caused it must pay to clean up as much oil as possible from both the water and the shoreline. STOP

Over the years, the State has taken various steps to prevent harm from oil spills. Recently, steps have been taken to set up programs to prevent harm from spills in the San Francisco Bay and in the L.A. area. The State wants to know whether people think this would be worth doing for the Central Coast. STOP

As you can see here, most of the Central Coast is rocky shoreline with some scattered sandy beaches.

Oil spills that harmed wildlife have happened here every few years. State and university scientists were asked to provide information about the effects of these past spills.

SHOW CARD C

This drawing shows the types of wildlife that Central Coast spills have harmed. It shows five types of birds and other types of small animals that live in or near the water. Take your time to look it over.

STOP UNTIL R IS FINISHED LOOKING AT CARD

SHOW CARD C AGAIN

The five birds shown here are the types of birds that past spills have harmed the most.

A-7. Do you happen to be familiar with any of these birds?

YES 1

NO . . . . . 2 } (SKIP TO A-9)
N/S 8

A-8. Which ones? **(CIRCLE THOSE MENTIONED)**

WESTERN GULL 1
PACIFIC LOON 2
RHINOCEROS AUKLET 3
COMMON MURRE 4
BRANDT'S CORMORANT 5

A-9. According to scientists, none of these birds are in any danger of becoming extinct.

The number next to each bird shows how many of them live in California. For example, there are about 290,000 Pacific Loons and 130,000 Western Gulls.

All five types of birds also live in other States. STOP

BOX 1

IF R ASKS IF WESTERN GULLS ARE THE SAME AS SEA GULLS, CHECK HERE ☐ AND SAY:

Western gulls are one of a dozen types of sea gulls.

SHOW CARD C AGAIN

Whenever oil washes up on the shoreline along the Central Coast, it harms many small animals and saltwater plants. Some are shown here.

They include clams, sea stars, crabs, mussels, kelp, and other seaweed.

None of these are in any danger of becoming extinct. STOP

Marine mammals—such as whales, seals, and dolphins—are not usually affected by the oil because they generally leave the area when a spill occurs. Fish also leave the area and are not affected. STOP

BOX 2

IF R ASKS WHAT HAPPENS TO SEA OTTERS OR MENTIONS HE/SHE THOUGHT SEA OTTERS WERE ALSO AFFECTED IN OIL SPILLS, CHECK HERE ☐ AND SAY:

Like other marine mammals, sea otters usually leave the area when a spill occurs. They have not usually been affected by past Central Coast spills.

Recently, the federal government passed a new law to help reduce the number of oil spills.

> Ten years from now, all oil tankers and barges will be required to have two outer hulls instead of the single-hull most of them have now. Double-hulls provide much more protection against oil leaking after an accident.

However, it will take ten years before all single-hulled tankers and barges can be replaced. Until then, spills are expected to happen every few years along the Central Coast, just as they have in the past, unless something is done. STOP

SHOW CARD D

This shows the total amount of harm to wildlife that state and university scientists expect will happen in the Central Coast area over the next ten years. It is based on studies scientists have made of past spills in this area. STOP

SHOW CARD D AGAIN

In the next ten years:

- scientists expect that a total of about 12,000 birds of various types will be killed by oil spills off the Central Coast.

  In addition, about 1,000 more birds are expected to be injured but survive.

- Also, many small animals and saltwater plants are likely to be killed along a total of about ten miles of shoreline. (STOP)

The harm from an oil spill is not permanent. Over time, waves and other natural processes break down the oil in the water and on the shoreline.

- Typically, within ten years or less after a spill, there will be as many of the affected birds as before the spill.

  The small animals and saltwater plants in the affected area recover somewhat faster, in about five years or less. (STOP)

A-10. Is there anything more that you would like to know about the harm oil spills are expected to cause off the Central Coast over the next ten years?

YES 1

NO . . . . . 2 **(SKIP TO A-11)**

A-10A. What is that?

_____

_____

_____

_____

A-11. If taxpayers think it is worthwhile, the State could prevent this harm by setting up a prevention program for this part of the coast. This program would be similar to those successfully used by other states, such as the State of Washington. It would last for ten years, until all tankers and barges have double-hulls. STOP

This program would do two things.

First, it would help prevent oil spills from occurring.

Second, if an oil spill does occur, it would prevent the oil from spreading and causing harm. STOP

Here is how a Central Coast program would prevent spills from occurring.

SHOW CARD E

- Oil spill prevention and response centers would be set up in three different locations along this part of the coast.

  Specially-designed ships, called escort ships, would be based at each center.

- An escort ship would travel alongside every tanker and barge as it sails along the Central Coast. This would help prevent spills in this area by keeping the tankers and barges from straying off-course and running into underwater rocks, other ships, or pipelines. STOP

SHOW CARD F

If any oil were spilled, here's how the program would keep it from spreading and causing harm.

- The crew of the escort ship would quickly put a large floating sea fence into the water to surround the oil. To keep it from spreading in rough seas, this fence would extend 6 feet above and 8 feet below the surface of the water.

- Then skimmers, like the one shown here, would suck the oil from the surface of the water into storage tanks on the escort ship.

  Other ships would be sent from the nearest prevention and response center to aid in the oil recovery and clean-up. STOP

A-12. Is there anything more you would like to know about how this prevention program would work?

YES 1

NO . . . . . 2 **(SKIP TO A-13)**

A-12A. What is that?

_____

_____

_____

BOX 3

IF R ASKS ABOUT HOW PROGRAM WOULD BE PAID FOR OR ABOUT PROGRAM COST, CHECK HERE ☐ AND SAY:

I will come to that in just a moment.

A-13. The money to pay for this program would come from both the tax-payers and the oil companies. Because individual oil companies cannot legally be required to pay the cost of setting up the program, all California households would pay a special one time tax for this purpose.

> This tax money would pay for providing the escort ships and setting up the three oil spill prevention and response centers along the Central Coast.

Once the prevention program is set up, all the expenses of running the program for the next ten years would be paid by the oil companies.

> This money would come from a special fee the oil companies would be required to pay each time their tankers and barges were escorted along the Central Coast.

Once the federal law goes into effect ten years from now, all tankers and barges will have double-hulls and this program would be closed down. STOP

BOX 4

IF R ASKS ABOUT PROGRAM COST, CHECK HERE ☐ AND SAY:

> I will come to that in just a moment.

IF R SAYS OIL COMPANIES SHOULD PAY ALL COSTS, CHECK HERE ☐ AND SAY:

> The State cannot legally force individual oil companies to pay for setting up the program. However, the oil companies can be required to pay a special fee each time one of their ships is escorted along the Central Coast. These fees will pay to keep the program operating over the next ten years.

We are interviewing people to ask how they would vote on this Central Coast prevention program if it were put on the ballot in a California election. STOP

There are reasons why you might vote for setting up this program and reasons why you might vote against it.

SHOW CARD G

The program would prevent harm from oil spills in the Central Coast area during the next ten years. Specifically, the program would:

- prevent the deaths of about 12,000 birds as well as the deaths of many small animals and saltwater plants along about 10 miles of shoreline, and

  prevent 1,000 more birds from being injured. STOP

On the other hand,

- the number of birds and other wildlife it would protect is small in comparison to their total numbers, and none are endangered. STOP

  Your household might prefer to spend the money to solve other social or environmental problems instead. STOP

  Or, the program might cost more than your household wants to spend for this.

REMOVE CARD G

SECTION B

If the Central Coast prevention program were put into place, it would cost your household a total of $5. You would pay this as a special one time tax added to your next year's California income tax.

B-1. If an election were being held today, and the total cost to your household for this program would be $5, would you vote for the program or would you vote against it?

FOR 1

AGAINST . . . . . 2 **(SKIP TO B-4)**

N/S . . . . . 8 **(SKIP TO B-5)**

BOX 5

IF R SAYS OIL COMPANIES SHOULD PAY ALL COSTS, CHECK HERE ☐ AND SAY:

The State cannot legally force individual oil companies to pay for setting up the program. However, the oil companies can be required to pay a special fee each time one of their ships is escorted along the Central Coast. These fees will pay to keep the program operating over the next ten years.

B-2. People have different reasons for voting for the Central Coast prevention program. What would the program do that made you willing to pay for it? **(PROBE: Was there something specific that the program would do that made you willing to pay for it?)**

_____

_____

_____

_____

_____

B-3. Occasionally, people vote for the program because they are concerned that oil spills may somehow harm human health. Suppose human health was definitely not affected and the program would only prevent harm to birds, small animals, and saltwater plants. Would you vote for or against the program if it cost your household $5?

VOTE "FOR" 1

VOTE "AGAINST" 2

N/S 8

SKIP TO SECTION C

B-4. Did you vote against the program because it isn't worth that much money to you, or because it would be somewhat difficult for your household to pay that amount, or because of some other reason?

ISN'T WORTH THAT AMOUNT 1

DIFFICULT TO PAY 2

OTHER REASON (SPECIFY) 3

_____

_____

_____

SKIP TO SECTION C

B-5. Could you tell me why you aren't sure about how you would vote? **(PROBE)**

_____

_____

_____

_____

SECTION C

Please think back to a few moments ago when I asked you whether you would vote for or against the program.

SHOW CARD H

C-1. At that time, did you think the harm from oil spills in the Central Coast over the next ten years would be about the same as that shown here, or a lot more or a lot less?

SAME 1
A LOT MORE 2
A LOT LESS 3
OTHER (SPECIFY) 4

_____

_____

_____

_____

N/S 8

SHOW CARD I

C-2. How serious did you consider this amount of harm to be . . .

Not serious at all 1
Not too serious 2
Somewhat serious 3
Very serious, or 4
Extremely serious? 5
N/S 8

SHOW CARD J

C-3. Did it seem to you that the prevention program I told you about would be completely effective at preventing harm from Central Coast oil spills, mostly effective, somewhat effective, not too effective, or not effective at all?

COMPLETELY EFFECTIVE 1
MOSTLY EFFECTIVE 2
SOMEWHAT EFFECTIVE 3
NOT TOO EFFECTIVE 4
NOT EFFECTIVE AT ALL 5
N/S 8

C-4. When you decided how to vote, did you think your household would have to pay the special tax for the program for one year or for more than one year?

ONE YEAR 1
MORE THAN ONE YEAR 2
N/S 8

C-5. Thinking about <u>everything</u> I have told you during this interview, <u>overall</u> did it try to push you to vote one way or another, <u>or</u> did it let you make up your <u>own</u> mind about which way to vote?

PUSHED ONE WAY OR ANOTHER 1

LET ME MAKE UP OWN MIND . . . . . 2 } **(SKIP TO SECTION D)**
N/S 8

C-6. <u>Which</u> way did you think it pushed you?

VOTE FOR THE PROGRAM 1
VOTE AGAINST THE PROGRAM 2
OTHER (SPECIFY) 3

_____

_____

_____

_____

N/S 8

C-6A. <u>What</u> was it that made you think that? **(PROBE: "Can you be more specific about what you have in mind?" "Anything else?")**

_____

_____

_____

SECTION D

Now I would like to ask you a few questions about your household's recreational activities.

D-1. Has anyone in your household ever driven along the Central Coast on Highway 1, the coast highway?

YES 1

NO . . . . . 2 } **(SKIP TO D-3)**
N/S 8

D-2. And, was this in the last five years?

YES 1
NO 2
N/S 8

D-3. In the past five years, has anyone in your household gone saltwater boating or saltwater fishing?

YES 1
NO 2
N/S 8

D-4. Does anyone in your household like to identify different species of birds?

YES 1
NO 2
N/S 8

D-5. During this past summer, about how many times did people in your household go to beaches anywhere along the California coast . . .

Never, ... 1
Once or twice, ... 2
Three to ten times, or 3
More than ten times? 4
N/S 8

SHOW CARD K

D-6. How often do you personally watch television programs about animals and birds in the wild . . .

Very often, ... 1
Often, ... 2
Sometimes, ... 3
Rarely, or 4
Never? 5
N/S 8

SHOW CARD L

D-7. Do you think of yourself as an . . .

Environmental activist, a 1
Strong environmentalist, a 2
Somewhat strong environmentalist, a 3
Not particularly strong environmentalist, or 4
Not an environmentalist at all? 5
N/S 8

Now, just a few questions about your background.

D-8. First, in total, how many years have you lived in California?

☐☐ YEARS

D-9. In what month and year were you born?

☐☐/☐☐
MONTH YEAR

D-10. What is the highest year of school you completed or the highest degree you received?

THROUGH 8th GRADE 01
9th, 10th, 11th, 12th GRADE (NO DIPLOMA) 02
HIGH SCHOOL EQUIVALENT (for example, GED) 03
HIGH SCHOOL GRADUATE (DIPLOMA) 04
SOME COLLEGE BUT NO DEGREE 05
ASSOCIATES DEGREE IN OCCUPATIONAL
OR VOCATIONAL PROGRAM 06
ASSOCIATES DEGREE IN ACADEMIC
PROGRAM 07
BACHELOR'S DEGREE (for example, BA, AB, BS) 08
MASTER'S DEGREE (for example, MA, MS, MEng,
MEd, MSW, MBA) 09
PROFESSIONAL SCHOOL DEGREE (for example, MD,
DDS, DVM, LLB, JD) 10
DOCTORATE DEGREE (for example, PhD, EdD) 11
REFUSED 97
N/S 98

D-11. Currently, how many adults in your household, including yourself, work for pay?

0 1 2 3 4 5 6 (OR MORE)

NUMBER WHO CURRENTLY WORK FOR PAY

SHOW CARD M

D-12. I'd like you to think about the income received last year by everyone in your household. Adding together all income for everyone in your household, which letter on this card best describes your household's total income for last year—1994—before taxes? Please include wages or salaries, social security or other retirement income, child support, public assistance, business income, and all other income.

LETTER A 01
LETTER B 02

LETTER C 03
LETTER D 04
LETTER E 05
LETTER F 06
LETTER G 07
LETTER H . . . . . 08 **(SKIP TO D-14)**
LETTER I 09
LETTER J 10
LETTER K 11
N/S 98
REFUSED 97

D-13. Did anyone in your household pay any California income taxes for last year, 1994, by having taxes withheld from wages, retirement income, or other money received, or has anyone in your household sent, or intend to send, tax money for last year to the State with a tax form?

YES 1
NO 2
N/S 8

SHOW CARD N

D-14. When you look ahead to the next few years, do you see your personal financial situation getting . . .

Much better, ...	1
A little better, ...	2
Staying about the same, ...	3
Getting a little worse, or	4
Much worse?	5
OTHER	6
N/S	8

D-15. Now that we're almost at the end of the interview and you have been able to think a bit more about the situation, I'd like to give you a chance to review your answer to the voting question.

You were asked if you would vote for or against a program that would prevent the harm that I showed you earlier on this card.

SHOW CARD O

If an election were being held today, would you vote <u>for</u> the program or <u>against</u> the program if it cost your household a one-time tax payment of $5?

FOR	1
AGAINST	2
N/S	8

D-16. There are different ways for people to pay for new programs to protect the environment.

One way is for the government to pay the cost. This will raise everyone's taxes. STOP

Another way is for businesses to pay the cost. This will make prices go up for everyone. STOP

If you had to choose, would you prefer to pay for new environmental programs ...

Through higher taxes, or 1
Through higher prices? 2
EITHER ONE/DON'T CARE WHICH 3
NEITHER 4
N/S 8

SHOW CARD P

D-17. Generally speaking, how much confidence do you have in the California state government? Would you say . . .

A great deal, ... 1
Some, ... 2
Hardly any, or 3
None? 4
N/S 8

END INTERVIEW TIME: ☐☐:☐☐ A.M. P.M.

D-18. What is your full name and phone number, in case my supervisor wants to check my work? (RECORD FULL NAME AND PHONE NUMBER ON RECORD OF ACTIONS. **DO NOT RECORD IT HERE**.)

RECORDED ON RECORD OF ACTION 1
NO PHONE 2
REFUSED 7

BOX S

IS THE SCREENER FOR THIS CASE COMPLETED?

YES . . . 1 THANK RESPONDENT FOR COOPERATION.

NO . . . . 2 SAY TO RESPONDENT: "I have just a few more questions I need to ask about the other adults in your household. Let me verify that there are (number from AS-1) people 18 or older living in this household."

RETURN TO SCREENER (PG 2). ASK QUESTIONS S-3 THROUGH S-5, ENUMERATION TABLE, AND S-12. THEN, THANK RESPONDENT FOR COOPERATION.

SECTION E
INTERVIEW EVALUATION QUESTIONS

PLEASE NOTE THE FOLLOWING ABOUT THE RESPONDENT BY CIRCLING THE NUMBER OF THE CORRECT RESPONSE:

E-1. SEX

MALE 1
FEMALE 2

E-2. RACE

WHITE, NOT HISPANIC 1
WHITE, HISPANIC 2
BLACK, NOT HISPANIC 3
BLACK, HISPANIC 4
ASIAN 5
OTHER (SPECIFY) 6
_____

E-3. TRANSFER THE RESPONDENT'S ZIP CODE FROM THE ADDRESS LABEL ON THE CALL RECORD FOLDER: ☐☐☐☐☐

E-4. What was the reaction of the respondent as you read **A-3 through A-13?** (This is the descriptive material including the maps and drawings).

		EXTREMELY	VERY	SOMEWHAT	SLIGHTLY	NOT AT ALL	N/S
a.	How distracted was the respondent?	1	2	3	4	5	8
b.	How attentive was the respondent?	1	2	3	4	5	8
c.	How interested was the respondent?	1	2	3	4	5	8

E-5. Did the respondent say anything suggesting that he or she had any difficulty understanding either the harm caused by Central Coast oil spills or the prevention program?

YES 1

NO . . . . . 2 **(SKIP TO E-6)**

E-5A. Please describe the difficulties.

_____

_____

_____

E-6. Did the respondent have any difficulty understanding the voting question, B-1?

YES 1

NO . . . . . 2 **(SKIP TO E-7)**

E-6A. Please describe the difficulties.

_____

_____

_____

E-7. When you asked B-1, did you feel the respondent was impatient to finish the interview?

YES 1

NO . . . . . 2 } **(SKIP TO E-8)**
NOT SURE 8

E-7A. How impatient was the respondent?

VERY IMPATIENT 1
SOMEWHAT IMPATIENT 2
A LITTLE IMPATIENT 3
NOT VERY IMPATIENT 4
NOT SURE 8

E-8. How serious was the consideration the respondent gave to the decision about how to vote?

EXTREMELY SERIOUS 1
VERY SERIOUS 2
SOMEWHAT SERIOUS 3
SLIGHTLY SERIOUS 4
NOT AT ALL SERIOUS 5
NOT SURE 8

E-9. Not counting you and the respondent, was anyone age 13 or older present when the respondent voted?

YES 1

NO . . . . . 2 **(SKIP TO E-10)**

OTHERS CAME IN AND OUT 3

E-9A. Do you think the other person(s) affected how the respondent voted or don't you know?

YES 1
NO 2
DON'T KNOW 8

E-10. Do you have any other comments about this interview?

_____

_____

_____

_____

Appendix A
Survey Show Cards and Graphics

1. Not important at all

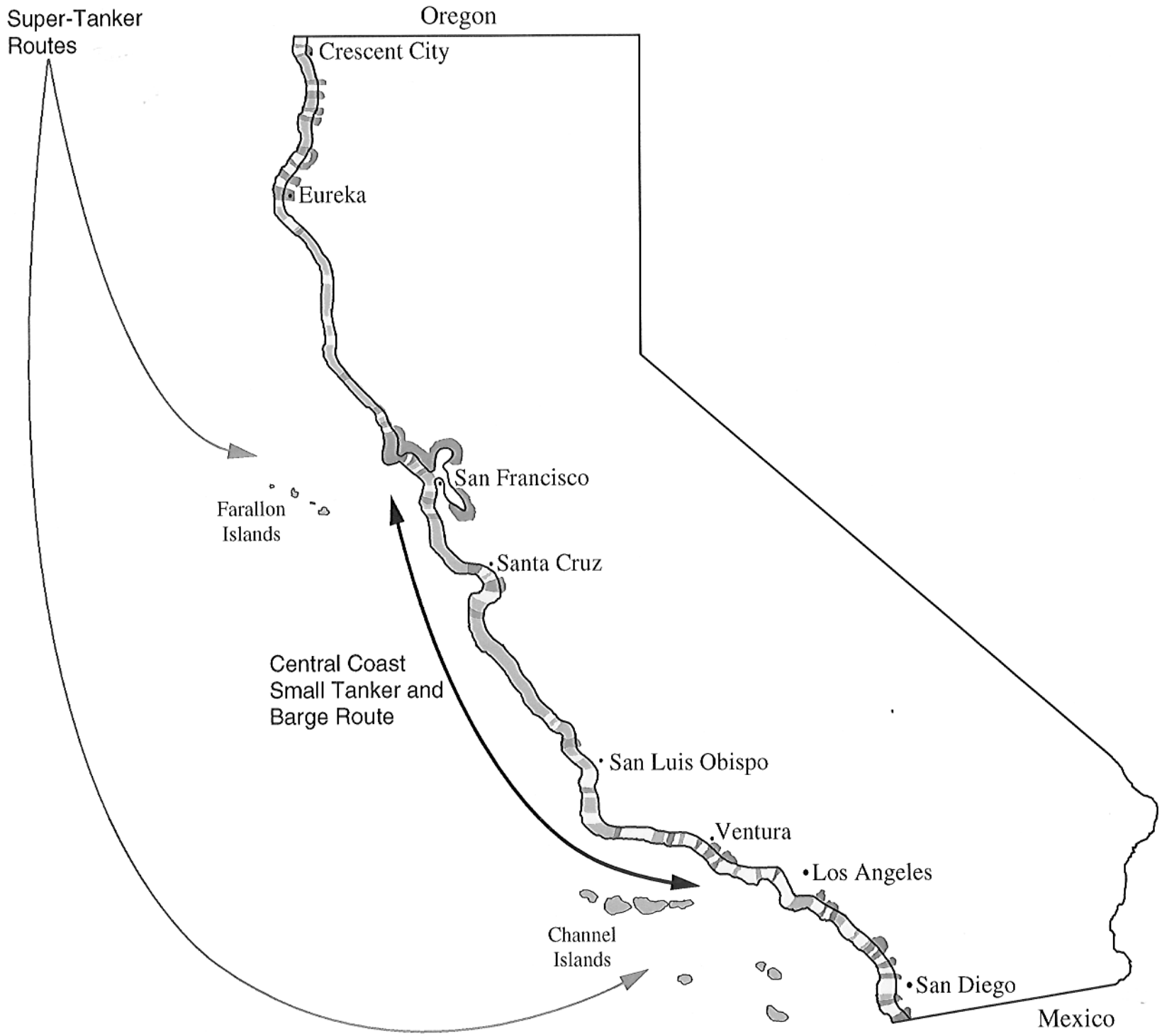
2. Not too important

3. Somewhat important

4. Very important

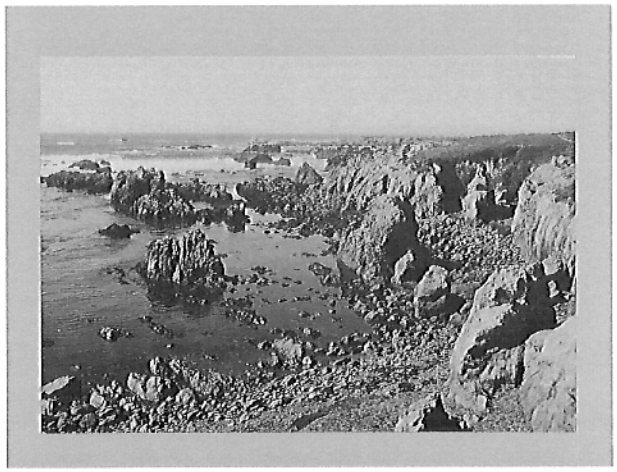
5. Extremely important

CARD A

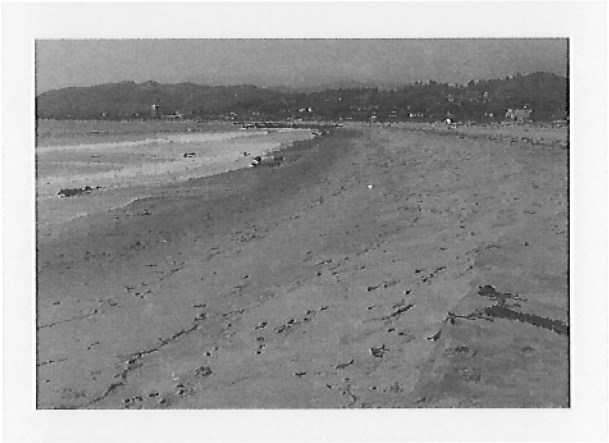

Super-Tanker
Routes
Oregon
Crescent City
Eureka
San Francisco
Farallon
Islands
Santa Cruz
Central Coast
Small Tanker and
Barge Route
San Luis Obispo
Ventura
Los Angeles
Channel
Islands
San Diego
Mexico

CARD B

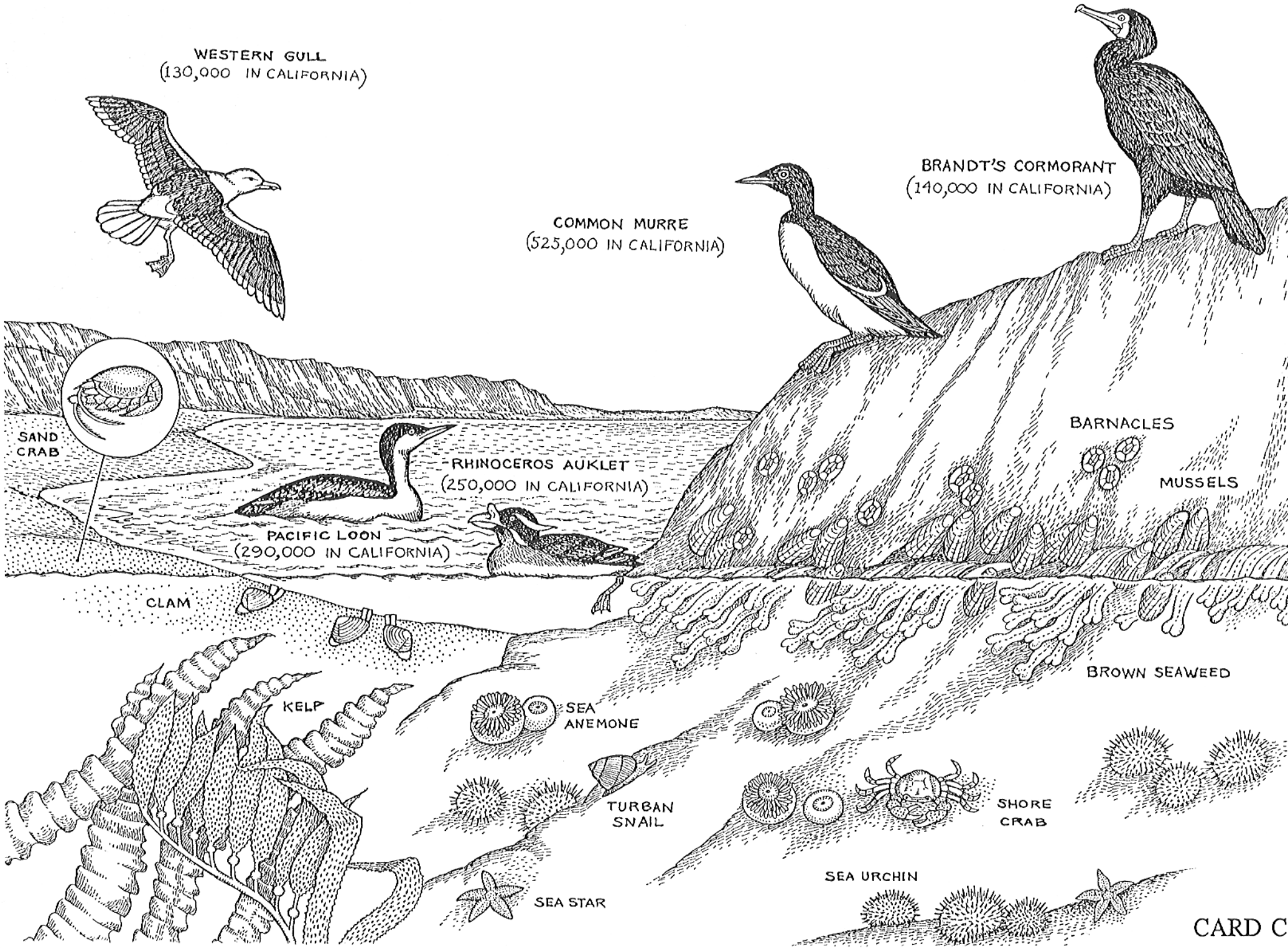
Saltwater marsh



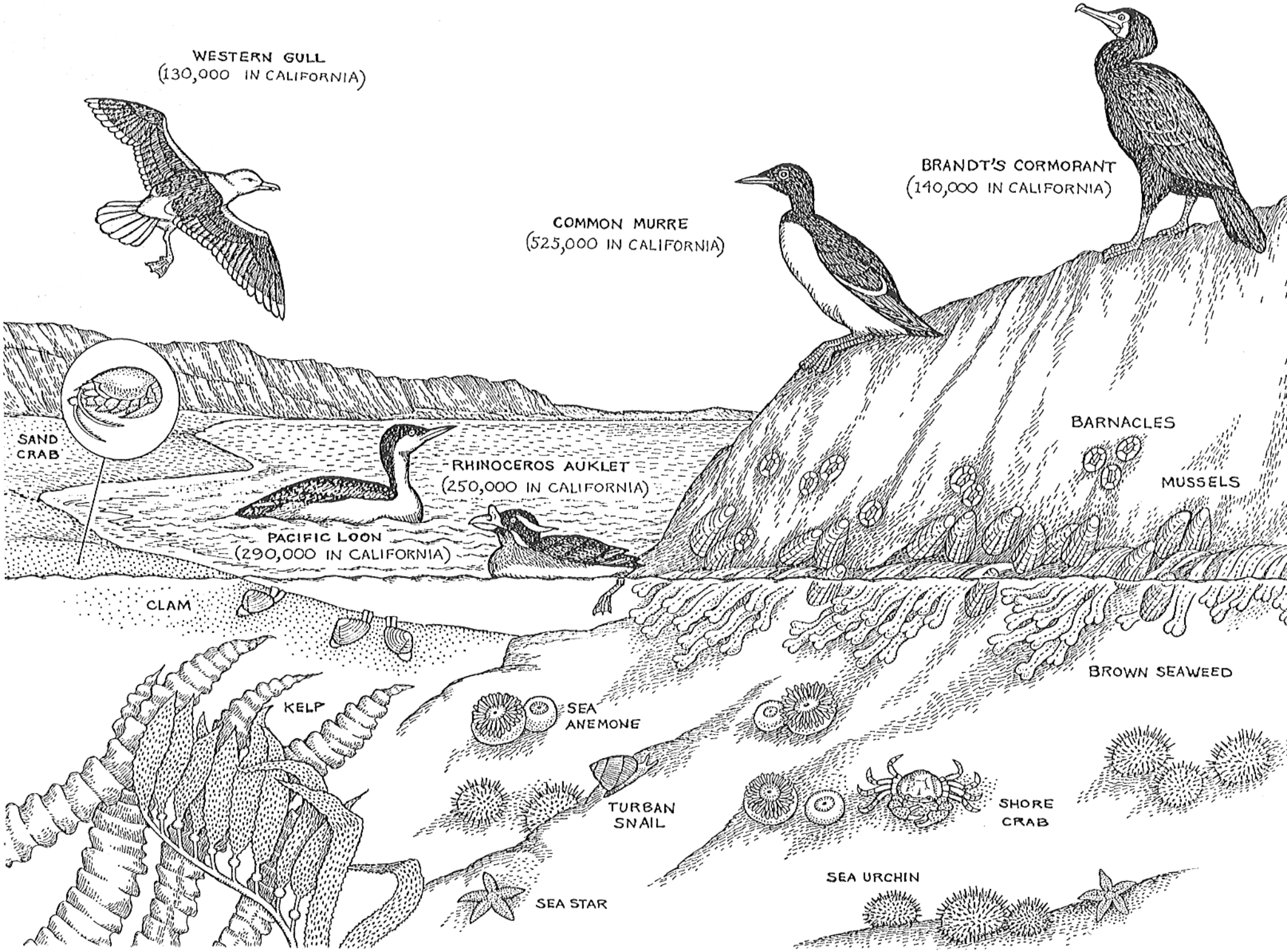
Rocky shoreline



Sandy beach


WESTERN GULL
(130,000 IN CALIFORNIA)
COMMON MURRE
(525,000 IN CALIFORNIA)
BRANDT'S CORMORANT
(140,000 IN CALIFORNIA)
SAND CRAB
RHINOCEROS AUKLET
(250,000 IN CALIFORNIA)
PACIFIC LOON
(290,000 IN CALIFORNIA)
BARNACLES
MUSSELS
CLAM
BROWN SEAWEED
KELP
SEA ANEMONE
TURBAN SNAIL
SHORE CRAB
SEA URCHIN
SEA STAR

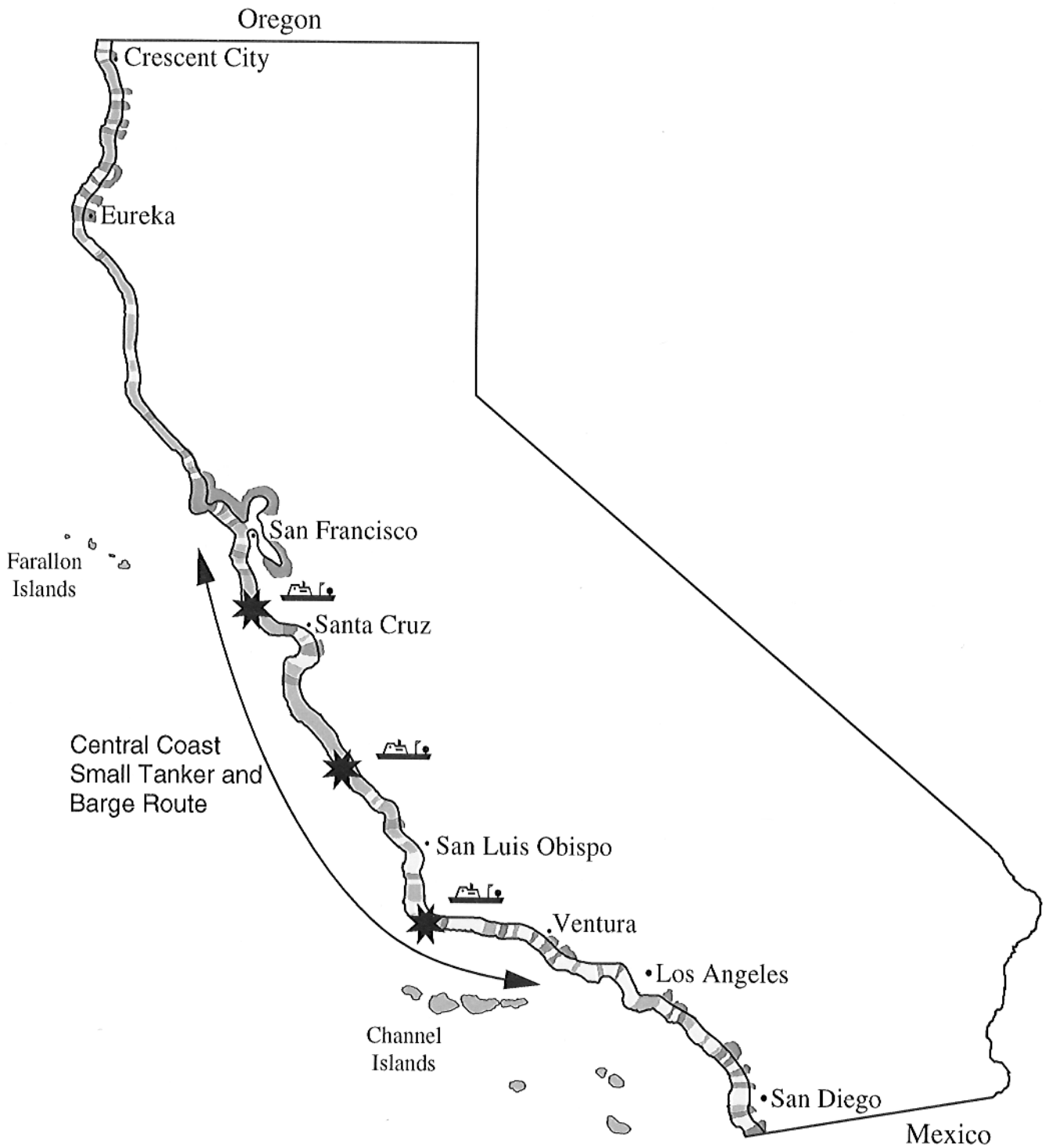
CARD C


WESTERN GULL
(130,000 IN CALIFORNIA)
BRANDT'S CORMORANT
(140,000 IN CALIFORNIA)
COMMON MURRE
(525,000 IN CALIFORNIA)
SAND CRAB
RHINOCEROS AUKLET
(250,000 IN CALIFORNIA)
PACIFIC LOON
(290,000 IN CALIFORNIA)
BARNACLES
MUSSELS
CLAM
BROWN SEAWEED
KELP
SEA ANEMONE
TURBAN SNAIL
SHORE CRAB
SEA URCHIN
SEA STAR

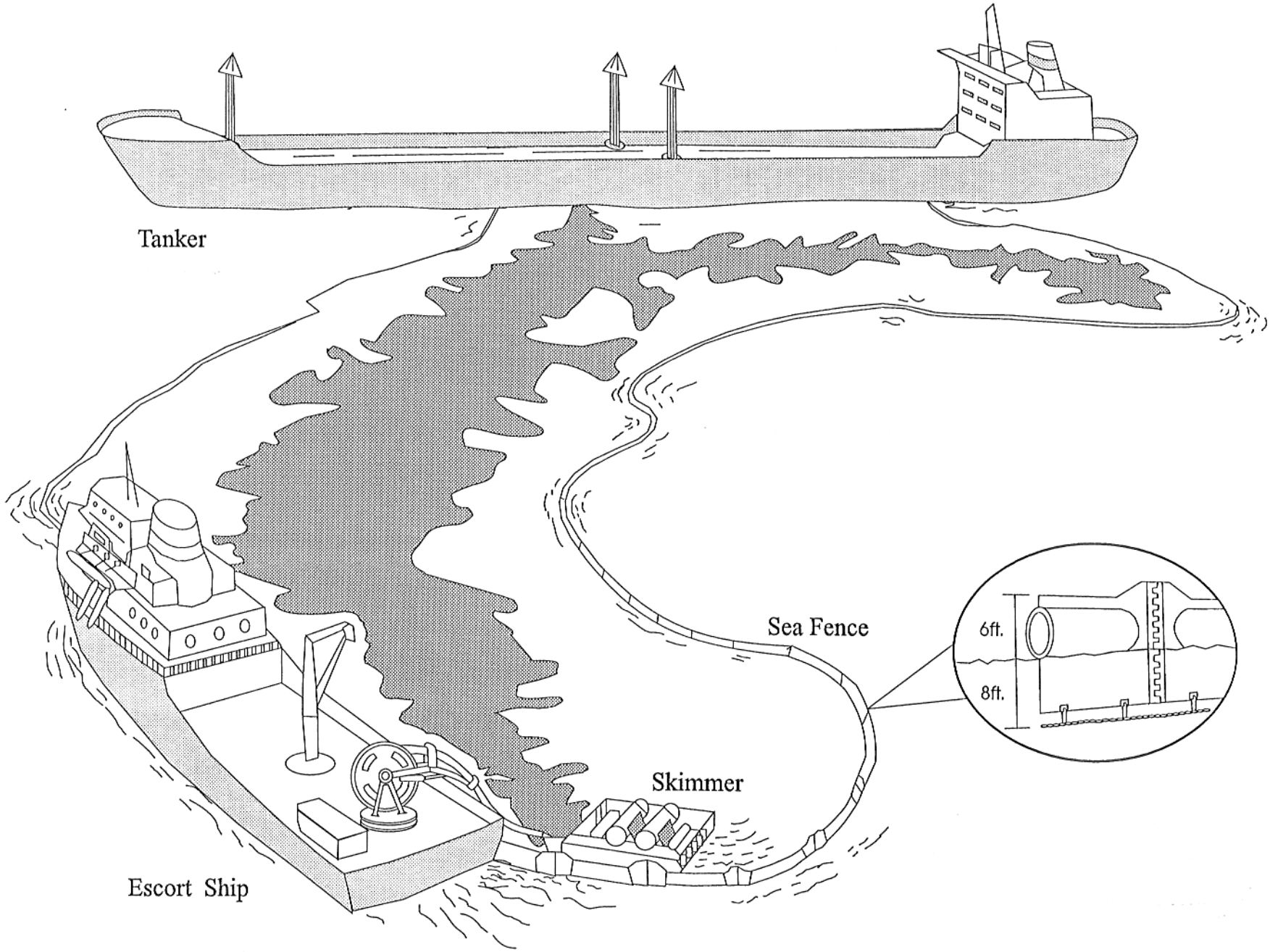
Expected Harm from Oil Spills off Central Coast

TOTAL HARM OVER NEXT 10 YEARS	TIME TO RECOVER AFTER EACH SPILL
12,000 birds killed **1,000 birds** injured	10 years
Many small animals and saltwater plants killed along **10 miles** of shoreline	5 years

CARD D


Oregon
Crescent City
Eureka
San Francisco
Farallon Islands
Santa Cruz
Central Coast Small Tanker and Barge Route
San Luis Obispo
Ventura
Los Angeles
Channel Islands
San Diego
Mexico

CARD E



CARD F

REASONS *FOR*

Over the next 10 years, the Central Coast prevention program would:

prevent deaths of about **12,000 BIRDS** as well as deaths of many small animals and saltwater plants along about **10 MILES** of shoreline, and

prevent **1,000 MORE BIRDS** from being injured.

REASONS *AGAINST*

The **NUMBER** of birds and other wildlife the program would protect is **SMALL** in comparison to their **TOTAL** numbers, and none are endangered.

Your household might prefer to spend the money on **OTHER PROBLEMS** instead.

The program might **COST MORE** than your household wants to spend for this.

Expected Harm from Oil Spills off Central Coast

TOTAL HARM OVER NEXT 10 YEARS	TIME TO RECOVER AFTER EACH SPILL
12,000 birds killed **1,000 birds** injured	10 years
Many small animals and saltwater plants killed along **10 miles** of shoreline	5 years

CARD H

Expected Harm from Oil Spills off Central Coast

TOTAL HARM OVER NEXT 10 YEARS	TIME TO RECOVER AFTER EACH SPILL
12,000 birds killed **1,000 birds** injured	10 years
Many small animals and saltwater plants killed along **10 miles** of shoreline	5 years

1. Not serious at all

2. Not too serious

3. Somewhat serious

4. Very serious

5. Extremely serious

1. Completely effective

2. Mostly effective

3. Somewhat effective

4. Not too effective

5. Not effective at all

CARD J

1. Very often

2. Often

3. Sometimes

4. Rarely

5. Never

CARD K

1. Environmental activist

2. Strong environmentalist

3. Somewhat strong environmentalist

4. Not particularly strong environmentalist

5. Not an environmentalist at all

CARD L

Total Before-Tax Income in 1994 for Everyone in Your Household

(Please include all income sources, including social security or other retirement, etc., for everyone in your household.)

A. Under $10,000
B. $10,000 to $19,999
C. $20,000 to $29,999
D. $30,000 to $39,999
E. $40,000 to $49,999
F. $50,000 to $59,999
G. $60,000 to $69,999
H. $70,000 to $79,999
I. $80,000 to $99,999
J. $100,000 to $149,999
K. $150,000 or more

CARD M

1. Much better

2. A little better

3. About the same

4. A little worse

5. Much worse

CARD N

Expected Harm from Oil Spills off Central Coast

TOTAL HARM OVER NEXT 10 YEARS	TIME TO RECOVER AFTER EACH SPILL
12,000 birds killed **1,000 birds** injured	10 years
Many small animals and saltwater plants killed along **10 miles** of shoreline	5 years

CARD O

1. Great deal

2. Some

3. Hardly any

4. None