HOW PEOPLE EVALUATE BUDGET PRIORITIES:
A CONJOINT ANALYSIS APPROACH

By

Chia-Nan Yeh

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Political Science

1998

ABSTRACT

HOW PEOPLE EVALUATE BUDGET PRIORITIES: A CONJOINT ANALYSIS APPROACH

By

Chia-Nan Yeh

The purpose of this study is to examine how people evaluate budget priorities by applying a method called *conjoint analysis*. Specifically, we evaluate how people make tradeoffs among four spending areas: education, welfare and health, crime and public safety, and natural resources and environmental control while holding other programs and tax rate constant. Traditional survey questions usually are addressed in a "wish list" type of format with unrealistic choices among incomplete sets of budget options, thus they actually measure people's "desires" rather than their "true preferences." In contrast, conjoint analysis asks respondents to make more "real world" decisions by forcing them to trade off competing interests and hence could offer researchers a way to collect more accurate information about voters' true preferences.

Conjoint analysis is a tradeoffs exercise in which respondents are forced to make preference decisions based on the relative importance of one service or policy over another. Conjoint analysis is defined as a method which identifies the joint effect of two or more independent variables on the ordering of a dependent variable.

Our findings reveal that respondents' policy preferences from traditional surveys and conjoint estimates are quite different; sometimes respondents' preference orders for a particular area even go in completely opposite direction. In the area of overall spending, for instance, traditional surveys show that voters most prefer a decrease in overall spending but conjoint estimates indicate that the status quo is virtually voters' first choice. The implication from our findings is that if policymakers make policy decisions based on inaccurate and incomplete information from traditional surveys, these policies cannot reflect peoples' true preferences and hence the quality of democracy might be discounted. Furthermore, in order to illustrate the applicability of conjoint analysis, simulations have been performed to predict the gubernatorial election and evaluate current welfare reform in Michigan. Lastly, reliability tests demonstrate that our conjoint estimates yield answers at least as reliable with previous studies on conjoint reliability.

To the memory of my brother, Chia-Lin

ACKNOWLEDGMENTS

I would like to thank those who made intellectual, emotional, and financial contributions to me. First, I wish to express my gratitude to my father, Lien-Fa Yeh, my mother, Pao-Lien Lu, my sister, Chia-Mei Yeh, and my parents in-law Ching-Tai Chen and Chiu-Fang Chao for their financial and emotional support.

I thanks to the members of my committee - Larry Heimann, Thomas Hammond, Paul Abramson, Paula Kearns, and Brian Silver. I sincerely appreciate Larry Heimann, who wore the hats of committee chair, employer, and friend. This dissertation would not been completed without his advice, assistance, and encouragement. I would also like to thank Thomas Hammond, who became my chair in the final stage, for his patience and guidance to correct my lousy writing and providing me insightful suggestions. In addition, I want to thank Roger Calantone of the business school for his suggestions on designing conjoint plan. Moreover, I gratefully acknowledge the assistance of Seon-Jou Kang, Brandon Prins, John Davis and Mark Hurwitz for helping me collect student data. Among my Taiwanese friends, I want to express my gratitude to Lu-Huei Chen, Jong-Tian Wang, Ching-Lung Chang, Min-Chieh Lin, Chien-Chin Chen for their encouragement and insightful suggestions.

Lastly, special thanks go to my wife, Nancy Chen, who shared everything with me during the process of writing my dissertation and now have to prepare for delivering our first baby. I appreciate your support more than I can say. I think my graduation will be the best gift for our first baby in this coming New Year.

TABLE OF CONTENTS

LIST OF TABLES ix

LIST OF FIGURES xi

INTRODUCTION 1

The Importance of Studying Budget Priorities 1

A New Survey Tool: Conjoint Analysis 4

Objectives of This Study 5

Organization of This Study 7

LITERATURE REVIEW 8

Public Opinion and Policy 8

Public Attitudes Toward Government Spending 11

Explanations of the Inconsistencies Between Policymakers and the Public Toward Government Spending 12

Existing Approaches to Evaluate Budget Priorities 14

Summary 24

THE LOGIC OF CONJOINT ANALYSIS 26

The Concept of Conjoint Analysis 26

Compositional vs. Decompositional Approaches 28

The Theory of Conjoint Analysis: An Information Integration Theory Approach 30

Design and Analysis of Conjoint Analysis 35

Why Conjoint Analysis Works 49

The Advantages and Limits of Conjoint Analysis 52

Classification of Conjoint Analysis 54

Current Applications of Conjoint analysis 56

Summary 56

RESEARCH DESIGN 58

Step by Step Procedure of Conducting a Conjoint Analysis 58

Summary 74

RESULTS AND DISCUSSIONS 76

Interpreting Conjoint Results from Voter Sample 76

Interpreting Conjoint Results from Student Sample 90

Comparison Between Traditional Survey Questions and Conjoint Results 98

Summary 103

GROUP DIFFERENCES IN CONJOINT ANALYSIS 105

Gender Differences 106

Generational Differences 112

Religious Differences 118

Party Differences 123

Children 128

Home Ownership 132

Summary 136

CONJOINT APPLICATIONS: SIMULATION OF GUBERNATORIAL ELECTION AND POLICY EVALUATION 137

Simulation of Gubernatorial Election in Michigan 137

Welfare Policy Evaluation 146

Summary 149

RELIABILITY AND VALIDITY TESTS 150

Why Reliability and Validity Check are so Important 150

Definitions of Reliability and Validity 153

Methods and Designs 157

Results and Discussions 159

Summary 173

SUMMARY AND CONCLUSIONS 175

The Need of New Survey Tool 175

Empirical Findings by Using Conjoint Analysis 177

Potential Applicability of the Conjoint Analysis 181

APPENDIX: CONJOINT SURVEY FORMS FOR STUDENT AND VOTER 183

BIBLIOGRAPHY 193

LIST OF TABLES

Table 3.1: The Relationship among a Profile, Attributes, and Levels 27

Table 3.2: The 2×2×2 Design and the Two One-Half Fractions 43

Table 3.3: A Graeco-Latin Square for the 2×3 Design 45

Table 3.4: A Full Profile with 2×3 Design in LINMAP Estimation 47

Table 4.1: Alternative Data Collection Methods 65

Table 5.1: Orthogonal Array for Voters (First Round of Survey) 78

Table 5.2: Part-worth Utilities from OLS and LINMAP (Voters) 80

Table 5.3: Overall Utilities of 10 Profiles for Voters (OLS) 88

Table 5.4: Goodness of Fit Measures for OLS and LINMAP 90

Table 5.5: Part-worth Utilities from OLS and LINMAP (Students) 91

Table 5.6: Overall Utilities of 10 Profiles for Students (OLS) 95

Table 5.7: Comparison of Overall Utilities Between Voters and Students 96

Table 5.8: Goodness of Fit Measures for OLS and LINMAP (Students) 98

Table 5.9 : Comparison of Budget Priorities Between Traditional Surveys and Conjoint Analysis (Voters) 100

Table 5.10 : Comparison of Budget Priorities Between Traditional Surveys and Conjoint Analysis (Student) 103

Table 6.1: Gender Differences in 4 Major Policy Areas 109

Table 6.2: Generational Differences in 4 Major Policy Categories 115

Table 6.3: Religious Differences in 4 Major Policy Categories 120

Table 6.4: Party Differences in 4 Major Policy Categories 125

Table 6.5: Children Differences in 4 Major Policy Categories 129

Table 6.6: Differences Between Home Owners and Renters in 4 Major Policy Categories 133

Table 7.1: Simulation Results for Gubernatorial Election (All Voters) 143

Table 7.2: Simulation Results for Gubernatorial Election (Independent Voters) 145

Table 7.3: Simulation Results for Welfare Policy Evaluation 147

Table 7.4: Simulation Results for Maximizing Voters' Utilities 148

Table 8.1: Reliability of Conjoint Estimates for Ranking Data (Voters) 161

Table 8.2: Reliability of Conjoint Estimates for Rating Data (Voters) 163

Table 8.3: Reliability of Conjoint Estimates for Ranking Data (Students) 164

Table 8.4: Reliability of Traditional Survey Questions for Ranking Data (Voters) 169

Table 8.5: Reliability of Traditional Survey Questions for Rating Data (Voters) 171

Table 8.6: Reliability of Traditional Survey Questions for Ranking Data (Students) 173

LIST OF FIGURES

Figure 3.1: Preference Models 41

Figure 3.2: Classification of Conjoint Analysis 55

Figure 5.1: Part-Worth Utilities for Voters (OLS) 81

Figure 5.2: Part-Worth Utilities for Voters (LINMAP) 81

Figure 5.3: Relative Importance of Attributes for Voters (OLS) 85

Figure 5.4: Relative Importance of Attributes for Voters (LINMAP) 85

Figure 5.5: Part-Worth Utilities for Students (OLS) 92

Figure 5.6: Part-Worth Utilities for Students (LINMAP) 92

Figure 5.7: Relative Importance of Attributes for Students (OLS) 93

Figure 5.8: Relative Importance of Attributes for Students (LINMAP) 94

Figure 6.1: Gender Differences in Education Spending 110

Figure 6.2: Gender Differences in Welfare Spending 110

Figure 6.3: Gender Differences in Crime Spending 111

Figure 6.4: Gender Differences in Natural Resources Spending 111

Figure 6.5: Generational Differences in Education Spending 116

Figure 6.6: Generational Differences in Welfare Spending 116

Figure 6.7: Generational Differences in Crime Spending 117

Figure 6.8: Generational Differences in Natural Resources Spending 117

Figure 6.9: Religious Differences in Education Spending 121

Figure 6.10: Religious Differences in Welfare Spending 121

Figure 6.11: Religious Differences in Crime Spending 122

Figure 6.12: Religious Differences in Natural Resources Spending 122

Figure 6.13: Party Differences in Education Spending 126

Figure 6.14: Party Differences in Welfare Spending 126

Figure 6.15: Party Differences in Crime Spending 127

Figure 6.16: Party Differences in Natural Resources Spending 127

Figure 6.17: Children Differences in Education Spending 130

Figure 6.18: Children Differences in Welfare Spending 130

Figure 6.19: Children Differences in Crime Spending 131

Figure 6.20: Children Differences in Natural Resources Spending 131

Figure 6.21: Home Differences in Education Spending 134

Figure 6.22: Home Differences in Welfare Spending 134

Figure 6.23: Home Differences in Crime Spending 135

Figure 6.24: Home Differences in Natural Resources Spending 135

CHAPTER 1

INTRODUCTION

The Importance of Studying Budget Priorities

The purpose of this study is to examine how people evaluate budget priorities by applying a technique called *conjoint analysis*. This topic is important for several reasons. First, the budget is definitely a significant policy issue. Since budgeting allocates scarce resources among rival groups, it appears regularly on the agenda of United States electoral politics, and candidates for public office routinely raise this concern in their campaign rhetoric. Indeed, one of the most distinctive features of government spending is its long-standing presence as an issue in U.S. politics. Other social programs and political controversies come and go over time, nevertheless, disagreements over the proper level of governmental spending for all programs remain fixed on the public agenda.

Second, one of the major characteristics of government budgeting is that decisions about how money will be spent are decided not by taxpayers but by their representatives. The distinction between payers and the deciders leads to two important features of public budgeting: public *accountability* and political *acceptability*. *Accountability* means to guarantee that all the money is spent as approved, and to report accurately to the public on how money was spent. *Acceptability* means that bureaucrats who make budget decisions are constrained by what the public wants (Rubin 1993: 133-34). However, a puzzle arises when we try to figure out what the public wants. Do policymakers and the public always have the same preferences for government spending? Unfortunately, evidence from survey research has shown that there appear to be significant inconsistencies between politicians and public opinion with respect to budget priorities. During the past decade, most political candidates and incumbent officeholders have requested reductions in public spending. Even so, public opinion on this issue seems to head in the opposite direction: while evidence from public opinion research has shown that the public also support cuts in overall spending for both federal and state governments (Sears and Citrin 1982; Peterson 1985; Modigliani and Modigliani 1987), when asked about cutting any particular program (especially social programs), most of the citizens oppose such actions (Ladd et al. 1979; Citrin 1979; Welch 1985).

Third, we know very little about the topic of public preferences toward government spending. While there have been some previous studies on this topic, the results and conclusions have varied widely. For instance, one explanation of the preference inconsistencies between politicians and the public is that people just do not understand the complexities of government budgeting (Croker 1981). An alternative explanation is that people want to eliminate the waste and inefficiency in government rather than a specific program (Ladd et al. 1979; Eismeier 1982). These reasons may contribute to the preference inconsistencies between politicians and the public; yet we believe that there is a deeper problem in the way we typically assess the voter's policy preference. Traditionally, researchers have depended on surveys for evaluating the "desires" of their target populations. Typical questions on a survey would usually be addressed in a "wish list" type of format with disaggregated questions and unbounded financial constraints. For example, two typical questions to evaluate government spending are: (1) "Government should provide many services, increase spending a lot" or

"Government should provide many fewer services, reduce spending a lot,"[1] and (2) Are you in favor of increasing government spending on _______ program, reducing it, or keeping it about the same?"[2] Facing such questions, respondents would usually want more of, perhaps, every service. One problem of this format is that the respondents will usually overstate their demand because they will not be charged for the increased level of service provision, and expressions of preference for "more" of a service will be expected to have only a small marginal impact upon themselves (Wilson 1983: 335-36). Therefore, what traditional survey measure is the people's "desire" rather than their "true preference."

Another problem of this format is that it provides respondents unrealistic choices among incomplete set of options. That is, single dimensional questions addressed by traditional surveys only ask people to evaluate the importance of a single attribute (factor or dimension) in isolation from the multitude of attributes that necessarily consist of a policy package or budget proposal. In the real decision process, however, the people's choice of a particular policy option is made on the basis of evaluating and combining multiple attributes. Some policy areas, such as environmental protection, are difficult to oppose when considered alone, but tend to lose support when compared with competing interests such as lost employment. Thus, traditional surveys cannot tell how people make tradeoffs among different attributes and hence may produce misleading information about people's preferences about government spending.

[1] NES survey question.
[2] CBS survey questions in 1980.

Because in the political world politicians always make tradeoffs among competing interests, it is extremely important to introduce a new survey tool which would allow people to choose from more realistic policy options. The new survey tool should define not only people's "desires," but more importantly, their actual "preferences." While the former can be addressed with a standard "wish list" type of question on the survey, the latter can only be addressed with precisely designed tradeoff type questions. For all these reasons, there is ample justification for a more detailed analysis of public attitudes toward budget priorities by using a new survey tool.

A New Survey Tool: Conjoint Analysis

One way to improve the limits of traditional survey is through the use of a technique called *conjoint analysis.* It provides respondents with more realistic choices and hence collects more accurate information about people's budget priorities. Currently, conjoint analysis and related techniques represent the most widely applied methodologies for measuring and analyzing consumer preferences (Carroll and Green 1995). However, this important technique has received very little attention in the policy analysis and public opinion literature.

Conjoint analysis is an exercise in making tradeoffs in which respondents are forced to make preference decisions based on the relative importance of one service or policy over another. Conjoint analysis is defined as a method which identifies the joint effect of two or more independent variables (attributes) on the ordering of a dependent variable (profile). A typical conjoint survey asks respondents to state their relative preference of various alternatives (profiles) by rank ordering the alternatives from most to least preferred. Researched products or services are sets of preferences called "profiles."

Each profile is defined by a specific set of attributes (factors). The primary outputs of a conjoint analysis are a set of intervally scaled "utilities" or "part-worths" for each level of each defined attribute. The part-worth utilities from conjoint analysis are then used to calculate the relative importance of attributes and overall utilities of profiles, and to simulate all kinds of what-if scenarios.

Objectives of This Study

The objectives in this study are as follows: First, we intend to provide an overview of conjoint analysis: the concept, theory, and design of conjoint analysis, the procedure to conduct conjoint analysis, why conjoint analysis works, the advantages and limits of conjoint analysis and classification of conjoint analysis and related techniques.

Second, to demonstrate the applicability of conjoint analysis, we will apply it to the topic of the public preferences toward government spending. Specifically, a conjoint study was conducted to evaluate people's budget priorities in Michigan. In the early 1990s, Michigan suffered a fiscal crisis due to economic slowdown, income tax change, tax base erosion, and spending pressure (Gold 1995: 301-305). Even though the fiscal crisis has largely been relieved recently, state government has had to fight spending pressures because the constitution of Michigan requires a balanced budget. A survey of public attitudes toward budget priorities in Michigan can serve two purposes: First, it may be used as a type of performance indicator to measure the public satisfaction with current services. Second, it can help policymakers to decide how to set the budget priorities across different policy areas. In this study, I will examine samples of Michigan voters[3] and undergraduate students[4] and assess their budget priorities for the state government spending in four general categories: *education*, *health and welfare*, *public safety and corrections*, and *natural resources and environmental control*, while holding other programs and tax rates constant.[5] The four categories explain 80.2 percent of total state expenditures in the 1996 fiscal year. Meanwhile, the four categories also represent the largest increase in spending items from 1990-96 fiscal years.

Third, group attitudes provide an underlying structure that organizes a broad range of policy opinions and candidates evaluations. In this study, I will discuss spending preference differences among different groups such as gender, age, parties, generations, and religions, etc.

Fourth, one of major advantages of conjoint analysis is that it provides the opportunity to conduct computer choice simulations to answer various "what-if" questions. In this study, I will show how the public preferences from conjoint analysis can be used in a simulation of the gubernatorial election and policy evaluation on welfare reform.

In addition to examining how people evaluate budget priorities, my last objective is to examine whether conjoint analysis can provide stable (reliabile) answers. Studies of public opinion have shown that people tend to have unstable responses to survey questions. If we ask the same respondents the same question in repeated interviews, probably only about half give the same answers. Furthermore, the literature has also shown that survey questions not only measure public opinion but also channel and shape

---

3 A random sample from the Meridian township voters (N=128).

4 A convenient sample from the undergraduate students of political science, Michigan State university (N=77).

it by framing issues, ordering alternatives, and setting the environment of the question (Zaller 1992). Hence, we will evaluate whether conjoint analysis can reduce two major problems in public opinion studies: response instability and response effects. Put another way, we will examine whether conjoint analysis can provide stable and valid results.

Organization of This Study

The design for this study on budget priorities is as following: Chapter 2 reviews public opinion toward government spending and the existing approaches and major findings on the study of public attitudes toward government spending. Chapter 3 introduces the methodology of conjoint analysis and the procedure to conduct it. Chapter 4 discusses the research design: the sample, questionnaire, and procedures. Chapters 5 and 6 examine the results from conjoint and traditional survey questions and analyze the preference differences among different groups. Chapter 7 demonstrates the applicability of conjoint analysis by simulating the gubernatorial election and policy evaluation on welfare reform. Chapter 8 checks the reliability of both conjoint and traditional survey questions and validity from conjoint results. Chapter 9 summarizes this study and discuss applicability of conjoint analysis for future research.

[5] For comparison purpose, five NES alike questions will be asked before conjoint questions.

CHAPTER 2

LITERATURE REVIEW

Democratic accountability requires that the public be reasonably well-informed about policy, and hence the public adjusts its preferences for "more" or "less" policy in response to what policymakers actually do. Therefore, to examine how people evaluate budget priorities, it is necessary to find out: Does public opinion really influence the policymaking? Do policymakers and the public always have the same preferences for government spending? If not, what cause the inconsistencies? Do current survey tools provide satisfactory answers about public preferences toward government spending? In this chapter, we review the relationship between public opinion and policy, public attitudes toward government spending, explanations of the inconsistencies between policymakers and the public toward government spending, and existing approaches to evaluate government spending.

Public Opinion and Policy

That public policy should be controlled by the public is a major requirement of democratic theory. In reality, we often evaluate the quality of democratic government by the responsiveness of policymakers to the preferences of the public and by formal opportunities for, and the practice of, citizen participation in political life. An important question is then to ask to what extent public opinion affects public policy? Some economics-style theories suggest that public opinion is a major influence on policy-making. Downs (1957) in his influential *An Economic Theory of Democracy* assumes that voter preferences can be arrayed on a single ideological continuum, from the political left to the political right, with citizens voting for the candidate closest to their own position on this ideological spectrum. In addition, Downs' model also assumes competition between two political parties or candidates for majority support. These assumptions drive the outcome: Parties and candidates converge toward the middle of the spectrum, with the winner being the candidate closest to the median voter. Since the winning program converges toward voter preference at the mid-point of the ideological continuum, the policy result renders an accurate representation of the composite position of the electorate as a whole.

Evidence from survey research does support the above idea. Monroe (1979,1983) shows that public policy in America is associated with majority preferences about two-thirds of the time. Page and Shapiro (1983) argue that major changes in people's preferences are followed by coherent changes in policy about two-thirds of the time. Farkas, Shapiro, and Page (1990) compare changes in public opinion in national surveys from 1973 to 1978 with changes in spending. Their findings suggest that public opinion affects spending on the environment, space, education, and drugs. Page and Shapiro (1992) characterize the public as composed of individuals whose preferences are fairly stable over time, particularly in the short-run. They also suggest that opinion change is effectively parallel across various subcategories of the American public. Stimson's (1991) study of a wide range of public opinion data further reveals that public preferences for policy in various domains are closely connected over time. Simply, preferences across a range of issues tend to move in the same liberal or conservative direction, producing what Stimson has called "policy sentiment" or "policy mood." That preferences for policy generally move in the same liberal or conservative direction denotes an underlying structure to those preferences. In addition, evidence from survey research also shows that the liberalism or conservatism of the public in the different states strongly influences whether states make liberal or conservative policies (Wright, Erickson, and McIver 1987; Erickson, Wright and McIver 1989 1992). Furthermore, some research suggests that aggregate voter turnout in elections is a predictor of respective welfare-benefit levels in the U.S and western democratic nations (Peterson and Rom 1989; Hicks and Swank 1992; Hill, Leighley, and Anderson 1995). Hill and Anderson (1995) provide a causal model which suggests that a reciprocal causal relationship between mass and elite preferences shapes state policy liberalism. In addition, party competition and lower-class mobilization enhance elite and policy liberalism.

However, not everyone agrees with the argument that the public's policy preferences have a substantial impact on what government does. Crigler (1990) claims that public opinion, the media, and Congress seldom attend to the same agenda. And other research on the agenda of Congress indicates that public opinion surveys seldom reach specific conclusions which offer guidance on the usually complicated policy matters before Congress. Kingdon (1995) argues that policy emerges from a "primeval soup" of ideas, speeches, bills, and proposals, where the contents of the soup come from a variety of sources, including public opinion. According to Kingdon, the mating of a policy solution and a policy problem may be due to nothing besides the efforts of policy advocates who attach their policy to an opportune problem. If policy proposals come from only a small segment of public opinion, then the other segments have fewer opportunities to redress their needs.

In sum, most evidence from public opinion research does show a close relationship between public opinion and policy. The public does adjust its preferences for "more" or "less" policy activity in response to what policymakers actually do. In contrast, although many studies have confirmed the influence of public opinion on policies, whether policymakers follow public opinion to make policies is not clear.

Public Attitudes Toward Government Spending

Evidence from public opinion has repeatedly shown that there are two remarkable and contradictory features of peoples' preferences toward government spending. First, it is known that most of citizens favor both tax reduction and increased levels of governmental services. For example, V.O. Key (1961) found in *Public Opinion and American Democracy* that a sizable fraction of the electorate wanted stronger social welfare protections while also favoring tax cuts "even if it means putting off some important things that need to be done." Key pointed out that both attitudes, though appearing contradictory, were real. As he wrote, "a simple calculus of self-interest makes simultaneous support of tax reduction and expansion of social welfare activities entirely consistent .... For the system as a whole, however, this type of opinion combination is irrational and creates problems in program making .... The balance of forces drives policymakers back toward concealed and indirect taxation, which may be regressive in its incidence" (168).

Sears and Citrin (1982) show that well over 70 percent of Californians before and after Proposition 13 (1978) thought their taxes were too high, yet majorities favored increases in spending for a wide variety services - mental health, police, fire, prisons and corrections, school, and transportation. Only for welfare did a majority of the public favor cutbacks. They conclude that people do, indeed, want "something for nothing" in public policies.

At the same time, however, there appears to be a inconsistency in the views of the public about reducing government spending. One the one hand, a majority of the public strongly favors cuts in government spending in general, when no specific programs are mentioned. One the other hand, when asked whether or not cuts should be made in specific programs, with only a few exceptions, a majority either favors maintaining current spending levels or prefers even more spending (Ladd et al. 1979; Bennett and Bennett 1990). Free and Cantril (1967), for example, argue that there appears to be an "ideological schizophrenia" within the U.S. public: people favor increasing spending on social programs, even while they repeatedly express demands for a smaller government.

Explanations of the Inconsistencies Between Policymakers and the Public Toward Government Spending

Several explanations have been addressed for the apparent distinction between taxing and spending preferences. The simplest explanation is that just as taxing and spending decisions have been institutionally divorced so they have been divorced in the minds of the public (Axelrod 1967; Brittan 1975). Another explanation, suggested by Downs (1963), is that because government forcibly provides complex bundles of different programs, taxpayers never reach equilibrium in their dealings with governments. That is, for each taxpayer there is some preferred reallocation of resources either within the public sector or between the public and private sectors, thus there is a tendency for every taxpayer to believe that taxes are too high. Finally, some argue that for some taxpayers dissatisfaction with tax levels may be rooted not in a belief that the overall tax burden is too great but rather that it is unfairly distributed (Kuttner 1980).

For the contradiction about reducing government spending, four explanations have been suggested by Crocker (1981). The first explanation involves the public's views about the relationship between inflation and government spending. Inflation consistently has been chosen as one of the most important problem facing the country since the end of Vietnam War. Moreover, the public seems to blame government spending for causing inflation. For instance, a series of NBC News/Associated Press surveys conducted in October, November, and December 1978, asked which of nine approaches was best to fight inflation. The most frequently named approaches were to cut government spending - 31 percent in October, 32 percent in November, and 25 percent in December.

The second explanation is that the public has little knowledge of the complexities of economics and government budgeting. Therefore, it is argued, the public holds contradictory views because most people lack an understanding of economic problems. For example, according to the April CBS News/New York Times Poll, 51 percent of those responding said that they did not know enough to have an opinion on Reagan's budget cutting proposals, and 52 percent said that they did not know enough about Reagan's tax cut proposal to have an opinion.

The third explanation, suggested by many polling experts, is that the public views the solution to be the elimination of waste and inefficiency in government. There is substantial evidence that, to many people, waste and inefficiency is the major problem with the government. In 1978 a Harris survey found that 94 percent of those surveyed agreed "that too much tax money is wasted by inefficient bureaucracy." A 1981 Gallup Poll found that the public believes that an average of 48 cents out of every dollar the federal government collects in taxes is wasted.

The last explanation is that the questions in polls simplify the issues and, consequently, lead to contradictory results. For instance, one might get different results if the questions on specific program reductions were asked in a way that indicated the tradeoff between increased spending and balancing the budget. Alternatively, one could argue that the questions on overall cuts in government spending do not inform respondents about what cuts are to be made.

One difficulty with the first two explanations is that neither give much credit to the public. If the public favors overall spending reductions and also favors maintenance of various social programs without increased taxes, how does it view the means of achieving both these ends? Similarly, the explanation of inefficient government also cannot solve the problem for policymakers. Students of public administration have learned that the motive of bureaucrats are driven by constraints, not profits, and public administration is different from business. Therefore, achieving the goal of efficient government is easier said than done. What we really should try is to use a better survey instrument to measure people's true preferences rather than their desires on different spending programs.

Existing Approaches to Evaluate Government Spending

Several studies have tried to explain the complexities of public preferences toward government spending by using all kinds of statistical techniques and other survey tools.

Eismeier (1982) examines the nature and determinants of individual budget priorities by using multiple discriminant analysis. Specifically, he uses multiple discriminant analysis to determine how respondents with different budget priorities differed according to several political, socioeconomic, and attitudinal variables, including tax levels. Eismeier finds that the growing dissatisfaction with taxes is rooted in the belief that government is not only too big but also inefficient and unresponsive. In other words, dissatisfaction with spending levels is due to people's lack of confidence in the performance of government. His second finding is that patterns of partisan and socioeconomic differences vary substantially across policy areas. However, he also points out that the inconsistent attitudes of budget priorities may not be as much at the individual level as it is at the aggregate level. That is, while the individual taxpayers' policy preferences may be internally consistent, they may be incompatible with the policy preferences of other taxpayers.

Wilson (1983) examines people's "sincere preferences"[1] on government spending through experimental surveys.[2] In particular, he estimates the frequency of the expression of "sincere preferences" for additional public spending - preferences which are accompanied by a willingness to pay additional taxes in order to obtain additional public goods or services. The format of the surveys is a modified n-prisoners' dilemma game which includes a series of questions about people's preferences toward specific policy areas. The survey design is best explained by the following example: "Do you think that

---

1 Individual expressing "true" or "real" preferences for increased, constant, or reduced government spending are considered to be expressing "sincere preferences."

2 The surveys were conducted on Eugene, Oregon (N=1115) and Tempe, Arizona (N=509) in 1977 and 1981 respectively.

the city should spend more, less, or the same amount on crime prevention?" There are three options for respondents.

1. If the answer is "spend more," the following question is asked: " Would you be willing to pay more in taxes in order to spend more on crime prevention?" If the answer is "yes," no further question. If the answer is "no," then proceed with question "Would you suggest that there be no change in service level if it means increased taxes or would you reallocate money from some other program? The answer will be either "no change" or " reallocate."

2. If the answer is "spend the same," no further question.

3. If the answer is "spend less," the following question is asked: "If the city were to reduce its spending in this area, what should it do with the money saved?" The answer will be either "reduce taxes" or " reallocate."

The structure of the survey was designed to permit individual respondents to overstate their preferences for the government spending ("would you like to see the city spend more on the provision of a service") and then attempt to hold them accountable for their preferences (are you willing to pay more in taxes in order to have more of the service").

The results from Wilson's study illustrate that approximately two-thirds or more of the demands for additional service expressed on the experimental survey constitute "sincere" preferences. The proportion of "sincere" preferences ranges from a low of 62 percent to a high of 84 percent in Tempe and from a low of 66 to a high of 78 percent in Eugene. The range of "insincere" preferences (those who decide they do not want the city to spend more on the provision of a good or service if it means as increase in taxes) varies from a high of 12 percent (Tempe) and 14 percent (Eugene) to a low of 6 percent (Tempe) and 5 percent (Eugene).

Welch (1985) explores the possibility of increasing spending level by means other than taxes, such as reallocation or nontax revenue. She creates a fourfold categorization called "willingness to pay" scale to demonstrate the willingness of individuals to raise revenue to support increased spending. The four categories are:

1. Taxers - those willing to increase taxes.
2. Probably realistic revenue raisers - those who were not willing to raise taxes but willing to adopt new user fees, increase revenue through intergovernmental aid, or reallocate funds to pay for increased service in priority areas.
3. Probably not realistic revenue raisers - those who prefer reallocation but wanted to spend more in more than half of the programs.
4. Free lunchers.

The findings from her study suggest that support for increasing government spending is high. Nevertheless, relatively few favor a "free lunch," increase spending without any reasonable means to pay for it. Most people are willing to raise additional revenue to pay for these services, or at least to reallocate from less preferred to more preferred services. From 72 percent to 88 percent of the respondents are found in one of the first two categories (taxers and probably realistic revenue raisers), where support for raising revenue in realistic ways is apparent. Yet, very few people think we can depend solely on greater efficiencies in government to pay for increased services. Therefore, she concluded: "The taxing-spending paradox may not be as great as previously thought" (p.316).

Sanders (1988) assesses the two opposite explanations of public attitudes on government spending: the spending attitudes of the public have been described as either too irrational, wanting something for nothing, or too rational, only supporting programs they benefited. He runs 16 logistic regressions that dependent variables are composed of each of the individual spending questions (programs) while independent variables include a composite index[3] and 6 demographic variables. From his findings, Sanders concludes that "some of the traditional views we have of the public's attitudes on public spending issues may be more myth than reality. The public may not simply be asking for something for nothing, nor do they seem to be saying 'help me.' The attitudes people have are more complex than that" (p.321). Specifically, he suggests that most citizens desire increasing spending on some of the programs and less on other programs. In fact, everyone desires to cut spending in at least one domain, and on average the number of programs people prefer to increase tends to balance out with those they want to cut. In addition, the role of self-interest is limited; most programs cannot rely solely on the support of those who benefit from them. Sanders points out that the two domains where self-interest plays the strongest role - food stamps and welfare - are also two of the three domains where most people want to cut. Therefore, he suggests that building support for a program needs to establish the legitimacy of the beneficiaries among the broad public.[4]

Jacob (1994) examines the nature, sources, and consequences of peoples' attitudes toward government spending by using a variation of Guttman scaling technique. He

---

3 The index is created as follows: First, respondents were asked if they received benefits from 7 of the programs (16 programs total in the survey). Second, a composite index was formed adding up the number of those seven areas in which an individual received support.

argues that most people have very limited ideas of political phenomena, so it is inappropriate to view all of the program-specific responses as equally valid representations of people's spending preferences. Therefore, he evaluates patterns of responses across programs by fitting a cumulative scaling model. According to this model, there is an underlying trait or dimension that ranges from a position "reduce spending in all programs (or a specific program)," at one extreme, to "increase spending in all programs (or a specific program)," at the opposite extreme. He uses 1984 NES postelection survey as his data in research (which is slightly different from the traditional NES survey questions). The survey includes the following questions: "If you had a say in making up the federal budget this year, which programs would you like to see increased and which reduced?" After that, the respondents were asked the following question about 10 program areas: "Should federal spending on _____ be increased, decreased, or kept about the same?"

Jacoby's analysis suggests that the public has coherently structured orientations toward spending on welfare-related programs. Yet, opinions on nonwelfare spending are highly varied and largely independent of each other. In addition, the general term "government spending" triggers individuals to think about welfare spending, even if the particular programs are not explicitly called to public attention. That is, most people seem to translate the phrase "government spending" into government spending on programs that will benefit the poor, blacks, and other disadvantaged people.

[4] Establishing legitimacy is easier with programs which offer benefits for most people (such as environment) or that most people might potentially need (such as Social Security or health). For a detailed discussion, see Sanders (1988, p.322).

Furthermore, individuals' responses to social welfare programs are generally quite consistent with their political predispositions. However, in the area of nonwelfare programs, people's preferences are very difficult to predict on a systematic basis. Put another way, Jacoby argues, "nonwelfare spending preferences are not political attitudes at all" (p.355).

Wlezien (1995) employs a thermostatic model to analyze public preferences for spending on defense and a set of five social programs. The thermostatic model refers to the public behaving like a thermostat: when the actual policy "temperature" differs from the favored policy temperature, the public will send a signal to adjust policy accordingly, and once sufficiently adjusted, the signal would stop. The public's preferences for spending in various policy areas are captured by the measures of net support for spending.[5] Time-series regression is used to estimate defense spending preference, five categories of social spending, and aggregate social spending.

From his study, Wlezien' points out that public preferences for spending in the specific policy areas do not generally reflect appropriations in those areas. Nonetheless, the public does adjust its preferences for more or less spending in the particular areas in response to what policymakers appropriate for social programs taken together. The reason for this difference, Wlezien suggests, is that specific information about spending for those programs is not regularly and widely available to the public. Thus, it may be that the public obtains only general information about social programs. The public still responds to spending for social programs, but in a general way. In general, he argues that

[5] The net support for spending is calculated by subtracting the percentage of people who think we are spending "too much" from the percentage of people who think we are spending "too little."

his findings are consistent with Eastonian model: there is a negative feedback of spending decisions on the public's preferences, whereby the public adjusts its preferences for more spending downward when appropriations increase, and vice versa, at least in the defense spending area and across a set of social spending areas.

Hansen (1998) examines public preferences over deficits, taxes, and spending by using tradeoff type of survey questions. He uses the data of 1995 NES Pilot Study which asked respondents 12 questions with a complete budget tradeoffs design. The budget choices in the survey were divided into two revenue categories: taxes and deficits and two spending categories: domestic and defense. Respondents then were asked to accept or reject a budget proposal that made them better off on one dimension (by raising domestic spending, increasing defense spending, lowering taxes, or reducing the deficits) but left them worse off on another (by raising taxes, enlarging the deficits, reducing domestic spending, or cutting back on defense). From his study, Hansen discovered that respondents overwhelmingly rejected the tradeoffs and revealed their preference for the policy status quo. He also evaluated the characteristics of individual preferences, such as their completeness, consistency, and coherence and found that peoples' preferences over the public budget were remarkably well structured: most people are able to determine their views, identify tradeoffs, and maintain consistent and coherent preferences.

Some overall conclusions can be drawn from these findings about public preferences toward government spending. First, the idea that the public wants "something for nothing" is not true if we can hold them accountable for their preferences. If the people favor more services, we could hold them accountable by asking them either to increase taxes or reallocate spending on various programs. These tradeoff procedures eliminate the problem with the traditional survey questions that respondents usually overstate their demand because they will not be charged for the increased level of service provision. Wilson's study reveals that more than two-thirds of the demand for additional service is "sincere" preferences - a willingness to pay more taxes to obtain more services. Welch discovered that most people are willing to raise taxes to pay for more of services, or reallocate funding from other different programs. Sanders' findings show that on average the number of programs people prefer with increases tends to balance out with those they want to cut. Hansen found that public opinion is remarkably well structured and overwhelmingly partial to the policy status quo. Thus, previous studies have demonstrated the limits of traditional surveys: they virtually measure people's desires rather than their true preferences. To obtain people's true preferences, we have to use tradeoff format of survey design and hence hold respondents accountable for their preferences.

Second, all these studies have tried either different statistical techniques or different survey tools to unveil the complexities of public attitudes toward budget priorities (Eismeier's discriminant analysis, Wilson's prisoners' dilemma format of survey, Welch's "willingness to pay" scale, Sanders's logistic regressions, Jacoby's scaling approach, and Wlezien's time-series analysis). Although these efforts are not small tasks, the effects of these applications are limited due to the problem of the data they used. Specifically, for those data which came from traditional surveys, the problems are caused by the fact that these questions only asked people to evaluate the importance of a single attribute rather than multiple attributes that necessarily consist of a policy package or budget proposal. For example, from traditional surveys we could obtain a respondent's preferences that he or she favors a budget proposal "increases in education, crime, and natural resources spending and cuts in welfare spending." However, we do not know how this respondent would evaluate a budget proposal "increases in education, crime, and natural resources spending and cuts in welfare spending" versus another budget proposal "increases in crime and natural resources spending and cuts in education and welfare spending." For those data which came from different survey tools, the limits lie in the fact that respondents only evaluate tradeoffs of partial (e.g., two-factors) attributes rather than all attributes which necessarily consist of a policy package or budget proposal. For instance, from Hansen's study we know that a respondent favors tax increases in order to cut the deficit and opposes cuts in domestic spending in order to increase defense spending. Even so, it is still very difficult to find out how a respondent chooses budget proposals with tax increases, lowering deficit, increases in domestic spending, and cuts in defense spending versus tax decreases, lowering deficit, decreases in domestic spending, and cuts in defense spending. In other words, we do not know how a respondent chooses his or her most preferred policy proposal from different policy packages.

In fact, in the real decisionmaking process, people's choices are made on the basis of evaluating and combining multiple attributes simultaneously. Lacy (1998) examines two models of individual-level responses to issue questions when respondents have nonseparable preferences. His results reveal that question-order effects occur on issues for which people have nonseparable preferences. Hence, the results from previous studies are not satisfactory because we obtained only fragmented and incomplete information about public opinion toward budget priorities. In contrast, conjoint analysis provides more accurate and complete information about people's preferences by producing part-worth utilities for each level of each attribute (i.e., utilities of cuts in domestic spending and increases in domestic spending). Part-worth utilities are then used to calculate the relative importances of attributes (how important an attribute is) and overall utilities of profiles (i.e., a profile with tax increases, lowering deficit, increases in domestic spending, and cuts in defense spending). In addition, we also could obtain people's preferences on different policy packages by simulating all kinds of policy combinations. This accurate information is very difficult to acquire from the traditional survey or from other survey tools previously discussed.

Finally, some survey formats (such as Hansen's tradeoff design) address questions in a direct and obtrusive way (such as raising taxes in order to increase domestic spending) which may cause respondents to underestimate their preference for a particular policy package because respondents were held accountable for the cost of increased levels. On the contrary, since conjoint analysis ask respondents rank ordering alternatives from most preferred to least preferred rather than ask respondents making tradeoffs between two competing answers (i.e. increasing in tax in order to increasing domestic spending), it is less obtrusive and hence respondents are more willing to reveal their true preferences.[6]

Summary

In general, the literature on public opinion has demonstrated that peoples' policy preferences have a substantial impact on what government do. In addition, previous studies have tried various statistical techniques and survey instruments to unveil the complexities of people's budget priorities; however, the effects of these applications are limited due to the constraints of traditional survey questions and different survey designs (i.e., Wilson's prisoners' dilemma format, Welch's "willingness to pay" scale, and Hansen's tradeoff format). Conjoint analysis forces respondents to make tradeoffs among various programs so that respondents evaluate all the attributes simultaneously (which necessarily consist of a budget proposal) rather than evaluate attributes one by one or pairs by pairs. Even though it is not perfect,[7] conjoint analysis does provide more accurate and complete information about people's preferences toward public spending when compared with the traditional survey.

---

[6] For the full profile data collection method only. Collecting data by a two-factor-at-time method is also obtrusive to respondents. See chapter 4 for a detailed discussion of the differences between full profile and two-factor-at-time methods.

[7] The limits of conjoint analysis will be discussed in chapter 3.

CHAPTER 3

THE LOGIC OF CONJOINT ANALYSIS

Studying budget priorities has both substantive and methodological significance. The purpose of this chapter is to discuss the concept, theory, design, and analysis of conjoint analysis, why conjoint analysis works, advantages and limits of conjoint analysis, classification of conjoint analysis, and current applications of conjoint analysis.

The Concept of Conjoint Analysis

Conjoint analysis is both a method of measurement and a method of analysis. It is a multivariate technique used specifically to understand how consumers decide preferences for products and services based on the simple premise: consumers evaluate the utility of a service or policy idea (real or hypothetical) by combining the separate amounts of utility provided by each attribute. The word "conjoint" has to do with the notion that the relative values of attributes considered jointly can be measured when they might not be measurable if taken one at a time (Barron 1977). Quite often respondents are asked to express the relative value of various alternatives to them by ordering the alternatives from most desirable to least desirable. The attempt in a conjoint analysis solution, then, is to assign values to the levels of each of the attributes so that the resulting values or utilities are as monotonic as possible with the input rank-order judgments.

In conjoint analysis, services and products are described by "profiles." Each profile is a combination of one arbitrarily selected level for each of the attributes. Attributes are the key dimensions (features) of services or policies, while levels are those specific points along the key dimensions (Hu 1996). The major objective of conjoint analysis is to identify the single profile which includes the most favored level for each of the attributes. Table 3.1 illustrates the relationship among a profile, attributes, and levels.

Table 3.1: The Relationship among a Profile, Attributes, and Levels

Profile X	**Attribute**	**Level**
A2	A	-- A1, **A2**, A3
B1	B	-- **B1**, B2, B3
C3	C	-- C1, C2, **C3**
D2	D	-- D1, **D2**, D3

This table can be explained as follows: Profile X is composed of the second level of attribute A (A2), the first level of attribute B (B1), the third level of attribute C (C3), and the second level of attribute D (D2). Hair et al. (1995) point out that there are two objectives of conjoint analysis: (1) to evaluate the contributions of determinant attributes and the values of these separate levels on the determination of consumer preferences, and (2) to build a valid model of consumer judgments that is useful in predicting consumer acceptance of any combination of attributes, even those not originally evaluated by consumers (556-559). In order to achieve these objectives, coefficients called "*utilities*" (or "*part-worths*") among different levels of attributes are first estimated, and then "*relative importance*" among attributes and "*profile utilities*" are developed to quantitatively measure preferences in consumer decisions. The meanings of the terms is as follows:

• *Level utility (part-worth):* a numerical expression of the value that a respondent places on each level of each attribute.

• *Attribute relative importance*: the computation of the relative importance for each attribute depends on the relative range between maximum and minimum level utilities within an attribute. This is based on the assumption that the larger the difference between maximum and minimum level utilities within attribute, the more determinative and salient the attribute is in the overall evaluation of profiles. The relative importance is described in a percentage term to reflect its weighted importance across all involved attributes. The

relative importance is defined as follows:

$$RI_j = \frac{Max(X_{ij}) - Min(X_{ij})}{\sum[Max(X_{ij}) - Min(X_{ij})]} \times 100\% \qquad (3.1)$$

Where

RI_j = the relative importance of attribute j.

Max (V_{ij}) = the maximum level (*i*) utility in attribute j.

Min (V_{ij}) = the minimum level (*i*) utility in attribute j.

• *Profile utility*: the overall utility of a profile calculated by summing all utilities of attribute levels defined in that profile.

Compositional vs. Decompositional approaches

The compositional or build-up model, such as regression or discriminant analysis, involves collecting respondent's judgments about each level of each attribute (e.g., a respondent's spending preference toward education, welfare, crime, and natural resources), and then combining these ratings to compose an overall preference.

In contrast, conjoint analysis uses a *decompositional* model. Knowing respondents' choice preferences between pairs of objects, and knowing the characteristics of the objects presented to respondents (as manipulated through an experimental design), respondents react to a set of "total" profile descriptions. These preferences are decomposed to determine how much utility is associated with each level of each attribute. Therefore, it is the job of the analyst to find a set of part-worths for the individual attributes , which given some type of composition rule (e.g., an additive one), are most consistent with the respondent's overall preferences. This avoids many of the subject artifacts and demand effects of compositional approaches.

The major advantages of the compositional approach include a strong and thoroughly tested conceptual foundation and ease of operationalization. However, these models have been criticized because they cannot accurately simulate decision processes in the real world. For example, they do not include trade-offs between attributes in the decision process (Johnson 1974), they tend to overweight less important attributes (Beckwith and Lehmann 1975), and they describe a more complicated decision process than probably exists in the real world (Olshavsky and Granbois 1979).

The decompositional models (such as conjoint analysis) tend to be strong where compositional methods are weak. Conjoint analysis provides a more realistic description of the decision process. Because people are evaluating a complete service or policy, they are able to incorporate the same attribute-discounting and trade-off processes they use when making actual decisions. The limit of conjoint methods is that they are cumbersome to administer due to the tedious procedures necessary to parameterize the models.

The Theory of Conjoint Analysis: An Information Integration Theory Approach

The seminal theoretical contribution to conjoint analysis was made by Luce, a mathematical psychologist, and Tukey, a statistician. Thereafter, a number of theoretical contributions (Krantz 1964; Tversky 1967) and algorithmic developments (Kruskal 1965; Carroll 1969, 1973; Young 1972; Shocker and Srinivasan 1977) appeared.

Information Integration Theory

Louviere (1988) provides a theoretical basis for conjoint analysis based on the information integration theory (IIT) developed by Anderson (1981,1982). Information integration theory allows us to study information processing as revealed by people's response to multiattribute options. Compared to other theories, Louviere argues that IIT is a better theory because it possesses a theory of errors; error theory is important because it permits one to falsify models.[1]

The basic assumptions of integration information theory are as follows:

1. The unknown and unobservable overall utility that a consumer possesses in his or her mind regarding the p-profile is linearly related to a consumer's response on a category-rating or ranking scale. That is,

$$U_p = a + bR_p + e_p \qquad (3.2)$$

where

U_p is the overall utility to measure of the p-th profile,

R_p is the observed response on a category-rating or ranking scale,

[1] IIT is similar in conceptual orientation to Social Judgment Theory (Adelman et al., 1974) and bears a strong resemblance in experimental format to axiomatic conjoint measurement such as Krantz (1964) and Tversky (1967).

e_p is a normally distributed error term with zero mean and constant variance σ^2, which satisfies assumptions of OLS (ordinary least squares).

2. The category-ranking scale used by a subject under approximate experimental instructions and task conditions approximates an interval measurement level.

3. A subject's response strategy reveals his or her decision model. The response strategy can be approximated by algebraic conjoint models amenable to experimental design and statistical parameterization.

Assumption 2 may be relaxed through the use of alternative response scales, such as binary scale (yes/no) or multiple choice scales (from most preferred to lest preferred, etc.).

Algebraic and Statistical Formalities

IIT assumes that simple algebraic conjoint models are valid approximations to the unobservable evaluation and decision process of consumers. However, the decision model that consumers "really use" is unknowable. A simple additive model is a good start to illustrate the conjoint method. The additive model can be expressed as follows with three determinant attributes:

$$U_p = C + W(X_{1p}) + W(X_{2p}) + W(X_{3p}) \qquad (3.3)$$

where U_p is the overall utility to measure of the p-th profile, C is an additive constant, and $W(X_{1p})$, $W(X_{2p})$, and $W(X_{3p})$ are part-worth utilities for the three attributes of the p-th profile.

Each level of each attribute may have a different part-worth utility. Different profiles can be described by combinations of different attributes with different levels. An additive model implies that consumers add the separate part-worth utilities to evaluate each profile. Assuming the levels of attributes 1, 2, and 3 as $X_{1a.}$ X_{2b}, and $X_{3c.}$ and the corresponding part-worth utilities, as $W(X_{1a})$, $W(X_{2b})$, and $W(X_{3c})$. Each a-th, b-th, c-th combination of levels thus represents one profile.

Hence, instead of using U_p, we will use U_{abc} with no loss of generality, but it means U_p. The total number of profiles that one can produce is the product of the a, b, and c levels. The total number of profiles is therefore a combinatorial problem defined by the set of possible combinations of the levels a, b, and c. Such problems are called factorial problems (which will be discussed later).

Now we rewrite the additive model for a single subject as follows:

$$U_{abc} = C + W(X_{1a}) + W(X_{2b}) + W(X_{3c}) \qquad (3.4)$$

where all terms were defined previously.

Recall that we cannot observe U_{abc} directly. Instead, we observe a subject's evaluation, R_{abc}, on a rating or ranking scale. Substituting Equation 3.2 and transposing, we get:

$$R_{abc} = C + W(X_{1a}) + W(X_{2b}) + W(X_{3c}) + e_{abc} \qquad (3.5)$$

where all terms were defined previously. Equation 3.5 suggests that a subject's rating or ranking are an additive function of the true but unknown part-worth utilities. The error assumptions of ITT imply that ANOVA or OLS can be used to test the additive model hypothesis of equation 3.4.

If a subject evaluates each of the a×b×c combinations more than once, there are sufficient observations to test equation 3.4. If equation 3.5 is true, there will be no nonadditivities in a subject's response data. That is, only the purely additive terms will be nonzero in a statistical analysis - there will be no significant interaction effects in the analysis of variance or multiple regression if equation 3.4 is true. This conclusion can be proved if we subtract the response to level 2 of attribute 1 from the response to level 1 of attribute 1, holding other attributes constant:

$$\begin{aligned} R_{1bc} - R_{2bc} &= [C + W(X_{11}) + W(X_{2b}) + W(X_{3c}) + e_{1bc}] \\ &\quad - [C + W(X_{12}) + W(X_{2b}) + W(X_{3c}) + e_{2bc}], \\ &= W(X_{11}) - W(X_{12}) + (e_{1bc} - e_{2bc}). \end{aligned} \tag{3.6}$$

Equation 3.6 shows that the difference in the part-worth utilities of level 1 and 2 of attribute 1 is a constant; therefore, curves of response data for each level of attribute 1 graphed against levels of attribute 2 (or 3) must be parallel. Responses to levels of one attribute are independent of responses to levels of other attributes, a property that generalizes to any number of attributes.

Therefore, if a consumer's decision process is additive, we expect to maintain the null hypothesis that no interactions in an analysis of variance or multiple regression. Thus, the additive model hypothesis can be falsified by using ANOVA or OLS to test the main and interaction effects of subject's evaluation. However, subjects must rate or rank each profile at least twice in order to have enough degrees of freedom to perform the necessary statistical tests. If an experimental (factorial) design is small (e.g. 2*2*2), multiple replications are doable. However, as the size of an experiment grows, by increasing the numbers of attribute numbers, by increasing the numbers of attribute levels, or by both, it becomes impractical to ask subjects to rank or rate multiple profiles.

Therefore, one must find other ways to solve this problem, such as fractional factorial design (which will be discussed later).

Now let us turn our attention to estimating the part-worth utilities - the $W(X_{1a})$, $W(X_{2b})$, and $W(X_{3c})$. Repeating equation 3.5 for convenience, we get:

$$R_{abc} = C + W(X_{1a}) + W(X_{2b}) + W(X_{3c}) + e_{abc} \qquad (3.7)$$

Considering the result of averaging 3.7 over attribute 2 and 3 (or subscripts b and c):

$$R_{a..} = C + W(X_{1a}) + W(X_{2.}) + W(X_{3.}) + e_{a...} \qquad (3.8)$$

$$= [C + W(X_{2.}) + W(X_{3.})] + W(X_{1a}) + e_{a...}$$

where dots (.) refer to the effect of averaging. Equation 3.8 constitutes a theorem that expresses that the marginal means ($R_{a..}$) of a subject's response data are equal to the unknown part-worth utilities ($W(X_k)$'s) up to a linear transformation. That is, the marginal response means are interval scale measures of the unknown part-worth utilities because a linear transformation is the only permissible transformation for interval scales.

This conclusion is true only if an additive model is "correct." Hence, if an additive model passes the statistical tests described earlier, one can use the marginal response means to represent the consumer's part-worth utilities for each attribute. In addition, the ($W(X_k)$'s) must be a function of the X_k's. (The different composition rules and preference models will be discussed in the next section). Therefore, one can test additive conjoint models and, if they cannot be rejected, one can infer relationships between consumer's positions on different profiles and part-worth utilities.

Design and Analysis of Conjoint Analysis

In conjoint analysis, the most widely used model is an additive main-effects (no interactions) part-worth model. This statement includes the use of different composition rules and preference models for consumers in conjoint analysis.

Choice of Composition Rules: Compensatory vs. Noncompensatory Models

The first step in conjoint analysis is to consider the principle ways in which attribute information of each respondent can be processed. Two major approaches, compensatory model and noncompensatory model, can be used in conjoint analysis.

1. *Compensatory models* are those cases in which all alternatives can be ultimately described in terms of single utility numbers that are commensurate with each other. The models are compensatory in the sense that a low value on some attribute can be compensated for a high value on some other attribute. In these models a single holistic value or utility is attached to each multidimensional profile. Tradeoffs between attributes are generated by different profiles that have equal utilities. The simplest compensatory model is the main-effects, additive utility model. The composition rule for this model is an additive one over all the attributes selected by the respondent; if the respondent selected attributes A, B, and C, then an additive composition rule would be expressed as A + B+ C (as seen in the Equation 3.3).

In conjoint analysis, the additive main-effects model is the most widely used. Except using additive model in conjoint analysis, there are numerous models available. For instance, the additive model belongs to the class of algebraic conjoint models known as simple polynomials (Krantz and Tversky 1971). The class of simple polynomials includes additive, multiplicative, distributive, and dual-distributive models. The multilinear model is the general form for these simple polynomials, and it can be expressed as follows for three determinant attributes:

$$U_{abc} = C_0 + C_1W(X_{1a}) + C_2W(X_{2b}) + C_3W(X_{3c}) \tag{3.9}$$
$$+ C_4W(X_{1a})\,W(X_{2b}) + C_5W(X_{2a})W(X_{3c})$$
$$+ C_6W(X_{2b})W(X_{3c}) + C_7W(X_{1a})\,W(X_{2b})W(X_{3c}),$$

where the C_0 is the origin of the utility scale and C_1 - C_7 are scaling constants. Equation 3.9 implies that attributes can be complements, substitutes, or independent. The multilinear form allows a general approach to study decision processes, because it contains completely multiplicative process (all attributes are complements), distributive processes (one attribute complements the other two, which are jointly independent); and dual-distributive processes (one attribute is independent of two others that are complements). These models can be written as follows:

$$U_{abc} = C_0 + C_1[W(X_{1a})W(X_{2b})W(X_{3c})], \text{ [multiplicative]} \tag{3.10}$$

$$U_{abc} = C_0 + C_1W(X_{1a})[C_2W(X_{2b}) + C_3W(X_{3c})], \text{ [distributive]} \tag{3.11}$$

$$U_{abc} = C_0 + C_1W(X_{1a}) + C_2W(X_{2b})W(X_{3c}), \text{ [dual-distributive]} \tag{3.12}$$

where all the terms are as defined previously.

Equation 3.10 - 3.12 are subsets of the multilinear form in which one or more C terms are equal to zero. Thus the multilinear form can be used as the general form for estimation and testing: all results that apply to this form also apply to equations 3.10 - 3.12. Although there are other composition rules available, but they are not often used in conjoint study.[2]

[2] For a discussion of these composition rules, see Green and Wind (1973), and Barron (1977).

2. *Noncompensatory models* do not allow tradeoffs between attributes. Comparisons are made on an attribute-by-attribute basis, and in general the multidimensional profiles are not collapsed into single utility numbers. For example, one attribute could be expressed in terms of pounds and another in terms of rating points. Since component values are applied to each attribute separately, the dimensions do not have to be commensurate.[3]

Preference Models

Utility preference models are the mathematical formulations that define the utility levels for each of the attributes. In practice, the attributes are modeled as either a linear (vector model), quadratic (ideal-point model), or piecewise (part-worth model) function (Green and Srinivasan 1978).

The Vector Model

The *vector model* is represented by a single linear function that assumes preference will increase as the quantity of attribute j increases (preference decreases if the function is negative). Preference for the j-th attribute is defined as:

$$S_i = \sum_{j=1}^{t} W_j X_{ij} \qquad (3.13)$$

where:

S_i = Preference for the stimulus object at level i,

$j = 1, 2, ..., t$ denote the set of attributes that have been chosen,

W_j = the individual's weight assigned to each of the j attributes. One weight is derived for each attribute,

X_{ij} = Level of the j-th attribute for the i-th level.

[3] For a discussion of noncompensatory models, see Yoon and Hwang (1995).

The vector model for the attribute with three levels would appear as a straight line, with the levels on the line (see Figure 3.1). The vector model requires that a single parameter be estimated for the variable treated as a vector. In contrast to the part-worth model (will be discussed later), the vector model defines the attribute variable not as a series of dummy variables but as a single linear variable where values are the measured values or levels associated with the attribute.

The Ideal Point Model

The *ideal point model* is operationalized as a curvilinear function that defines an optimum or ideal amount of an attribute. The ideal point model is appropriate for many qualitative attributes, such as those associated with taste or smell. For instance, too much sweetness may be less than optimal, while just the right amount is highly preferred.

The ideal point model establishes an inverse relationship between preferences and the weighted distance d_i^2 between the location of the *i*-th level and the individual's ideal point, I_j. The *ideal point model* is expressed as :

$$d_i^2 = \sum_{j=1}^{t} W_j (X_{ij} - I_j)^2 \qquad (3.14)$$

where:

$j = 1, 2, ..., t$ denote the set of attributes that have been chosen,

W_j = the individual's weight assigned to each of the j attributes. One weight is derived for each attribute,

X_{ij} = Level of the *j*-th attribute for the *i*-th level,

I_j = the individual's ideal point, j.

The ideal point model for the attribute with three levels would appear as a curve with the center of the curve higher than either end (see Figure 3.1). The highest point is the ideal quantity of the attribute. Mathematically, the implications of specifying each of the models ultimately extends to the number of parameters that must be estimated. For the ideal point model, 2t parameters must be estimated (W_j and I_j), and for the part-worth model, (*i*-1) parameters must be estimated, where *i* is specified to be the number of levels for each of *j* attributes.

The Part-Worth Model

The *part-worth model* is the simplest of the utility estimation models. This model represents attributes by a piecewise linear curve (see Figure 3.1). This curve is formed by a set of straight lines that connect the point estimates of the utilities for the attribute levels.

The *part-worth model* posits that

$$S_i = \sum_{j=1}^{t} f_j(X_{ij}) \qquad (3.15)$$

where

S_i = Preference for the stimulus object at level *i*,

j = 1, 2, ..., t denote the set of attributes that have been chosen,

X_{ij} denote the level of the *j*-th attribute for the *i*-th stimulus (level),

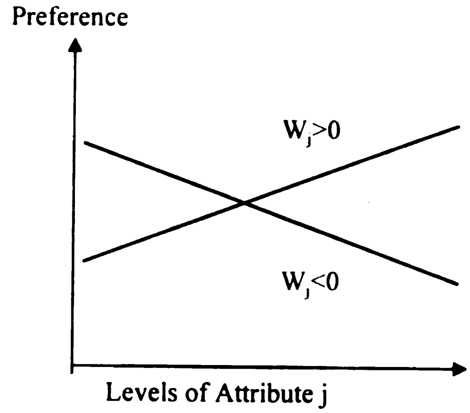
f_j is the function denoting the part-worth utility of different levels of the stimulus object, X_{ij} for the *j*-th attribute.

The *part worth model* reflects a utility function that defines a different utility (part-worth) value for each of the *i* levels of a given attribute. Because of design considerations, most conjoint studies constrain the number of levels to be less than 6, though in reality, the number of levels varies from 2 to 9.

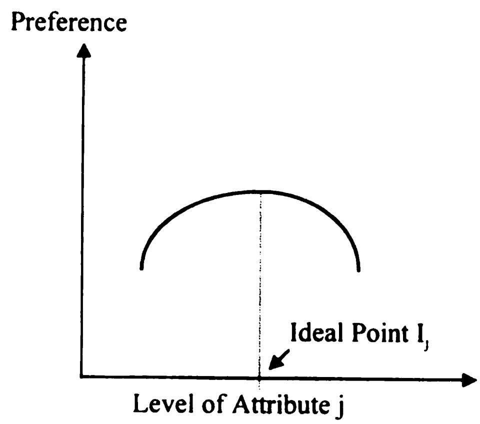
The implications of specifying a given preference model (linear, ideal point, or part-worth) extend beyond the actual shape of the preference curve being modeled. Each model requires that a different number of parameters be estimated. The part-worth model requires that each level of an attribute be defined by a dummy variable distinct column within the design matrix. As would be expected, a total of i-1 dummy variables are required to estimate i levels.

Figure 3.1: Preference Models

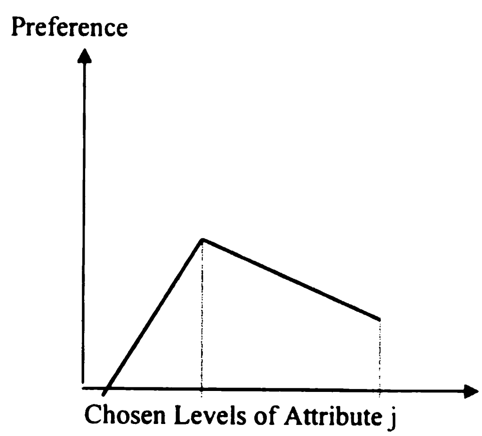
1. Vector Model



2. Ideal Point Model



3. Part-Worth Model



Fractional Factorial Design

Factorial design involves combinations of levels of determinant attributes. When the size of an experiment grows, by increasing the numbers of attribute numbers, by increasing the numbers of attribute levels, or by both, it becomes impractical to ask respondents to rank or rate all the profiles. The usual solution for this factorial problem is to use fractional factorial design. A fractional factorial design is a sample of treatments selected from a complete factorial design. We should be cautious of factorial designs because, without all possible combinations of attribute levels (complete factorials), information is lost. Not only is information lost in that tests of some effects cannot be conducted, but information lost is confounded with information obtained.[4]

In using fractional factorial design, one is willing to trade off the measurement of all possible interactions effects in order to obtain a smaller number of replicates in which, for example, all single-factor (main) effects and two-factor interactions can still be estimated without confounding. By using this class of designs one assumes that all higher-order interactions (as three-factor and beyond) are negligible. The most common type of fractional factorial involves the case in which all main effects and all two factors interactions can be estimated on an uncorrelated (orthogonal) basis. However, one assumes the higher interactions (involving three or more factors) are negligible and can be ignored. For instance, let us consider a two-attribute case in which both attributes have two levels. Table 3.2 illustrates both the complete factorial and two one-half

[4] Confounded means that some effects are correlated (often highly or perfectly correlated) with other effects. This is the multicollinearity problem in econometrics.

fractions of a 2 × 2 × 2 design. Levels of the three two-level attributes are represented as -1 and +1.

Table 3.2: The 2 × 2 × 2 Design and the Two One-Half Fractions

Main Effects			Two-Way Interactions			Three-Way Interactions
X_1	X_2	X_3	$X_1 X_2$	$X_1 X_3$	$X_2 X_3$	$X_1 X_2 X_3$
-1	-1	-1	1	1	1	-1
-1	-1	1	1	-1	-1	1
-1	1	-1	-1	1	-1	1
-1	1	1	-1	-1	1	-1
1	-1	-1	-1	-1	1	1
1	-1	1	-1	1	-1	-1
1	1	-1	1	-1	-1	-1
1	1	1	1	1	1	1
				Fraction A		
-1	-1	1	1	-1	-1	1
-1	1	-1	-1	1	1	1
1	-1	-1	-1	-1	-1	1
1	1	1	1	1	1	1
				Fraction B		
-1	-1	-1	1	1	1	-1
-1	1	1	-1	-1	1	-1
1	-1	1	-1	1	-1	-1
1	1	-1	1	-1	-1	-1

Table 3.2 shows that both one-half fractions contain columns that are duplicates of one another. Specifically, X_1 is exactly negatively related with $X_2 \times X_3$ in Fraction A, and exactly positively correlated in Fraction B. Therefore X_1 is perfectly correlated with the $X_2 \times X_3$, and the effect of X_1 can be estimated only if $X_2 \times X_3$ is equal to zero. Otherwise, the effect represented by the column X_1 includes both the effects of X_1 and $X_2 \times X_3$, and these effects cannot be separated. Likewise, X_2 is perfectly confounded with the $X_1 \times X_3$ interaction. Like X_1, neither the X_2 nor the X_3 estimates can be interpreted

unless both $X_1 \times X_2$ and $X_1 \times X_3$ are exactly zero. This is the same logic in regression analysis; when interpreting a single coefficient, we will say "holding other variables constant." Lastly, the three-way interactions column labeled $X_1 X_2 X_3$ is perfectly correlated with the intercept or grand mean.

It would not be difficult to ask respondents to rate or rank 8 combinations. However, with the increase of attributes, levels, or both in a design, it will be impractical to execute it. For instance, let us consider a design with five attributes and each of attribute with three levels: it is a $3\times3\times3\times3\times3$ experiment with total 243 combinations. Obviously, it would be impossible to carry out an evaluation study of this magnitude and it would cause confusion and fatigue among the respondents. In such cases, one might wish as few combinations (profiles) as possible. With fractional factorial design, the 3^5 design can be reduced to 16 combinations.[5]

Orthogonal Arrays

In using fractional factorial design, the researcher is willing to trade off information about higher types of interactions for convenience and lower cost in constructing and testing fewer experimental combinations. Among fractional factorial designs, orthogonal arrays are most widely used in conjoint analysis. Before we discuss orthogonal arrays, let us turn our attention to the latin square design. Latin square design is a type of fractional factorial in which high parsimony in number of combinations is achieved by neglecting all interaction effects. In many evaluation-type of experiments

[5] Some useful fractional factorial design catalogs include Hahn and Shapiro (1966), Conner and Zelen (1959), and Anderson (1984).

this may be sufficiently accurate, particularly if one is permitted to transform the original response data monotonically before estimating the model's parameter values.

Let us consider a simple example that combining two (orthogonal) latin squares in order to obtain a graeco-latin square, as illustrated below for two factors, each at three levels:

Table 3.3: A Graeco-Latin Square for the 2×3 Design

		Factor B		
		1	2	3
	1	A_1B_1	A_2B_3	A_3B_2
Factor A	2	A_2B_2	A_3B_1	A_1B_3
	3	A_3B_3	A_1B_2	A_2B_1

Note that each pair A_i, B_j appears exactly once in the Table 3.3 and that each A_i and B_j separately appears once in each row and column.

Orthogonal arrays build on the preceding notion (illustrated by the graeco-latin square) by developing even more highly fractionated designs in which, nevertheless, all main effects can be estimated on an unconfounded basis, assuming that all interaction effects can be neglected. Orthogonal arrays represent the most parsimonious (in the sense of lowest number of combinations) set of designs available for main-effect parameter estimation (Green 1974).

In sum, to use fractional factorial design one must assume that the unobserved effects, which are confounded with effects of interest, are equal to zero. If such assumptions are untenable or if one is uncertain about their satisfaction in a particular design being considered, one should consider alternative designs with different confoundment properties.[6]

Estimation Methods

To analyze input data, there are four major estimation methods available: MONANOVA (Kruskal 1965), PREFMAP (Carroll 1973), LINMAP (Shocker and Srinivasan 1977), and OLS regression. Because LINMAP and OLS are two most popular estimation methods for nonmetric and metric data respectively, the discussion will be focused on LINMAP and OLS methods.

LINMAP Estimation

LINMAP is one of the most popular nonmetric conjoint estimation procedures. The use of linear programming enables LINMAP to obtain global optimum parameter estimates, whereas there is no guarantee that other procedures would achieve global optimums. In addition, in LINMAP, attribute weights can be considered to be non-negative, and part worth functions can be constrained to be monotone or of the ideal point type; such constraints cannot be readily imposed for other approaches.

To show an intuitive explanation of the LINMAP procedure, consider two attributes X_1 and X_2 with three levels, respectively. Let a_0, a_1 and a_2 be the part-worths for X_1 and b_0, b_1 and b_2 be the part-worths for X_2. The full profile design with all 9 possible cards is presented below.

---

[6] For a discussion of fractional factorial design, see Green, Carroll, and Carmone (1978), and Louviere (1988), Chapter 2.

Table 3.4: A Full Profile with 2×3 Design in LINMAP Estimation

Card #	Attribute X_1	Attribute X_2	Utility	Rank
1	0	0	$a_0 + b_0$	9
2	0	1	$a_0 + b_1$	6
3	0	2	$a_0 + b_2$	3
4	1	0	$a_1 + b_0$	8
5	1	1	$a_1 + b_1$	5
6	1	2	$a_1 + b_2$	2
7	2	0	$a_2 + b_0$	7
8	2	1	$a_2 + b_1$	4
9	2	2	$a_2 + b_2$	1

The utility of the profile (card) is equal to the sum of the part-worths of its attributes. The last column displays a respondent's rank order for the profiles, with Card 9 being the most preferred and Card 1 the least preferred.

Since Card 9 is the most preferred, its utility should be greater than that of all other cards, i.e.

$$a_2 + b_2 > a_0 + b_{0,}\ a_0 + b_{1,}\ a_0 + b_{2,}\ a_1 + b_{0,}\ a_1 + b_{1,}\ a_1 + b_{2,}\ a_2 + b_{0,} \text{ and } a_2 + b_{1.}$$

Likewise, the fact that Card 6 is preferred to Card 1, 2, 3, 4, 5, 7, and 8 implies that

$a_1 + b_2 > a_0 + b_{0,}\ a_0 + b_{1,}\ a_0 + b_{2,}\ a_1 + b_{0,}\ a_1 + b_{1,}\ a_2 + b_{0,}$ and $a_2 + b_{1.}$

Similar constraints result from comparing, in turn, the third, fourth, fifth, sixth, seventh, eighth, and ninth ranked cards to other cards of lower preference. Note that LINMAP uses ordinal information about preferences, even if the input data consists of ratings, which are (nominally) measured on an interval scale.

LINMAP uses the mathematical technique of linear programming to determine the values for a_0, $a_{1,}$ $a_{2,}$ b_0, b_1 and b_2 in such a way that the paired comparisons are satisfied as best as possible, taking into account the extent to which any inequalities are violated (Srinivasan and Shocker 1973).

OLS Estimation

The OLS regression approach to conjoint analysis offers a simple, yet robust method of deriving alternative forms of respondent utilities. It is commonly available, inexpensive to use, and easy to interpret estimation method. The attractiveness of the OLS model is in part a result of the ability to scale respondent choices using rating scales, rather than rankings. The OLS procedure has the important advantage of providing standard errors for the estimated parameters (Green and Srinivasan 1978). The ability to implement designs having larger numbers of attributes and levels (through fractional factorial designs) has made this methodology the de facto standard for conjoint analysis. The objective of OLS conjoint analysis is to produce a set of additive part-worth utilities that identify each respondent's preference for each level of a set of product attributes. In application, the OLS model solves for utilities by using a dummy matrix of independent variables. Each independent variable indicates the presence or absence of a particular level. The dependent variable is the respondent's evaluation of one of the profiles described by the independent variables. This model is expressed:

$$U_p = C + \sum_i \sum_j w_{ij} X_{ij} + e_p \qquad (3.16)$$

Where:

U_p = the overall utility to measure of the *p*-th profile,

C = constant or intercept,

W_{ij} = the part-worth utility of *i*-th level *j*-th attribute estimated in the regression,

X_{ij} = dummy variable representing the presence or absence of attribute *j*, level *i* in alternative X, and

e_p = error term for profile P.

Why Conjoint Analysis Works

Conjoint analysis came from the psychometric tradition which is rigorous and idealistic, whereas its adoption by market research community has been approximate and pragmatic. Although the professional success of conjoint practitioners attests to how well it works, we do not have a clear account of why it works. Huber (1987) provides the following reasons for the success of conjoint analysis:

1. Conjoint asks respondents make tradeoffs which are similar to those in the market.

A conjoint task is valuable because it forces the respondent to evaluate conflicting attributes, as between the type of TV and its price. People usually try to avoiding making such decisions by searching for unambiguous solutions involving dominance or following a relatively clear heuristic. However, the market also requires such decisions and people make them when they must in the marketplace or the conjoint task.

2. The simplification in conjoint analysis may mirror that in the market.

It would be misleading to think that conjoint analysis uses a small number of attributes to simplify decision process if people use a very different kind of simplification (e.g. evaluating all factors) in the marketplace. However, there is evidence that people's decisionmaking is based on remarkably few dimensions or factors (Olshavsky and Granbois 1979). If this is true, then conjoint analysis may indicate those few attributes on which people base their decisions.

In addition, to the extent that conjoint analysis is used to predict aggregate shares, it does not matter whether an individual's choice of attributes is unstable over time. As long as conjoint analysis reflects an unbiased choice of attributes at the time, the aggregate shares will also be unbiased. The criteria for conjoint analysis to work at the aggregate level are considerably less strict than for individuals.

3. Conjoint uses orthogonal arrays.

An orthogonal design is that the levels of different attributes across profiles are uncorrelated. Such design makes sure that an estimate of one attribute is unaffected by the estimate of the other attributes. Most econometric models could suffer moderate levels of multicollinearity without much harm. However, respondents often simplify the conjoint task by focusing on one or two variables; this leads to substantial difficulties (i.e. wrong sign) if any attributes in the design are correlated. This explains an inherent advantage in an orthogonal array. For an orthogonal array, the main-effect estimate for each attribute is independent of the others, whereas in the correlated case, the independence does not hold. If attributes are correlated, misspecification leads to biased estimates. The particular misspecification that so often occurs in conjoint is simplification, where a number of attributes are effectively ignored. With orthogonal arrays, the estimated coefficients for the attributes remain unbiased. In the correlated case, misspecification results in distortions in the coefficients for both the attributes focused on and those ignored. Therefore, orthogonal arrays play an important role of increasing the robustness of conjoint analysis by making it less likely that coefficients have counter-intuitive signs.

4. Conjoint simulations can explain heterogeneous preferences in a market.

Typically, part-worth functions are estimated at the individual level, and then these are aggregated to produce estimates of market share under various conditions or scenarios. These simulations implicitly reject the notion that one homogeneous customer can explain choices in the marketplace. Instead, one is forced to deal with each customer having an idiosyncratic preference function, or at least with an explicit clustering in which strongly differing tastes are represented in different clusters. This practice of preserving the heterogeneity of individuals in simulations facilitates the representation of two important properties of markets that are difficult to achieve with other market research techniques. These are the properties of *differential substitution* and *dominance*.

Differential substitution refers to the notion that a new competitor in a market tends to take share differentially from those brands with which it is most similar. For example, Diet Pepsi took most share from Classic Pepsi and Coke, and had relatively little impact on the lemon-lime soft drink (i.e., 7-Up) category. Dominance refers to the idea that a brand that is equal on most attributes but slightly worse than its competitor on others gets a very low share.

Researchers understand these two phenomena and expect simulations to reflect them in positioning and new-product studies. Unfortunately, most models of market research have had difficulty explaining differential substitution and dominance. In contrast, both phenomena arise naturally out of a conjoint simulation. For instance, differential substitution occurs because individuals who like Pepsi tend to like Coke. Generally, changes in any brand will have a greater impact on similar brands than dissimilar ones. Dominance is represented in a conjoint simulation since a brand that is dominated consistently loses out to that competitor and achieves almost no share.

In summary, conjoint analysis works because it is derived from a task that forces respondents to trade off attributes in ways that may mirror actual buying behavior. The orthogonal arrays provide not only efficacy but also strong degree of robustness against misspecification. Finally, the preservation of the utility function at the level of individual or segment permits us to simulate a market that behaves in ways we expect.

The Advantages and Limits of Conjoint Analysis

Advantages of Conjoint Analysis

Several advantages of conjoint analysis are noteworthy:

1. Under conjoint analysis, respondents perform a very realistic task - making trade-offs between different products and services. This avoids many of the *subject artifacts* and *demand effect* of traditional *compositional* approaches in that respondents' separate evaluations of each attribute level are aggregated to form estimates of respondent preferences (Conjoint experiment 1997).

2. Conjoint analysis employs a *decompositional* model which differs from almost all other multivariate methods in which it can be carried out at the level of the individual respondent. This implies that the researcher can develop a separate model for predicting the preferences of each respondent. Most other methods of analysis take measures from each respondent but can only perform an analysis using all respondents combined.

3. Most multivariate methods require the assumption of linear relationships between the predictor and the dependent variables (preference). Conjoint analysis is not hampered by this assumption. It can make separate predictions for the effects of each level of the dependent variables and does not assume that the dependent variable increases (or decreases) in equal amounts for each unit change in the predictor variables. Thus conjoint analysis can handle any type of non-linear relationship, even the complex curvilinear relationship in which one value is positive, the next negative and the third positive again (Hair et al. 1995).

4. Most importantly, individual level models can be used to simulate a market, that is, to predict how a group of respondents would respond to a product or service with a particular set of attributes and values.

5. In a study of the residential choice behavior, Lerman and Louviere (1978) compared the specification of conjoint analysis with 27 different utility specifications. The results showed that each of the 27 specifications was statistically inferior to the specification of conjoint study. They suggested that conjoint analysis can provide information that not only upgrades the predictive power of an econometric choice model, but this information probably could not have been obtained in any other way.

Limits of Conjoint Analysis

Some disadvantages of conjoint analysis are:

1. Some profiles may seem artificial or unlikely to respondents. This is because the use of orthogonal arrays. For instance, in the case of orthogonal arrays a necessary and sufficient condition that the main effects of any two factors be uncorrelated is that each level of one factor occurs with each level of another factor with *proportional* frequencies. In order to make factors with proportional frequencies, therefore, one must create some unrealistic and unlikely profiles.

2. Compared to traditional surveys, filling out a conjoint survey is demanding for respondents, especially when attributes, levels, or both become larger and larger. The conjoint results acquired under such conditions may not be representative of the real life behavior of the individual where one may have more time and motivation to deliberate on the choice from among a small set of options.

3. Respondents may not consider all attributes when evaluating hypothetical situations. Usually respondents will simplify their decisionmaking by focusing on a couple of the most important attributes.

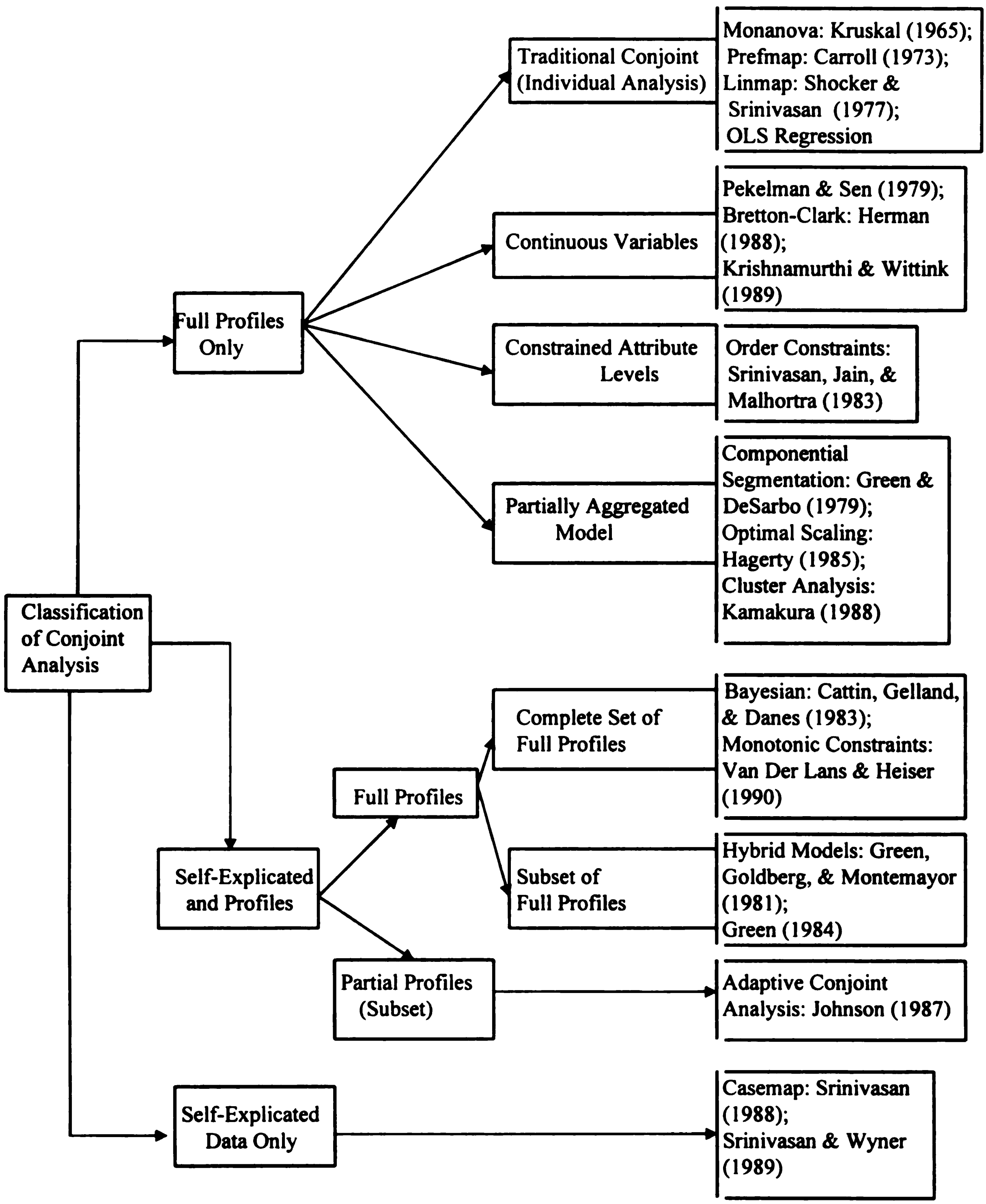
4. Attributes are often limited to only a few levels to keep the number of treatment combinations from becoming too large. Generally one will reduce attributes to six or less than six to avoid the problem of information overload.

5. Conjoint analysis is costly. Since the personal interview is most widely used approach, it is very expensive to collect a large sample size.

Classification of Conjoint Analysis

Figure 3.2 (adapted from Carroll and Green 1995) provides a classification of the various types of conjoint analysis. Conjoint analysis can be conducted in two major methods: *regular or traditional conjoint analysis,* and *hybrid conjoint analysis.* In addition, a related approach - *self-explicated model* - often is used in similar studies. Regular conjoint analysis is based on a *decompositional* method to derive a set of utilities for all individual attribute levels defined in profiles chosen by a special design. In the self-explicated model, the respondent first evaluates the levels of each attribute on a 0-10 (say) desirability scale. Then, the respondent then is asked to allocate, say, 100 points across the attributes so as to reflect their relative importance. Part-worths are obtained by multiplying the importance weights with attribute-level desirability ratings. The hybrid conjoint contains not only those terms in the regular conjoint but also a *compositional* term derived from respondents' self-explicated reactions to a set of attributes and their levels (Green and Srinivasan 1990; Carroll and Green 1995).

Figure 3.2: Classification of Conjoint Analysis



Current Applications on Conjoint Analysis

Conjoint analysis has been extensively applied to marketing research for the following purposes: new product/concept development, pricing, market segmentation, competitive analysis, repositioning, advertising, and distribution (Wittink and Cattin 1989; Wittink, Vriens, and Burhenne 1996). A few researchers have noticed that conjoint analysis could be applied to the studies of policy analysis and public opinion (Green and Srinivasan 1978; Louviere 1988; Shukla and Bruno 1991; Johnson 1995), but the literature has paid very little attention to this important method. There have been reported applications to health care system and services (McClain and Rao 1974; Parker and Srinivasan 1976; Wind and Lawrence 1976; Chakraborty, Ettenson, and Gaeth 1994), water quality-cost alternatives (Whitmore and Cacadias 1974), education policy and school administration (Sparling 1977; Hooley 1981; Ramsing 1982; Zufryden 1983; Shukla 1990), public bus service preferences, residential choice behavior (Louviere 1988: 63-67), competing values in public opinion (Shamir and Shamir 1995), and evaluation utilization (Johnson 1995).

Since it is versatile and applicable to a wide variety of problems and there is computer software available, conjoint analysis is a very promising method which could be applied to the study of policy analysis and public opinion.

Summary

In a nutshell, conjoint analysis is a powerful alternative to traditional surveys. It requires tradeoffs that are similar to those in the real decision process. It employs the best elements of experimental design: it is decompositional rather than compositional; it uses orthogonal arrays to reduce combination, it can simulate preferences under various scenarios. The result from conjoint analysis is that we obtain more accurate information about people's preferences. Thus, conjoint analysis has great potential to be applied to the study of policy analysis and public opinion.

CHAPTER 4

RESEARCH DESIGN

The design for this study can be divided into two parts. In the surveys we first ask respondents the traditional survey questions and then conjoint questions. The data for this study were collected from a convenient sample of 77 undergraduate students in Michigan State university and a random sample of 128 Meridian Township voters in Michigan. (A detailed description will be discussed in Step 8 of the procedure).

Step by Step Procedure of Conducting a Conjoint Analysis

Green and Srinivasan (1978), Churchill (1995), and Hu (1996) all provide some suggestions or guidelines for conducting a conjoint survey. Based on the literature review and the purpose of this study, we provide a step by step procedure as follows: (1). Select Attributes and Determine Attribute Levels, (2). Select Preference Model and Composition Rule, (3). Determine Data Collection Method, (4). Construct fractional factorial design, (5). Select Stimulus Presentation Method, (6). Select Measurement Scale for the Dependent Variable, (7). Collect Data, (8). Select the Estimation Method, (9). Interpret the Results and Decide on Aggregation of Judgments, and (10). Check Reliability and Validity.

1. Select Attributes and Determine Attribute Levels

The first step in the process involves deciding on the attributes to be used and specifying the actual levels for each attribute when constructing the profiles.

The Principles of Selecting Attributes

When choosing attributes, researchers should be guided by the principles that the attributes used should be both *feasible* and *important* to individuals. Feasible attributes are those that the organization can do something about, in that the organization has the technology or other resources to make the changes that might be indicated by consumer preferences. Important attributes are those that actually affect consumer choice. The attributes actually used can be determined by managerial judgment or by any of the other techniques which are productive at the exploratory research stage of an investigation, including depth interviews with individual subjects, focus groups, analysis of insight-stimulating examples, and so on. In any single conjoint analysis study, only a handful of the attributes that could be used will be used, so it is important that they are carefully selected. When the number of attributes that need to be varied exceed reasonable limits with respect to data collection problems, a series of pilot studies can be conducted. In a typical conjoint analysis, the number of attributes actually used is less than six.

The Principles of Deciding Attribute Levels

The number of levels for each attribute has a direct influence on the number of attributes respondents will be asked to evaluate and consequently on the burden placed on each subject. In general, we wish to minimize that burden; however, at the same time, we wish to end up with good estimates of the utility of each attribute level. Our ability to generate good estimates requires that the number of attributes be relatively large versus the number of parameters that need to be estimated, and the number of parameters in turn depends on the preference model being used. Thus, it is an essential question to ask: Is the model linear in the sense that more or less of the attribute can be expected to be most desired, or is it nonlinear in a systematic way, or could there be a very nonsystematic relationship between preference and attribute levels?

Another factor that affects the choice of attribute levels is their effect on consumer choice. Using levels which are similar to those in existence increases respondents' belief in the task and the validity of their preference judgments. In addition to that, using attribute levels which are outside the range normally encountered decreases the believability of the task for respondents but can increase the accuracy by which the parameters can be estimated statistically. The general recommendation is to make the ranges for the various attributes somewhat larger than what is normally found but not so large as to make the options unbelievable.

Attributes and Levels Used in This Study

Based on the feasibility and importance considerations, in the current study four determinant attributes are used. They are four largest spending categories in the state government of Michigan: *education*, *health and welfare*, *public safety and corrections*, and *natural resources and environmental control.* Each of these attributes has three spending levels: increase spending 5 percent, decrease spending 5 percent, or maintain current spending levels.[1] The four categories explain 80.2% of total state expenditures (actual spending) in the 1996 fiscal year. Meanwhile, the four categories also represent the largest increase in spending items from 1990-96 fiscal years.[2] Based on Governor

[1] We consider the most widely discussed issue "taxes vs. government spending." However, adding one more attribute "taxes" will lead to a increase in minimum profiles from 9 to 16; not even mention double the minimum profiles (32). Such large profiles will definitely cause the problem of information overload. Since this a pilot study, we think it is better to keep the format as simple as possible. Therefore, we only use 4 determinant attributes in this study.

[2] Calculated based on the report of the Department of Management and Budget in Michigan.

Engler's platform,[3] the four categories should be an appropriate indicator of his budget priorities. Furthermore, the four categories include some of the most salient issues of today such as education, Medicaid, health care, crime, and environmental quality. Studies have showed that respondents usually decide and answer issues most salient to them (Sniderman 1993: 221). Hence, focusing on these salient issues can help respondents answer questions more easily and reduce the information overload problem.

In addition, a *pretest* of 10 cases was conducted to decide whether the design is reasonable and appropriate and whether respondents can correctly interpret the attributes and levels. For instance, we attempted to determine whether the breakdown of the expenditures was reasonable and whether the levels of each factor were appropriate.

2. Select Preference Model and Composition Rule

Mentioned earlier, generating good estimates requires that the number of attributes be relatively large versus the number of parameters that need to be estimated, and the number of parameters depends on the preference model being used. In other words, we should decide which preference models will be applied: vector, ideal point, or part-worth models?

However, choosing a preference model is not so simple as it appears to be (linear, quadratic, or piecewise). First, Green and Srinivasan (1978) point out that the complexity (and number of parameters required) increases from vector to part-worth. This results in the precision of estimated parameters decreasing from vector to part-worth because more parameters must be estimated from the same number of observations. On the other hand,

[3] According to Governor John Engler's State of the State Address (1998), his next steps to make Michigan first in the 21st century includes cutting taxes, protecting the environment, education, strengthen families,

ability to fit an arbitrary pattern of true preferences increases from vector to part-worth (i.e., the vector model can fit only a linear pattern of true preferences, whereas the part-worth model can fit linear, quadratic, or any other pattern). Since these factors tend to go in opposite directions, no a priori recommendations can be made on which model to use. Later studies have proved that simple models predict *individual preferences* better than complex models (Pekelman and Sen 1979; Cattin and Punj 1984).

In addition to selecting a preference model, whether *interactions* should be included in the preference model also affect the model's ability to accurately predict people's preferences. This issue involves deciding the composition rule, that is, whether the simple additive model is enough to accurately predict preferences or we should consider using multiplicative, distributive, or dual-distributive model. Once again, according to previous studies, the conclusions are that the simpler models perform better than the more complex models (Akaah and Korgaonakar 1983; Neslin 1981, and Goldberg 1967).

Even though these findings seem to favor simple models, Hagerty (1986) showed that simple models did predict better for *individuals' preferences* but complex models predict better for *market share*. Now it seems that the choice among preference models depends on the trade-off between the more complex interaction models and their costs. Specifically, the major obstacle for using a complex model is the increased burden placed on a respondent by the need for a large design matrix to obtain the degrees of freedom necessary to estimate a more complex model. In order to obtain more accurate

---

and fighting crime.

prediction, for instance, Hagerty's analysis used up to 32 profiles. Such design definitely will cause respondents' tendencies to become careless when rating or ranking profiles.

Based on above considerations, the *additive part-worth model* is used in this study for several reasons. First, so far there is no theory can indicate that people's preferences for the four spending categories (education, welfare, crime, and natural resources) are linear, curvilinear, or piecewise. Nevertheless, the part-worth model provides the greatest flexibility in allowing different shapes for the preference function along each of the attributes. Second, since the design in this study only has three levels for each attribute, the part-worth model is indistinguishable from the quadratic (ideal point) form (Hagerty 1986). Third, since no theory shows that there are interactions among the four spending categories, an additive model should be enough to accurately predict preferences. In practice, conjoint studies rarely apply models beyond the additive form because the simple model is easier to apply and the simpler model (additive) performs better than the more complex models such as the multiplicative or distributive models (Akaah and Korgaonakar 1983). Hence, the additive part-worth model should work well in this study.

3. Determine Data Collection Method

Two major data collection procedures are used in conjoint analysis: *full profile approach* and *two-factor-at-a-time procedure.*

The *full-profile* approach (or *concept evaluation*) uses the complete set of factors (see Table 4.1 for a comparison of full profile and two-factor-at-a-time procedures). Respondents provide rank orders (or rating) of preference for product concepts which differ simultaneously with respect to all attributes been studied. The full profile approach has the advantages of greater "realism," since respondents are choosing among concepts which are more elaborately specified, and at least theoretically, of being able to quantify interactions among attribute. The major drawback of this procedure is the possibility of information overload and the resulting temptation on the part of the respondent to simplify the experimental task by ignoring variations in the less important factors or by simplifying the factor levels themselves. The conjoint results acquired under such conditions may not be representative of the real life behavior of the individual where one may have more time and motivation to deliberate on the choice from among a small set of options. Because of the problem of information overload, the full profile procedure is usually limited to six factors in any specific sort.[4]

The *two-factor-at-a-time procedure* (or paired comparison) considers factors on a two-at-a-time basis. The respondent is asked to rank the various combinations of each pair of attribute levels from most preferred to least preferred. The *two-factor-at-a-time* approach is simple to apply and it can reduce information overload on the part of the respondent. However, two-factor-at-a-time approach sacrifices realism; people are usually unclear as to what should be assumed about the t-2 attributes that are not being considered in a specific evaluation task. In addition, there is a tendency for respondents either to forget where they are in the table or to adopt patternized types of responses, such as always attending to variations in one attribute before considering the other.

Furthermore, studies have shown that the full profile approach usually overstates the importance of the most preferred attribute while the two-factor-at-a-time approach

[4] To overcome the information overload problem, one could use "bridge-type" factors design. See Green and Srinivasan (1978) for a discussion.

usually overstates the importance of the least preferred attribute (Green and Srinivasan 1978).

In this study, the full profile approach is used to collect data because it provides greater realism and entails fewer judgments by the respondent (when implemented by various kinds fractional factorial design). In addition, our design only has four factors with three levels in each attribute, hence the problem of information overload should be reduced to a minimal or negligible level.

Table 4.1: Alternative Data Collection Methods

Full Profile Approach	Profile A
Education	5%
Health & Welfare	-5%
Crime & Public Safety	5%
Natural Resources	0%

Two-Factor-at-a-Time

Health & Welfare	Education -5%	Education 0%	Education 5%
-5%	7	3	1
0%	8	5	2
5%	9	6	4

Note: 1 denotes the most preferred combination and 9 denotes the least preferred combination.

4. Construct Fractional Factorial Design

As previous discussed, it would be too costly to carry out a conjoint survey if we use the full set of profiles. For example, there are four attributes (factors) in this study: education, welfare and health, public safety and corrections, and natural resources. Every attribute has three levels: decrease 5% spending, status quo, and increase 5% spending. If we array all possible combinations of the four attributes, a total of $3\times3\times3\times3 = 81$ alternatives would have to be tested. Obviously, it would be too costly to carry out an evaluation study of this magnitude and it would cause confusion and fatigue among the respondents. An appropriate way is to use various types of *fractional factorial* design to reduce the number of combinations to a manageable size while at the same time maintaining orthogonality. The most common type of fractional factorial design involves the case in which all main effects (single factor) and all two factor interactions can be estimated on an uncorrelated (orthogonal) basis. However, one assumes the higher interactions (involving three or more factors) are negligible and can be ignored.[5] In this study, the *main-effects* plan is adopted which permits one to estimate only the main effects, assuming negligible interactions (Green 1974).

Orthogonal main-effect plans can be built for symmetrical factorial experiments involving $(s^n - 1)/(s - 1)$ factors, each having s levels, with s^n combinations, where s is a prime or the power of a prime number (Addelman 1962). In this study, $s=3$, $n = 2$ and $(s^n - 1)/(s - 1) = 4$. So, we only need s^n $(3^2) = 9$ combinations for this study. If we used any fewer than this number of profiles, we would not be able to derive a set of utilities. And if we used only the minimum number, we could not check the respondent's internal consistency in evaluating the profiles (Curry 1997). Therefore, it is generally recommended that we include 1.5 to 2 times the minimum number of profiles in the ranking or rating task. A necessary and sufficient condition that the main effects of any two factors be uncorrelated is that each level of one factor occurs with each level of another factor with *proportional* frequencies (Addelman, 1962).

For comparison purposes, 10 profiles will be used in a random sample while 18 profiles will be used in the convenient sample. Such design is based on the consideration

[5] For a discussion of fractional factorial design, see Green, Carroll, and Carmone (1978) and Louviere (1988), Chapter 2.

that 10 profiles are more manageable in a mail survey (voter sample). On the other hand, it is not too difficult to ask students rank ordering 18 profiles in the classrooms (see Appendix for both voter and student survey forms). Therefore, in the voter sample we cannot check respondent's internal consistency while in the student sample we could check respondent's internal consistency.

5. Select Stimulus Presentation Method

Step 5 in the process involves selecting the form of presentation of the stimuli and the nature of the judgments to be secured from subjects. The presentation of the hypothetical stimuli in the full profile approach has involved variations and combinations of three basic approaches: *verbal description*, *paragraph description*, and *pictorial representation*. The paragraph description is used in both voter and student samples.

6. Select Measurement Scale for the Dependent Variable

Related to the issue of the form of presentation of the stimuli is the nature of the judgments that will be secured from respondents; that is, the measurement scale for the dependent variable. The two most common approaches measured respondents' preferences for each alternative are: nonmetric (rank order, paired comparison) or metric (rating scales assuming approximately interval scale or ratio scales). Some of the main reasons advanced by those using rank-order judgments are their ease of use by subjects, ease of administration, and a desire to keep the judgment task as close as possible to a consumer's behavior while actually making decision. In addition, ranking data are likely to be more reliable, since it is easier for a respondent to say which he or she prefers something more as compared to expressing the magnitude of his or her preference. Those using rating scales believe that they are less time-consuming, are more convenient for respondents to use, and are easier to analyze (through the ordinary least squares).

For comparison purposes, we used both ranking and rating approaches to collect data. In all the first round of surveys, ranking is used. In the second round of voter surveys, half of the respondents were asked to rate 12 profiles while another half of the respondents were asked to rank order 10 profiles.

7. Select the Estimation Method

Step 7 in the execution of a conjoint analysis involves selecting the technique by which the input data will be analyzed. Four estimation methods are available for this study: MONANOVA (Kruskal 1965), PREFMAP (Carroll 1973), LINMAP (Shocker and Srinivasan 1977), and OLS regression. For the comparison purpose, both LINMAP and OLS were used to estimate all input data. LINMAP is one of the most popular nonmetric conjoint estimation procedures. The use of linear programming enables LINMAP to obtain global optimum parameter estimates, whereas there is no guarantee that other procedures would achieve global optimums. The OLS regression approach to conjoint analysis offers a simple, yet robust method of deriving alternative forms of respondent utilities. It is commonly available, inexpensive to use, and easy to interpret estimation method. The attractiveness of the OLS model is a result of the ability to scale respondent choices using rating scales.

8. Collect Data

Samples

The data for this study were collected from a convenient sample of 77 undergraduate students in Michigan State university and a random sample of 128 Meridian Township voters in Michigan.

Purposive (nonrepresentative) sample: A purposive sample of undergraduate students were collected from the Michigan State University. In July 1997, we collected 77 completed surveys from the students enrolled in the second summer school session. In August 1997, we obtained 22 completed follow-up surveys out of the 79 students.[6] Even though a nonrepresentative sample cannot be generalized to make inferences about the distribution of attitudes in the total population, several advantages make it attractive: (1) it is less expensive, (2) it can collect a large number of cases at relatively small cost, (3) researchers can use questions that reflect their theoretical interests, and (4) small subcultural groups or other special populations can be studied (Abramson 1983: 23-26).

Random Sample: A random sample of 128 registered voters were collected from the Meridian Township in Michigan by using random numbers created by *uniform distribution* through the software called *Crystal Ball.*[7] To maximize the response rate, the subjects were first contacted by telephone and asked to participate. Several days later, the respondents who agreed to participate received a survey package by mail.[8] The package included a cover letter, the questionnaires (both traditional and conjoint

---

[6] Originally, we planed to collect the follow-up sample in a two month interval. However, the summer session only lasts one and a half months, hence the data were collected in August (one month interval).
[7] Crystal Ball is software for business managers. It provides all kinds of simulations for decision making.
[8] It is extremely difficult to carry out conjoint survey by telephone. The best way to conduct conjoint survey will be the personal interview. Due to the budget constraints, we could only afford the mail survey.

questions), a postage-paid return envelope, and a lottery ticket as an incentive. The cover letter thanked them for participating and asked that they returned the completed survey within 10 days. In the late November 1997, the telephone solicitation netted 421 completed contacts, of which 251 agreed to participate. Later, we obtained the total number 128 completed surveys, a response rate of 50 percent of those who agreed to participate or 30 percent of the original completed contacts.

Approximately two months later, we sent questionnaires to those who previously agreed to participate in our follow-up surveys. For comparison purposes, we divided 88 agreed subjects into two groups (44 subjects in each groups). In the first group, we asked respondents to rank order 10 profiles but with a different orthogonal array. In the second group, we asked respondents to rate 12 profiles in a 0-100 points scale. We obtained 34 completed surveys from ranking data while 27 completed surveys from rating data.

Format

In both student and voter surveys, we first asked 5 traditional questions before conjoint questions. In student surveys, traditional questions were presented with the format of unbounded financial resources, such as "Generally speaking, do you think the State of Michigan should: increase overall spending, decrease overall spending, or keep overall spending at current levels?" and "In the area of _________ (i.e., education), you would prefer to see the State of Michigan: increase spending in the area of education, decrease spending in the area of education, or maintain education spending at current levels?" In voter surveys, traditional questions were presented by asking respondents expressing their own ideal spending levels (except first general question), such as "In the area of _________ (i.e., education), you would prefer to see the State of Michigan:

increase spending in the area of education by _____%, decrease spending in the area of education by _____%, or maintain education spending at current levels?"

In the conjoint questions, we asked respondents to rank order 18 profiles (student sample) or 10 profiles (voter sample) from the most preferred to the least preferred. In the rating survey (second round for voter), we asked respondents to rate 12 profiles in a 0-100 points scale. This rating approach will allow respondents filling in tied scores. In the end of the surveys, respondents were asked to fill in the demographic information in order to analyze group differences.

9. Interpret the Results and Decide on Aggregation of Judgments

Step 9 involves interpreting the results and deciding if the responses from individual subjects will be aggregated and, if so, how?

The Principles of Making Aggregation Judgments

Although it is possible to derive the utilities for each level of each attribute at the individual level, individual results are very difficult for researchers to use for developing marketing strategy. The other extreme is to pool the results across all respondents and then to estimate one overall utility function. The middle ground is to form segments from groups of respondents in such a way that the models for the groups will have predictive power close to that found for the individual-level models while having some clear marketing strategy implications for researchers. This raises the question of how these groups should be formed. Most typically it is done by forming segments that are homogeneous with respect to the benefits that the respondents want from the product or service. Operationally, this often translates into estimating utilities for the individual-level models and then clustering respondents into groups that are homogeneous with respect to the utilities assigned to the various levels of the individual attributes.

An attractive feature of conjoint analysis is that it allows market share predictions for selected product alternatives. For example, a common choice rule is the *first choice* rule, which assumes that each respondent chooses the object with the highest predicted preference. Given the estimated utilities for each level of each attribute, the researcher then can investigate which of several product options being considered is likely to appeal most to respondents and what may also be the share of preference for each of other options. In this study, we ran several simulations for gubernatorial election in Michigan and welfare policy evaluation. In addition, the researcher can also attempt to link individual part-worth utilities with selected consumer characteristics, such as whether Democrats have higher utilities for an increase in welfare expenditures than Republicans or not.

Interpreting Results

We analyzed the results by comparing traditional surveys results with conjoint result: how different are respondents' preferences in traditional and conjoint surveys? We also examined the differences among different groups such as gender, religion, party identification, etc. under traditional and conjoint results. In addition, we simulated the gubernatorial election in Michigan to see which candidates (John Engler or Jeffrey Fieger) would win the election based on their positions on the four policy areas. We also evaluated respondents views of current welfare reform in Michigan: Do voters in Meridian Township satisfy the current welfare spending level?

10. Check Reliability and Validity

Test of Reliability

Conjoint analysis can provide more accurate information on people's preference. But can it also provide reliable results? Public opinion studies have shown that people tend to have unstable responses to survey questions. The major explanations for the response instability are: (1) people do not have meaningful beliefs (Converse 1964), or (2) people have true attitudes but survey tools cannot detect it due to measurement errors (Achen 1975). Thus, it is interesting to know if a different survey tool (conjoint analysis) can provide better reliability.

In this study, all the surveys asked respondents traditional survey questions and conjoint questions. In conjoint analysis, several different tests were used to check reliability. Reibstein, Bateson, and Boulding (1988) provide four types of reliability for conjoint analysis:

(1) *Reliability Over Time* asks: "Would the answers be the same at different time?

(2) *Reliability Over Stimulus Set* asks: "Would the answers be the same if a different set of profiles had been used?"

(3) *Reliability Over Attributes Set* ask: "Would the utilities for a given set of attributes have been the same if these attributes had been included in a study with other attributes?"

(4) *Reliability Over Data Collection Procedure* ask: "Would the answers obtained have been the same if a different data collection procedure had been used?"

In our study, we evaluate the following three criteria: reliability over time, reliability over stimulus set, and reliability over data collection method. First, we test reliability by using *alternate forms method with spaced testing*: after finishing a conjoint analysis task, a subset of the respondents will be approached after a period of the time (one or two months) and asked for preference judgments on a second set of profiles. The second profiles ("alternate form") would also have *n* descriptions from the same factor levels as the first set but would avoid duplication of profiles from the original one. Product moment correlation and Spearman rank order correlation are used to test reliability.

Tests of Validity

Except for checking reliability, we also test validity (only in student sample) by the following ways:

1. Cross-validity: the ability of the model to predict the ranking (or first choice) of a holdout set of profiles (only for the student sample). In addition, data used for the reliability tests also provide a method of cross-validation. The parameters of the preference function estimated from the first set of preference data can be used to predict the preferences for the second set. The predicted preferences can be correlated to the actual to obtain a measure of cross-validity.

2. Predictive validity: predictive validity is measured by holdout profiles (student sample only). Holdout profiles are additional profiles which the respondent evaluates beyond those generated by actual conjoint design. In the student sample, we include three holdout profiles in our design. In particular, we use first 15 profiles (Profile 1-15) to predict last three profiles (Profile 16-18).

Summary

In this study, we provided a 10 step procedure to conduct conjoint analysis: (1) Four determinant attributes are were chosen: *education*, *health and welfare*, *public safety* *and corrections*, and *natural resources* - and each of these attributes with three spending levels: increase spending 5 percent, decrease spending 5 percent, or maintain current spending levels, (2) An additive part-worth model was selected, (3) A full profile approach was utilized to collect data, (4) 10 profiles were evaluated in voter surveys while 18 profiles were evaluated in student surveys, (5) A paragraph description was used to present stimulus, (6) Rank order and rating were chosen to measure dependent variables, (7) Data were collected from a convenient sample of 77 undergraduate students in Michigan State university and a random sample of 128 Meridian Township voters in Michigan, (8) LINMAP and OLS were utilized to estimate part-worth utilities, (9) Traditional and conjoint results were compared, part-worth utilities differences among groups were analyzed, and simulations for gubernatorial election and policy evaluation were executed, and (10) Product moment correlation and Spearman rank order correlation were adopted to test reliability while holdout choices were used to test validity.

CHAPTER 5

RESULTS AND DISCUSSIONS

In the student sample, respondents were asked to rank order 18 profiles in both first and second round of surveys but with a different orthogonal array in the second round of survey. In the voter sample, respondents were asked to rank order 10 profiles in the first round of survey. In the second round of the survey, half (N = 44) of voters were asked to rank order 10 profiles while another half (N = 44) were asked to rate 12 profiles in a 0 - 100 points scale. Both ranking and rating data were estimated by ordinary least squared regression (by SPSS 8.0) and LINMAP (by CONJOINT LINMAP). In this chapter we first discussed the part-worth utilities, relative importances of attributes, overall utilities of profiles, and goodness-of-fit measures obtained from Meridian Township voters and undergraduate students in Michigan State University. The logic of LINMAP and OLS has been discussed in Chapter 3. Since OLS is widely used procedure in political science, the interpretations of conjoint analysis will be emphasized on OLS results. Then, the results from both traditional survey and conjoint analysis were compared by using different simulation models.

Interpreting Conjoint Results from Voter Sample

Interpreting the Part-worth Utilities from Voter Sample

The additive part-worths model was used in both voter and student samples. For each individual, the part-worths (weights or regression coefficients) were estimated by OLS. The purposes of part-worths were used to determine (1) the importance of levels of attributes, (2) the relative importances of attributes, and (3) the overall utilities of profiles.

The OLS regressions were estimated at the individual level, for example, resulting in 128 equations for voters (first round of survey). Furthermore, these 128 "individual equations" were averaged, resulting in an "aggregate equation" for voters. Then, the aggregate equation is used to calculate the relative importances of attributes and estimate the overall utilities for 10 profiles.

Recalled Equation 3.16 (in Chapter 3), the OLS model can be expressed:

$$U_p = C + \sum_i \sum_j w_{ij} X_{ij} + e_p$$

Where:

U_p = the overall utility to measure of the p-th profile,

C = constant or intercept for profile P,

W_{ij} = the part-worth utility of i-th level j-th attribute estimated in the regression,

X_{ij} = the dummy variable representing the presence or absence of attribute $_j$, level $_i$ in alternative X, and

e_p = error term for profile P.

For illustration, let us consider the example of voters in the first round of survey with an orthogonal array of 10 profiles (see Table 5.1). Now for W_{ij} and X_{ij}, the i = 0, 1, and 2 (0 = -5%, 1 = 0%, and 2 = +5%) while j = 1, 2, 3, and 4 (1 = education, 2 = welfare/health, 3 = crime/public safety, and 4 = natural resources/environmental control).

Table 5.1: Orthogonal Array for Voters (First Round of Survey)

Profile	Education	Welfare	Crime	Natural
1	0	0	0	0
2	0	1	1	2
3	0	2	2	1
4	1	0	1	1
5	1	1	2	0
6	1	2	0	2
7	2	0	2	2
8	2	1	0	1
9	2	2	1	0
10	1	0	2	0

From Table 5.1, for instance, the Profile 1, 4, and 7 can be written as:

$$U_1 = C + W_{01}X_{01} + W_{02}X_{02} + W_{03}X_{03} + W_{04}X_{04} + e_1 \qquad (5.1)$$

$$U_4 = C + W_{11}X_{11} + W_{02}X_{02} + W_{13}X_{13} + W_{14}X_{14} + e_4 \qquad (5.2)$$

$$U_7 = C + W_{21}X_{21} + W_{02}X_{02} + W_{23}X_{23} + W_{24}X_{24} + e_7 \qquad (5.3)$$

In these cases, different profiles represent different spending preferences and can be expressed as follows:

Profile 1 indicates that a voter prefers cutting spending across-the-board, Profile 4 shows that a voter favors a decrease in welfare spending and maintaining spending education, crime and natural resources at current levels, and Profile 7 reveals that a voter prefers increasing spending in education, crime, and natural resources while cuts in welfare.

Part-worth Utilities of Attributes

Table 5.2 reveals the voter's part-worth utilities from OLS and LINMAP estimations. Although the scaling on each set of part-worth utilities differs due to computational differences, the utility estimates from each technique are quite similar.[1] This can be seen by comparing the %ΔUtility columns in Table 5.2, which shows the percent change in part-worths that come from either spending increases or decreases when compared with the current level of spending.

In the area of education (see Figure 5.1 and 5.2) , for example, both OLS and LINMAP show that voters suffer a significant utility loss from decreases in education spending (utility drop from 0.357 to -1.302 in OLS and from 3.091 to -10.833 in LINMAP) and a mild support in utility from increasing education spending (utility increase from 0.357 to 0.945 in OLS and from 3.091 to 7.741 in LINMAP). In the areas of welfare, crime, and natural resources, voters are better off when spending on status quo while worse off when either spending cuts or increased. The results indicate that in general Meridian Township voters are satisfied with most of current spending levels in Michigan.

[1] Cattin and Wittink (1976) report that the difference between OLS and LINMAP is about 0.03 units in terms of Pearson's rho, with OLS being the better approach when the attribute weights are normally distributed (compensatory model) and LINMAP being the preferred method when the attribute weights show a lexicographic structure (noncompensatory model). Jain et al. (1978) report that MONANOVA, LINMAP, LOGIT, and OLS produce roughly the same level of cross-validity, with LOGIT and LINMAP being slightly preferred procedures.

Table 5.2: Part-worth Utilities from OLS and LINMAP (Voters)

		OLS		LINMAP	
Factor	***Level***	***Utility***	***%ΔUtility***	***Utility***	***%ΔUtility***
	-5%	-1.302	-464.5%	-10.833	-450.5%
Education	0%	0.357	-----	3.091	-----
	5%	0.945	164.5%	7.741	150.4%
Relative Importance		52.44%		50.44%	
	-5%	-0.052	-115.0%	-0.672	-121.6%
Welfare/Health	0%	0.344	-----	3.105	-----
	5%	-0.292	-185.0%	-2.433	-178.4%
Relative Importance		14.83%		15.04%	
	-5%	-0.250	-178.2%	-2.680	-200.4%
Crime/Public Safety	0%	0.320	-----	2.668	-----
	5%	-0.070	-121.8%	0.012	-99.6%
Relative Importance		13.31%		14.52%	
	-5%	-0.447	-216.2%	-4.587	-265.2%
Natural Resources	0%	0.385	-----	2.776	-----
	5%	0.062	-83.8%	1.810	-34.8%
Relative Importance		19.42%		20.00%	
Constant		5.521		N/A	
N=		128		128	

Figure 5.1: Part-Worth Utilities for Voters (OLS)

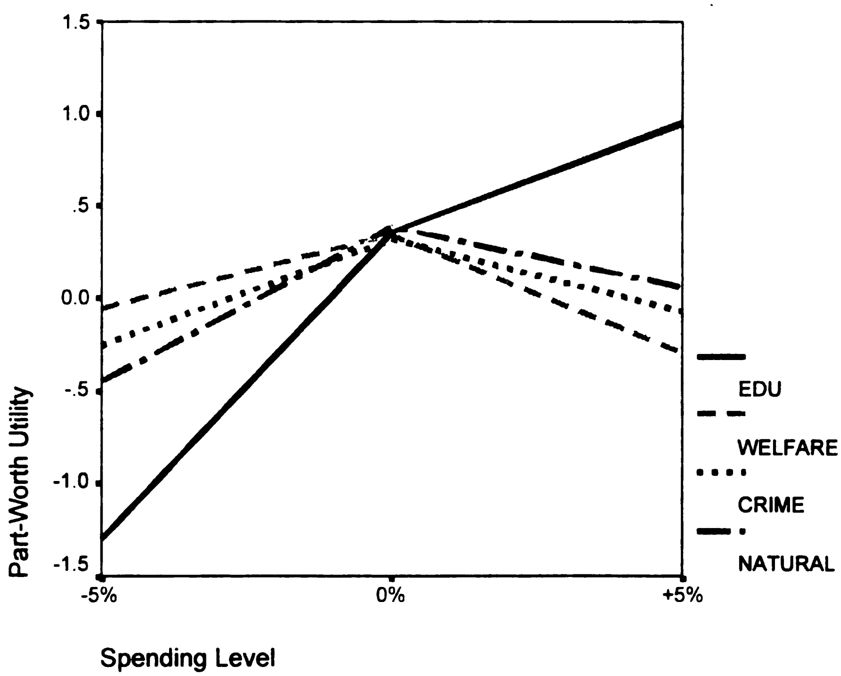
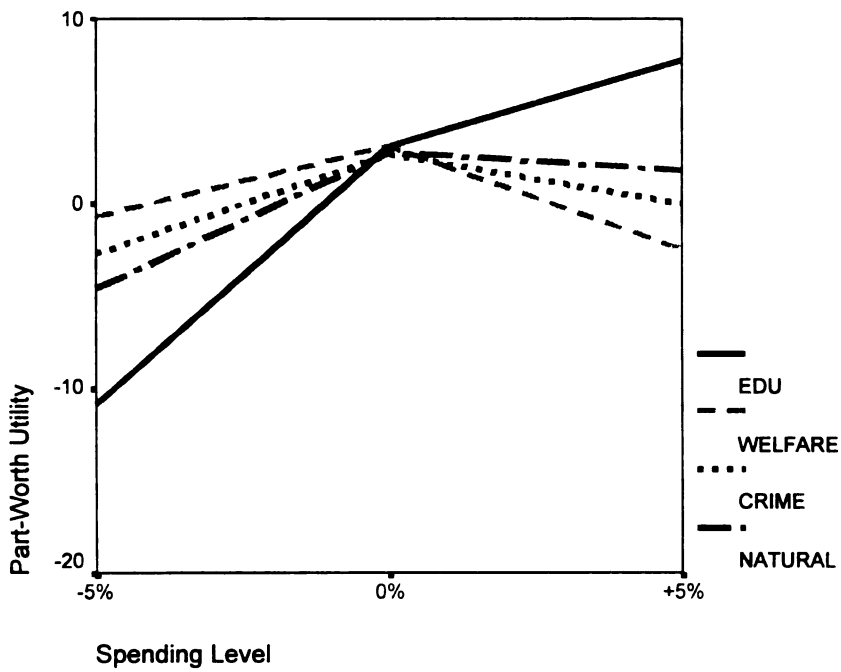


Figure 5.2: Part-Worth Utilities for Voters (LINMAP)



General Notes on Interpreting Part-worth Utility

The part-worth model indicates that the attribute levels are categorical and that no assumption is made about the relationship between the attribute and the rank and scores. The input data are scaled so that for each attribute, the utilities of the levels sum to zero. In other words, the average utility for the levels of the attribute is zero (i.e., the average utility for education is: -1.302 + 0.357 + 0.945 = 0 in OLS), and the actual part-worths indicate how much better or worse than average that level is. It should be noted that these utilities are all relative to the other levels of the attribute. That is, while we can state that level A is preferred to level B, and we can also state the amount that it is preferred by, we cannot say in any absolute way that level B is disliked. For instance, if a level of a attribute has a negative average utility, this only indicates that the level is less preferred than the average level for that attribute, and not that it is disliked in any absolute sense.

For example, consider the area of education with the following spending levels: +100%, +50%, and +5%. In this case, it is safe to bet that +5% would have a negative utilities for most respondents. However, this does not imply that they dislike +5%. If we use different levels: +5%, 0%, and -5%, it is possible that the same respondents would have highly positive utilities for +5% - even though their absolute "liking" for +5% did not change across the two studies.

It should also be mentioned that interpretation of conjoint utilities is true of all types of conjoint analysis, no matter how they are scaled (i.e., how you define levels). For example, one method of scaling sets the utility for the "worst" level of a attribute to zero. However, this does not mean that the lowest level has no utility to the respondent, just that it is X utiles worse than the next lowest level. Furthermore, it does not imply that the worst level of each attribute has the same utility, even though they all receive a utility of zero.

Relative Importance of Attributes

The second purpose of part-worth utility is to determine the relative importance of attributes. Relative importance is based on the assumption that the larger the difference between maximum and minimum level utilities within attribute, the more determinative and salient the attribute is in the overall evaluation of profiles. The relative importance of an attribute can be calculated from Equation 3.1:

$$RI_{j} = \frac{Max(X_{ij}) - Min(X_{ij})}{\sum[Max(X_{ij}) - Min(X_{ij})]} \times 100\%$$

Where

RI_j = relative importance of attribute j.

Max (V_{ij}) = maximum level (*i*) utility in attribute j.

Min (V_{ij}) = minimum level (*i*) utility in attribute j.

For example, the relative importance for education in OLS is:

$$RI_1 = \frac{0.945 - (-1.302)}{0.945 - (-1.302) + 0.344 - (-0.292) + 0.320 - (-0.250) + 0.385 - (-0.447)} \times 100\%$$

$$= 52.44\%$$

As seen in Table 5.2, Figure 5.3 and Figure 5.4, for the attributes and levels tested, Meridian Township voters view education as the most important attribute in their evaluation of budget priorities with a relative importance rating of 52.44 percent in OLS and 50.44 percent in LINMAP. For the attributes and levels tested, although natural resources is the second most important attribute (19.42 percent in OLS and 20 percent in LINMAP), its rating is only slightly higher than welfare and health (14.83 percent in OLS and 15.04 in LINMAP) and crime and public safety (13.31 percent in OLS and 14.52 in LINMAP). One thing should be mentioned: Because the relative importance is determined by the range rather than direction of levels, it should be noted that the highest relative importance rating could go either way: the highest attribute's relative importance rating could be the most desirable or least desirable attribute.

Figure 5.3: Relative Importance of Attributes for Voters (OLS)

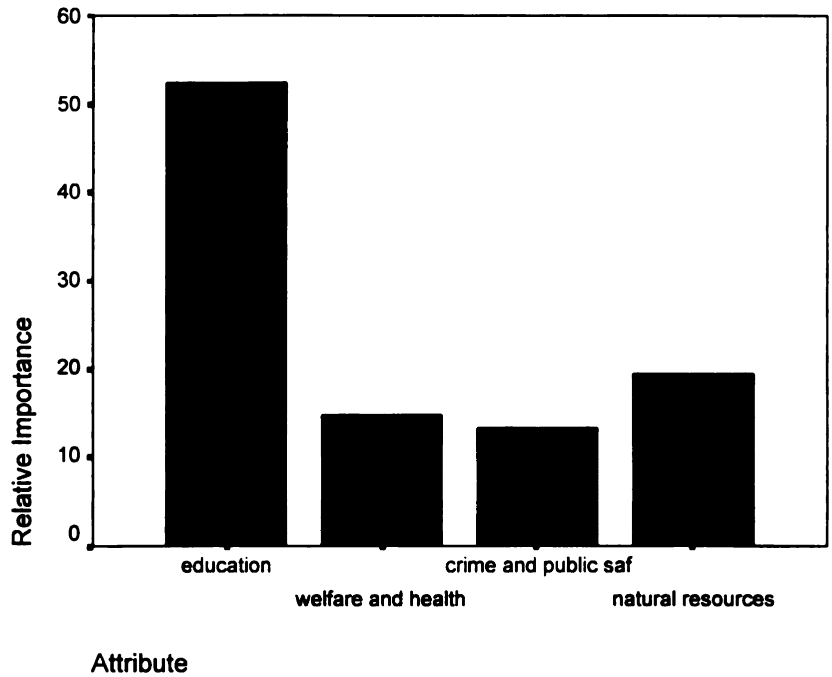
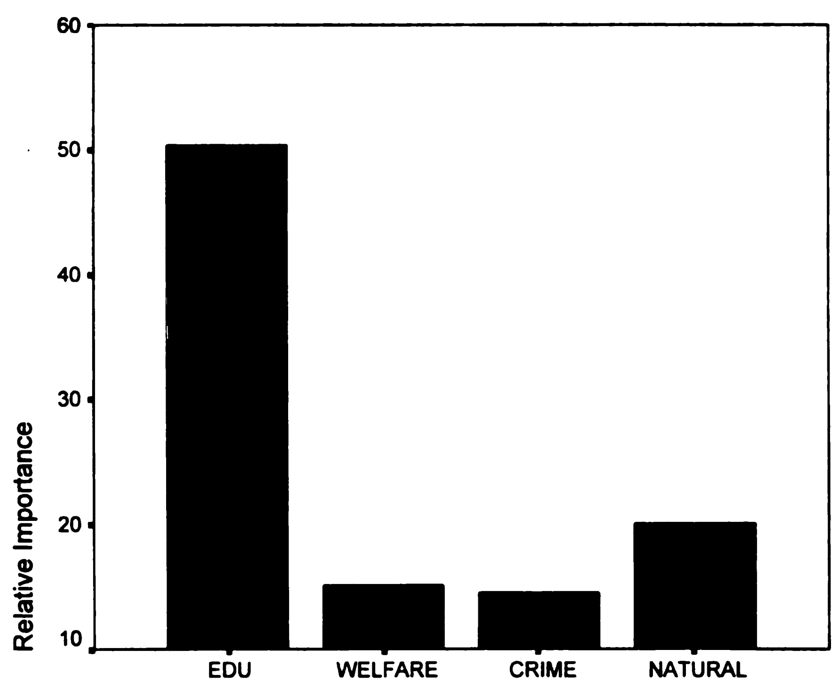


Figure 5.4: Relative Importance of Attributes for Voters (LINMAP)



General Notes on Interpreting Relative Importance of Attributes

Researchers usually use average utilities to compute the importance of each attribute, as seen in equation 3.1. Although this form of analysis appears valuable and informative, there are several problems inherent in it. Using the example in Table 5.2, importance is typically interpreted as follows: in OLS, education has the greatest impact on voter's preference (52.44%), followed by natural resources (19.42%), welfare and health (14.83%), and crime (13.31%). This statement may be incomplete and misleading: If we changed the range of levels of an attribute, we would expect its range of utilities to change and therefore its relative importance. For instance, if we tested spending levels of natural resources from 0% to +5% (omitting -5%), the utility range for natural resources would have been just 0.323 utiles, making it the least important among the attributes. Therefore, it is more precise to qualify the earlier statement of relative importance by saying that *for the attributes and levels tested*, education has the greatest impact on voter's preference, followed by natural resources, welfare and health, and crime. The correct way to compute the importances is to compute them for each respondent and then average across respondents. Since all attributes had identical levels, the relative importances in this study did not suffer the problems of misleading and incomplete information.

Overall Utilities of Profiles

The third purpose of part-worth utilities is to decide the overall utilities of profiles. To illustrate how to derive the overall utility, recall Equation 5.1 to 5.3 (Profile

1, 4, and 7). We could obtain the overall utilities by replacing W_{ij} and X_{ij} with estimates from the Table 5.2.

U_1 = 5.521 - 1.302*Edu - 0.052*Welf - 0.250*Crime - 0.447*Natural

= 3.470

U_4 = 5.521 + 0.357*Edu - 0.052*Welf + 0.320*Crime + 0.385*Natural

= 6.531

U_7 = 5.521 + 0.945*Edu - 0.052*Welf - 0.070*Crime + 0.062*Natural

= 6.406

Note: Edu, Welf, Crime, and Natural are equal to 1 because they are dummy variables (1= presence, 0 = absence).

Table 5.3 reveals the overall utilities for 10 profiles and the simulations from three models in OLS[2] (there is more detailed discussion in Chapter 7). The *maximum utility model* is simply the probability of choosing a profile as the most preferred. The *BTL (Bradley-Terry-Luce) model* computes the probability of choosing a profile as the most preferred by dividing the profile's utility by the sum of all the simulation total utilities. The *logit model* is similar to BTL but uses the natural log of the utilities instead of the utilities. The overall utilities show the profile 8 has the highest overall utility with a score 6.9 while profile 1 has the lowest utility with a score 3.5. This means that among the 10 profiles, voters most prefer an increase in education spending, cuts in crime spending, and maintaining welfare and natural resources at current levels while voters least prefer cuts in across-the-board. Likewise, from the three simulation models we

[2] We only report the OLS results because LINMAP produce the similar results in maximum utility model but cannot use BTL and logit models.

know that the profile 8 is the most preferred option with a probability 100 percent in maximum utility model, 12.63 percent in the BTL model, and 27.89 percent in the logit model.

Table 5.3: Overall Utilities of 10 Profiles for Voters (OLS)

Profile	Edu	Welf	Crime	Natur	Overall Utility	Max Utility (%)	BTL (%)	Logit (%)
1	-5%	-5%	-5%	-5%	3.5	0.00	6.31	0.86
2	-5%	0%	0%	5%	4.9	0.00	8.99	3.78
3	-5%	5%	5%	0%	4.2	0.00	7.71	1.87
4	0%	-5%	0%	0%	6.5	0.00	11.88	18.47
5	0%	0%	5%	-5%	5.7	0.00	10.37	8.08
6	0%	5%	-5%	5%	5.4	0.00	9.82	5.94
7	5%	-5%	5%	5%	6.4	0.00	11.65	16.30
8	5%	0%	-5%	0%	6.9	100.00	12.63	27.89
9	5%	5%	0%	-5%	6.0	0.00	10.99	11.37
10	0%	-5%	5%	-5%	5.3	0.00	9.65	5.44

Measures of Goodness of Fit

SPSS (OLS) provides Pearson's R and Kendall's tau statistics to indicate how well the model fits the data. Pearson's R and Kendall's tau statistics are correlations between the observed and estimated preferences based on different assumptions and formulae.[3] In many conjoint analyses, the number of parameters is close to the number of profiles rated, which will artificially inflate the correlation between observed and estimated scores. Therefore, these coefficients may be very high in the voter sample. As seen in Table 5.4, both Pearson's R and Kendall's tau are 1.000, a perfect correlation between observed and estimated scores. The perfect correlations demonstrate that the conjoint model fits the data well. With the caveats we just mentioned, it is better to look

[3] Pearson's R (product moment correlation) is the standard correlation statistics reported in all OLS results. Kendall's tau is the difference between proportions of the pairs in "right order" and "wrong order." A tau of 1.0 indicates a perfect rank order, a tau of -1.0 indicates a perfect negative relationship, and a value of zero indicates an unrelated ordering.

at other ways to measure goodness of fit and this will be discussed later in the chapter of reliability and validity tests.

In a goodness of fit measure, LINMAP checks how the respondent evaluated each pair of stimuli (profiles).[4] Then it uses the respondent's utility function to predict how he or she will evaluate each of these pairs, and compares it to the actual results. Finally, it calculates the percentage of all the pairs that were violated, for example, not consistent with the predictions. In Table 5.4, we see 96.9 percent of respondents violated 0-5 percent pairs and there are only 4 (3.1 percent) respondents violated 5 percent or more pairs comparisons (default threshold is 15 percent). Likewise, the result indicates that the conjoint model fits the data well. It should be noted that the percentage of pairs violated will typically be much larger when rating scales are used as opposed to a ranking procedure. Ratings typically contain a numbers of ties. Paired comparisons involving these tied levels lead to violations, since their predicted utilities are rarely equal.

[4] LINMAP used to measure goodness of fit by producing a Index of fit (C^*). Index of fit is a poorness of fit measure and can be interpreted as the ratio of the average poorness of fit (unexplained variation) to the average goodness of fit (explained variation). However, the latest version CONJOINT LINMAP only reports the percentage of paired comparisons that were violated.

Table 5.4: Goodness of Fit Measures for OLS and LINMAP

OLS			
Pearson's R	1.000	Sig.=	0.000
Kendall's tau	1.000	Sig.=	0.000

LINMAP	Distribution of Violations (%)	
Pairs Violated (%)	N	Percent
0 to 5	124	96.9%
6 to 10	3	2.3%
11 to 15	1	0.8%
N=	128	

Interpreting Conjoint Results from Student Sample

Part-worth Utilities of Attributes

The results from students are quite similar to those from voters. The part-worth utilities and the relative importance of attributes from student sample are captured in Table 5.5. A very sharp linear function in education spending shows that students strongly favor an increase in spending while strongly against cuts in spending (see Figure 5.5 and 5.6). The results from students are similar to voters but have a steeper linear utility function. For students, this result is reasonable; however, in all other programs the students also show quite similar preferences with the voters: all other three policy areas have very flat utility functions. In the areas of crime and natural resources, students prefer spending levels at the status quo and are against either spending increases or cuts. The spending increases in crime and natural resources are almost indistinguishable from status quo; in OLS the utility decrease from 0.434 to 0.410 in the crime spending and from 0.477 to 0.470 in natural resources spending. One difference from voters is welfare spending; students have slightly higher utility when welfare spending increases (from 0.176 to 0.548 in OLS and from 1.489 to 2.657 in LINMAP) while voters strongly oppose increases in welfare spending (see Figure 5.1 and 5.2).

Table 5.5: Part-worth Utilities from OLS and LINMAP (Students)

		OLS		LINMAP	
Factor	***Level***	***Utility***	***%ΔUtility***	***Utility***	***%ΔUtility***
	-5%	-3.324	-858.9%	-21.169	-759.7%
Education	0%	0.438	-----	3.209	-----
	5%	2.886	558.9%	17.960	459.7%
Relative Importance		60.98%		60.42%	
	-5%	-0.724	-511.6%	-4.146	-378.4%
Welfare/Health	0%	0.176	-----	1.489	-----
	5%	0.548	211.5%	2.657	78.4%
Relative Importance		12.50%		10.50%	
	-5%	-0.843	-294.5%	-5.889	-260.9%
Crime/Public Safety	0%	0.434	-----	3.661	-----
	5%	0.410	-5.5%	2.228	-39.1%
Relative Importance		12.54%		14.74%	
	-5%	-0.947	-298.6%	-6.040	-286.2%
Natural Resources	0%	0.477	-----	3.244	-----
	5%	0.470	-1.4%	2.796	-13.8%
Relative Importance		13.99%		14.33%	
Constant		8.110		N/A	
N=		77		77	

Figure 5.5: Part-Worth Utilities for Students (OLS)

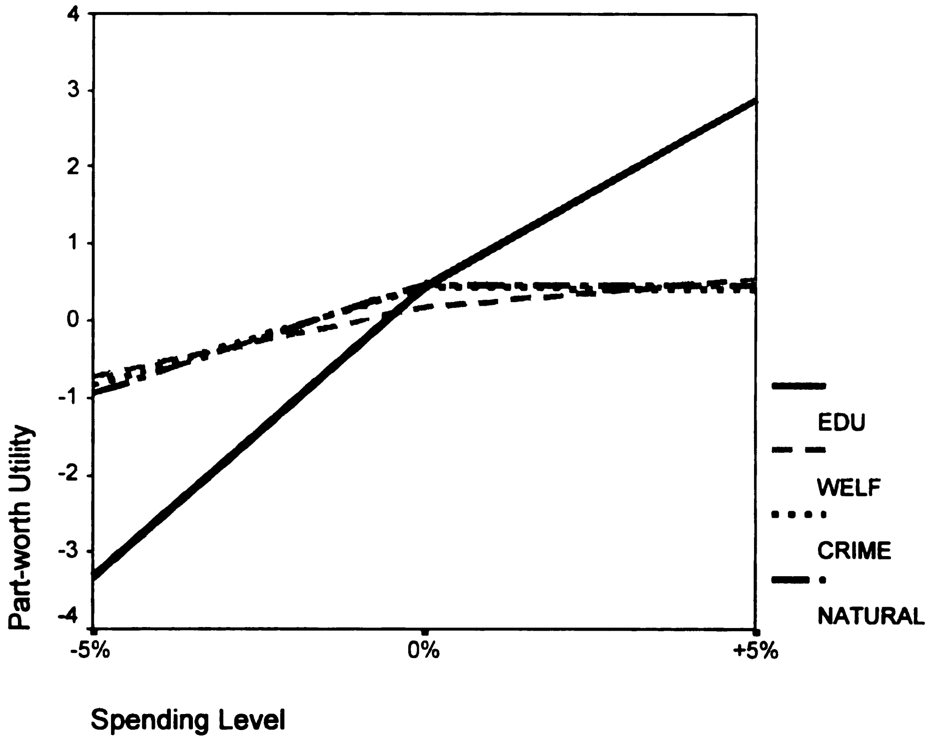
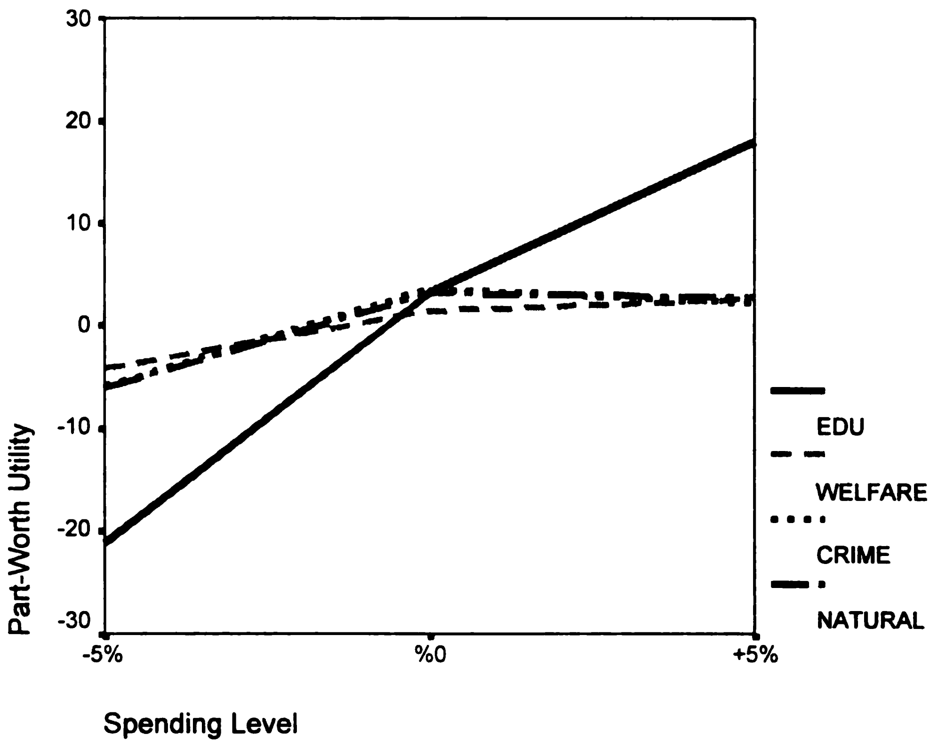


Figure 5.6: Part-Worth Utilities for Students (LINMAP)



Relative Importance of Attributes

From Table 5.5 and Figure 5.7 and 5.8, we see that the relative importances from students are similar to those from voters. The relative importance scores (for the attributes and levels tested) show that education has the greatest impact on students' preferences (60.98%), followed by natural resources (13.99%), crime (12.54%), and welfare and health (12.50%) in OLS (same preference sequence in LINMAP). Without a doubt, education is the most important issue for students; the relative importance score is 8.54 percent higher than voters. In the spending areas of natural resources, crime, and welfare have slightly different ratings, but the utility gaps between students and voters are indistinguishable.

Figure 5.7: Relative Importance of Attributes for Students (OLS)

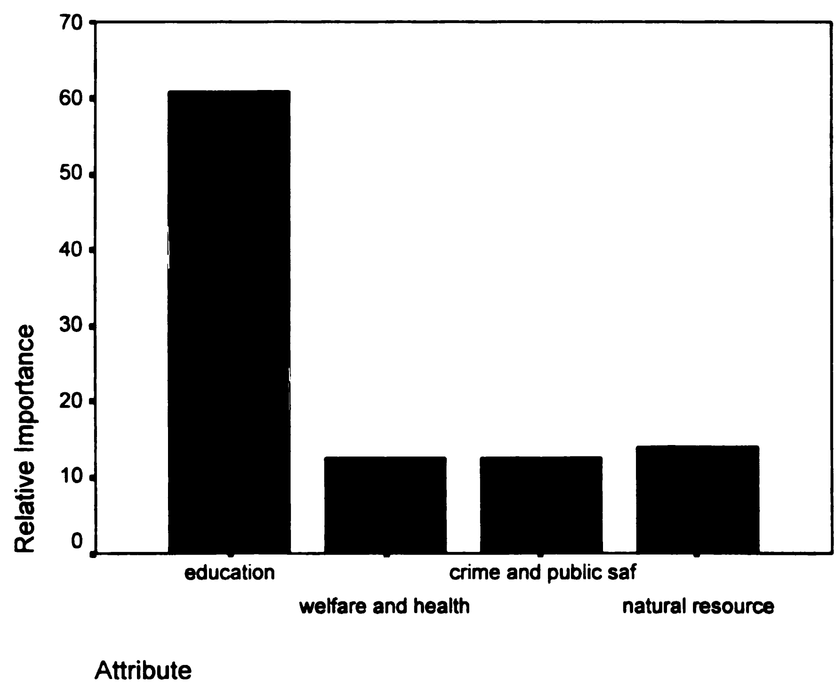
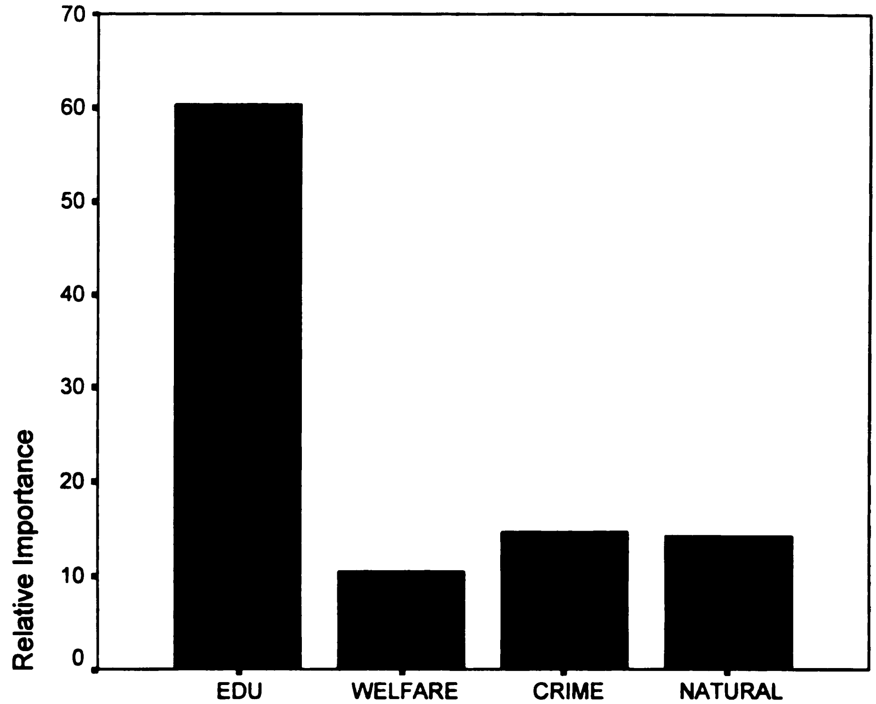


Figure 5.8: Relative Importance of Attributes for Students (LINMAP)



Overall Utilities of Profiles

The overall utilities and the three simulation results from student sample are captured in Table 5.6. Among the 18 profiles, profile 17 has the highest overall utility with score 12.50 while profile 1 has the lowest overall utility with score 2.3 in OLS. Thus, among the 18 profiles, students most prefer increases in education and welfare while maintaining crime and natural resources at the current levels. All three simulation models also show that profile 17 has the highest probability among the 18 choices: 100% in maximum utility, 8.53 percent in BTL and 39.68 percent in Logit models.

Table 5.6: Overall Utilities of 10 Profiles for Students (OLS)

Profile	Edu	Welf	Crime	Natur	Overall Utility	Max Utility (%)	BTL (%)	Logit (%)
1	-5%	-5%	-5%	-5%	2.3	0.00	1.56	0.00
2	-5%	0%	0%	5%	5.9	0.00	4.02	0.05
3	-5%	5%	5%	0%	6.2	0.00	4.26	0.08
4	0%	-5%	0%	0%	8.7	0.00	5.98	0.96
5	0%	0%	5%	-5%	8.2	0.00	5.61	0.56
6	0%	5%	-5%	5%	8.7	0.00	5.98	0.95
7	5%	-5%	5%	5%	11.2	0.00	7.64	10.78
8	5%	0%	-5%	0%	10.8	0.00	7.40	7.63
9	5%	5%	0%	-5%	11.0	0.00	7.56	9.55
10	5%	-5%	-5%	-5%	8.5	0.00	5.81	0.75
11	-5%	-5%	5%	0%	4.9	0.00	3.39	0.02
12	0%	-5%	0%	5%	8.7	0.00	5.98	0.95
13	-5%	0%	0%	-5%	4.4	0.00	3.05	0.01
14	0%	0%	-5%	0%	8.4	0.00	5.73	0.66
15	5%	0%	5%	5%	12.1	0.00	8.26	26.53
16	0%	5%	5%	-5%	8.6	0.00	5.86	0.81
17	5%	5%	0%	0%	12.5	100.00	8.53	39.68
18	-5%	5%	-5%	5%	5.0	0.00	3.40	0.02

In addition, it is interesting to know what the first choice would be if we provide students or voters the same 10 or 18 profiles. To compare the preferences of students and voters on the same basis, we ran a simulation on the student sample by simulating 10 profiles used in the first round of voter sample.[5] As seen in Table 5.7, voters' first choice is Profile 8 with overall utility 6.9 while students' first choice is Profile 7 with overall utility 13.2. We also see that the three highest overall utilities of profiles (Profile 7, 8, and 9) in the student sample are almost twice as those in the voter sample (Profile 8, 4, and 7). Looking at the Profile 7, 8, 9 in the voter sample and Profile 8, 9, 10 in the student sample, we noticed all these profiles have an increase in education spending. These results reveal the students' decision pattern of choosing the profiles with increases

[5] Of course, it could be done the other way, by running a simulation on voter sample using 18 profiles in the student sample.

in education spending first, then evaluating other programs. Basically, the simulation can be used in various "what-if" situations to achieve a particular research interest.

Table 5.7: Comparison of Overall Utilities Between Voters and Students

					Voters			
Profile	Edu	Welf	Crime	Natur	Overall Utility	Max Utility (%)	BTL (%)	Logit (%)
1	-5%	-5%	-5%	-5%	3.5	0.00	6.31	0.86
2	-5%	0%	0%	5%	4.9	0.00	8.99	3.78
3	-5%	5%	5%	0%	4.2	0.00	7.71	1.87
4	0%	-5%	0%	0%	6.5	0.00	11.88	18.47
5	0%	0%	5%	-5%	5.7	0.00	10.37	8.08
6	0%	5%	-5%	5%	5.4	0.00	9.82	5.94
7	5%	-5%	5%	5%	6.4	0.00	11.65	16.30
8	5%	0%	-5%	0%	6.9	100.00	12.63	27.89
9	5%	5%	0%	-5%	6.0	0.00	10.99	11.37
10	0%	-5%	5%	-5%	5.3	0.00	9.65	5.44

					Students			
Profile	Edu	Welf	Crime	Natur	Overall Utility	Max Utility (%)	BTL (%)	Logit (%)
1	-5%	-5%	-5%	-5%	2.7	0.00	2.92	0.00
2	-5%	0%	0%	5%	7.3	0.00	7.78	0.10
3	-5%	5%	5%	0%	7.2	0.00	7.67	0.09
4	0%	-5%	0%	0%	10.0	0.00	10.74	1.67
5	0%	0%	5%	-5%	9.0	0.00	9.65	0.60
6	0%	5%	-5%	5%	10.4	0.00	11.13	2.38
7	5%	-5%	5%	5%	13.2	100.00	14.11	38.73
8	5%	0%	-5%	0%	13.0	0.00	13.89	31.47
9	5%	5%	0%	-5%	12.7	0.00	13.63	24.75
10	0%	-5%	5%	-5%	7.9	0.00	8.47	0.20

Measures of Goodness of Fit

In Table 5.8, we see that both Pearson's R (0.996) and Kendall's tau (0.976) are slightly lower than those in the voter sample. This is because we use 18 profiles rather than 10 profiles in the survey and hence increase the chance of a respondent making mistakes and reduce the chance of a respondent making consistent judgments. The model for students, however, still fits the data well. In addition to the Pearson's R and Kendall's tau, in this sample we use 3 holdout profiles. Holdout profiles are additional profiles which the respondent evaluates beyond those generated by actual conjoint design. Specifically, we use first 15 profiles (Profile 1-15) to predict last three profiles (Profile 16 - 18). Holdout profiles have a numbers of uses. They can help assess the predictive accuracy of the conjoint study; ensure that data are collected for particular purposes; or test specific hypotheses. As seen in Table 5.8, the Kendall's tau for 3 holdouts is 1.00 with confidence level 94%. The result exhibits that this model performs well and has a very good ability to predict choices.

The results from LINMAP tell a little different story. There are 50.6 percent of respondents (adding pairs violated from 6 to 25) violated more than 5 percent pairs and 5.2 percent respondents (adding pairs violated from 16 to 25) violated more than 15 percent pairs (default threshold). The results are reasonable because in the student sample we have 18 rather than 10 profiles which increase the chance of violating pairs of profiles.

In order to check holdout profiles, LINAMP provides two measures: average correlation and average percentage of paired violation. The average correlation is an index which measures the degree of agreement between the predicted rankings based on the respondent's utility function, and the respondent's actual ranking. The measure "percentage of pairs violated" also uses ranked results of the predicted and actual holdout data. As seen in Table 5.8, for holdout profiles, we have an average correlation 0.78 and average percentage of paired violation 13 percent. Even though these are not perfect, the results from the student sample are satisfactory.

Table 5.8: Goodness of Fit Measures for OLS and LINMAP (Students)

	OLS		
Pearson's R	0.996	Sig.=	0.000
Kendall's tau	0.976	Sig.=	0.000
Kendall's tau for 3 Holdouts	1.000	Sig.=	0.059

	LINMAP		
	Distribution of Violations (%)		
Pairs Violated	N	Percent	
0 to 5	38	49.4%	
6 to 10	27	35.1%	
11 to 15	8	10.4%	
16 to 20	3	3.9%	
21 to 25	1	1.3%	
Average Correlation			0.78
Average % Paired Violated			13%
N=	77		

Comparison Between Traditional Survey Questions and Conjoint Results

In this study, one of the major purposes is to compare results between traditional surveys and conjoint analyses. In particular, we like to know what the "response effect" of conjoint study is. In order to compare conjoint results with traditional surveys, we ran several simulations to make conjoint results comparable to those in traditional surveys. To compare conjoint results with the first question, "Generally speaking, do you think the State of Michigan should: increase overall spending, decrease overall spending, or keep overall spending at current levels," we ran three simulations with increasing spending across-the-board (+5%, +5%, +5%, +5%), maintaining spending at current levels across-the-board (0%, 0%, 0%, 0%), and decreasing spending across-the-board (-5%, -5%, -5%, -5%). In order to compare conjoint results with Question 2, 3, 4, and 5: "In the area of _______ (education, welfare, crime, or natural resources), you would prefer to see the

State of Michigan increase spending, decrease spending, or maintain spending at current levels," we ran simulations with the specified attribute level (+5%, 0%, or -5%) while holding other attribute constant (by keeping other attribute levels as 0%). For instance, we simulated three profiles with (+5%, 0%, 0%, 0%), (0%, 0%, 0%, 0%), and (-5%, 0%, 0%, 0%) to represent respondents who favor an increase in education spending, keeping education spending at the current level, and cuts in education spending in traditional survey questions. In traditional surveys, we report percentages that respondents answered in every spending level (increase, status quo, and decrease) from question 1 to 5. In the conjoint study, we report results from two simulation models (BTL and LOGIT) with probability in each profile. In this fashion, we could make conjoint results comparable to those obtained from traditional surveys.

Voter Sample

Table 5.9 reveals the differences between traditional surveys and conjoint analysis. The columns "General" represent the actual percentage of respondents who favor an increase, the status quo, or a decrease in overall spending or particular policy areas. The columns "Conjoint" show the simulation results from two probability models (BTL and LOGIT). The "General" columns can be viewed as the "direct response" from respondents (respondents only evaluate single dimensional questions) while "Conjoint" columns can be viewed as the "deliberate response" (respondents make trade-offs among four attributes). In rows, we report traditional questions and profiles for overall spending and four policy areas, followed by percentage (actual answers) or probability (simulation results), and choice order made by respondents (1,2, and 3).

Table 5.9 : Comparison of Budget Priorities Between Traditional Surveys and Conjoint Analysis (Voters)

	General	Conjoint		General	Conjoint		General	Conjoint	
		BTL	LOGIT		BTL	LOGIT		BTL	LOGIT
Overall	*Increase*	(5%,5%,5%,5%)		*Status Quo*	(0%,0%,0%,0%)		*Decrease*	(-5%,-5%,-5%,-5%)	
%/Prob.	*28%*	37.23%	31.19%	*22%*	41.82%	66.77%	*50%*	20.95%	2.10%
Choice	*2*	2	2	*3*	1	1	*1*	3	3
Education	*Increase*	(5%,0%,0%,0%)		*Status Quo*	(0%,0%,0%,0%)		*Decrease*	(-5%,0%,0%,0%)	
%/Prob.	*47%*	38.13%	60.19%	*41%*	35.15%	33.44%	*12%*	26.73%	6.36%
Choice	*1*	1	1	*2*	2	2	*3*	3	3
Welfare	*Increase*	(0%,5%,0%,0%)		*Status Quo*	(0%,0%,0%,0%)		*Decrease*	(0%,-5%,0%,0%)	
%/Prob.	*22%*	31.86%	24.04%	*42%*	35.07%	45.38%	*36%*	33.07%	30.58%
Choice	*3*	3	3	*1*	1	1	*2*	2	2
Crime	*Increase*	(0%,0%,5%,0%)		*Status Quo*	(0%,0%,0%,0%)		*Decrease*	(0%,0%,-5%,0%)	
%/Prob.	*34%*	32.98%	30.20%	*47%*	34.95%	44.59%	*20%*	32.07%	25.21%
Choice	*2*	2	2	*1*	1	1	*3*	3	3
Natural	*Increase*	(0%,0%,0%,5%)		*Status Quo*	(0%,0%,0%,0%)		*Decrease*	(0%,0%,0%,-5%)	
%/Prob.	*44%*	33.65%	33.54%	*43%*	35.30%	46.32%	*13%*	31.05%	20.15%
Choice	*1*	2	2	*2*	1	1	*3*	3	3

As seen in the Table 5.9, in the question of overall spending, voters' first choice is cuts in overall spending (50%), followed by an increase in overall spending (28%), and maintaining overall spending at the current level (22%). However, based on tradeoffs among four policy areas, the differences between traditional surveys and conjoint analysis are quite dramatic (see the "choice" row). In conjoint analysis, the simulation models could be interpreted as: the profile with the highest probability voters will choose is maintaining overall spending across-the-board (41.82% in BTL and 66.77% in Logit), followed by increasing overall spending across-the-board (37.23% in BTL and 31.29% in Logit), and decreasing overall spending across-the-board (20.95% in BTL and 2.10% in

Logit). Therefore, in conjoint simulation, voters' first choice is maintaining overall spending at the status quo rather than a decrease in overall spending (in traditional question).

In the areas of education, welfare, and crime, the simulation results are in accordance with traditional surveys: they share the same preference orders, even though the simulation probabilities are different from the percentage that respondents actually chose. In the area of education, for instance, voters most prefer an increase in spending (47% in actual choice and 38.13% in BTL), followed by maintaining spending at the current level (41% in actual choice and 35.15% in BTL), and a decrease in spending (12% in actual choice and 26.73% in BTL); same preference orders can be found in welfare and crime spending. In the area of natural resources, however, the results from traditional survey and conjoint study are quite different. Voters' first choice on budget priorities is an increase in natural resources spending (44%), followed by the status quo (43%), and a decrease in spending (13%), while simulation models reveal that a profile with highest probability voters will choose is the status quo on spending (35.30% in BTL), followed by an increase in spending (33.65 in BTL), and a decrease in spending (31.05% in BTL). Again, we see the different results between conjoint simulations and traditional survey questions. These preference differences are due to the fact that conjoint analysis forces respondents to make tradeoffs among different attributes while the traditional survey lets respondents evaluate single but incomplete budget options. Hence, the results from conjoint analysis are more accurate and complete than traditional survey.

In a nutshell, the results from conjoint study and traditional survey are consistent in the areas education, welfare, and crime while inconsistent in the areas of overall spending and natural resources.

Student Sample

The results from the student sample demonstrate more contradictory preference differences between traditional surveys and conjoint study. In Table 5.10, we see that the results from traditional survey and conjoint study are inconsistent in all five questions. For instance, the students' spending preferences in the traditional survey and conjoint analysis go in opposite directions in the overall spending question: students' first priority is cuts in overall spending (51%), followed by an increase in overall spending (40%), and the status quo on overall spending (9%). On the contrary, the profile with highest probability from student sample is an increase in overall spending (51.33% in BTL), followed by the status quo (39.09% in BTL), and cuts in overall spending (9.58% in BTL). In the areas of education, welfare, crime, and natural resources, we also see inconsistencies between traditional survey questions and conjoint simulations.

In sum, while respondents only evaluate single dimensional questions in traditional surveys, they are forced to make tradeoffs among various attributes in the conjoint exercises. The informaiton we obtain from conjoint analysis is therefore more accurate and complete than traditional surveys (as we seen in the Table 5.9 and 5.10).

Table 5.10 : Comparison of Budget Priorities Between Traditional Surveys and Conjoint Analysis (Students)

	General	Conjoint		General	Conjoint		General	Conjoint	
		BTL	LOGIT		BTL	LOGIT		BTL	LOGIT
Overall	*Increase*	(5%,5%,5%,5%)		*Status Quo*	(0%,0%,0%,0%)		*Decrease*	(-5%,-5%,-5%,-5%)	
%/Prob.	*40%*	51.33%	97.04%	*9%*	39.09%	2.96%	*51%*	9.58%	0.00%
Choice	*2*	1	1	*3*	2	2	*1*	3	3
Education	*Increase*	(5%,0%,0%,0%)		*Status Quo*	(0%,0%,0%,0%)		*Decrease*	(-5%,0%,0%,0%)	
%/Prob.	*75%*	44.00%	95.81%	*0%*	34.31%	4.12%	*25%*	21.70%	0.07%
Choice	*1*	1	1	*3*	2	2	*2*	3	3
Welfare	*Increase*	(0%,5%,0%,0%)		*Status Quo*	(0%,0%,0%,0%)		*Decrease*	(0%,-5%,0%,0%)	
%/Prob.	*31%*	35.16%	51.44%	*22%*	34.11%	36.46%	*47%*	30.73%	12.09%
Choice	*2*	1	1	*3*	2	2	*1*	3	3
Crime	*Increase*	(0%,0%,5%,0%)		*Status Quo*	(0%,0%,0%,0%)		*Decrease*	(0%,0%,-5%,0%)	
%/Prob.	*36%*	34.24%	38.62%	*12%*	34.94%	48.73%	*52%*	30.82%	13.00%
Choice	*2*	2	2	*3*	1	1	*1*	3	3
Natural	*Increase*	(0%,0%,0%,5%)		*Status Quo*	(0%,0%,0%,0%)		*Decrease*	(0%,0%,0%,-5%)	
%/Prob.	*38%*	35.79%	52.08%	*6%*	35.09%	41.67%	*56%*	29.11%	6.24%
Choice	*2*	1	1	*3*	2	2	*1*	3	3

Summary

In this chapter we analyzed the part-worth utilities, relative importances of attributes, and overall utilities of profiles. In general, voters and students are satisfied with the current spending levels in most areas. The only exception is in the education area; both voters and students reveal that they favor an increase in education spending. Such results may reflect the demographic characteristics in our samples. From the Meridian Township voter sample, over 95 percent voters have at least a college degree, and from the student sample, education undoubtedly is the most important issue for students. Thus, we were not surprised that both voters and students support an increase in education spending. For welfare, crime, and natural resources spending, both voters and students exhibit that they are satisfied with the current spending levels (although students have slightly higher utility for welfare spending increase).

In comparisons between traditional surveys and conjoint analyses, the results reveal significant differences between traditional surveys and simulations from conjoint results: sometimes the preference orders for a particular area even go in completely opposite direction. These results are striking because most studies use results obtained from traditional surveys to interpret, analyze, evaluate, and predict all kinds of political issues in fields of public opinion and policy analysis. If the information from traditional surveys are inaccurate or incomplete, all research results based on these surveys will be less persuasive and misleading. In addition, if policymakers make policy decisions based on the traditional survey results, these policies will not reflect people's true preferences and hence we discount the quality of democracy. Therefore, it is important to introduce new survey tools in order to collect more accurate and complete information about public opinion. Conjoint analysis produces more accurate and complete information about people's budget priorities. The "more accurate" means that we obtain respondents' true preferences when holding them accountable. In Chapter 2 we have shown that respondents will not overstate their preferences when holding them accountable. The "more complete" means that respondents are forced to evaluate all factors (which necessarily consists of a budget proposal) rather than single dimensional consideration. Although conjoint analysis is not the only way to conduct a survey, it is indeed a better way to collect more accurate and complete data and information.

CHAPTER 6

GROUP DIFFERENCES IN CONJOINT ANALYSIS

Group attitudes provide an underlying structure that forms a broad range of policy opinions and candidates' evaluations. Groups help people structure the political world in two ways. First, the political environment itself is structured in terms of groups; in the realm of politics the relevance of groups derives from their centrality to governmental institutions, the policy process, and election campaigns. Thus, the distribution of the vote across different social groups has been a regular characteristic in the discussion and interpretation of election results. From the early Columbia University studies (Berelson, Lazarsfeld, and McPhee 1954; Berelson, Lazarsfeld, and Gaudet 1944) to the analyses of the most recent elections (Abramson, Aldrich, and Rhode 1998; Miller and Shanks 1996), basically all studies provide at least some information about how voters' sociodemographic characteristics relate to their political attitudes and candidate preferences.

Second, while groups obviously play an important role in the aggregate patterns of electoral politics, they are also a central part of individual political cognition. Individuals have a predisposition to rely on cues and shortcuts when making choices and forming preferences, and groups are ideal for this purpose in the realm of politics. Groups help orient individuals to their social world; they provide standards, supply information, define friend and foe. In numerous situations groups consists of frames of reference that assist individuals in interpreting their social experiences. Such guidance can be particularly powerful in the political world (Brady and Sniderman 1985; Dalton 1988: 151-52).

The literature on public opinion has shown that at a given point in time, different social groups may differ markedly from each other in the level of their policy preferences; that is, in the proportions by which they favor or oppose various policies. Group preferences can vary because of different interests, different social, economic, political , and cultural environments, diverse experiences, or differing positions in society, any one of which can lead to different perceptions, wants, and values (Page and Benjamin 1992: 285-88).

The purpose of this chapter is to analyze group differences among gender, generations, religion, party identification, children, and home owner by evaluating their part-worth utilities, relative importance of attributes, and overall utilities of profiles. Since our data solely came from the Meridian Township with a very limited sample size (N=128), our results cannot be generalized to make inferences about the distribution of attitudes in the U.S. To provide some references, we first report previous studies on group attitudes. Then, we analyze group differences in terms of part-worth utilities and relative importances of attributes among Meridian Township voters.

Gender Differences

Studies on public opinion have shown that there were persistent gender differences on policy preferences, differences which remain after variables other than gender are controlled for (Smith 1984; Shapiro and Mahajan 1984; Page and Benjamin 1992).

The most important gender differences are concerned with force and violence. Smith (1984) reported that men are more likely than women to support violent options. In addition, gender differences also found in "compassion" issues concerned with policies which aid the poor, the unemployed, the sick, and others in need, or which protect the well-being of people in general. Women were more supportive than men of such policies; Shapiro and Mahajan (1984) indicated that gender difference for these policies is over 3 percent, less than half that of the force issues.

Various researchers have explained these differences both by the way women have been socialized as caretakers and nurturers and by the rise of a women's political consciousness (Stoper and Johnson; Gilligan 1982). Women may have extended the scope of their more traditional roles to encompass the problems of society as a whole. Therefore, women might be concerned about the issues that include policies to help those in need and also a somewhat distinct set of regulatory policies (e.g., environmental regulation, consumer protection, other safety issues) which are intended to protect the public. In regulation and protection issues, Shapiro and Mahajan (1984) stated that there is no gender difference for spending on the environment, although the overall average may conceal a change which occurred: by the 1980s perhaps 5 percent more women than men supported spending on these programs. In addition, there were striking differences for particular types of regulations and safeguards. Women have shown more support for the 55 mile per hour speed limit, jail terms for drunk drivers, and so on.

For Meridian Township voters, the policy preferences of gender are captured in Table 6.1 and Figures 6.1 to 6.4. In the area of education (see Figure 6.1), the utility functions for both women and men are linear but women are much better off with an increase in education spending: for an increase 5 percent in education spending, women feel better off by 0.787 utility than men while for a 5 percent decrease in education spending, women feel worse off by 0.611 utility than men. The relative importance ratings show that both women and men have identical preference orders: education is the most important issue for both women and men, followed by welfare, natural resources, and crime. Nevertheless, the relative importance also indicates that women are more concerned with education than men, with an 18.46 percent relative importance rating difference.

In the area of welfare and health (see Figure 6.2), women favor welfare spending at the status quo but suffer a significant utility loss with cuts in welfare spending (utility drops from 0.431 to -0.578). On the contrary, men have the highest utility gain with cuts in welfare spending (utility gains from 0.280 to 0.333), yet have a dramatic utility drop with a spending increase (utility drop from 0.280 to -0.612).

For crime and natural resources issues (see Figure 6.3 and 6.4), both women and men reveal similar patterns of utility functions: they prefer the status quo while dislike either increase or decrease in spending. In addition, in the area of natural resources, both women and men suffer a dramatically utility loss with a decrease in spending (utilities drop from 0.320 to -0.430 in women and from 0.433 to -0.460 in men).

Table 6.1: Gender Differences in 4 Major Policy Areas

Attribute	Level	Female Utility	Male Utility	Difference
	-5%	-1.655	-1.045	-0.611
Education	0%	0.255	0.432	-0.177
	5%	1.400	0.613	0.787
Relative Importance		58.24%	39.78%	18.46%
	-5%	-0.578	0.333	-0.911
Welfare/Health	0%	0.431	0.280	0.152
	5%	0.147	-0.612	0.760
Relative Importance		19.24%	22.68%	-3.44%
	-5%	-0.217	-0.275	0.057
Crime/Public Safety	0%	0.215	0.397	-0.182
	5%	0.002	-0.122	0.124
Relative Importance		8.24%	16.11%	-7.87%
	-5%	-0.430	-0.460	0.030
Natural Resources	0%	0.320	0.433	-0.113
	5%	0.110	0.027	0.083
Relative Importance		14.29%	21.42%	-7.13%
Constant		5.575	5.482	
N=		54	74	

Figure 6.1: Gender Differences in Education Spending

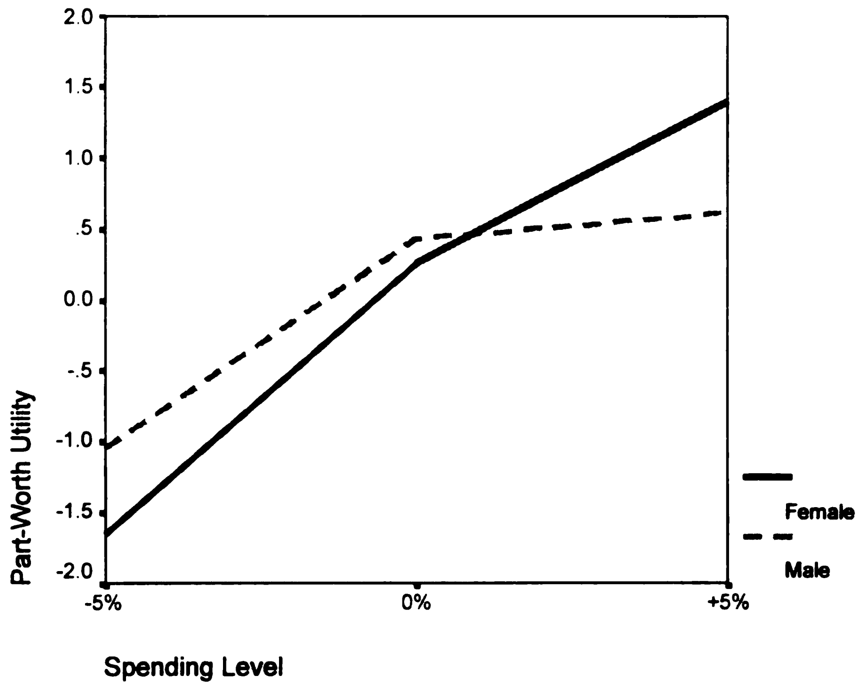


Figure 6.2: Gender Differences in Welfare Spending

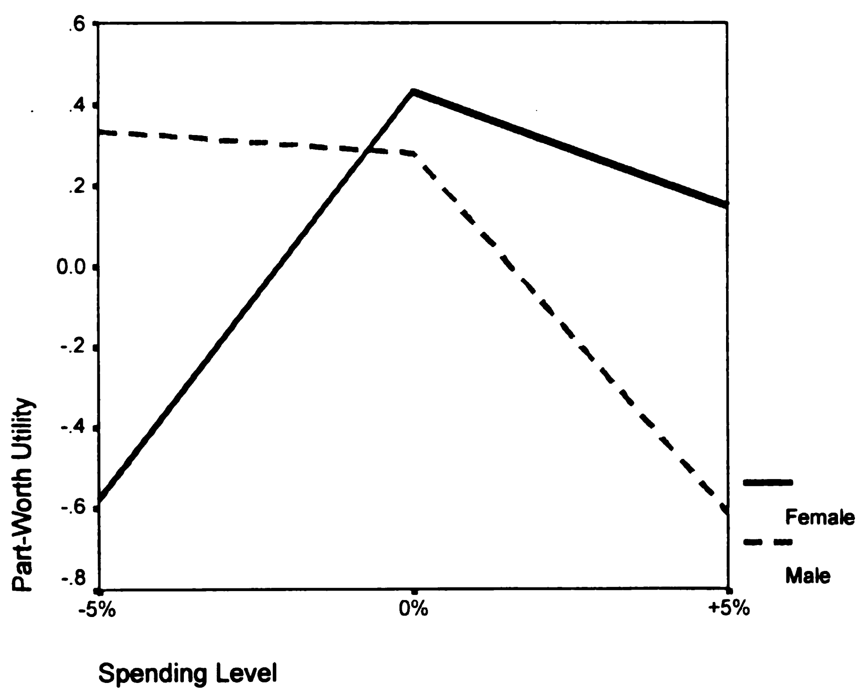


Figure 6.3: Gender Differences in Crime Spending

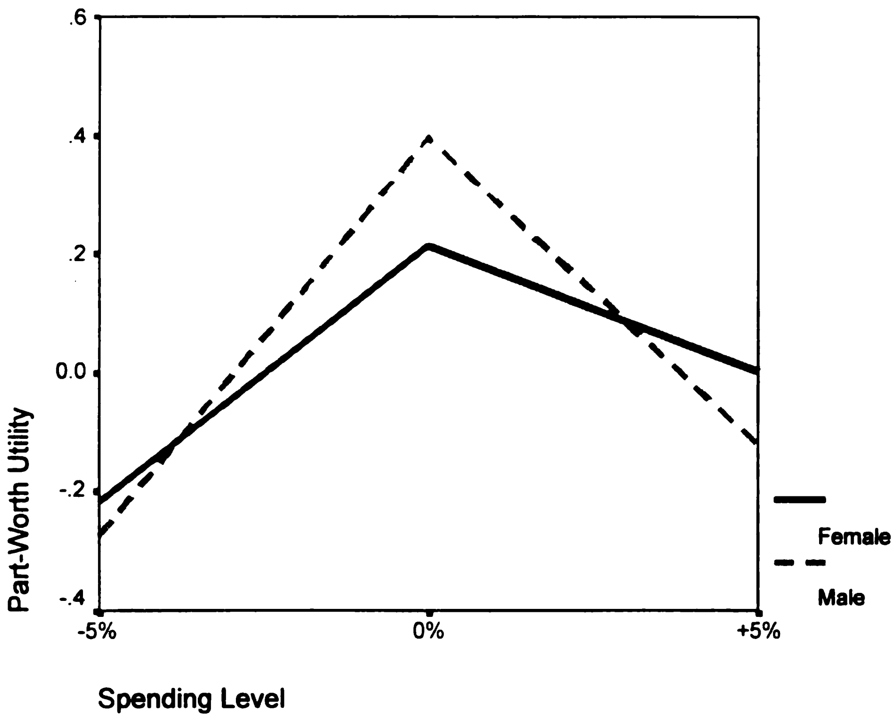
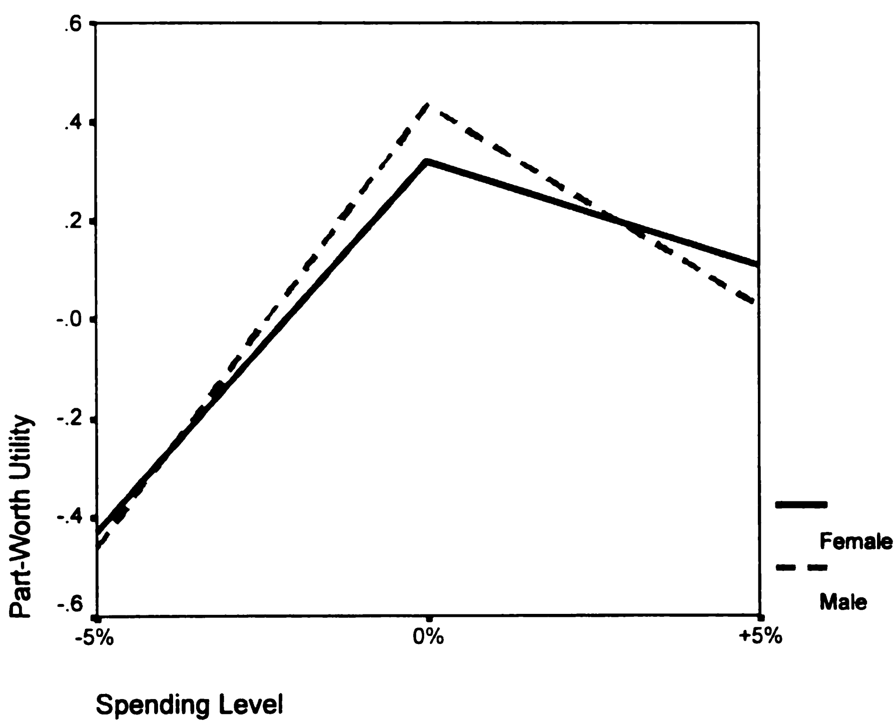


Figure 6.4: Gender Differences in Natural Resources Spending



Generational Differences

Generational change is a continuing process which transforms all societies gradually. There are at least two possible reasons for the generational change. One reason involves “life cycle”: age brings with it different economic and social circumstances and physical conditions, so that peoples’ attitudes and opinions tend to change as they grow older. The other reason involves “cohorts”: people in different age groups have lived through different experiences - the Great Depression, World War II, Vietnam War, periods of economic affluence or stringency - which led them, when they were young and flexible, to form attitudes that they tend to carry through the remainder of their lives. On the other hand, during a given period all age groups may be affected by the same events and social and economic conditions, and by the same interpretations offered through the media. When there are no strong generational or life cycle effects, all age groups move about the same way in response to such “period effects” (Braungart and Braungart 1989; Abramson 1989: 236-242).

Page and Benjamin (1992) argued that both life cycle and cohort differences have sometimes appeared between the opinions of younger and older Americans over the last fifty years. The older people have tended to be sensitive about the threats to their Social Security pensions, though in more general terms the youngest age group has actually been just as supportive of government spending on Social Security and Medicare, and the elderly have not sought income transfers to themselves at the expense of transfers to the young. Likewise, since the older people feel more vulnerable to street crime, they tended to favor various types of protection. Page and Benjamin (1992) showed that the elderly are more supportive of gun control than the middle-aged (p.303).

In addition, Page and Benjamin pointed out that new adult cohorts have typically been the most supportive of the social welfare policies and benefits that were already in place in the society in which they grew up. This has occurred in affluent European welfare states as well as in the U.S. Moreover, the younger people have tended to be more supportive of environmental protection (p. 303).

Although there is no consensus on the definition of political generations, Miller proposed an appropriate definition and classification of political generations (Miller, 1992).[1] According to Miller, a generation is composed of a temporally bounded group of people who have in common formative experiences that may have occurred over a substantial time interval. The cohorts are defined by the election years in which individuals were first eligible to vote for president. According to Miller, there are six sets of cohorts or generations:

A. Pre-New Deal: (1) Pre-1920; (2) Republican Normalcy (1920-1928),

B. New Deal: (3) Roosevelt Era (1932-1944); (4) New Deal Consolidated (1948-1964),

C. Post-New Deal: (5) Years of Turmoil (1968-1976); (6) Reagan Era (1980-1988).

The cohorts within each generation seemed more like each other and less like those other generations.

The oldest generations, the Pre-New Deal generation, are made up of a population whose first votes, and then whose presumed periods of early adult or preadult politicization, preceded the first Roosevelt election in 1932. The New Deal generation

[1] Alternatively, we obtained the similar conjoint results by using generational definitions based on William Strauss and Neil Howe's book, Generations: The History of American's Future, 1584 to 2069. Their definitions of generations are as follows: G.I.: birth years 1901 to 1924, Silent: birth years 1925 to 1942, Boomer: birth years 1943 to 1960, and 13er: birth years 1961 to 1981.

was socialized predominantly under Democratic presidential leadership. It consists of those whose first years of voting eligibility may have included the first Roosevelt election in 1932 or any of the subsequent elections up through 1964. The Post-New Deal generation is composed of the still-going set of population whose early adult experiences included the turmoil of the late 1960s, the Vietnam War, the counterculture, and the Watergate scandals of the Nixon Administration. This Post-New Deal generation is defined as open-ended and still-growing; it therefore includes as well those socialized by the political dramas associated with the presidencies of Carter, Reagan, Bush, and Clinton. Since there is no Pre-New Deal generation in our sample, we examined New Deal and Post-New Deal generations with four generation groups.

Table 6.2 reveals the differences in policy priorities for different generations. For the education issue (see Figure 6.5), most cohorts have a linear utility function favoring an increase in education spending. The only exception is the New Deal 2 (New Deal Consolidated) cohort which prefers spending at the status quo. For the relative importance, education is the most important issue for all cohorts. In the areas of welfare and health (see Figure 6.6), all cohorts prefer maintaining spending at the current level. While most cohorts suffer significant utility loss from increase in welfare spending, New Deal 2 suffers serious utility loss from cuts in welfare spending. For the spending of crime and public safety (see Figure 6.7), most cohorts have flat utility functions favoring spending at the status quo except the New Deal 1 (Roosevelt Era). New Deal 1 (the oldest cohort in this study) has linear utility function favoring increase in crime spending with slightly utility increase (from 0.491 to 0.602) but suffering a great deal of utility loss (from 0.491 to -1.093) with cuts in crime spending. The crime issue is also the second most important issue for New Deal 1 (education is the first one). In the area of natural resources (see Figure 6.8), most cohorts (New Deal 1, New Deal 2, and Post-New Deal 1) favor spending at the status quo but have dramatic utility loss with cuts in natural resources spending. Post-New Deal 2 (the youngest cohort in this study) is the only exception in that they favor either an increase in natural resources spending or at the status quo (identical utility score 0.305).

Table 6.2: Generational Differences in 4 Major Policy Categories

Attribute	Level	New Deal 1	New Deal 2	Post-New Deal 1	Post-New Deal 2
	-5%	-1.551	-0.780	-1.602	-1.019
Education	0%	0.227	0.589	0.313	0.321
	5%	1.324	0.190	1.289	0.697
Relative Importance		52.01%	35.75%	58.28%	46.21%
	-5%	-0.023	-0.487	-0.041	0.204
Welfare/Health	0%	0.324	0.387	0.357	0.295
	5%	-0.301	0.099	-0.315	-0.499
Relative Importance		11.31%	22.82%	13.55%	21.39%
	-5%	-1.093	-0.219	-0.243	-0.087
Crime/Public Safety	0%	0.491	0.433	0.33	0.187
	5%	0.602	-0.214	-0.088	-0.1
Relative Importance		30.65%	17.02%	11.55%	7.75%
	-5%	-0.065	-0.562	-0.369	-0.61
Natural Resources	0%	0.199	0.372	0.455	0.305
	5%	-0.134	0.19	-0.086	0.305
Relative Importance		6.03%	24.41%	16.63%	24.65%
Constant		5.426	5.567	5.519	5.519
N=		8	22	64	34

New Deal 1: Roosevelt Era (1932-1944)
New Deal 2: New Deal Consolidated (1948-1964)
Post-New Deal 1: Years of Turmoil (1968-1976)
Post-New Deal 2: Reagan Era (`1980-1988).

Figure 6.5: Generational Differences in Education Spending

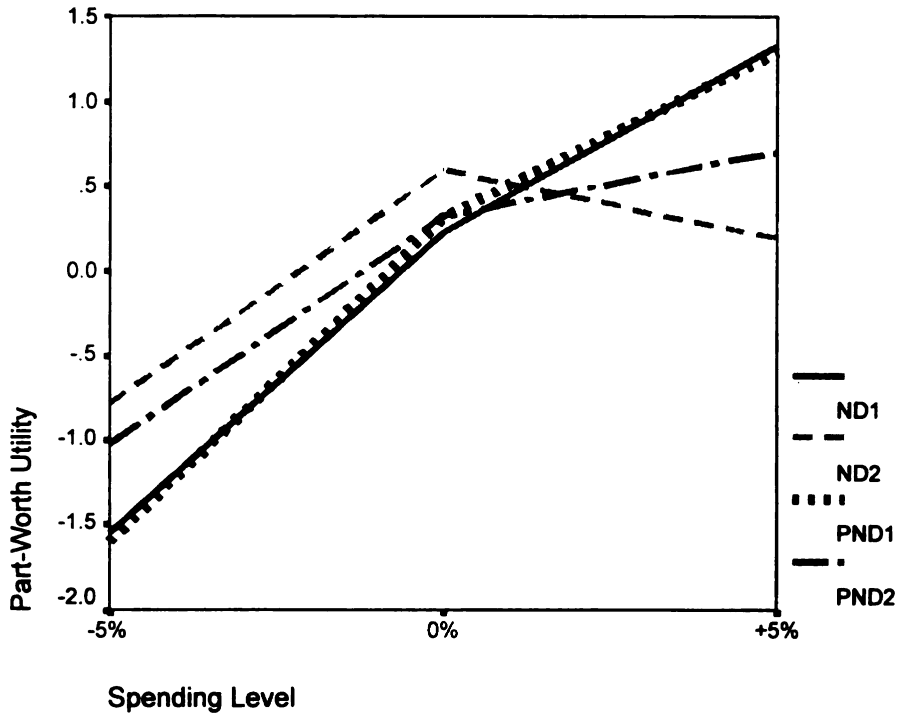


Figure 6.6: Generational Differences in Welfare Spending

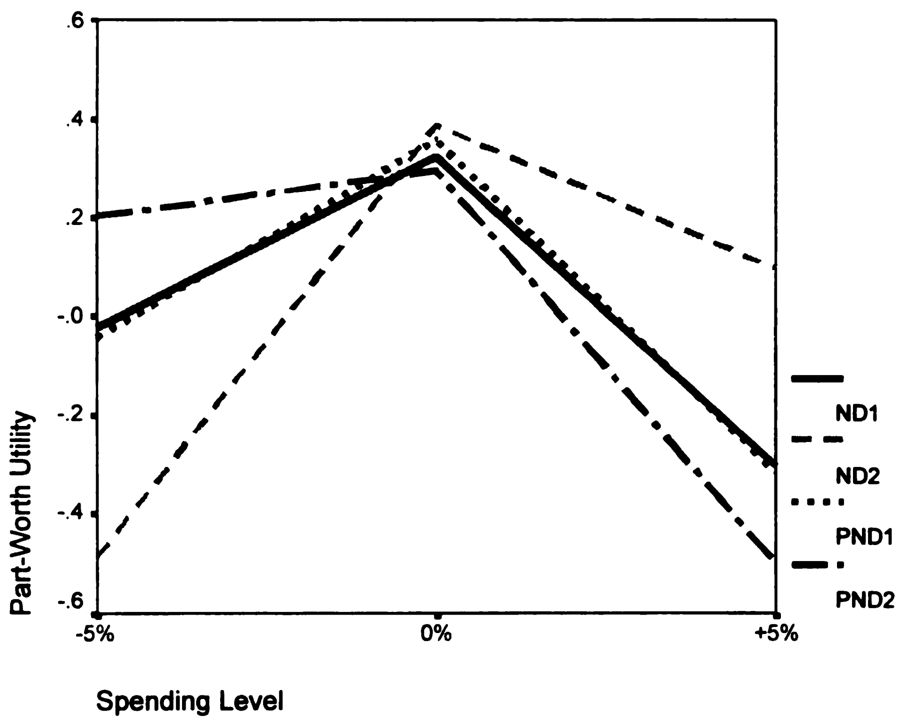


Figure 6.7: Generational Differences in Crime Spending

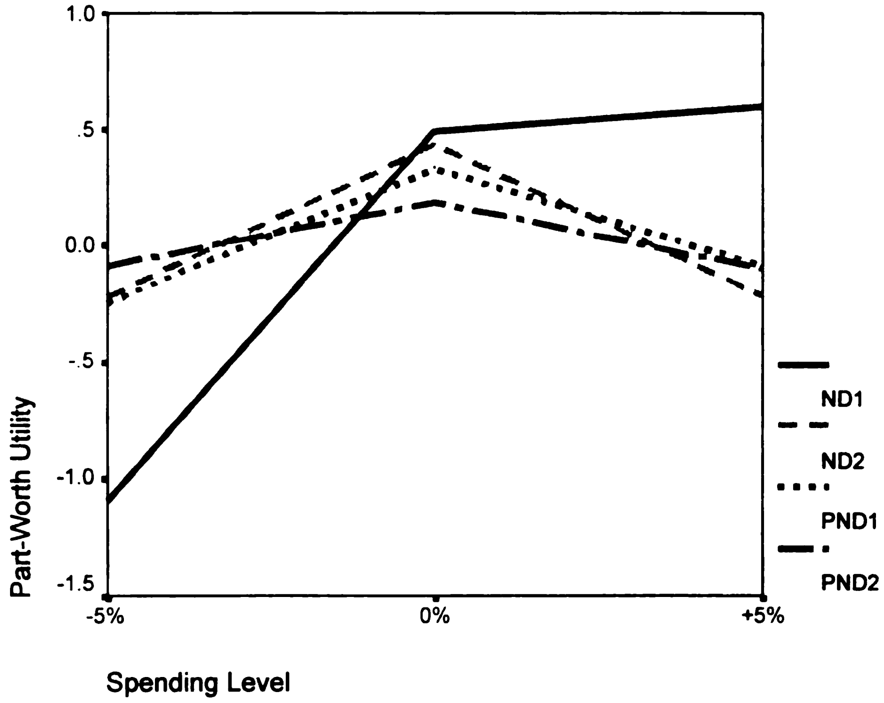
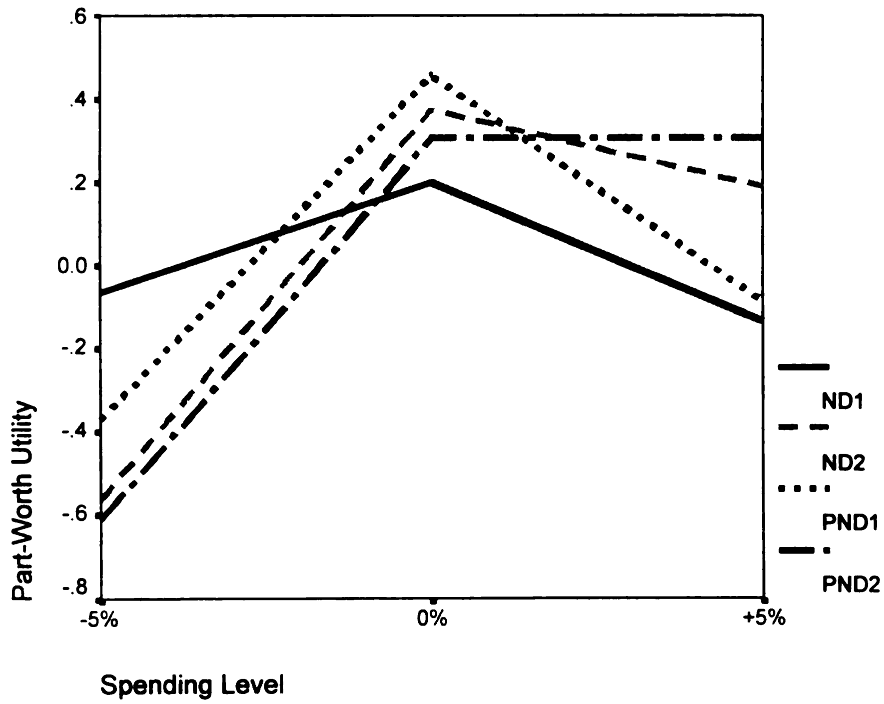


Figure 6.8: Generational Differences in Natural Resources Spending



Religious Differences

It is well known that religion plays an important role in American Politics. After reviewing numerous studies of voting behavior, Converse (1974) concludes that "religious differentiation intrudes on partisan political alignment in an unexpectedly powerful degree wherever it conceivably can" (p. 734). American religions differ in their positions on a number of policy questions. In some cases these differences are related to differing party loyalties due to the fact that religion has been one of the enduring bases of party alignments. But many differences follow from distinctive religious and cultural beliefs and experiences. According to Wald (1997), there are six religious categories: Roman Catholic, Jewish, Mainline Protestant, Evangelical Protestant, Black Protestant, and Secular. Previous studies have identified some basic political tendencies of these religious groups. American Jews have long been identified with left-wing political causes. Seculars tend to be young, mobile, well-educated, affluent, and to live in urban or metropolitan areas. Because such characteristics are usually associated with unconventional thinking, it has been thought that the seculars tend to share the liberal political outlook of Jews. Historically associated with the Republican party and conservative positions on many issues, the mainline Protestants have been faced with growing liberal sentiment from clergy and denominational leaders. In contrast, the evangelicals have been encouraged to move in a conservative direction by some of their most vocal pastors.

Like Jews, Catholics coalesced around the Democratic party in the 1930s, and the Catholics' commitment to the party was reinforced in 1960 when the Democrats conferred their presidential nomination on a Catholic candidate. Some observers have doubted the depth of this commitment, arguing that attachment to the more liberal political party was a transient stage, reflecting the immigrant and working class background of Catholics. Under conditions of economic parity with Protestants, it has been suggested, Catholics will embrace the political conservatism that appears to be a more natural outgrowth of Church doctrine. Although many Catholics have moved into the economic mainstream and the Republicans have courted them aggressively (using the abortion issue among others), it remains open to debate whether the Catholic masses have moved to the right side of the political spectrum (Wald 1997: 172-76).

For religious differences, we could find some clues from the 1992 NES surveys. For the question of maintaining or increasing government service, religious preferences from the more government services to less government services are: Black Protestant, Secular, Jewish, Roman Catholic, Evangelical Protestant, and Mainline Protestant. About the welfare spending issue, religious preferences from spending more to spending less are: Black Protestant, Secular, Jewish, Roman Catholic, Mainline Protestant, and Evangelical Protestant.

Table 6.3 exhibits the religious differences from Meridian Township voters. In the area of education (see Figure 6.9), all subgroups have almost identical preferences - linear utility functions favoring an increase in education spending. Education is also the most important issue for Seculars, Evangelical Protestants, and Mainline Protestants (second most important issue for Catholics). For welfare spending (see Figure 6.10), most subgroups favor spending at the status quo except the Catholics, which have a negative utility function favoring cuts in welfare spending. The welfare issue is also the most important issue for Catholics. In the area of crime (see Figure 6.11), the Mainline

Protestants slightly prefer an increase in crime spending (utility increase from 0.171 to 0.192) while other subgroups favor spending at the current level. For the natural resources issue (see Figure 6.12), all subgroups favor spending at the status quo but Seculars have similar utilities between the status quo (0.521) and an increase in spending (0.453). In sum, we did not find any significant religious differences in four spending areas.

Table 6.3: Religious Differences in 4 Major Policy Categories

Attribute	Level	Secular	Catholic	Mainline	Evangelical	Other
	-5%	-1.505	-1.005	-1.422	-1.342	-1.299
Education	0%	0.403	0.366	0.310	0.532	0.341
	5%	1.102	0.638	1.112	0.810	0.958
Relative Importance		49.55%	31.80%	61.16%	48.24%	51.95%
	-5%	-0.426	0.628	-0.268	-0.104	-0.047
Welfare/Health	0%	0.230	0.460	0.363	0.386	0.340
	5%	0.196	-1.088	-0.096	-0.281	-0.293
Relative Importance		12.48%	33.22%	15.23%	14.95%	14.56%
	-5%	-0.009	-0.362	-0.363	-0.099	-0.265
Crime/Public Safety	0%	0.256	0.603	0.171	0.537	0.300
	5%	-0.247	-0.241	0.192	-0.438	-0.035
Relative Importance		9.56%	18.66%	13.38%	21.86%	12.99%
	-5%	-0.973	-0.443	-0.193	0.077	-0.497
Natural Resources	0%	0.521	0.400	0.230	0.295	0.394
	5%	0.453	0.043	-0.037	-0.372	0.103
Relative Importance		28.41%	16.32%	10.22%	14.95%	20.50%
Constant		5.624	5.469	5.496	5.493	5.524
N=		40	28	45	11	4

Figure 6.9: Religious Differences in Education Spending

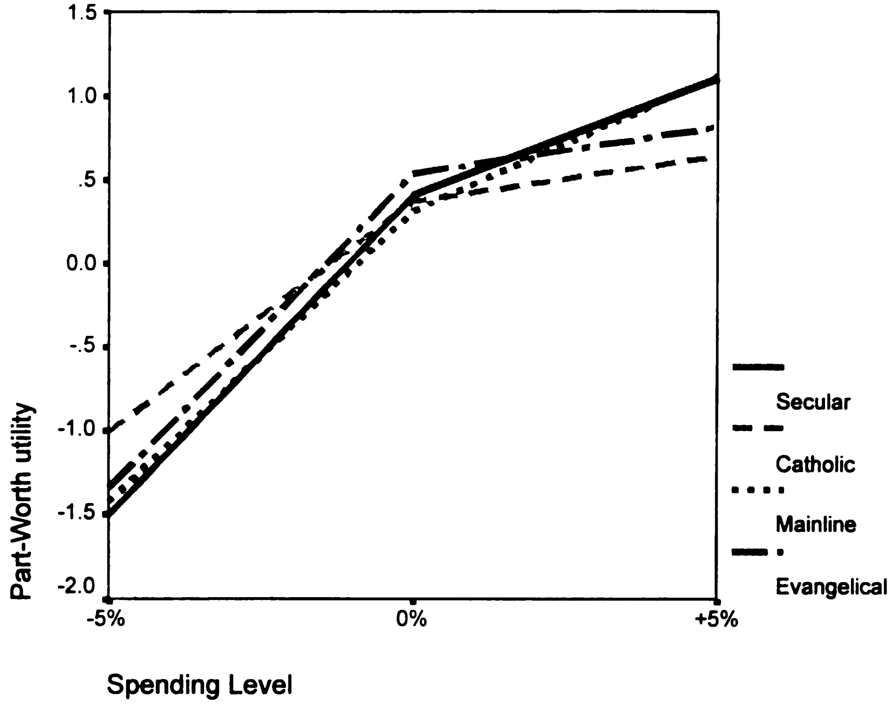


Figure 6.10: Religious Differences in Welfare Spending

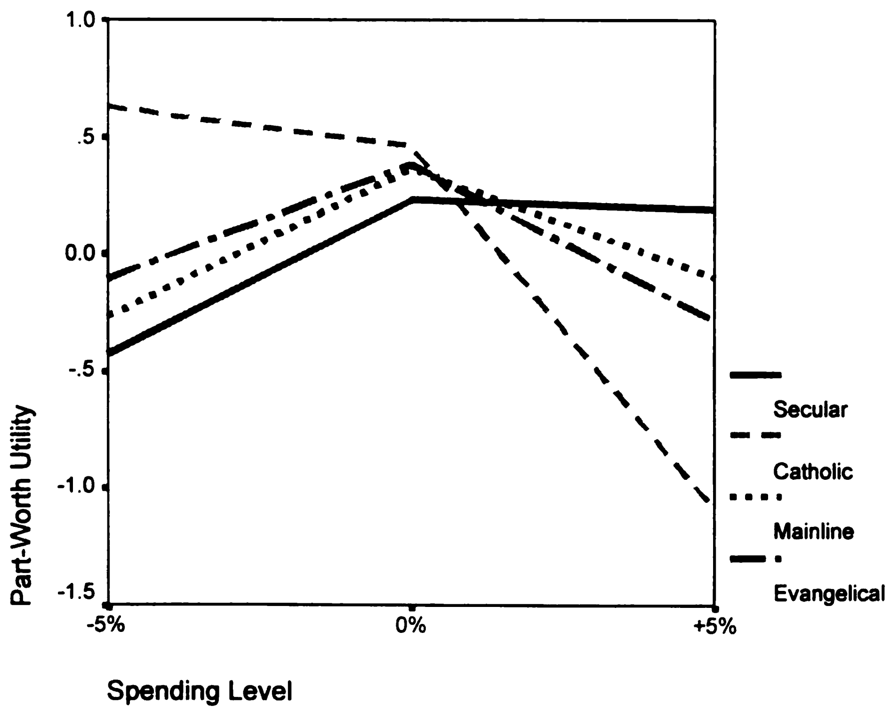


Figure 6.11: Religious Differences in Crime Spending

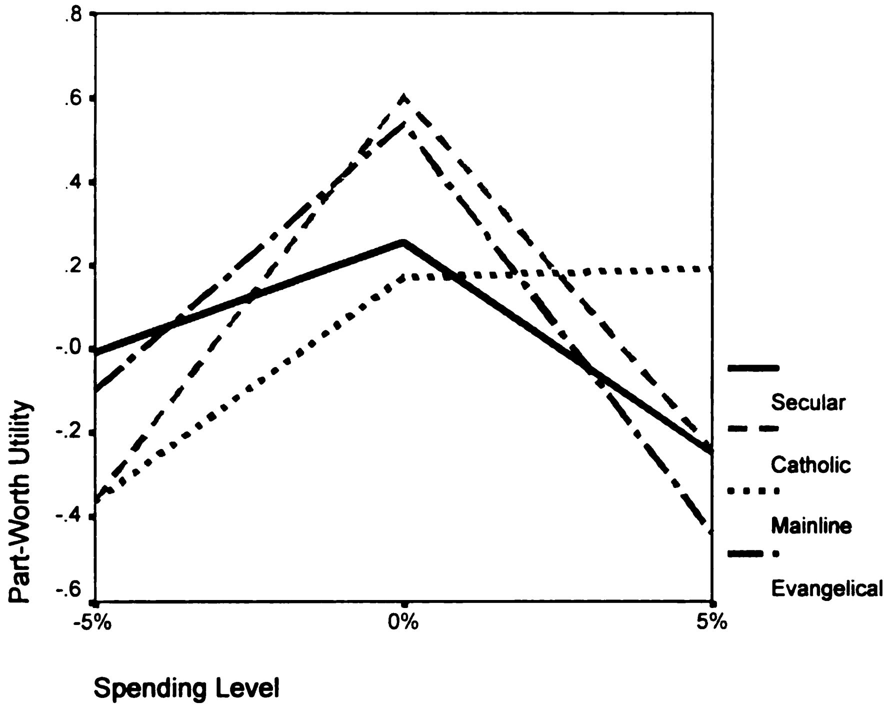
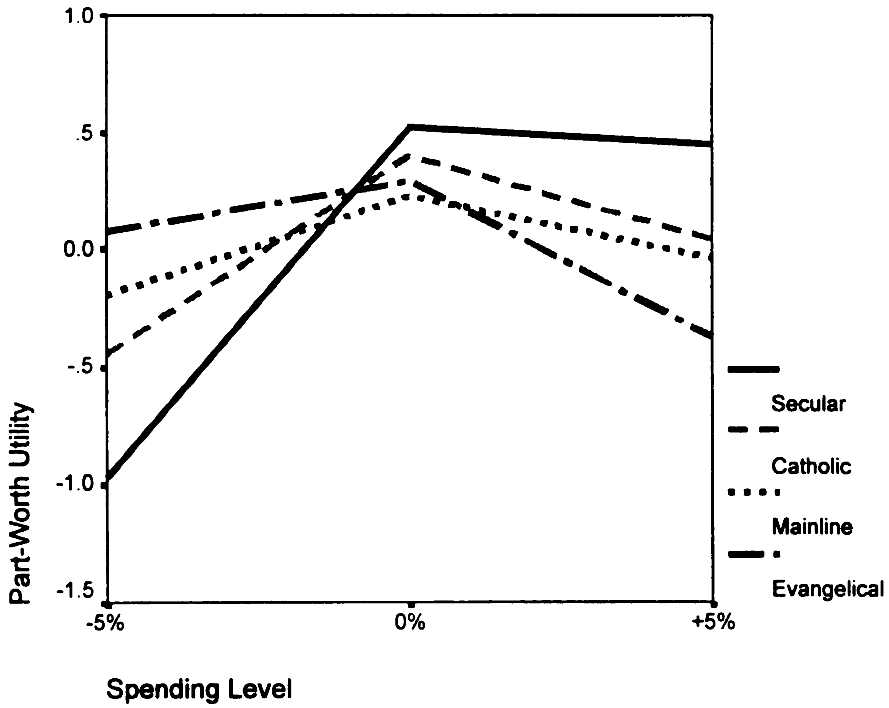


Figure 6.12: Religious Differences in Natural resources Spending



Party Differences

The policy preferences of Democrats and Republicans tend to differ in ways related to various economic, ethnic, and religious bases of the two parties as well as to their different receptivity to messages from party leaders.

Since the New Deal the Democrats have enjoyed more support among working-class and low-income people, and among Catholics, and Jews, blacks, and various white ethnic groups, whereas the Republican have prevailed among the white and relatively wealthy Anglo-Saxon Protestants. Over this period, white Southerners and some Catholics and ethnics have gradually drifted away from their Democratic loyalty, and the large ideological differences of 1930s and 1940s have diminished. Still, the parties in the 1990s continued to differ from each other in similar ways.

In accord with their working-class status and relatively low income, Democrats have tended much more often than Republicans to favor a broad range of government social welfare programs, such as education, for example, and assistance with medical care, jobs, income, and housing.

In addition, Page and Benjamin (1992) argue that party groups are actually the only ones for which they regularly found cases of divergent movements of opinion, with the opinions of party groups opinions becoming more different from each other over time. Such divergent partisan trends are related to party leadership of public opinion, especially in the area of foreign policy. For instance, two years after the Republican President Eisenhower took office in 1952, Republicans had tended to move from their isolationist posture toward favoring active U.S. involvement abroad, while Democrats' opinions stayed much the same. Similar situations occurred in Watergate, Vietnam War, etc. (Pp.307-311).

Table 6.4 demonstrates the utility differences of four policy areas among party groups. For education spending (see Figure 6.13), Democrats have a linear utility function favoring an increase in education spending while Republicans are satisfied with the current spending on education. The utility function for Democrats looks like a straight line that favoring an increase in education spending with its highest utility score 2.017 while opposing cuts in education spending with lowest utility score -2.098. The utility for Independents also indicates a linear function favoring an increase in education spending. Democrats and Independents view education as the most important issues with relative importance scores 54.37 percent and 54.07 percent respectively, Republicans only view education as the third important issue in the relative importance score.

In the area of welfare and health, Democrats are the only group favoring increased spending: they are better off when spending increases (utility increase from 0.104 to 0.639) while worse off when spending decreases (utility decrease from 0.104 to -0.744). In contrast, both Independents and Republicans favor spending at the status quo but Republicans suffer a serious utility loss when spending increases; the utility decreases from 0.534 to -1.054. To what extent that Republicans oppose an increase in welfare spending? If we look at the relative importance score, we will find out that Republicans strongly object to increase spending: welfare is the number one issue for Republicans with a relative importance score 33.75 percent.

For the crime and public safety issues, (see Figure 6.15), Democrats are the only group favoring an increase in crime spending. Both Republicans and Independents are satisfied with the current spending and suffer a utility loss when either spending increases or decreases. However, before we say that Democrats prefer increasing crime spending, we should also notice that the crime issue is the least important area for Democrats with a relative importance score 5.02 percent only.

For the natural resources spending (see Figure 6.16), Democrats prefer an increase in natural resources spending while both Republicans and Independents favor the current spending level, although Independents' utility is almost indistinguishable between the status quo and an increase in spending.

Table 6.4: Party Differences in 4 Major Policy Categories

Attribute	Level	Democrat	Independent	Republican
	-5%	-2.098	-1.630	-0.459
Education	0%	0.081	0.354	0.566
	5%	2.017	1.276	-0.106
Relative Importance		54.37%	54.07%	21.78%
	-5%	-0.744	-0.124	0.520
Welfare/Health	0%	0.104	0.327	0.534
	5%	0.639	-0.203	-1.054
Relative Importance		18.28%	9.86%	33.75%
	-5%	-0.212	-0.442	-0.133
Crime/Public Safety	0%	0.043	0.378	0.482
	5%	0.169	0.064	-0.349
Relative Importance		5.02%	15.27%	17.66%
	-5%	-0.919	-0.740	0.128
Natural Resources	0%	0.148	0.378	0.567
	5%	0.771	0.361	-0.695
Relative Importance		22.33%	20.81%	26.81%
Constant		5.641	5.545	5.414
N=		38	39	51

Figure 6.13: Party Differences in Education Spending

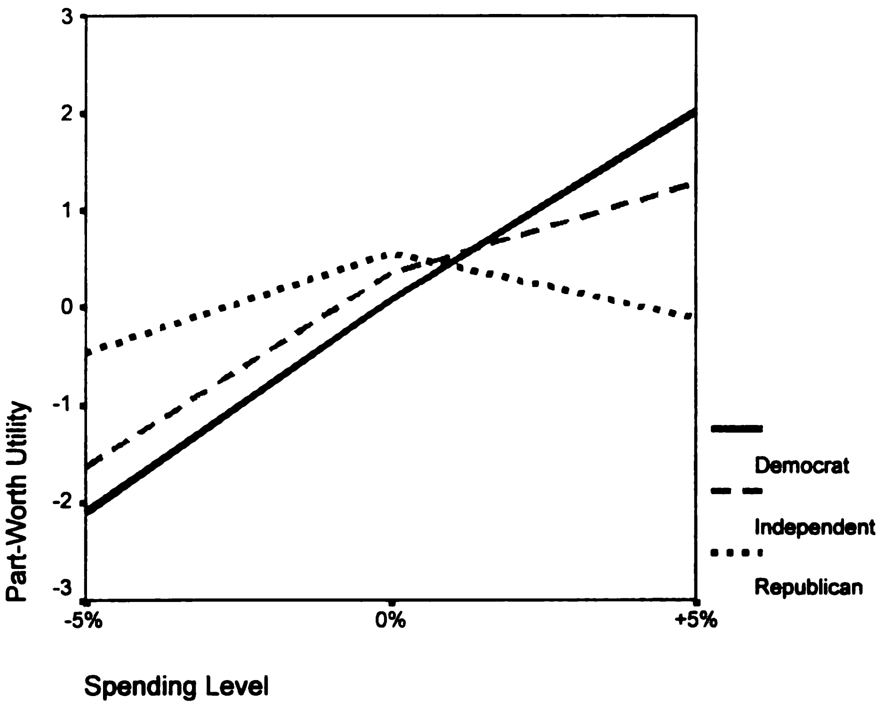


Figure 6.14: Party Differences in Welfare Spending

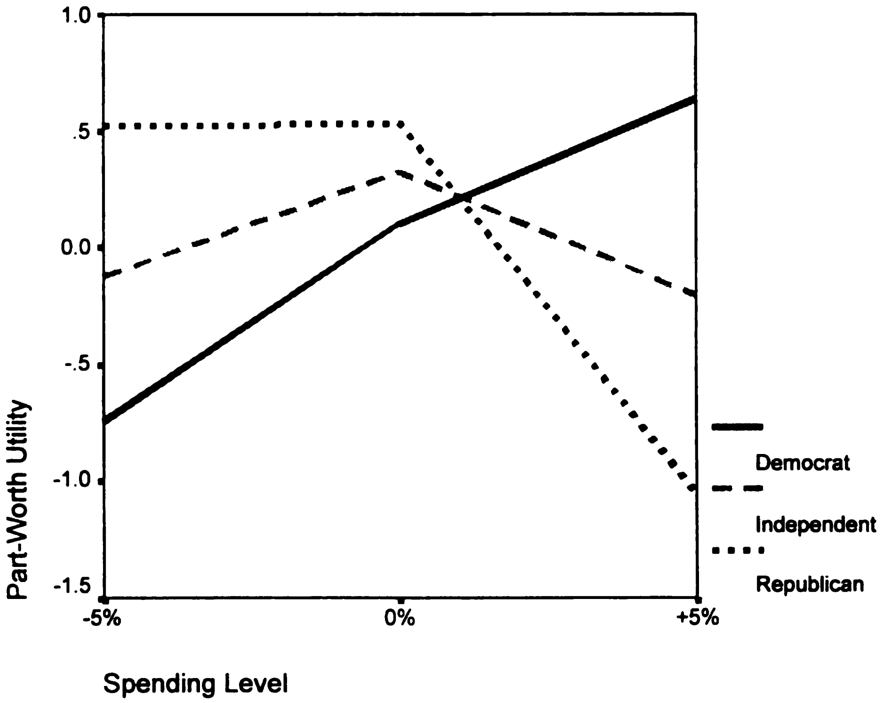


Figure 6.15: Party Differences in Crime Spending

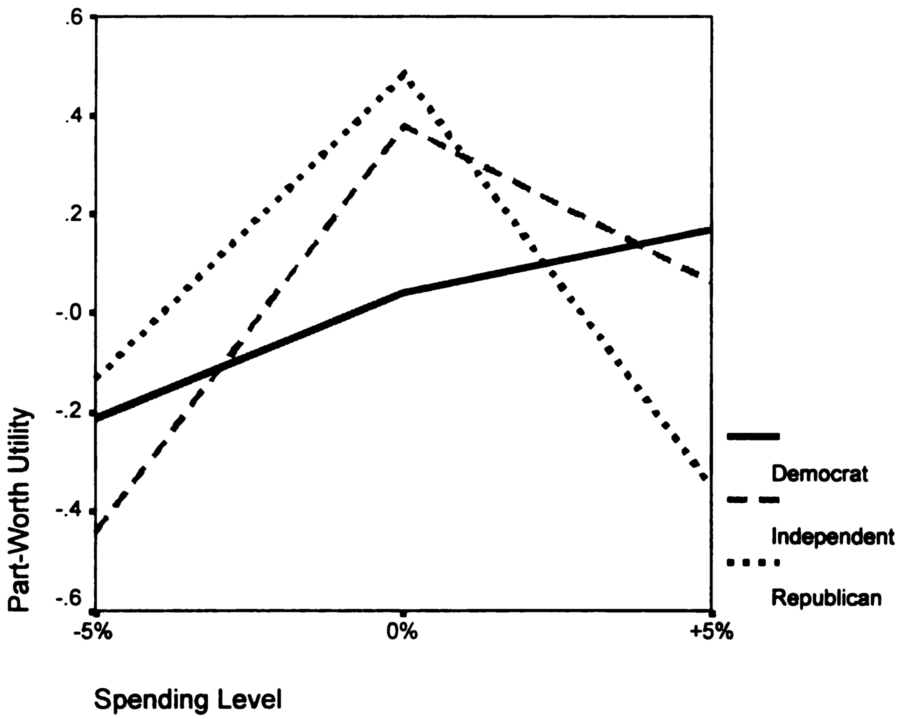
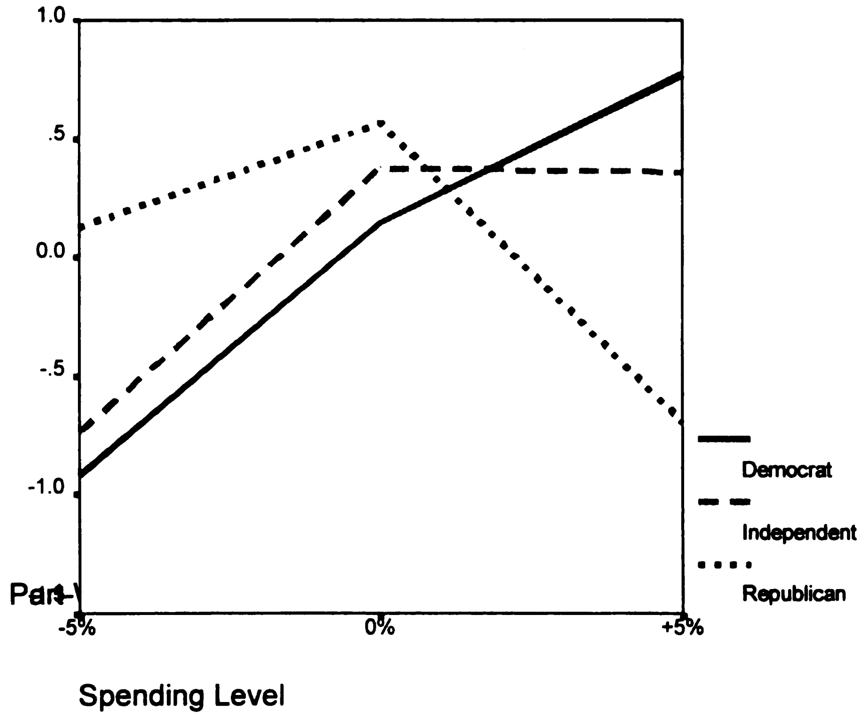


Figure 6.16: Party Differences in Natural Resources Spending



Children

From the Table 6.5 (also see Figure 6.17), we see whether having children at home or not only makes a minor difference in the area of education: both groups have linear utility functions favoring spending increases while people with children at home have slightly higher utility gain when an increase in education spending. Similarly, both groups view education as the most important issue and the group with children at home has relative higher relative importance score (55.51 percent in "with children at home" vs. 44.65 percent in "without children at home"). However, in the areas of welfare, crime, and natural resources (see Figure 6.18, 6.19 and 6.20), there is no significant difference between the two groups; both groups prefer the status quo than either increases or cuts in spending.

Table 6.5: Children Differences in 4 Major Policy Categories

Attribute	Level	Utility W/ Children	Utility W/O Children	Difference
	-5%	-1.523	-1.060	-0.462
Education	0%	0.384	0.329	0.055
	5%	1.139	0.732	0.407
Relative Importance		55.51%	44.65%	10.86%
	-5%	0.185	-0.311	0.496
Welfare/Health	0%	0.353	0.333	0.02
	5%	-0.538	-0.022	-0.515
Relative Importance		18.57%	16.04%	2.53%
	-5%	-0.229	-0.274	0.044
Crime/Public Safety	0%	0.308	0.333	-0.025
	5%	-0.079	-0.06	-0.02
Relative Importance		11.21%	15.11%	-3.90%
	-5%	-0.268	-0.644	0.376
Natural Resources	0%	0.438	0.328	0.11
	5%	-0.169	0.317	-0.486
Relative Importance		14.72%	24.21%	-9.49%
Constant		5.478	5.569	-0.091
N=		67	61	

Figure 6.17: Children Differences in Education Spending

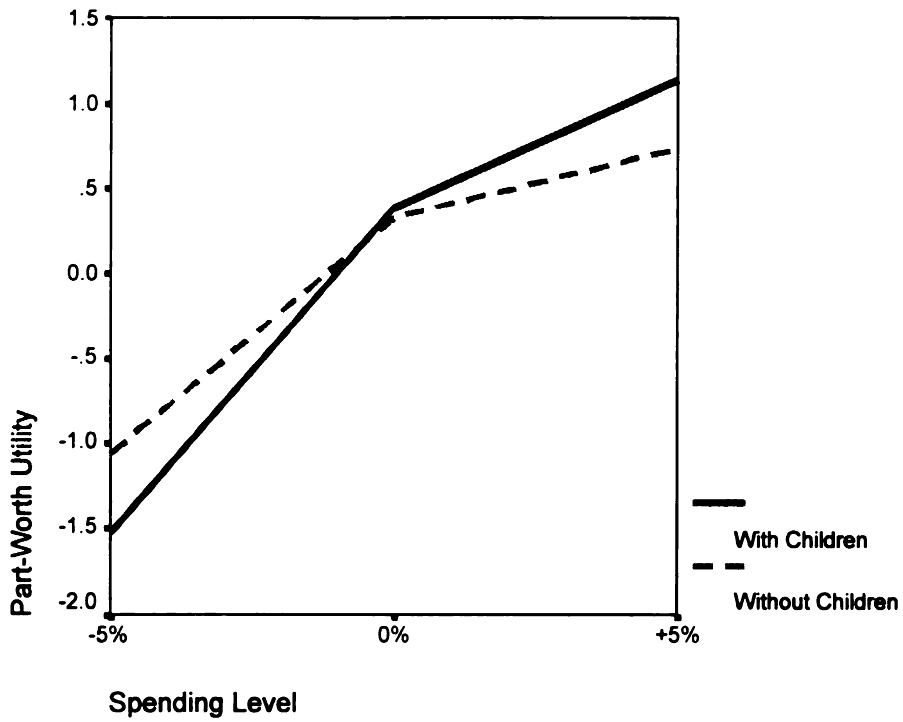


Figure 6.18: Children Differences in Welfare Spending

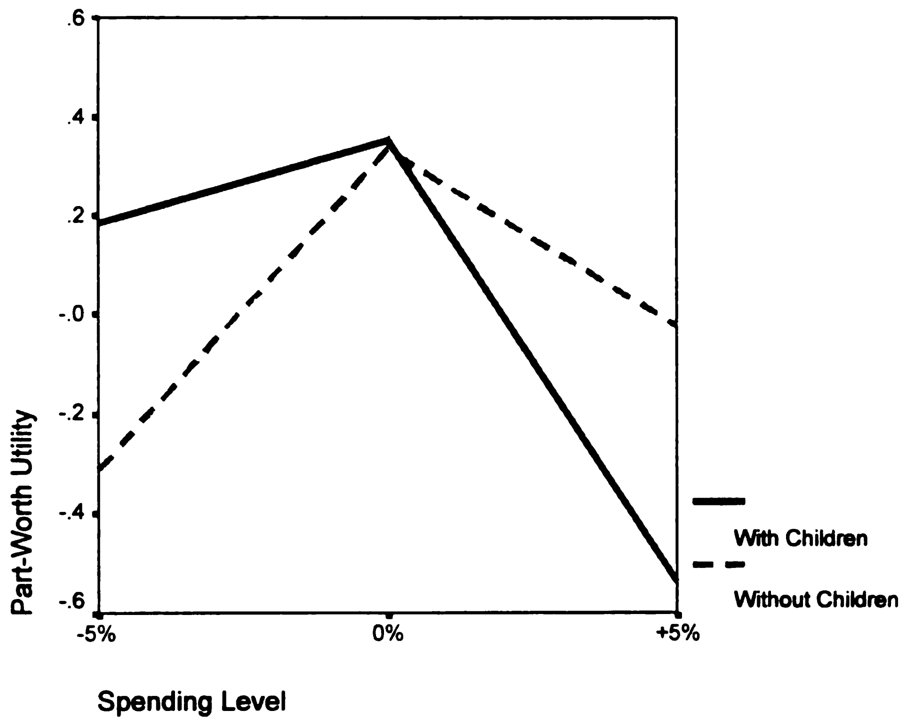


Figure 6.19: Children Differences in Crime Spending

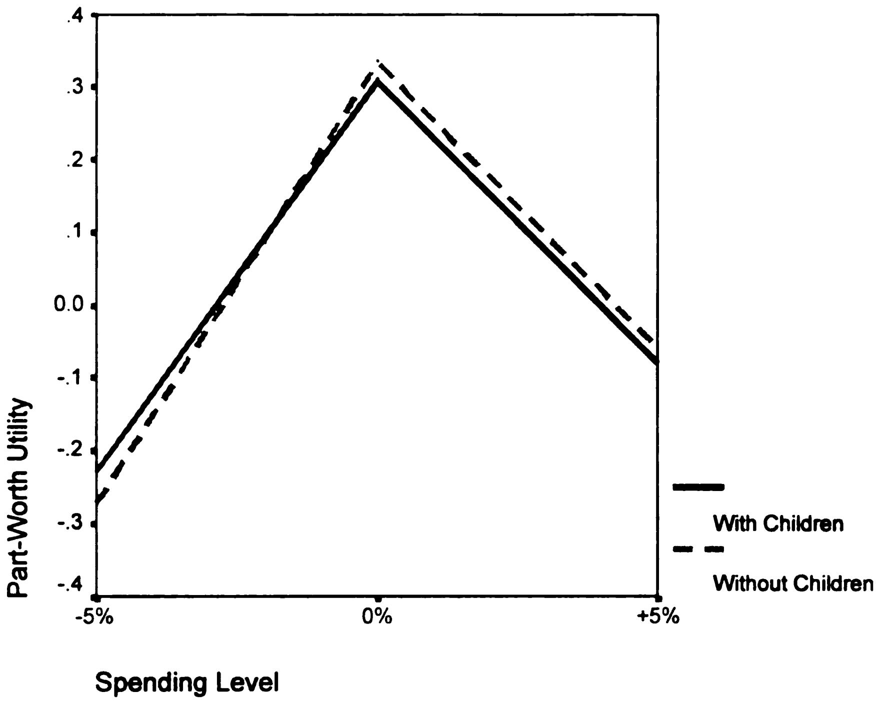
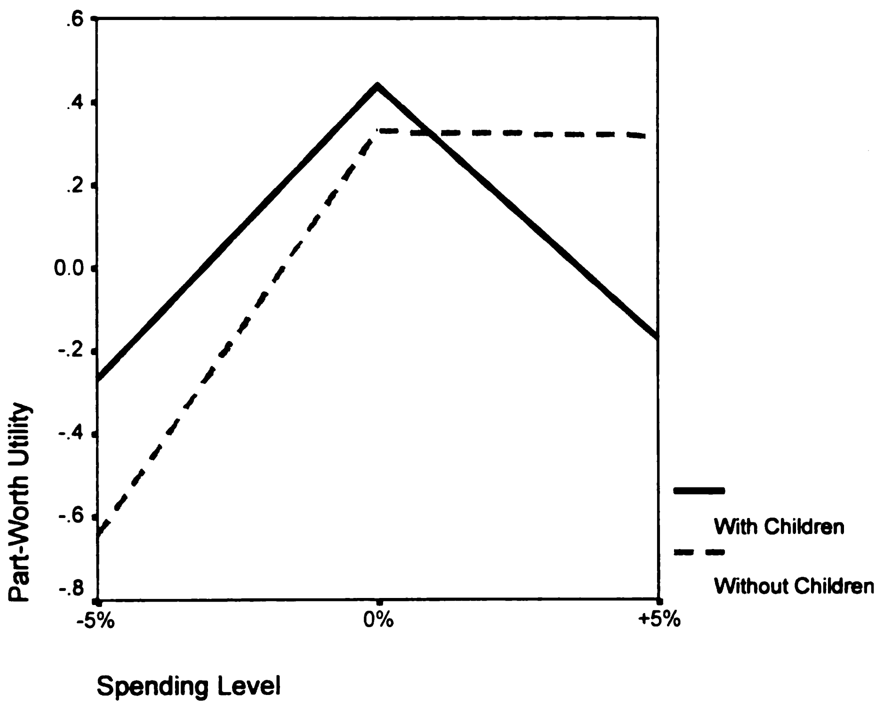


Figure 6.20: Children Differences in Natural resources Spending



Home Ownership

Home ownership represents a traditional indicator of social economic status in the United States. Therefore, we could view home owners as middle or high income families while viewing renters as low income families. People of low income are considerably more likely to favor social welfare spending while people of middle or high income are more likely to be against social welfare spending. In other issues, class cleavages are usually weak indicators for various spending preferences. Table 6.6 exhibits the utility differences between home owners and renters. First we notice that there is no difference on education spending; both groups have a linear utility function favoring spending increases (see Figure 6.21). For welfare and crime spending (see Figure 6.21 and 6.22), both home owners and renters are satisfied with current spending levels. The only difference between home owners and renters occurs in the area of natural resources: renters prefer spending increases (utility increase from 0.142 to 0.198) while home owners favor spending at the current level. Since there are only 15 cases in the renter category (while 113 cases in home owner category), we are not surprised by the results that no significant utility difference is found between home owners and renters.

Table 6.6: Differences Between Home Owners and Renters in 4 Major Policy Categories

		Utility		Difference
Attribute	**Level**	**Home Owner**	**Renter**	
	-5%	-1.276	-1.503	0.227
Education	0%	0.369	0.272	0.097
	5%	0.907	1.231	-0.324
Relative Importance		50.67%	54.79%	-4.12%
	-5%	-0.003	-0.417	0.414
Welfare/Health	0%	0.329	0.453	-0.124
	5%	-0.326	-0.036	-0.290
Relative Importance		15.20%	17.45%	-2.25%
	-5%	-0.285	0.009	-0.293
Crime/Public Safety	0%	0.347	0.120	0.227
	5%	-0.062	-0.128	0.067
Relative Importance		14.65%	4.97%	9.68%
	-5%	-0.422	-0.640	0.218
Natural Resources	0%	0.417	0.142	0.275
	5%	0.004	0.498	-0.493
Relative Importance		19.48%	22.97%	-3.49%
Constant		5.512	5.591	-0.080
N=		113	15	

Figure 6.21: Home Differences in Education Spending

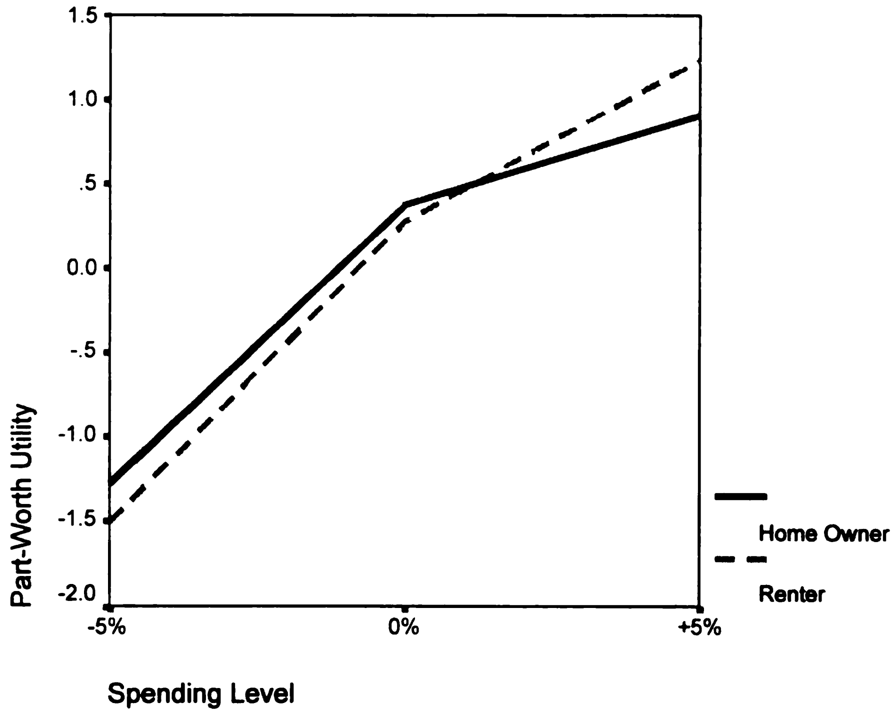


Figure 6.22: Home Differences in Welfare Spending

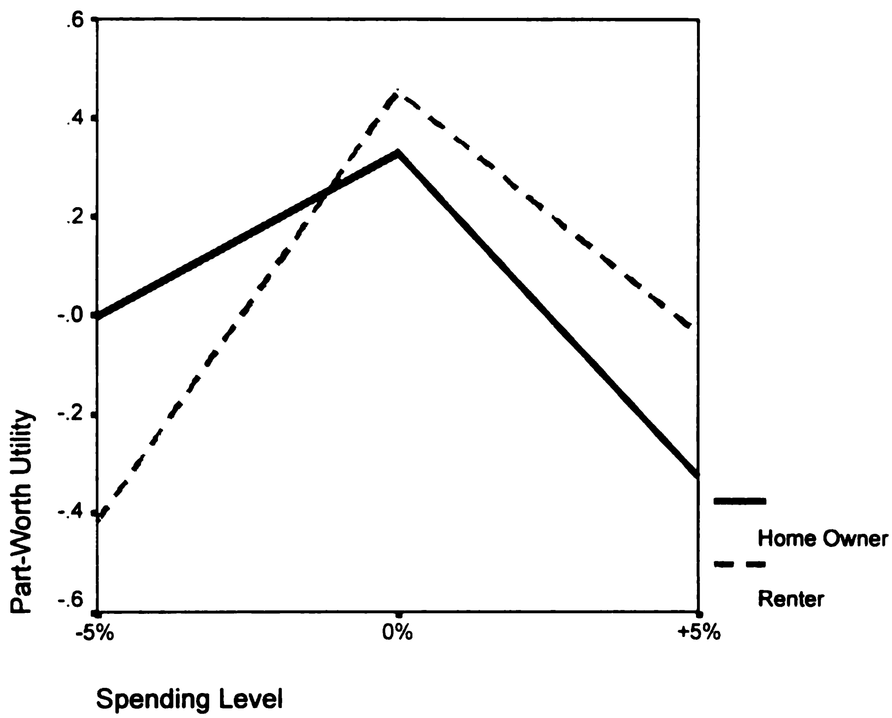


Figure 6.21: Home Differences in Crime Spending

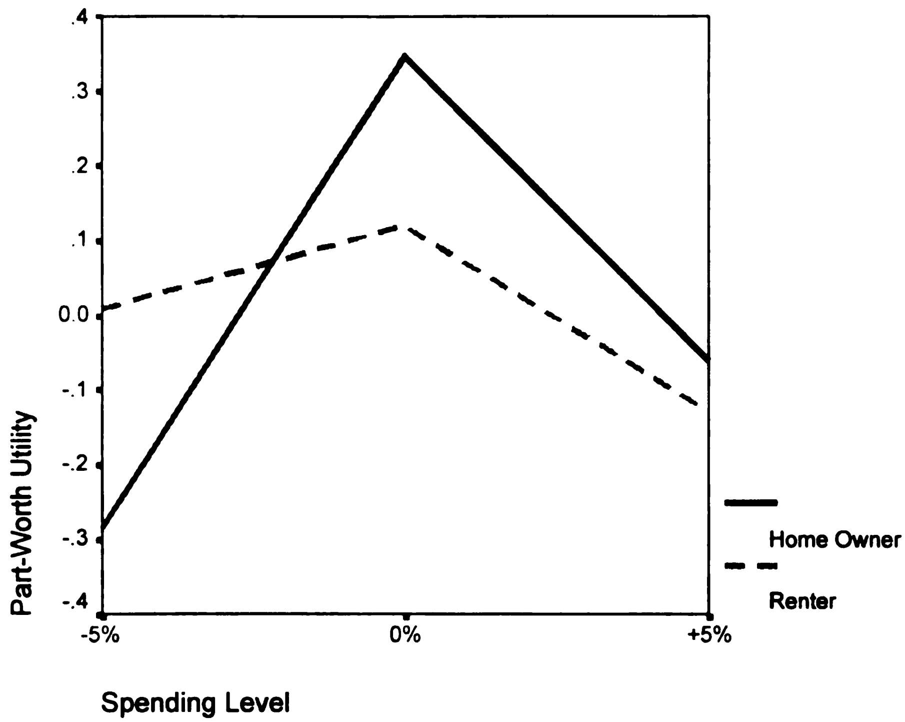
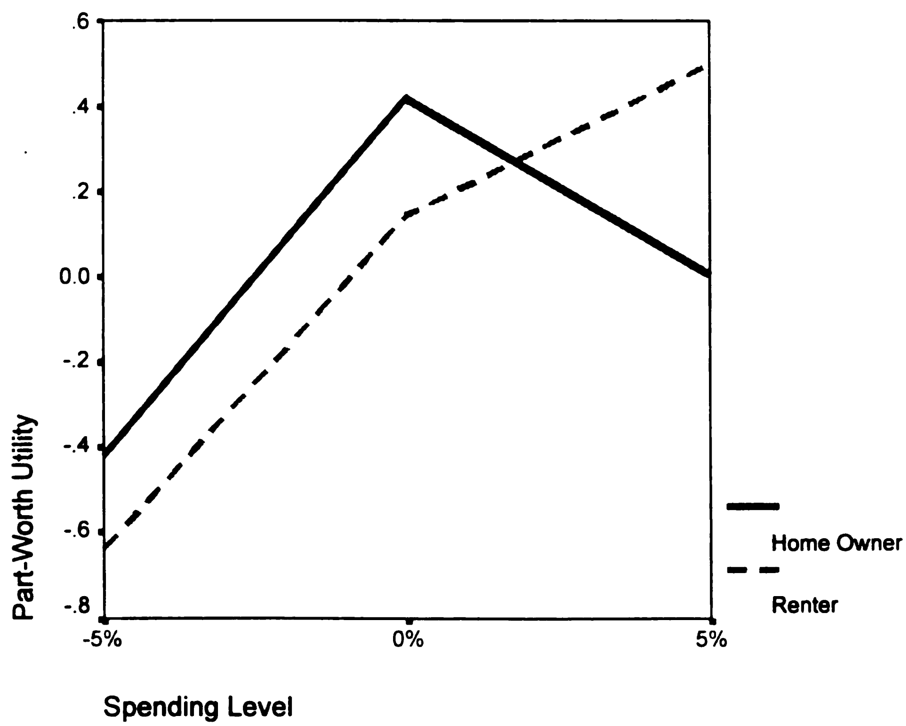


Figure 6.24: Home Differences in Natural Resources Spending



Summary

In this chapter we examined the group differences in terms of part-worth utilities and relative importances of attributes. Among Meridian Township voters, we found some similar group differences with previous studies: women are more concerned with compassion issues such as education and welfare than men; the New Deal 1 generation (the oldest cohort in this study) favors an increase in crime; Post-New Deal 2 (the youngest cohort in this study) favors either an increase in natural resources spending or the status quo; Democrats have a linear utility function favoring education spending increased while Republicans are satisfied with current spending level on education; Democrats also prefer an increase in welfare spending while Republicans favor spending at the status quo but suffer a dramatic utility loss with an increase in welfare; people with children at home have a slightly higher utility increase when education spending increases.

As we noted before, our data came from the local area with a limited sample size (N=128), thus our conjoint results cannot be generalized to inference about group attitudes in the U.S. Despite these limitations, in this study we have shown Meridian Township voters' policy preferences are based on their trade-off judgments rather than on single dimensional consideration of incomplete budget options.

CHAPTER 7

CONJOINT APPLICATIONS: SIMULATIONS OF GUBERNATORIAL ELECTION AND POLICY EVALUATION

One of major advantages of conjoint analysis is that it provides an opportunity to conduct a computer simulation of choices by voters. A computer simulation can answer various "what-if" scenarios and determine the likely results of attribute/attribute level tradeoffs for the conceptualized services or products (e.g., different budget proposals in our study). Choice simulations reveal respondent preferences for the specific services defined by the researcher. In this study, first we ran simulations to identify which candidate (John Engler/Republican or Jeffrey Fieger/Democrat) would win the gubernatorial election in Michigan and then show how the underdog candidate could surpass his opponent by changing his policy positions. Second, we show how candidates can target a segment of voters (i.e., Independents) in order to win the gubernatorial election. Third, we illustrate a policy evaluation in welfare reform by simulating several policy alternatives against the current welfare spending level. Lastly, we demonstrate how a governor can distribute revenue to a single spending area in order to maximize voters' utilities.

Simulation of Gubernatorial Election in Michigan

Simulating the gubernatorial election is related to the topic of *issue voting*. Issue voting includes two major evaluations: *retrospective* voting and *prospective* voting. Retrospective voting examines the incumbent's performance on policy, and how likely is it that his opponent would outperform him? Prospective voting examines the policy platforms addressed by the candidates and evaluates which candidate's policy promises are most similar to what the voter believes the government should be doing (Niemi and Weisberg, 1993: 98-99).

An important question is to what extent the voters' voting behavior is based on policy issues rather than on their social backgrounds or party loyalties. The authors of *The American Voter* argue that the voters are often ill informed about public policy and may not be able to vote based on issues. They reported that in 1956 on 16 different issues only 18 to 36 percent of the electorate satisfied what the authors considered three prerequisites for issue voting: (1) being familiar enough with the issues to have an opinion, (2) knowing what the government is doing on the issue, and (3) seeing the parties differ on the issue. Converse (1964) concluded that some issues were marked by "nonattitudes," with respondents answering questions basically randomly simply to satisfy the overinsistent interviewer; this further emphasized the relative unimportance of issues in voting. However, Abramson *et al.* (1995) found that there is a reasonably high relationship (from 41% to 80%, depending on how many prerequisites for issue voting are satisfied) between the voter's issue position and the candidate he or she supported on various issues from 1972 to 1992 presidential election (pp. 179-189). They argue that failure to meet the prerequisites for issue voting does not mean that the voter has ill-formed preferences. On the contrary, the voter's ability to see differences between candidates varies because the political environment changes from election to election (p. 182).

In our gubernatorial election simulation, voters actually evaluate candidates based on both the prospective voting and retrospective voting. Specifically, voters will evaluate both governor John Engler's performance and his platform. On the other hand, voters will tend to focus more on Jeffrey Fieger's platform because he is the challenger (although voters will still possibly evaluate the performances of the past Democrat governors). Since it is not necessary to meet the prerequisites for issue voting, we can run gubernatorial election simulations without knowing how many criteria the Meridian Township voters can meet.

Three Choice Simulation Models

Researchers generally use three types of models to simulate market behavior based on conjoint data: *maximum utility (first choice) model*, *Bradley-Terry Luce/BTL (average choice) model*, and *logit (share of preference) model*. Each model estimates the respondent's preferences by combining their utilities with a set of competing profiles on the study attributes. None of these models can support absolute statements such as: "Proposal A will get 20% votes" or "Proposal C will beat Proposal B in 2-to-1 votes." Conjoint models produce only relative results and are best used for ranking alternative courses of action. For example, they can tell us that Proposal A is expected to be least preferred or that voters will preferred Proposal C rather than Proposal B (Curry 1997a).

Maximum utility (first choice) model identifies the profile with the highest utility as the profile of choice. This profile is selected and receives a value of 1; ties receive a 0.5 value. After the process is repeated for each respondent's utility set, the cumulative "votes" for each profile are evaluated as a proportion of the votes or respondents in the sample. Although it looks conceptually and computationally simple, the first choice model has one major drawback. That is, it tends to produce more extreme results than are observed in actual buying situations; it overestimates preferences for the most preferred profiles and underestimates preferences for those that are least preferred. This occurs because first choice models make no allowance for respondent error (Curry 1997b).

The Bradley-Terry-Luce (average choice) model estimates choice probability in a different fashion. It computes the probability of choosing a profile as the most preferred by dividing the profile's utility by the sum of all the simulation total utilities. The BTL model can be expressed as follows:

$$P(p) = \frac{U_p}{\sum U_p} \qquad (7.1)$$

where:

P(p) = probability for choosing Profile P; p = 1,..., p, and

U_p = overall utility for Profile P.

Unlike the Maximum utility model, which assigns a respondent's entire preference to the profile with the highest utility, the Share of Preference (logit) model splits the probability of purchase among all competing products or according to their utilities. It is critical to ask: why should the logit model be better? First it recognizes that consumers do not always purchase the product for which they have the highest utility. Second, many products and services represent low-involvement purchases for consumers; for these products consumers often have no clear first choice.

Share of preference (logit) models use an assigned choice probability that is proportional to an increasing monotonic function of the alternative utility. The choice probabilities are computed by dividing the logit value for one product by the sum for all other products in the simulation. The logit transformation for converting a subject's utility for profile A, (U_A), into a preference share, P(A), for profile A in a market made up of two products, A and B, is:

$$P(A) = \frac{e^{bU_A}}{e^{bU_A} + e^{bU_B}} \qquad (7.2)$$

The preference share for Profile B is:

$$P(B) = \frac{e^{bU_B}}{e^{bU_A} + e^{bU_B}} \qquad (7.3)$$

where:

P(A) = probability for choosing Profile A; P(B) = probability for choosing Profile B,

e = 2.7183,

b = a measure of the individual's involvement in the product category, and

U_A = overall utility for Profile A; U_B = overall utility for Profile B.

This expression can be generalized to any number of profiles by adding a term in the denominator for each additional profile.

Researchers use the logit model because it has three intuitively desirable properties. First, the higher a profile's utility, the greater is its preference share. Second, the individual's preferences across all profiles sum to 1. And third, the greater the individual's involvement in the category (i.e. the larger b becomes), the more distinct are the preference shares.

Although the logit model generally produces better estimates than the first choice model, it overestimates preference for similar products. If we add identical (or very similar) products to our model, the total preference share for those products is artificially inflated (the "red bus/blue bus" problem). This is because the independence from irrelevant alternatives (IIA) assumption[1] implied by the logit rule.

In addition, the use of BTL and logit models usually involves arbitrary scaling assumptions; that is, the results are not invariant over linear transformations on the utilities estimated by conjoint analysis (Green and Srinivasan 1990).

In summary, while the literature shows that the first choice model can provide the best overall validation, choice behavior models (BTL and logit) have a strong probabilistic component which will be very useful because (1) consumers do not always choose the alternative with the highest utility, and (2) consumers sometimes have a low involvement for some services or products.

Simulation Results for Gubernatorial Election

In this simulation, first we identify two candidates' positions on 4 major policy categories. Since Engler is the incumbent governor, we assume that his policy positions are maintaining all spending areas at current levels. On the other hand, based on the Democrats' policy preferences in Meridian Township, we could assume that Democrat candidate Jeffrey Fieger's policy positions are increase in spending in education, welfare, and natural resources while maintaining crime spending at current level. A simulation is then executed to identify which candidate will win the gubernatorial election in Meridian Township. After we obtain the results, we run another simulation to see how the

[1] The independence of irrelevant alternatives problem arises from the assumption that the relative probability of choosing two existing alternatives is unaffected by the presence of additional alternatives. However, this is not the case in reality; adding an irrelevant alternative changes the probabilities assigned to all categories. For an example, see Kennedy (1992), p244. Empirical evidence presented by Kamakura and Srivastava (1984) suggests that the IIA assumption is reasonable at the individual level.

underdog candidate can surpass his opponent by changing his policy positions. Simulation results include profile utilities and results from three choice models.

Simulations results for two candidates John Engler and Jeffrey Fieger are captured in Table 7.1. In the Simulation 1, John Engler's policy positions are maintaining all spending areas at the status quo while Jeffrey Fieger proposes a platform with increasing spending on education, welfare, and natural resources but maintaining crime spending at the status quo. Simulation results show that profile utilities for Engler and Fieger are 6.9 and 6.6 respectively. The maximum utility model shows that Engler will win the election with probability 100 percent. The BTL and logit models also reveal that Engler has a higher probability to win: 51.37% (Engler) vs. 48.63% (Fieger) in the BTL model and 59.16% (Engler) vs. 40.84% (Fieger) in the logit model.

Table 7.1: Simulation Results for Gubernatorial Election (All Voters)

	Simulation 1		Simulation 2		Simulation 3	
	Profile		Profile		Profile	
Attribute	Engler	Fieger	Engler	Fieger	Engler	Fieger
Education	0%	5%	0%	5%	5%	5%
Welfare/Health	0%	5%	0%	0%	0%	0%
Crime/Public Safety	0%	0%	0%	0%	0%	0%
Natural Resources	0%	5%	0%	5%	0%	5%
	Simulation Results		Simulation Results		Simulation Results	
Candidate	Engler	Fieger	Engler	Fieger	Engler	Fieger
Profile Utility	6.9	6.6	6.9	7.2	7.5	7.2
Max Utility	100%	0.00%	0.00%	100%	100%	0.00%
BTL	51.37%	48.63%	49.06%	50.94%	51.10%	48.90%
Logit	59.16%	40.84%	43.42%	56.58%	58.00%	42.00%

Can Fieger surpass his opponent Engler by changing his policy positions? Simulation 2 provides the answer. In the Simulation 2, Fieger changes his policy positions by proposing another platform with increasing spending in education and natural resources but maintaining welfare and crime spending at current levels while Engler has identical policy positions with Simulation 1. Now the results from Simulation 2 indicate that the profile utilities for Engler and Fieger are 6.9 and 7.2 individually. The maximum utility model reveals that Fieger will win the election in this simulation. Similarly, the BTL and logit models also indicate that Fieger has a higher probability to win: 50.94 (Fieger) vs. 49.06% (Engler) in the BTL model and 56.58% (Fieger) vs. 43.42% (Engler) in the logit model.

Suppose that Engler knew Fieger's campaign strategies. How can Engler regain his advantage position? In Simulation 3, Engler proposes another platform with an increase in education while maintaining welfare, crime, and natural resources at current levels while Fieger's policy positions are same as Simulation 2. Now we see Engler has profile utility 7.5 while Fieger has profile utility 7.2 - Engler beats Fieger again by changing his policy positions. All three simulation models also show that Engler will win the election or has higher probability to win. If we continue the simulation exercise, this process simply generates a majority rule cycle. However, our purpose here is to demonstrate how we can apply conjoint analysis to the gubernatorial election rather than examine candidates' campaign strategies.

Targeting A Particular Group

In addition to simulating in the aggregate level (all voters), conjoint analysis can target a specific group for the purpose of segmentation. For instance, we could target different parties, religious groups, gender, generations, etc., to find out what their policy preferences are and which candidate they will support. In order to demonstrate this segmentation, let us consider a situation that assuming two parties will support only their own candidates. Then, in this case gaining the support from independent voters may be the major concern for both candidates. One way to attract independent voters is to propose a platform that meets the voters' needs. In order to do so, candidates can change their policy positions on four major spending categories. Table 7.2 reveals the simulation results for independent voters.

Table 7.2: Simulation Results for Gubernatorial Election (Independent Voters)

	Simulation 4		Simulation 5	
	Profile		Profile	
Attribute	Engler	Fieger	Engler	Fieger
Education	0%	5%	5%	5%
Welfare/Health	0%	5%	0%	5%
Crime/Public Safety	0%	0%	0%	0%
Natural Resources	0%	5%	0%	5%
	Simulation Results		Simulation Results	
Candidate	Engler	Fieger	Engler	Fieger
Profile Utility	7.0	7.4	7.9	7.4
Max Utility	0.00%	100%	100%	0.00%
BTL	48.69%	51.31%	51.79%	48.21%
Logit	40.74%	59.26%	63.34%	36.66%

Note: N=35

In Simulation 4, both candidates have their original policy positions. Fieger obviously has the advantage: all three simulation models show that Fieger will win the election. The profile utilities for Fieger and Engler are 7.4 and 7.0 respectively. Fieger has the advantage that 100 percent probability in maximum utility model and 51.31

percent in BTL and 59.26 percent in logit models. Under this condition, Engler becomes the underdog and has to propose another platform. After changing his policy positions by increasing education spending while maintaining welfare, crime, and natural resources at current levels (Simulation 5), now Engler surpasses Fieger with overall utility 7.9 (Engler) vs. 7.4 (Fieger). In this fashion, when candidates intend to gain support from a particular group, they can target this group by simulating different scenarios - changing their policy positions in this case (if the position changes do not violate their basic political beliefs or their party's ideological positions).

Welfare Policy Evaluation

One of the major achievements of John Engler is his welfare reform. In his website, John Engler's achievement in welfare reform are: "statewide, nearly half of all welfare recipients who can work are working and earning a paycheck, and more than 160,000 families have achieved independence from cash assistance. The state welfare caseload has declined for four years in a row, and since 1991, the number of recipients has fallen by more than 325,000 -- more than any other state."[2] One way to evaluate whether welfare reform is a success or not is to examine voters' satisfaction with the current welfare spending level. In order to so, we could compare three welfare spending levels with an increase in spending, a decrease in spending , and maintaining spending at the current level. As shown in Chapter 5, we could operationalize three welfare spending levels by using three profiles with an increase in welfare spending while maintaining other spending areas at status quo (0%, +5%, 0%, 0%), a decrease in welfare spending

[2] Michigan John Engler 1991-1997 Accomplishments.
http://www.migov.state.mi.us/issues/engleraccomp9197.html

while maintaining other spending areas at status quo (0%, -5%, 0%, 0%), and maintaining all spending areas across-the-board ((0%, 0%, 0%, 0%).

Looking at the Table 7.3, we found that Meridian Township voters are very satisfied with the current welfare spending level - with a profile utility 6.9 (Profile B) while opposing either an increase in welfare spending (Profile C's utility 6.3) or a decrease in welfare spending (Profile A's utility 6.5). The BTL and logit models also indicate the current welfare spending level (Profile B) has a higher probability to gain support. Based on the simulation, we can conclude that Engler's welfare reform is very successful, at least within Meridian Township.

Table 7.3: Simulation Results for Welfare Policy Evaluation

Attribute	Profile A	Profile B	Profile C
Education	0%	0%	0%
Welfare/Health	-5%	0%	5%
Crime/Public Safety	0%	0%	0%
Natural Resources	0%	0%	0%

Simulation Results

	Profile A	Profile B	Profile C
Profile Utility	6.5	6.9	6.3
Max Utility	0.00%	100%	0.00%
BTL	33.07%	35.07%	31.86%
Logit	30.58%	45.38%	24.04%

Now let us consider another situation. Due to the low unemployment rate and the economic resurgence in Michigan, Department of Budget and Management Office estimates that there is extra revenue at hand this year. Assuming Governor John Engler decides to distribute extra revenue to a single spending area in order to maximize voters' utilities. What is the possible way to distribute the money? How can governor John

Engler find out the voters' true preferences rather than their desires? Conjoint analysis can serve this purpose of finding the best combination of different spending levels by running several simulations. Since we know voters are satisfied with the current welfare spending level, we could increase the spending level either on education, crime, or natural resources. Specifically, we run three simulations with profiles that increasing education spending while keeping other spending areas at the status quo (+5%, 0%, 0%, 0%), increasing crime and public safety spending while keeping other spending areas at the status quo (0%, 0%, +5%, 0%), or increasing in natural resources spending while maintaining other spending areas at the status quo (0%, 0%, 0%, +5%).

The simulation results are captured in Table 7.4. Obviously, an increase in education spending (7.5 utility for Profile D) is the most favorite option for the Meridian Township voters, followed by increasing the crime spending (6.5 utility for Profile E), and an increase in natural resources spending (6.0 utility for Profile F). Thus, in order to maximize the voters' utilities, Governor Engler should increase education spending when there is an extra revenue at hand.

Table 7.4: Simulation Results for Maximizing Voters' Utilities

Attribute	Profile D	Profile E	Profile F
Education	5%	0%	0%
Welfare/Health	0%	0%	0%
Crime/Public Safety	0%	5%	0%
Natural Resources	0%	0%	5%

Simulation Results

	Profile D	Profile E	Profile F
Profile Utility	7.5	6.5	6.0
Max Utility	100%	0.00%	0.00%
BTL	37.53%	32.65%	29.81%
Logit	62.92%	23.67%	13.41%

Summary

In this chapter, we demonstrate the potential applicability of conjoint analysis by simulating the gubernatorial election in Michigan and evaluating the political support of welfare policy. We acknowledge that the voters do not vote for a particular candidate just according to their spending preferences on the four policy areas nor will candidates' campaign strategies solely rely on changing their policy positions. However, the simulation exercise in the current study at least can provide candidates more accurate information about the voters' true preferences rather than their wishes or desires. Similarly, evaluating spending levels may not be the only way to conduct a policy evaluation. Nevertheless, conjoint analysis provides a new way to evaluate voters' satisfaction with all kinds of policy issues. In conclusion, conjoint analysis has great potential applicability to the field of public policy and public opinion.

CHAPTER 8

RELIABILITY AND VALIDITY TESTS

In this chapter, we examine the reliability and validity of conjoint analysis and traditional survey questions. First, we point out the importance of reliability and validity checks and two major explanations of unstable public opinion. Focus is then given to definitions of reliability and validity by adopting a generalizability theory approach. After evaluating four types of reliabilities, attention is given to the measures we use: product moment correlation and Spearman rank order correlation. Lastly, we report the reliability results from conjoint analysis and traditional survey questions.

Why Reliability and Validity Checks Are so Important

Research on public opinion usually proceeds on the assumption that people have well organized attitudes on major political issues. Thus, surveys are passive instruments to measure these attitudes. The typical view is that when people say they favor X, they just recall a preexisting state of feeling favorably toward X. However, accumulated studies on public opinion has shown that people tend to have unstable answers to survey questions.

Random Response Variance

One explanation of the response instability is that people simply do not have meaningful attitudes. Converse (1964) finds significant intertemporal instability in response to a panel study that shows a sample of Americans the same set of issue questions in 1956, 1958, and 1960. His findings show that the correlation between individuals' positions on important policy issues in 1956 and 1960 is quite low. The highest correlation on all issues was slightly less than 0.5. The correlation on most issues fell in the range of 0.3 to 0.4. While the public is seemingly fickle in its issue positions, partisanship remains fairly stable, exhibiting an intertemporal correlation slightly above 0.7. Converse also shows that the correlation between an individual's responses in 1956 and 1958 and 1958 and 1960 is roughly equal to the correlation between 1956 and 1960, suggesting that fluctuations in individual responses are not the result of steady changes over time; rather, "the only change that occurred was random change" (p.243). Thus, he argues that "large portions of an electorate simply do not have meaningful beliefs, even on issues that have formed the basis for intense political controversy among the elite for substantial periods of time" (p. 245). However, facing the interviewers, respondents still politely choose an answer but in an almost apparently random fashion.

Achen (1975, 1983) challenges Converse's original findings by showing that measurement error in survey instruments can account for the over-time fluctuations in individual responses (see also Erikson 1979; Judd and Milbrun 1980). Achen argues that survey questions are stochastic or probabilistic measures of true attitudes. Each question contains a stochastic component, and temporal changes in individual responses are likely the product of random fluctuations in the error. Respondents are not to be indicted for displaying unstable attitudes or preferences: "measurement error is primarily a fault of the instruments, not of the respondents" (Achen 1975: 1229). Therefore, he concluded that even when people's "survey responses" fluctuate greatly, the members of the public have underlying "true attitudes" that are overwhelmingly stable.

Zaller (1992) challenges both approaches and suggests that people have conflicting opinions on most issues and these conflicting opinions might lead them to support the issue either way. He argues that most people simply do not possess preformed attitudes at the level of specificity demanded in surveys. Instead, they carry around in their heads a mix of only partially consistent ideas and considerations (ambivalence axiom). When questioned, they average across the opinions that happened to be salient at the moment of response (response axiom) and that have been recently thought about (accessibility axiom), and use them to decide among the options offered. Therefore, Zaller argues that responses from individuals do not reflect anything that can be described as either true attitudes or random guessing, rather, they simply reflect the thoughts that are most accessible in their memories at the moment of response.

Response Effects

Aside from the *random response variance* that has been attributed to measurement error, there exists *systematic variance* from artificial response effects. The literature on response effects has shown that survey questions not only measure public opinion but also shape and channel it by the way in which they frame issues, order the options, change the tone of wording, add more options (i.e., the "don't know" option), and other contexts of the question.[1] For instance, a 1976 NORC experiment reported by Turner and Krauss (1978) asked on one form about whether spending in each of 11 categories (e.g., space exploration, environment, defense) should be increased, followed by a single question on whether the federal income tax paid by the respondent was too high. On the other form the tax question preceded the spending items. In the latter case, 14% more respondents said taxes were too high than did so when the spending questions

---

[1] For a detailed discussion of response effects, see Schuman and Presser (1996), Questions and Answers in Attitude Surveys.

had been posed first, a difference significant beyond the .001 level of confidence. These findings have led researchers to a conclusion that respondents do not simply reveal preexisting attitudes on surveys; instead, to some extent, people are using the questionnaire to decide what their attitudes are.

If we know that people change their opinions over time and that their answers depend on seemingly trivial changes in the order of questions, then the notions of responsible electorate (Key 1966), reasoning voters (Popkin 1991), and rational publics (Page and Shapiro 1992) become difficult to support. On the other hand, if people are using the questionnaires to decide what their attitudes are, then can a different survey tool (conjoint analysis in this study) produce a more accurate information or valid answer? No matter which view we accept, examining the reliability and validity of survey instruments has become an important and necessary procedure in the study of public opinion.

Definitions of Reliability and Validity

Validity refers to the degree to which instruments truly measure the constructs which they are intended to measure. A necessary but not sufficient condition for validity of measures is that they are reliable. Reliability can be defined generally as the degree to which measures are free from error and therefore yield consistent results. The relationship between reliability and validity can be distinguished as follows:

$$X_0 = X_t + X_s + X_r$$

Where:

X_0 = an observed score,

X_t = the true score,

X_s = systematic error, and

X_r = random error.

A measure is said to be valid when $X_0 = X_t$. On the contrary, reliability refers to the consistency and reproducibility of the observed results. A measure is said to be reliable when there is no variation in X_0 due to chance or random errors, that is, $X_r = 0$. Therefore, reliability is a necessary, but not sufficient, condition for validity (Bateson et al. 1987).

To apply this equation to our additive, main-effects part-worth model, it can be written as follows (recall Equation 3.16):

$$U_p = C + \sum_i \sum_j w_{ij} X_{ij} + e_p \qquad (3.16)$$

Where:

U_p = the overall utility to measure of the p-th profile,

C = constant or intercept,

W_{ij} = the part-worth utility of i-th level j-th attribute estimated in the regression,

X_{ij} = dummy variable representing the presence or absence of attribute j, level i in alternative X, and

e_p = error term for profile P.

Theoretically, a measure is valid if the part-worths W_{ij} represents a true measure of the respondent's underlying part-worth for i-th level of the j-th attribute. Likewise, it would be reliable if the W_{ij} included no variation due to random error. However, this would not preclude variations due to systematic error.

Operationally, the distinction between validity and reliability blurs because it is impossible to measure X_t, the true measurement. Instead of using surrogates or approximations, the most widely used approach to this problem is the multitrait multimethod matrix suggested by Campbell and Fiske (1959). Traditionally, reliability check has been done on multi-item scales. There are three basic methods: test-retest, internal consistency, and alternative forms. The basic difference among the three methods lies in what the scale is to be correlated with to compute the reliability coefficient. In test-retest, the identical set of measures is applied to the same subjects at two different times. The two sets of obtained scores are then correlated. In internal consistency, a measurement scale is applied to subjects at one point in time; subsets of items within the scale are then correlated. In alternative forms, two similar sets of items are applied to the same subjects at two different times. Scale items on one form are designed to be similar (but not identical) to scale items on the other form. The resulting scores from the two administrations of the alternative forms are then correlated. The clear assumptions here are that there is a single construct called "reliability" and that there are various ways of measuring it (Peter 1979).

Reliability and Validity Checks: A Generalizability Approach

One major problem not explicitly addressed by traditional approaches to reliability is that measurement error can come from many sources, and each definition of error changes the meaning of the "reliability coefficient." For example, the components of variance which are "true" and "error" in computing a test-retest correlation are different from those for an internal consistency estimate. One approach to simultaneously analyzing multiple sources of variance in a measurement procedure has been addressed - generalizability theory.

Generalizability theory is based in part on the concept of sampling. However, the primary focus is not on the sampling of the people from populations of people, but rather on sampling conditions of measurement from universes of possible measurement conditions. Generalizability theorists argue that the traditional approaches to reliability do not address explicitly the sources of the measurement error being assessed. They argue there is no such thing as a reliability score; rather the score must specify the conditions of measurement over which reliability has been measured (Rentz 1987).

By adopting a generalizability perspective, a number of different reliability measures can be computed from conjoint analysis (Reibstein et al. 1988):

Reliability Over Time

Reliability over time is assessing by taking one measure, then repeating it with the same individuals and the same measurement devices at a subsequent time. The only aspect varied is the time of administration; everything else is held constant. The question is whether the W_{ij} at time t is the same as that at time t + x where t is some time lag.

Reliability Over Attribute Set

This approach examines whether the part-worths for a given attribute level for an individual depends on the other attributes or levels in the stimuli. Operationally, the tests involve assessing the stability of part-worths computed for attributes that are common when other attributes in the stimuli are varied.

Reliability Over Stimulus Set

The use of fractional factorial design in conjoint analysis reduces the problem of information overload for respondents. However, it also raises a new kind of reliability question: "reliability over stimulus set" addresses whether the derived part-worths are sensitive to the fractional factorial used for estimation. If the part-worths are influenced by the fractional factorial chosen, then using such designs would have to be reconsidered.

Reliability Over Data Collection Procedure

Conjoint analysis includes a large range of data collection procedures. In general, these variants can be categorized along three dimensions: type of data, data-gathering procedure, and dependent variable. Reliability over the data collection method procedure addresses whether the part-worths would be different if a different methodological variant were used.

Methods and Designs

This section looks at the reliability of conjoint estimates and traditional questions. In the student sample, we collected the follow-up survey in a one month period.[2] The traditional questions were identical in the first and second round of surveys and were presented to students with the format of unbounded resources, such as "Generally speaking, do you think the State of Michigan should: increase overall spending, decrease overall spending, or keep overall spending at current levels?" and "In the area of _________ (i.e., education), would you prefer to see the State of Michigan: increase spending in the area of education, decrease spending in the area of education, or maintain

[2] Because the summer session only last one and a half month, we could not collect the data in a two month period like we did for the voter sample.

education spending at current levels?" In conjoint questions, we asked students to rank order 18 profiles in both surveys but with a different fractional factorial design in the second round of survey.

In the voter sample, the data for reliability tests were collected in a two month period. In both first and second rounds of the surveys, traditional questions were presented by asking respondents expressing their own ideal spending levels (except first general question), such as "In the area of _________ (i.e., education), would you prefer to see the State of Michigan: increase spending in the area of education by _____%, decrease spending in the area of education by _____%, or maintain education spending at current levels?" In the conjoint questions, we asked respondents to rank order 10 profiles from the most preferred to the least preferred in the first round of survey. Later, in the follow-up surveys, we divided 88 subjects into two groups (44 subjects in each groups). In the first group, we asked respondents to rank order 10 profiles but with a different fractional factorial design while in the second group, we asked respondents to rate 12 profiles in a 0-100 points scale.

In conjoint analysis, we measure 3 types of reliability from above design: reliability over time,[3] reliability over stimulus set, and reliability over data collection method. The use of test-retest with different fractional factorial design is to measure reliability over time and reliability over stimulus set while the use of two data collection methods (ranking and rating) is to measure reliability over data collection method.

---

[3] Strictly speaking, we do not test reliability over time directly if we adopt the definition above. Since we test reliability in a one or two month period with different fractional factorial designs, this measure actually provides more rigorous reliability test and hence can be viewed as an indirect way to measure reliability over time.

Both OLS and LINMAP were used to estimate the part-worths from ranking and rating data. *Pearson product moment correlation* of the part-worth utilities from the two tasks provided a measure of reliability. In addition, the part-worths W_{ij} estimated from the first set of stimuli were used to predict utilities to the stimuli in the second setting U_{2p}; the *Spearman rank order correlation* was used to check the reliability from the two tasks. This procedure was then repeated with the part-worths estimated from the second set of stimuli, predicting utilities to the stimuli in the first setting U_{1p}. Parker and Srinivasan (1976) refers this method as *cross-validation.*

It should be noted that measures examining the utilities of the stimuli, U_p, are actually assessing the reliability of the model. Measures based on the W_{ij} are assessing the reliability of the part-worths. The W_{ij}'s have the utmost importance in policy decisions and are used as the basis of individual-level simulations. Reproducibility of the output of the model is a less stringent test because, as a result of the compensating nature of the model, it is possible to have stable U_p's with less stable W_{ij}'s (Bateson et al. 1987).

Results and Discussions

Reliability of Conjoint Analysis

Voter Sample: Ranking Data

Table 8.1 illustrates the part-worths correlations for ranking data in the voter sample between 2 time administrations. In the first column, we compare the product moment correlations between first and second round utilities for both OLS and LINMAP. The mean product-moment correlation for 32 subjects is 0.638 in OLS (median = 0.630) and 0.596 in LINMAP (median = 0.601).[4] Since LINMAP uses a different algorithm (linear programming) to calculate part-worths, the minor difference is under our expectation. In addition, using OLS (SPSS) we were able to use the part-worth utilities derived from a respondent's first round survey to predict his or her rankings of 10 profiles in the second round (cross-validation). Comparing the predicted values in the first round with actual responses in the second round produces a Spearman rank order correlation of 0.663 (Corr1). The procedure then was reversed (using second round utilities to predict first round results), the Spearman's rank order correlation was 0.594 (Corr2). The results from rank order correlations (grand mean = 0.629) are in accord with those from product moment correlations (grand mean = 0.617).

Compared to other conjoint reliability research, is conjoint reliability in this study good enough? Bateson et al. (1987) summarize previous reliability studies in conjoint analysis: they report that product moment correlations range from 0.49 to 0.92 (mean = 0.747) with time lag from 1 day to 2 month and sample size from 8 to 116, while Spearman rank order correlations are from 0.51 to 0.90 (mean = 0.717) with time lag from 1 day to 1 month and sample size from 41 to 68. However, all these studies evaluate the reliability in the areas of consumer behavior and marketing research rather than peoples' political attitudes. Therefore, both results in our study demonstrate a reasonably high reliability for a survey involving political issues.

[4] We also report the median because most studies on conjoint reliability report only the median.

Table 8.1: Reliability of Conjoint Estimates for Ranking Data (Voters)

	Pearson Product Moment Correlation		Spearman Rank Order Correlation	
ID	OLS	LINMAP	Corr1	Corr2
13	0.503	0.415	0.624	0.314
16	0.477	0.531	0.612	0.273
22	0.933	0.935	0.939	0.967
23	0.464	0.590	0.733	0.382
24	0.323	0.281	0.370	0.248
26	0.331	0.366	0.292	0.273
32	0.526	0.433	0.503	0.407
36	0.946	0.935	0.879	0.927
38	0.939	0.867	0.918	0.939
43	0.640	0.416	0.527	0.673
45	0.719	0.683	0.794	0.733
46	0.887	0.891	0.909	0.891
50	0.436	0.513	0.304	0.176
53	0.684	0.749	0.758	0.406
56	0.274	0.110	0.427	0.212
61	0.621	0.580	0.709	0.544
65	0.848	0.815	0.770	0.823
66	0.506	0.543	0.612	0.515
71	0.669	0.651	0.736	0.673
72	0.854	0.680	0.855	0.839
77	0.227	0.092	0.418	0.243
88	0.692	0.708	0.721	0.624
90	0.967	0.955	0.960	0.964
95	0.749	0.603	0.758	0.713
102	0.908	0.825	0.906	0.918
106	0.454	0.317	0.539	0.370
119	0.616	0.388	0.620	0.693
122	0.601	0.382	0.552	0.624
133	0.626	0.716	0.770	0.648
139	0.824	0.847	0.802	0.823
164	0.634	0.666	0.485	0.564
166	0.522	0.599	0.426	0.602
N=	32	32	32	32
Mean=	0.638	0.596	0.663	0.594
Median=	0.630	0.601	0.715	0.624
Grand Mean=	0.617		0.629	

Note: Corr1 measures the correlation of responses using their first round utilities to predict a subject's responses in the second round. Corr2 reverses this process, using the second round utilities to predict a subject's responses in the first round.

Voter Sample: Rating Data

Table 8.2 reveals the reliability tests of conjoint analysis for rating data in the voter sample. The mean product moment correlation for 27 subjects is 0.716 in OLS (median = 0.762) and 0.657 (median = 0.708) in LINMAP. Furthermore, the cross-validation from Spearman rank order correlation is 0.684 (Corr1) and 0.689 (Corr2) respectively. Both reliability scores show that rating data have even more reliability than ranking data. Since ratings usually produce tied scores and hence should yield lower reliability scores, the higher correlations in rating sample could be caused by specific sample characteristics. If we look at the reliability scores in Table 8.1, we see that there are a couple of outliers with very low reliability scores. Respondents 56 and 77, for instance, only had the product moment correlation 0.274 and 0.227 in OLS respectively. These low reliability scores indicate that both respondents answered the surveys carelessly or have very low involvement in the survey. In contrast, the rating sample shows a more homogenous product moment correlation and rank order correlation. Hence, rating sample exhibits higher reliability scores than ranking sample in this study.

As we have shown, the reliability test we used here also provides a way to identify careless respondents and hence we could remove them from the sample or give less weight in the simulation in order to obtain more accurate information or predict market share.[5]

[5] Another way to identify careless respondents is to use holdout cases. We will discuss this later.

Table 8.2: Reliability of Conjoint Estimates for Rating Data (Voters)

	Pearson Product Moment Correlation		Spearman Rank Order Correlation	
ID	OLS	LINMAP	Corr1	Corr2
6	0.777	0.776	0.673	0.637
34	0.813	0.787	0.818	0.827
111	0.767	0.739	0.818	0.809
138	0.822	0.612	0.772	0.779
141	0.737	0.663	0.782	0.685
159	0.630	0.767	0.745	0.699
171	0.779	0.764	0.818	0.839
173	0.747	0.541	0.648	0.602
175	0.761	0.775	0.758	0.790
176	0.926	0.805	0.915	0.916
178	0.762	0.754	0.729	0.774
179	0.846	0.890	0.879	0.895
180	0.729	0.628	0.564	0.594
181	0.320	0.390	0.358	0.378
185	0.823	0.800	0.782	0.737
188	0.550	0.550	0.430	0.382
192	0.547	0.392	0.382	0.559
219	0.635	0.429	0.697	0.708
224	0.829	0.857	0.867	0.860
225	0.973	0.898	0.967	0.947
227	0.834	0.583	0.636	0.674
230	0.401	0.433	0.321	0.322
232	0.597	0.462	0.564	0.611
239	0.892	0.774	0.891	0.895
242	0.666	0.708	0.503	0.559
243	0.398	0.318	0.345	0.322
244	0.781	0.657	0.806	0.799
N=	27	27	27	27
Mean=	0.716	0.657	0.684	0.689
Median=	0.762	0.708	0.745	0.708
Grand Mean=	0.687		0.686	

Student Sample: Ranking

The reliabilities of conjoint analysis for the student sample are captured in the Table 8.3. The reliability correlations from student sample indicate higher reliability scores than those from both voters' sample: the product moment correlation is 0.733 in OLS and 0.741 in LINMAP individually while the Spearman rank order correlation is

0.707 (Corr1) and 0.709 (Corr2). The higher reliability scores in the student sample are expected because we used 18 rather than 10 profiles and hence respondents have more chances to correct mistakes and make consistent responses. Nevertheless, we should notice that this is not to say that the more profiles, the better reliability; too many profiles, i.e., over 25 profiles, will cause the fatigue effects and the problem of information overload for respondents.

Table 8.3: Reliability of Conjoint Estimates for Ranking Data (Students)

	Pearson Product Moment Correlation		Spearman Rank Order Correlation	
ID	OLS	LINMAP	Corr1	Corr2
5	0.947	0.931	0.938	0.956
8	0.713	0.707	0.744	0.693
12	0.620	0.639	0.508	0.564
13	0.886	0.881	0.884	0.817
22	0.850	0.854	0.776	0.849
24	0.699	0.689	0.666	0.693
25	0.972	0.962	0.979	0.979
26	0.904	0.922	0.886	0.884
33	0.648	0.610	0.658	0.638
38	0.650	0.743	0.499	0.444
39	0.418	0.498	0.220	0.322
40	0.596	0.501	0.655	0.633
43	0.844	0.922	0.791	0.806
50	0.864	0.836	0.826	0.855
52	0.851	0.891	0.872	0.815
63	0.498	0.551	0.494	0.482
65	0.933	0.973	0.893	0.901
71	0.776	0.680	0.814	0.823
73	0.811	0.839	0.764	0.764
74	0.707	0.746	0.686	0.696
76	0.799	0.784	0.808	0.818
77	0.149	0.139	0.190	0.165
N=	22	22	22	22
Mean=	0.733	0.741	0.707	0.709
Median=	0.788	0.765	0.770	0.785
Grand Mean=	0.737		0.708	

Validity of Conjoint Analysis

Cross-Validation

Data used for testing reliability also provide a method of cross-validation. The overall profile utilities estimated from the first round are used to predict the preferences for the second round. The predicted overall utilities in the first round are then correlated with the actual overall utilities in the second round to obtain a measure of cross-validation (Corr1). The procedure is then reversed by predicting from round 2 to round 1, thus completing a double cross-validation (Corr2). We found that the cross-validity scores are satisfactory in both voter and student samples: in Table 8.1 and 8.2, we obtained the mean cross- validation scores (rank order correlation) 0.629 for ranking data and 0.686 for rating data; in Table 8.3, the mean cross-validation score (rank order correlation) is 0.708 for the student sample.

Holdout Profiles

Except for the measure of cross validation, the most widely used measure to test validity is to include holdout profiles in the design. Holdout profiles are additional profiles which the respondent evaluates beyond those generated by actual conjoint design. The value of using holdout profiles is twofold. First, in the field holdout profiles are useful in identifying inconsistent respondents whose responses are unlikely to correspond to their behavior. These respondents can then be removed from the sample or given less weight in the simulation. Second, it permits an immediate assessment of the relevance of conjoint analysis to choice. Where a conjoint model appears not to correspond to choice, it can be changed or improved. This permits the testing of different forms of conjoint analysis and leads to continuous improvement in its predictive validity (Johnson 1997;

Huber 1987). In the student sample, we include three holdout profiles in our design. In particular, we use the first 15 profiles (Profile 1-15) to predict last three profiles (Profile 16-18). As shown in Table 5.8 (Chapter 5), the Kendall's tau for 3 holdouts in the student sample is 1.00 with a confidence level 94%. The validity score from holdout profiles indicates that the model's predictive validity is satisfactory.

Besides, holdout profiles can be used to check validity for the specific group. For example, one of the major party differences for Democrats and Republicans is their attitude toward welfare spending; Democrats prefer an increase in welfare spending while Republicans prefer cuts in welfare spending. In order to check the validity of the two parties' responses on budget priorities, we could choose two profiles which represent the two parties' preferences and see how well the prediction is. In order to do so, we choose two profiles as holdouts from Profile 10 -18:

Democrat (Profile 17)

- Education (+5%), Welfare (+5%), Crime (0%), Natural Resources (0%)

Republican (Profile 10)

- Education (+5%), Welfare (-5%), Crime (-5%), Natural Resources (-5%)

We should point out that the holdout profiles may not completely reflect the two parties' policy positions because we need to choose profiles beyond the minimum combinations (9 in this case).[6] Therefore, we have to choose holdout profiles between Profile 10 to 18. In this case, we use the Profile 10 and 17 as holdout profiles. The

[6] In addition, due to the necessary and sufficient condition for orthogonal array, some profiles may be unrealistic. A necessary and sufficient condition (for orthogonal array) that the main effects of any two factors be uncorrelated is that each level of one factor occurs with each level of another factor with *proportional* frequencies

results show that both party groups predict holdout profiles well: the Kendall's tau for two holdouts are 1.0 in both Democrat and Republican sample.

Although holdouts require less additional data to test validity, it is a relatively weak measure of validity: the validity score in holdout profiles is obtained from the same data in a single survey design rather than from a second set of profiles or from respondent's actual behavior (e.g., consumer's buying behavior). Hence, in our study cross-validation is a better and more appropriate measure of validity because it uses different complete sets (fractional factorial design) to test validity.

Reliability of Traditional Survey Questions

Voter Sample: Ranking Data

Table 8.4 illustrates the reliability of traditional questions for ranking data in the voter sample. The first row "identical answer" indicates the situation when a respondent provides identical answers in the both time administrations. The second row "sign violation" shows the situation where a respondent provides answers with different directions in one of six following answers: from positive to zero, from positive to negative, from zero to positive, from zero to negative, from negative to zero, and from negative to positive. The row of "same direction" indicates when a respondent provide answers with same direction (i.e., from positive to positive) but with different intensity. The row "combined" shows the combined percentage between "identical answer" and "same direction" (we count the row "combined" as reliable responses). Then we calculate the grand mean by averaging the combined percentages for four spending areas and "identical answer" in overall spending.

Table 8.4: Reliability of Traditional Survey Questions for Ranking Data (Voters)

	Overall	Education	Welfare	Crime	Natural	Mean
Identical Answer	65.63%	43.75%	43.75%	43.75%	37.50%	46.88%
N=	21	14	14	14	12	
Sign Violation	34.38%	28.13%	28.13%	34.38%	25.00%	30.00%
N=	11	9	9	11	8	
Same Direction	N/A	28.13%	28.13%	21.88%	37.50%	28.91%
N=	N/A	9	9	7	12	
Combined	N/A	71.88%	71.88%	65.63%	75.00%	71.09%
N=	N/A	23	23	21	24	
Grand Mean	70.00%					
Total N=	32	32	32	32	32	32
		Sign Violation: Differences b/t 1st and 2nd Surveys				
Mean		5.01%	6.37%	7.07%	2.21%	
Range		8.00%	24.00%	8.70%	4.30%	
		Same Direction: Differences b/t 1st and 2nd Surveys				
Mean		14.36%	5.93%	10.43%	54.48%	
Range		94.00%	8.40%	36.00%	468.00%	

In the overall spending, the result reveals that respondents have the highest consistency in the category "identical answers": 65.63 percent (21 out of 32 cases) voters provide identical answers in ranking sample. In other spending areas, the category "identical answer" drops about 10 percent or more (i.e., identical answers in education, welfare, and crime are 43.75 percent and natural resources is 37.59 percent). On average, in ranking data 46.88 percent voters give identical answers. In addition, in all four areas (education, welfare, crime, and natural resources), on average 28.91 percent voters offer answers with the same directions. If we count the category "identical answer" and "same direction" as "reliable responses," we obtain a score (grand mean) 70 percent of consistent responses. However, we should notice that in general 30 percent of the respondents have sign violations. Among them, respondents gave crime issue most inconsistent answers with 34.88 percent sign violation.

Since we asked respondents to fill in their preferred spending levels in traditional questions, we could also examine the "intensity" of their policy preferences by looking at the preferred spending level differences between first and second round of surveys. For instance, a respondent's preferred spending levels in education may be +100% in the first round but only +5% in the second round. The preferred spending level difference is 95% in this case. In the row "mean difference," we averaged the preferred spending level differences between the first round and second round of surveys in the four policy areas for all respondents. In the education area, i.e., there are 9 cases in the "same direction" and their preferred spending level differences are: 3.2%, 8%, 95%, 5%, 5%, 1%, 3%, 1%, and 8% respectively. Therefore, the "mean difference" for education spending is 14.36%. In the "range difference," we calculated the range of preferred spending levels in each spending area. For instance, the "range difference" for education spending is 94% (95% - 1%). We repeat the procedure and then obtain the "mean difference" and "range difference" from both categories "sign violation" and "same direction."

In the category "sign violation," the inconsistency between the first and second round is reasonable: mean percentage differences are from a low 2.21 percent to a high 7.07 percent while the range differences are from a low 4.3 percent to a high 24 percent. On the contrary, in the category "same direction," we see significant preferred spending level differences between the first and second round of surveys: the mean differences are from a low 5.93 percent to a high 54.48 percent while range differences are from a low 8.4 percent to a high 468 percent.

Voter Sample: Rating Data

The reliability of traditional survey questions for rating data is captured in Table 8.5. First we notice that respondents provide highly consistent responses in the area of overall spending: 85.19 percent of voters offer identical answers. Nevertheless, we also see the percentages of identical answer drop approximately 40 percent in the individual spending areas. Among them, education issue has the lowest percentage of identical answer (25.93 percent) but also has lowest sign violation (25.93 percent) because 48.15 percent respondents provide answers in the same directions. Welfare and crime issues have highest inconsistent responses with 40.74 percent sign violation in both areas. The reliability scores in rating data are very close to those in the ranking data. On average, 69.63 percent respondents gave consistent responses in rating data (combining categories identical answer and same direction).

Table 8.5: Reliability of Traditional Survey Questions for Rating Data (Voters)

	Overall	Education	Welfare	Crime	Natural	Mean
Identical Answer	85.19%	25.93%	40.74%	44.44%	44.44%	48.15%
N=	23	7	11	12	12	
Sign Violation	14.81%	25.93%	40.74%	40.74%	29.63%	30.37%
N=	4	7	11	11	8	
Same Direction	N/A	48.15%	18.52%	14.81%	25.93%	26.85%
N=	N/A	13	5	4	7	
Combined	N/A	74.07%	59.26%	59.26%	70.37%	65.74%
N=	N/A	20	16	16	19	
Grand Mean	69.63%					
Total N=	27	27	27	27	27	27
		Sign Violation: Differences b/t 1st and 2nd Surveys				
Mean Difference		5.40%	4.69%	7.09%	9.31%	
Range Difference		6.20%	8.00%	29.00%	49.00%	
		Same Direction: Differences b/t 1st and 2nd Surveys				
Mean Difference		6.29%	5.50%	5.25%	23.29%	
Range Difference		23.00%	14.00%	7.00%	129.00%	

Looking at the preferred spending levels differences, we find the similar responses with ranking data: preferred spending level differences in the sign violation category are relatively lower than those from the same direction category with mean difference from 4.69 to 9.31 percent and range difference from 8 to 49 percent. In contrast, in the category "same direction," we see significant preferred spending level differences between the first and second round of surveys: the mean differences are from a low 5.25 percent to a high 23.29 percent while range differences are from a low 7 percent to a high 129 percent.

Why do people show small preferred spending level differences when their answers are inconsistent while fairly significant preferred spending level differences appear when their answers are consistent but with different intensity? One possible reason for this may be just due to chance because of the small case numbers in each category. However, we see the significant differences between the category "sign violation" and "same direction" in both ranking and rating data. Maybe there are some psychological explanations for this phenomena; however, we do not know what is the "correct" answer in the current study. No matter what reasons might be, further study is necessary to unveil this question that people's responses tend to show small differences when they provide inconsistent answers while significant differences when they provide consistent answers but with different intensity.

Student Sample: Ranking Data

Table 8.6 reveals the reliability of traditional survey questions for the student sample. Since we only asked students to choose spending levels from increase, decrease, or stay about the same in all spending areas, there are two categories in Table 8.5: identical answer and sign violation. Similar to voter's reliability in traditional survey questions, students offer highly consistent answers in the area of overall spending with identical answer 72.73 percent. Moreover, students offer most consistent responses in the area of education: 90.91 percent students (20 out of 22 cases) provide identical answers in the second round of survey. Unlike voters, students offer inconsistent answers in the areas of natural resources and welfare with sign violations of 45.45 percent and 31.82 percent respectively. In general, 72.73 percent of the students provide reliable responses with identical answers in five traditional questions.

Table 8.6: Reliability of Traditional Survey Questions for Ranking Data (Students)

	Overall	Education	Welfare	Crime	Natural	Mean
Identical Answer	72.73%	90.91%	68.18%	77.27%	54.55%	72.73%
N=	16	20	15	17	12	
Sign Violation	27.27%	9.09%	31.82%	22.73%	45.45%	27.27%
N=	6	2	7	5	10	
Total N=	22	22	22	22	22	22

Summary

Compared to previous conjoint reliability studies (with a mean product moment correlation = 0.747 and a rank order correlation = 0.717), our conjoint estimates have demonstrated satisfactory reliability scores in both voter and student samples. In ranking data of voter sample, we obtained 0.617 of mean product moment correlation and 0.629 of mean rank order correlation. In rating data of voter sample, conjoint estimates provided reliability scores 0.737 of mean product moment correlation and 0.708 mean rank order correlation. In the student sample, conjoint results showed 0.737 of product moment correlation and 0.708 of rank order correlation.

Since we used different measures to test reliability,[7] we could not compare the reliability between conjoint analysis and traditional survey. However, reliability of traditional surveys at least could be served as an reference of sample characteristics. From our traditional surveys questions, voters provided 70 percent "reliable" answers in ranking data while 69.93 percent "reliable" answers in rating data. Similarly, students

[7] Conjoint reliability is measured by product moment correlation and rank order correlation while reliability in traditional survey is measured by percentage of identical answers and same directions that respondents provided.

gave 72.73 percent "reliable" answers in ranking data. These results revealed a higher than average reliability.[8]

In sum, before we can be fully confident of conjoint methods, further study needs to be done to completely evaluate the reliability of conjoint estimates over time, over different stimulus sets, and over different data collection methods - especially for those political issues that tend to suffer greater numbers of stochastic shocks. Since this is the first reliability and validity tests in conjoint analysis focused on political attitude, the results we obtained here are encouraging and satisfactory.

[8] For example, Zaller (1992) showed that respondents gave 45% to 55% reliable answers both times on issues "American relations with Russia" and "Level of government services" (pp. 580-581).

CHAPTER 9

SUMMARY AND CONCLUSIONS

The contributions of this study are both methodological and substantial. First, we have examined people's budget priorities by asking them to make tradeoffs among four spending areas: education, welfare and health, crime and public safety, and natural resources and environmental control. Second, we have introduced a new survey tool - conjoint analysis, which yields more accurate information about people's preferences, to the studies of public opinion and policy analysis. Now let us consider three broad questions about this study. Why do we need a new survey tool? What do we learn from the conjoint study? What is the potential applicability of conjoint analysis?

The Need of New Survey Tool

The quality of democratic government depends on the responsiveness of policymakers to the preferences of the public. In order to do so, policymakers need to collect accurate and complete information about the public's true preferences. Studies from survey research have shown that there are significant inconsistencies between policymakers and public opinion with respect to public policy. Free and Cantril (1967), for example, found that the public exhibited a combination of philosophical opposition to welfare state policies and operational support for virtually the full range of welfare programs. They argued that this "schizoid" belief pattern made it extremely difficult for the government to engage in rational policy-making.

One of the reasons for this inconsistency is due to the limits of traditional surveys: when a traditional survey asks voters to answer a set of disaggregated questions unbounded by resource constraints, what it gets is the public's "wish list" rather than their true preferences. The "wish list" of result is one of the major reasons why we often see the "schizoid" belief pattern from normal surveys. In addition, some policy areas, such as environmental protection, when considered alone, it is difficult to oppose, but it tends to lose support rather quickly when compared with competing interests such as lost employment. Because in the political world policymakers always have to make tradeoffs among competing interests, introducing a new survey tool is extremely important. The new survey tool should define not only people's "desires" but, more importantly, their true "preferences." While the former can be addressed with a traditional wish list type of question on the survey, the latter can only be addressed with precisely designed tradeoff type questions.

We believe that conjoint analysis is such a tool and can offer us a way to collect more accurate information about voters' true preferences. Conjoint analysis can yield more accurate and complete information because it is different from traditional surveys in two aspects. First, conjoint analysis is a trade-off exercise in which respondents are forced to make preference decisions based on the relative importance of one policy or service over another. It forces the respondents to make more "real world" decisions by looking at complete descriptions of an array of policies or services and select the one(s) they would choose in an actual "buying" environment. Specifically, a conjoint task is valuable because it forces the respondent to evaluate competing interests, as between more government services and tax increases. People usually try to avoiding making such decisions by searching for unambiguous solutions involving dominance or following a relatively clear heuristic. However, people make tradeoffs when they must "buy a product" in the market or when they participated in the conjoint task.

Second, conjoint analysis employs a *decompositional* model which differs from almost all other multivariate methods in that it can be carried out at the level of the individual respondent. This means that the researcher can develop a separate model for predicting the preference of each respondent. As shown in Chapter 7, we could simulate all kinds of "what-if" scenarios to achieve a particular research purpose. In contrast, most other methods of analysis take measures from each respondent but can only perform an analysis using all respondents combined.

It is important to note that we are not saying that conjoint analysis is the only way to conduct survey or even desirable in every situation. But when dealing with policy issues that require voters to make tradeoffs among competing interests, conjoint analysis is definitely one of best candidates. Of course, since conjoint analysis is a more difficult and time-consuming task than traditional surveys, this additional information comes at price. Nonetheless, when evaluating voters' preferences on competing policies, usually you get what you pay for and thus we believe that the additional investment in conjoint analysis is worthwhile.

Empirical Findings by Using Conjoint Analysis

We have mentioned many advantages of conjoint analysis, but what do we learn from our conjoint study? First, let us look at what we learned from both the traditional survey and the conjoint study. In traditional surveys, we only know respondents favoring an increase or a decrease in overall spending, or respondents' preference orders among three spending choices - increase, decrease or stay about the same. For instance, we could acquire the individual preference for a particular area, such as a respondent preferring an increase in education spending. However, we do not know how respondents make tradeoffs among these spending areas and the weights for a particular spending level. In contrast, in conjoint analysis we obtained respondents' true preferences through the part-worth utilities for each attribute level, the relative importances of attributes, the profile overall utilities, and the possibility to simulate all kinds of "what-if" scenarios. We have argued that conjoint analysis produces more accurate and complete information about people's budget priorities. The "more accurate" is possible because we hold respondents accountable when they express their preferences (as discussed in Chapter 2). The "more complete" is possible because respondents are forced to evaluate all factors (which necessarily consists of a budget proposal) rather than the single dimensional consideration.

From our samples, we pointed out that both voters and students are satisfied with most of the current spending levels except education - they favored an increase in education spending. However, the most striking part was that we obtained inconsistent spending preferences between traditional survey and conjoint estimates. In traditional survey questions, for instance, we found that voters most preferred a decrease in overall spending, followed by an increase in overall spending, and maintaining overall spending at the current level. In contrast, in conjoint estimates (voters making tradeoffs among attributes), we showed that voters most favored maintaining overall spending at the current level, followed by an increase in overall spending, and a decrease in overall spending. These results revealed that there appeared significant inconsistencies between voters' answers based on single dimensional considerations and tradeoff exercises.

Similarly, in the area of natural resources, we found that in traditional survey questions voters most preferred an increase in natural resources spending, followed by maintaining natural resources spending at the current level, and a decrease in natural resources spending. In contrast, in conjoint estimates, voters most favored maintaining natural resources spending at the current level, followed by an increase in natural resources spending, and a decrease in natural resources spending. Similar inconsistent results also can be found in the student sample.

There is one important implication from above results. Most studies in public opinion and policy analysis use data obtained from traditional surveys to interpret, analyze, and evaluate all kinds of political issues and phenomena. If information obtained from traditional surveys are inaccurate or incomplete, all research results based on these surveys will be less persuasive and misleading. Therefore, if policymakers make policy decisions based on the traditional survey results, these policies cannot reflect people's true preferences and hence the quality of democracy is discounted. Indeed, some political analysts have used this contradictory survey data to blame the voters for their apparent inconsistency while others accused politicians who told voters that was possible to "have their cake and eat it too" (Free and Cantril 1967; Kuttner 1980; Sears and Citrin 1982).

Second, our analysis of group differences in conjoint analysis also supported some previous findings in the studies of public opinion. For example, we found that women were more concerned with compassion issues such as education and welfare than men; the generation New Deal 1 (the oldest cohort in this study) favored an increase in crime spending because of their vulnerability to street crime; Post-New Deal 2 (the youngest cohort in this study) preferred either natural resources spending increased or the status quo; Democrats favored an increase in education spending while Republicans were satisfied with current spending on education; Democrats preferred increasing spending on welfare spending while Republicans favored spending at the status quo but suffered a serious utility loss when welfare spending increases; people with children at home had a slightly higher utility increase with education spending increased.

We acknowledge that our sample size (N=128) is limited and hence these results cannot be generalized to make inferences about the distribution of attitudes in the U.S. Nevertheless, these findings are insightful because the results we obtained are based on the respondents' tradeoff exercise rather than on their single dimensional evaluation (unrealistic choices among incomplete sets of options). In other words, these group preferences in conjoint estimates can be viewed as their true preferences rather than their wishes or desires.

Third, we conducted the first reliability and validity tests in conjoint analysis which focused on political attitudes, and the results are encouraging. Although from our voter and student samples we cannot provide a definite answer about which survey tool is better, we know the conjoint methods yield answers at least as reliable with traditional survey questions. One thing should be mentioned: before we can fully confident of conjoint methods, further study needs to be done to completely evaluate the reliability of conjoint estimates over time, over different stimulus sets, and over different data collection methods - especially for those political issues that tend to suffer greater numbers of stochastic shocks.

Lastly, we have demonstrated the applicability of conjoint analysis by various simulations. Specifically, we showed that conjoint simulations can be served as predicting the gubernatorial election in Michigan, targeting a particular group (independent voters in this study), evaluate welfare policy, and distributing extra revenue to maximize voter's utility. These useful "what-if" simulations cannot be done through the use of traditional surveys. Hence, in the fields of public opinion and policy analysis, the potential for conjoint application is promising.

Potential Applicability of the Conjoint Analysis

As shown in Chapter 7, we have demonstrated the applicability of conjoint analysis in different ways. Beyond these examples, conjoint analysis has great potential to be applied in the areas of public opinion and policy analysis. One possible topic in the area of public opinion is the concepts of materialism and postmaterialism. Inglehart (1971) operationalized the concepts of materialist and postmaterialist by asking respondents to rank two most important goals from the following four goals faced in the nation: (1) maintaining order in the nation, (2) giving the people more say in important political decision, (3) fighting rising prices, and (4) protecting freedom of speech. Those respondents who choose (1) and (3) are considered materialists and those who choose (2) and (4) are considered postmaterialists. Respondents who select one materialist value and one postmaterialist value are classified as having mixed value priorities.

In order to apply conjoint methods to this topic, we could dichotomize each of the four questions to be either present or not. Since there are only 4 attributes with 2 levels in each attribute, we could use a full set of profiles ($2\times2\times2\times2 = 16$) to estimate the part-worth utilities without causing the fatigue effect. What we can get from this conjoint design are the responses in which respondents making tradeoff among these values. In addition, the conjoint method is better than ranking because it is less obtrusive and can avoid the problem of ranking.[1]

In addition to the application in public opinion, conjoint analysis provides a tremendous potential for conducting cost-benefit analysis for public policies, measuring people's willingness to pay for public goods, engaging in program evaluation, and exploring the decision-making process. Since policy-making and political decisions always involves trade-offs among competing interests, the potential applicability for conjoint analysis is encouraging and promising.

[1] Ranking forces respondents who may appreciate values equally to choose among them in their rank ordering. But even if people have such meaningful rankings, when standing before actual decisions they often balance and trade off values. For instance, while people may prefer freedom over equality, they may end up opting for a solution that trades off some freedom over equality. The ranking task does not catch this but conjoint analysis can. For a detailed discussion of the ranking problem, see Shamir and Shamir (1995).

APPENDIX

CONJOINT SURVEY FORMS FOR STUDENT AND VOTER

Subject Consent Information

We appreciate your willingness to test a new survey instrument designed to measure voter preferences. In order to ensure that you are fully informed and consent to participate in this effort, please read through the following statements:

- Please understand that you are in no way obligated to participate in this survey.
- You may discontinue your participation at any time without penalty.
- If you indicate your agreement at the end of this survey, you may be reconnected by a researcher for a one-time, follow-up interview. This follow-up interview is strictly optional and you may decline to participate at any time without penalty. If you are not selected for a follow-up interview or choose not to participate at that time, your involvement will be limited to this one survey session.
- All responses are held in strict confidence. Subjects will remain anonymous in any report of research findings. On request and within these restrictions results may be made available to subjects.
- If you have any questions or concerns about this survey, please contact:
  Professor Larry Heimann
  Department of Political Science
  Michigan State University
  Ph: 517-353-3281 Fax: 517-432-1091
  e-mail: heimann@pilot.msu.edu

You indicate your voluntary agreement to participate in this research by completing and returning this questionnaire.

* Note: Subject consent form was presented to respondents in all surveys.

Overview on the State of Michigan Budget

Education expenditures covers all state contributions for elementary, secondary, and higher education.

Health & Welfare expenditures covers state spending on social services, mental health, Medicaid, and other health expenses.

Crime & Public Safety expenditures covers state spending on police services, correctional facilities like prisons and jails, and other public safety expenses.

Natural Resources expenditures includes costs of natural resource management, environmental protection, and environmental clean-up costs.

Currently, Michigan expenditures are as follows[1]:

Education	$11.26 billion	38.2%
Health & Welfare	$9.83 billion	33.4%
Crime & Public Safety	$1.86 billion	6.3%
Natural Resources	$0.65 billion	2.2%
All other spending	$5.87 billion	19.9%
Total Spending	$29.47 billion	100.0%

Assume that "all other spending" is fixed at its current level. Any net cuts in spending will be either placed in a "rainy-day" account or used to cut some form of taxes in Michigan. Any net increase in spending will be funded by an increase in some form of taxation in Michigan. The exact form such cuts or increases would take have not yet been determined.

* Note: overview on the State of Michigan budget was presented to respondents in all surveys.

[1] Based on the actual spending in the 1996 fiscal year. Data came from the Department of Management and Budget.

Survey Form (Student)

General Questions

1. Generally speaking, do you think the State of Michigan should:
   - ☐ increase overall spending,
   - ☐ decrease overall spending,
   - ☐ or keep overall spending at current levels?

2. In the area of education, you would prefer to see the State of Michigan:
   - ☐ increase spending in the area of education,
   - ☐ decrease spending in the area of education,
   - ☐ or maintain education spending at current levels?

3. In the area of welfare and health, you would prefer to see the State of Michigan:
   - ☐ increase spending in the area of welfare and health,
   - ☐ decrease spending in the area of welfare and health,
   - ☐ or maintain welfare and health spending at current levels?

4. In the area of crime and public safety, would you prefer to see the State of Michigan:
   - ☐ increase spending in the area of crime and public safety,
   - ☐ decrease spending in the area of crime and public safety,
   - ☐ or maintain crime and public safety spending at current levels?

5. In the area of natural resources, would you prefer to see the State of Michigan:
   - ☐ increase spending in the area of natural resources,
   - ☐ decrease spending in the area of natural resources,
   - ☐ or maintain natural resources spending at current levels?

* Note: general questions were presented to both student surveys.

Survey Instructions

Below is a list of 18 gubernatorial candidates, A – J, and their positions on state budget expenditures in four policy areas. In order to evaluate these candidates and their positions, please follow these steps:

1. In the space on the left hand side of each candidate, please put a "V" down for each candidate you would consider voting for.

2. After reviewing all the candidates you find acceptable, place a "1" on the right-hand side of the candidate you most prefer. After that, rank order the remaining acceptable candidates.

3. After reviewing all the candidates you would not vote for, place an "18" on the right-hand side of the candidate you least prefer. After that, place a "17" next to the candidate who represents your second-least-preferred position. Continue this process until all 18 candidates are rank-ordered.

After you have finished this, please turn to the next page for a few brief demographic questions. All individual responses will be held in strict confidence. *Thank you very much for taking the time to fill this survey out.*

* Note: survey instructions were presented to respondents in all surveys, except the profiles for voter have only 10 rather than 18 and rating survey has different instructions.

Budgetary Proposals for Gubernatorial Candidates
(First Round for Student: Rank Order)

	__A__	__B__	__C__	__D__	__E__
Education	-5%	-5%	-5%	0%	0%
Health & Welfare	-5%	0%	5%	-5%	0%
Crime & Public Safety	-5%	0%	5%	0%	5%
Natural Resources	-5%	5%	0%	0%	-5%

	__F__	__G__	__H__	__I__	__J__
Education	0%	5%	5%	5%	5%
Health & Welfare	5%	-5%	0%	5%	-5%
Crime & Public Safety	-5%	5%	-5%	0%	-5%
Natural Resources	5%	5%	0%	-5%	-5%

	__K__	__L__	__M__	__N__	__O__
Education	-5%	0%	-5%	0%	5%
Health & Welfare	-5%	-5%	0%	0%	0%
Crime & Public Safety	5%	0%	0%	-5%	5%
Natural Resources	0%	5%	-5%	0%	5%

	__P__	__Q__	__R__
Education	0%	5%	-5%
Health & Welfare	5%	5%	5%
Crime & Public Safety	5%	0%	-5%
Natural Resources	-5%	0%	5%

Background Information (for statistical analysis purposes only)

Sex: ☐ Male ☐ Female

Year in which you were born: 19☐☐

Education:

☐ 8th grade or less ☐ Some high school
☐ High school graduate or GED ☐ Junior college/Technical school graduate
☐ Some college (1-3 yr.) ☐ College graduate (4 yr.)
☐ Some post-graduate study ☐ Graduate degree

Race: ☐ African-American or Black ☐ Asian or Pacific Islander
☐ Native American ☐ White or Caucasian
☐ Hispanic or Latin-American ☐ Other ________________

Religion:

☐ No religious group ☐ Roman or Orthodox Catholic
☐ Islamic ☐ Jewish
☐ Mainline Protestant ☐ Evangelical Protestant
☐ Buddhist ☐ Other non-christian

Political: Please place on an "X" on the location which best represents your party affiliation

x	x	x	x	x	x	x
Strong Democrat	Moderate Democrat	Weak Democrat	Independent	Weak Republican	Moderate Republican	Strong Republican

Are you currently married? ☐ Yes ☐ No

Do you have children living with you? ☐ Yes ☐ No
If so, what is the age of the oldest child? _________ Youngest child? _________

Do you own your current home? ☐ Yes ☐ No

In a couple of months we'd like to recontact some of the people we've surveyed for a short interview. If you would be willing to participate, please indicate in the space below.

I would be willing to be recontacted: ☐ Yes ☐ No
Please send me a Michigan Lottery ticket: ☐ Yes ☐ No

* Note: background information was presented to respondents in all first round of surveys.

Budgetary Proposals for Gubernatorial Candidates

(Second Round for Student: Rank Order)

	__A__	__B__	__C__	__D__	__E__
Education	-5%	-5%	-5%	-5%	0%
Health & Welfare	-5%	-5%	5%	0%	0%
Crime & Public Safety	-5%	-5%	5%	5%	5%
Natural Resources	-5%	0%	0%	-5%	-5%

	__F__	__G__	__H__	__I__	__J__
Education	0%	5%	0%	5%	-5%
Health & Welfare	0%	-5%	5%	5%	0%
Crime & Public Safety	0%	5%	-5%	0%	0%
Natural Resources	0%	5%	-5%	-5%	5%

	__K__	__L__	__M__	__N__	__O__
Education	0%	0%	5%	0%	5%
Health & Welfare	-5%	-5%	5%	5%	-5%
Crime & Public Safety	5%	0%	5%	-5%	0%
Natural Resources	5%	0%	0%	5%	-5%

	__P__	__Q__	__R__
Education	5%	5%	-5%
Health & Welfare	0%	0%	5%
Crime & Public Safety	-5%	-5%	0%
Natural Resources	0%	5%	5%

Survey Form (Voter)

General Questions

1. Generally speaking, do you think the State of Michigan should:
   - ☐ increase overall spending,
   - ☐ decrease overall spending,
   - ☐ or keep overall spending at current levels?

2. In the area of education, you would prefer to see the State of Michigan:
   - ☐ increase spending by _______ %,
   - ☐ decrease spending by _______ %,
   - ☐ or maintain education spending at current levels?

3. In the area of welfare and health, you would prefer to see the State of Michigan:
   - ☐ increase spending by _______ %,
   - ☐ decrease spending by _______ %,
   - ☐ or maintain welfare and health spending at current levels?

4. In the area of crime and public safety, would you prefer to see the State of Michigan:
   - ☐ increase spending by _______ %,
   - ☐ decrease spending by _______ %,
   - ☐ or maintain crime and public safety spending at current levels?

5. In the area of natural resources, would you prefer to see the State of Michigan:
   - ☐ increase spending by _______ %,
   - ☐ decrease spending by _______ %,
   - ☐ or maintain natural resources spending at current levels?

* Note: general questions were presented to both voter surveys.

Budgetary Proposals for Gubernatorial Candidates
(First Round for Voter: Rank Order)

	__A__	__B__	__C__	__D__	__E__
Education	-5%	-5%	-5%	0%	0%
Health & Welfare	-5%	0%	5%	-5%	0%
Crime & Public Safety	-5%	0%	5%	0%	5%
Natural Resources	-5%	5%	0%	0%	-5%

	__F__	__G__	__H__	__I__	__K__
Education	0%	5%	5%	5%	0%
Health & Welfare	5%	-5%	0%	5%	-5%
Crime & Public Safety	-5%	5%	-5%	0%	5%
Natural Resources	5%	5%	0%	-5%	-5%

Budgetary Proposals for Gubernatorial Candidates
(Second Round for Voter: Rank Order)

	__A__	__B__	__C__	__D__	__E__
Education	-5%	-5%	-5%	0%	0%
Health & Welfare	-5%	0%	5%	-5%	0%
Crime & Public Safety	-5%	0%	5%	0%	5%
Natural Resources	-5%	5%	0%	0%	-5%

	__F__	__G__	__H__	__I__	__K__
Education	0%	5%	5%	5%	0%
Health & Welfare	5%	-5%	0%	5%	-5%
Crime & Public Safety	-5%	5%	-5%	0%	5%
Natural Resources	5%	5%	0%	-5%	-5%

Survey Instructions (Second Round for Voter: Rating)

In this section, we are going to ask you to evaluate each of the 12 candidates using a 100 points scale, where 100 points means that this is your ideal candidate (on these four issues) and 0 points means that this candidate and the policy position he/she represents is a complete anathema to you. Please put your rating in the appropriate box. At the end of the
survey are two brief yes or no questions comparing this format with the earlier format in which you rank ordered the candidates.

Thank you very much for your time in complete this survey. Your participation is very important to the success of this project.

Budgetary Proposals for Gubernatorial Candidates

(Second Round for Voter: Rating)

	__A__	__B__	__C__	__D__	__E__	__F__
Education	-5%	5%	5%	0%	5%	0%
Health & Welfare	0%	-5%	0%	5%	0%	-5%
Crime & Public Safety	0%	5%	-5%	-5%	5%	0%
Natural Resources	5%	5%	0%	5%	0%	0%
Candidate Rating	pts	pts	pts	pts	pts	pts

	__G__	__H__	__I__	__J__	__K__	__L__
Education	0%	-5%	-5%	0%	5%	0%
Health & Welfare	-5%	-5%	5%	0%	5%	0%
Crime & Public Safety	5%	-5%	5%	5%	0%	0%
Natural Resources	-5%	-5%	0%	-5%	-5%	0%
Candidate Rating	pts	pts	pts	pts	pts	pts

Format Questions:

1. Compared with the first survey where you rank ordered 10 candidates, did you find this 100 point rating method easier to complete? Yes No

2. Compared with the first survey where you rank ordered 10 candidates, did you find this 100 point rating method took more or less time to complete? Yes No

BIBLIOGRAPHY

BIBLIOGRAPHY

Abramson, Paul R. 1983. *Political Attitudes in America: Formation and Change*. W.H. Freeman and Company.

-----. 1989. Generations and Political Change in the United States." *Research in Political Sociology* 4: 235-280.

-----, and Charles W. Ostrom, Jr. 1991. "Macropartisanship: An Empirical Reassessment." *American political Science Review* 85: 181-92.

-----, John Aldrich, and David Rohde. 1995. *Change and Continuity in the 1992 Elections*. Revised Edition, CQ Press.

Achen, Christopher. 1975. "Mass Political Attitudes and the Survey Response." *American Political Science Review*, 69: 1218-31.

-----. 1983. "Toward Theories of Data: The State of Political Methodology." In *Political Science: The State of the Discipline*, ed. Ada Finifter. Washington, DC: The American Political Science Association. pp. 69-94.

Addelman, Sidney. 1962. "Orthogonal Main-Effect Plans for Asymmetrical Factorial Experiments." *Technometrics* 4: 21-46.

Akaah, Ishmael and Pradeep K. Korgaonkar. 1983. "An Empirical Comparison of the Predictive Validity of Self-Explicated, Huber-Hybrid, Traditional Conjoint, and Hybrid Conjoint Models." *Journal of Marketing Research* 20: 187.

Anderson, Norman H. 1981. *Foundation Measurement of Information Integration Theory*. New York: Academic Press.

-----. 1982. *Methods of Information Integration Theory*. New York: Academic Press.

Axelrod, Robert. 1967. "The Structure of Public Opinion on Policy Issues." *Public Opinion Quarterly* 31: 51-60.

-----. 1972. "Where the Votes Come From: an Analysis of Election Coalitions." *American political Science Review* 66: 11-20.

Barron, Hutton F. 1977. "Axiomatic Conjoint Measurement." *Decision Science* 8: 548-559.

Bateson, John E, David Reibstein, and William Boulding. 1987. "Conjoint Analysis Reliability and Validity: A Framework for Future Research." *Review of Marketing.*

Beck, Paul Allen. 1974. "A Socialization Theory of Partisan Realignment." *In The Politics of Future Citizens*, ed. Richard Niemi, et al. San Francisco: Jossey-Bass.

Beckwith, N., and Lehamnn, D. 1975. "The Importance of Halo Effects in Multiattribute Attitude Models." *Journal of Marketing Research* 12: 265-275.

Bennett, Linda., and Stephen Earl Bennett. 1990. *Living with Leviathan: Americans Coming to Terms with Big Government.* Lawrence: University Press of Kansas.

Berelson, Bernard, Paul Lazarsfeld, and William McPhee. 19544. *Voting: A Study of Opinion Formation in a Presidential Campaign.* Chicago: University of Chicago Press.

Bojamic, David C and Roger J.Calantone. 1990. "A Contribution Approach to Price Bundling in Tourism." *Annual of Tourism Research, 17*: 528-540.

Brady, Henry and Paul Sniderman. 1985. "Attitude Attribution: A Group Basis for Political Reasoning." *American Political Science Review* 79: 1061-1078.

Braungart, Richard and Margaret Braungart. 1989. "Political Generations." *Research in Political Sociology* 4: 281-319.

Brittan, Samuel. 1975. "The Economic Contradictions of Democracy." *British Journal of Political Science* 5: 129-59.

Campbell, Augus, Philip E. Converse, Warren E. Miller, and Donald E. Strokes. 1960. *The American Voter.* University of Chicago Press.

Campbell, Donald R. and Donald W. Fiske. 1959. "Convergent and Discriminant Validation by the Multitrait-Multimethod Matrix." *Psychological Bulletin*, 56: 81-105.

Carmines, Edward G. and Harold W. Stanley. 1992. "The Transformation of the New Deal Party System: Social Groups, Political Ideology, and Changing Partisanship Among Northern Whites, 1972-1988." *Political Behavior* 14: 213-240.

Carroll, J. Douglas. 1969. "Categorical Conjoint Measurement." paper presented at Meeting of Mathematical Psychology, Ann Arbor, MI, August.

-----. 1973. "Models and Algorithms for Multidimensional Scaling, Conjoint Measurement, and Related Techniques." in *Multiattribute Decisions in Marketing*, P.E Green, Y. Wind, eds. Hindsdale, IL: Dryden Press, 335-37; 341-48.

-----. and P. E. Green. 1995. "Psychometric Methods in Marketing Research: Part I, Conjoint Analysis." *Journal of Marketing Research* 32: 385-91.

Cattin, , Philippe, and Dick Wittink. 1976. "A Monte Carlo Study of Metric and Nonmetric Estimation Methods for Multiattribute Models." Research Paper No. 341, Graduate School of Business, Standard University.

Chakraborty, Goutam, R. Ettenson, and G. Gaeth. 1994. "How Consumers Choose Health Insurance." *Journal of Health Care Marketing* 14: 21-33.

Citrin, Jack. 1979. "Do People Want Something for Nothing?" *National Tax Journal* 32: 113-29.

Conjoint Experiment. [Online] Available http://www.hwc.ca/datapcb/Communc/generic/09stud51.htm, March 10, 1997.

Conner, W. S. and M. Zelen. 1959. *Fractional Factorial Experimental Designs for Factors at three levels*. Applied Math Series S4, Washington, DC: National Bureau of Standard.

Converse, Philip E. 1964. "The Nature of Belief System in Mass Publics." In *Ideology and Discontent*, ed. David Apter. New York: Free Press.

-----. 1974. "Some Priority Variables in Comparative Electoral Research." In *Electoral Behavior*, ed. Richard Rose, 727-745. New York: Free Press.

Croker, Royce. 1981. "Federal government Spending and Public Opinion." *Public Budgeting & Finance* Autumn: 25-35.

Curry, Joseph. 1997a. *Conjoint Analysis: After the Basics*. [Online] Available: http:/www.swatooth.com/pages, May 24, 1997.

-----. 1997b. *Understanding Conjoint Analysis: Predicting Choice*. [Online] Available: http:/www.swatooth.com/pages, May 24, 1997.

Dalton, Russell. 1988. *Citizen Politics in Western Democracies*. NJ: Chatham House Publishers, Inc.

Department of Management and Budget of Michigan. 1997. *Executive Budget of Michigan, Fiscal Year 1998*.

Downs, Anthony. 1957. *An Economic Theory of Democracy*. New York: Harper.

Eismeier, Theodore J. 1982. "Public Preferences about Government Spending." *Political Behavior* 4: 133-45.

Erikson, Robert S. 1979. "The SRC Panel Data and Mass political Attitudes." *British Journal of Political Science* 9: 89-114.

-----, Gerald C. Wright, Jr, and John P. McIver. 1989. "Political Parties, Public Opinion, and State Policy in the United States." *American Political Science Review* 83: 728-50.

Free, Lloyd A., and Hadley Cantril. 1967. *The Political Beliefs of Americans: A Study of Public Opinion.* New York: Simon and Schuster.

Gillian, Carol. 1982. *In a Different Voice*. Cambridge, MA: Harvard University Press.

Gold, Steven D. 1995. *The Fiscal Crisis of the States: Lessons for the Future*. Georgetown University Press.

Goldberg, Lewis R. 1965. "Diagnostics vs. Diagnostic Signs: The Diagnosis of Psychosis vs. Neurosis from the MMPI." *Psychological Monographs: General and Applied* 79: 1-27.

Green, Paul E. 1974. "On the Design of Choice Experiments Involving Multifactor Alternatives." *Journal of Consumer Research* 1: 61-68.

-----, and V. Srinivasan. 1973. *Multiattribute Decisions in Marketing: A Measurement Approach*. The Dryden Press.

-----, and V. Srinivasan. 1978. "Conjoint Analysis in Consumer Research: Issues and Outlook." *Journal of Consumer Research* 5: 103-23.

-----. and V. Srinivasan. 1990. "Conjoint Analysis in Marketing: New Developments with Implications for Research and Practice." *Journal of Marketing* 54: 3-19.

-----. and Yoram Wind. 1975. "New Way to Measure Consumers' Judgments." *Harvard Business Review* July-August: 107-117.

-----, Carroll, J. Douglas, and Carmone, Frank J. 1978. *"Some New Types of Fractional Factorial Designs for Marketing Experiments*." in Research in Marketing Vol. 1, ed. J. N. Sheth, Greenwich, CT: JAI Press.

Hagerty, Michael R. 1986. "The Cost of Simplifying Preference Models." *Marketing Science* 5: 298-321.

Hahn, G.J. and S. S. Shapiro. 1966. *A Catalog and Computer Program for the Design and Analysis of Orthogonal symmetric and Asymmetric Fractional Factorial Experiments*. Technical Report 66-C 165. Schenectady, New York: General Electric Research and Development Center.

Hair, J. F., Anderson, R.E., Tatham, R.L. and Black, W. C. 1995. *Multivariate Data Analysis: with Readings, 4th Ed.* Englewood Cliffs, NJ: Prentice Hall.

Hansen, John M. 1998. "Individuals, Institutions, and Public Preferences over Public Finance." *American Political Science Review* 92: 513-31.

Hicks, Alexander M. and Duane H. Swank. 1992. "Politics, Institutions, and Welfare Spending in Industrialized Democracies." *American Political Science Review*, 86: 658-74.

Hill, Kim Quaile and Jan E. Leighley. 1992. "The Policy Consequences of Class Bias in State Electorates." *American Journal of Political Science* 36: 351-65.

-----, Jan E. Leighley, and Angela Hinton-Anderson. 1995a. "Lower Class Mobilization and Policy Linkage in the U.S. States." *American Journal of Political Science* 39: 75-86.

-----, and Angela Hinton-Anderson. 1995b. "Pathways of Representation: A Causal Analysis of Public Opinion-Policy Linkages." *American Journal of Political Science* 39: 924-35.

Hooley, Graham. 1981. "Modeling the Student University Choice Process." *European Research* 9: 158-1570.

Hu, Clark. Conjoint Analysis Library. [Online] Available http://www.nevada.edu/~huc/html/welcome.html. November 6, 1996.

Huber, Joel. 1987. "Conjoint Analysis: How We Got Here and Where We Are." *Sawtooth Software Conference Proceedings.*

Inglehart, Ronald. 1971. "The Silent Revolution in Europe: Intergenerational Change in Post-Industrial Societies." *American Political Science Review* 65: 991-1071.

Jain, Arun, Frankin Acito, Naresh Malhotra, and Vijay Mahajan. 1979. "A Comparison of the Internal Validity of Alternative Parameter Estimation Methods in Decompositional Multiattribute Preference Models." *Journal of Marketing Research* 16: 313-22.

Jacoby, William G. 1994. "Public Attitudes Toward Government Spending." *American Journal of Political Science* 38: 336-61.

Jelen, Tred G. and Clyde Wilcox. 1995. *Public Attitudes Toward Church and State*. M.E. Sharpe, Inc: New York.

Johnson, Richard. 1974. "Trade-off Analysis of Consumer Values." *Journal of Marketing Research* 11: 121-127.

-----. 1997. *Including Holdout Choice Tasks in Conjoint Studies*. [Online] Available: http:/www.swatooth.com/pages, May 24, 1997.

Johnson, R. Burke. 1995. "Estimation An Evaluation Utilization Model Using Conjoint Measurement and Analysis." *Evaluation Review*, 19, 313-338.

Judd, Charles and Michael Milburn. 1980. "The Structure of Attitude systems in the General Public: Comparison of Structural Equation Model." *American Sociological Review* 45: 627-43.

Kamlet, Mark S. and David C. Mowery. 1987. "Influences on Executive and Congressional Budgetary Priorities, 1955-1981." *American Political Science Review*, 81: 155-78.

Kamakura, Wagner and Rajendra K. Srivastava. 1984. "Predicting Choice Shares Under Conditions of Brand Interdependence." *Journal of Marketing Research* 21: 420-34.

Kennedy, Peter. 1992. *A Guide to Econometrics. Third Edition*. The MIT Press: Cambridge, Massachusetts.

Key, V.O. 1952. *Politics, Parties, and Pressure Groups*. New York: Knopf.

-----. 1959. "Secular Realignment and Party System." *Journal of Politics* 21: 198- 210.

-----. 1961. *Public Opinion and American Democracy*. New York: Knopf.

-----. 1966. *The Responsible Electorate: Rationality in Presidential Voting, 1936-60.* Cambridge, MA: Harvard University Press.

Kiewiet, D. Roderick, and Mathew D. McCubbins. 1985. "congressional Appropriations and the Electoral Connection." *The Journal of Politics* 47: 59-82.

Kingdon, John. 1995. *Agenda, Alternatives, and Public Policies*. 2nd Edition. Harper Collins College Publishers.

Krantz, David H. 1964. "Conjoint Measurement: The Luce-Tukey Axiomatization and Some Extensions." *Journal of Mathematical Psychology* 1: 248-77.

Kruskal, Joseph B. 1965. "Analysis of Factorial Experiments by Estimating Monotone Transformations of the Data." *Journal of the Royal Statistical Society, Series B,* 27: 251-63.

Kuttner, Robert. 1980. *Revolt of the Haves: Tax Rebellions and Hard Times.* New York: Simon and Schuster.

Lacy, Dean. 1998. "A Theory of Nonseparable Preferences in Survey Responses." Paper presented at the Annual Meeting of the Midwest political Science Association, Chicago, IL, April 23-25, 1998.

Ladd, Everett Carl, Mayilyn Potter, Linda Basilick, Sally Daniels, and Dana Suszkiw. 1979. "The Polls: Taxing and Spending." *Public Opinion Quarterly* 43: 126-35.

Lazarsfeld, Paul, Bernard Berelson, and Hazel Gaudet. 1944. *The People's Choice.* New York: Duell, Sloan and Pearce.

Leigh, Thomas W., D. MacKay, and J. Summers. 1984. "Reliability and Validity of Conjoint Analysis and Self-Explicated Weights: A Comparison." *Journal of Marketing Research* 21: 456-62.

Lerman, S.R. and J.J. Louviere. 1978. "On the Use of Direct Utility Assessment to Identify the Functional Form of Utility and Destination Choice Models." *Transportation Research Record* 673: 78-86.

LINMAP, 1989. *Conjoint LINMAP manual.* Bretton-Clark.

Louviere, Jordan J. 1988. *Analyzing Decision Making: Metric Conjoint Analysis.* Sage Publications.

Marchant-Shapiro, Theresa and Kelly D. Patterson. 1995. "Partisan Change in the Mountain West." *Political Behavior* 17: 359-378.

McClain, John and Rao, Vithala. 1974. "Trade-offs and Conflicts in Evaluation of Health systems Alternatives: A Methodology for Analysis." *Health Services Research,* 9: 35-52.

McLean, R. and V. Anderson. 1984. *Applied Factorial and Fractional Designs.* New York: Marcel Deckker.

Miller, Warren E. 1992. "Generational Changes and Party Identification." *Political Behavior* 14: 333-352.

-----, and Merrill Shanks. 1996. *The New American Voter*. Cambridge: Harvard University Press.

Modigliani, Andre and Franco Modigliani. 1987. "The Growth of the Federal Deficit and the Role of Public Attitudes." *Public Opinion Quarterly* 51: 459-80.

Monroe, Alan D. 1979. "Consistency Between Public Preferences and National Policy Decisions." *American Politics Quarterly* 7: 3-19.

-----. 1983. "American Party Platforms and Public Opinion." *American Journal of Political Science* 27: 27-42.

Niemi, Richard and Herbert Weisberg. 1993. *Classics in Voting Behavior*. Washington, D.C.: CQ Press.

Neslin, Scott A. 1981. "Linking Product Features to Perceptions: Self Stated versus Statistically Revealed Importance Weights." *Journal of Marketing Research* 18: 80-86.

Olshavsky, R., and Granbois D. 1979. Consumer Decision Making - Fact or Fiction?" *Journal of Consumer Research* 6: 93-100.

Page, Benjamin I and Robert Y. Shapiro. 1983. "Effect of Public Opinion on Policy." *American Political Science Review* 77: 175-190.

-----. and Robert Y. Shapiro. 1992. *The Rational Public: Fifty Years of Trends in Americans' Policy Preferences*. Chicago: University of Chicago press.

Parker, Barnett and V. Srinivasan. 1976. "A Consumer Preference Approach to the Planning of Rural Primary Health Care Facilities." *Operations Research* 24: 991-1025.

Pekelman, Dov and Subrata K. Sen. 1979. "Improving Prediction in Conjoint Measurement." *Journal of Marketing Research* 16: 211.

Peter, Paul J. 1979. "Reliability: A Review of Psychometric Basics and Recent Marketing Practices." *Journal of Marketing Research* 16: 6-17.

Peterson, Paul E. 1985. "The New Politics of Deficits." In *The New Directions in American Politics*, ed. John E. Chubb and Peterson, Paul E. Washington, DC: The Brookings Institution.

-----, and Mark Rom. 1989. "American Federalism, Welfare Policy, and Residential Choices." *American Political Science Review*, 83: 711-28.

Popkin, Samuel. 1991. *The Reasoning Voter*. Chicago: University of Chicago Press.

Ramsing, K and J. Wish. "What do Library Users Want? A Conjoint Measurement Technique May Yield the Answer," *Information Processing and Measurement* 18: 237-242.

Reibstein, David, J. Bateson, and W. Boulding. 1988. "Conjoint Analysis Reliability: Empirical Findings." *Marketing Science* 7: 271-86.

Rentz, Joseph O. "Generalizability Theory: A Comprehensive Method for Assessing and Improving the Dependability of Marketing Measures." *Journal of Marketing Research* 16: 19-28.

Rubin, Irene S. 1993. *The Politics of Public Budgeting. Second Edition*. Chatham House Publishers, Inc.

Sanders, Arthur. 1988. "Rationality, Self-Interest, and Public Attitudes on public Spending." *Social Science Quarterly* 69: 311-24.

Schick, Allen. 1990. *The Capacity to Budget*. The Urban Institute Press.

Schuman, Howard and Stanley presser. 1996. *Questions and Answers in Attitude Surveys*. Sage Publications.

Sears, David O., and Jack Citrin. 1982. *Tax Revolt: Something for Nothing in California*. Cambridge: Harvard University Press.

Shamir, Michal and Jacob Shamir. 1995. "Competing Values in Public Opinion: A Conjoint Analysis." *Political Behavior* 17:107-113.

Shapiro, Robert Y. and Harpreet Mahajan. 1986. "Gender Differences in Policy Preferences: A Summary of Trends from the 1960s to the 1980s." *Public Opinion Quarterly* 50: 42-61.

Shocker, Allan D. and V. Srinivasan. 1977. "LINMAP (Version II): A FORTRAN IV Computer Program for Analyzing Ordinal Preference (Dominance) Judgments Via Linear Programming Techniques for Conjoint Measurement." *Journal of Marketing Research* 14: 101-103.

Shukla, P.K. 1990. *Measuring Faculty Preferences for Collective Bargaining: A Conjoint Analysis with Extensions to Selected Illustrations in Educational Administration*, unpublished dissertation, UCLA.

-----, and James Bruno. 1991. "Use of Conjoint Analysis and Marketing Approaches in Education Survey." *Education* 112: 451-58.

Smith, Tom W. 1984. "The Polls: Gender and Attitudes Toward Violence." *Public Opinion Quarterly* 48: 384-96.

Sniderman, Paul M. 1993. "The New Look in Public Opinion Research." in Ada Finifter ed. *Political Science: The State of the Discipline II.* American Political Science Association.

Sparling, J. 1977. Conjoint Scaling as a Decision Aid in Curriculum Development. Paper presented at 1977 AERA National Conference.

SPSS. 1997. *SPSS Conjoint 8.0 Manual.* SPSS, Inc.

Stoper, Emily and Roberta A. Johnson. 1977. "The weaker Sex and the Better Half: The Idea of Women's Moral Superiority in the American Feminist Movement." *Polity* 10: 192-217.

Strauss, William and Neil Howe. 1991. *Generations: The History of American's Future, 1584 to 2069.* William Morrow and Company, Inc: New York.

Srinivasan, V. and Allan D. Shocker. 1973. "Linear Programming Techniques for Multidimensional Analysis of Preferences." *Psychometrika* 38: 337-69.

Stanley, Harold W. and Richard G. Niemi. 1991. "Partisanship and Group support, 1952-1988." *American Politics Quarterly* 19: 189-210.

Stimson, James A. 1991. *Public Opinion in America: Moods, Cycles, and Swings.* Boulder: Westview Press.

Sundquist, James L. 1983. *Dynamics of the party System: Alignment and Realignment of Political Parties in the United States*, Revised Version. Washington, DC: Brookings Institution.

The National Election Studies, Center for Political Studies, University of Michigan. The NES Guide to Public Opinion and Electoral Behavior. [Online] Available http://www.umich.edu/~nes/nesguide/nesguide.htm. Ann Arbor, MI: University of Michigan, Center for Political Studies [producer and distributor], 1995-1998.

Turner, C.F. and E. Kruss. 1978. "Fallible Indicators of the Subjective State of the Nation." *American Psychologist* 33: 456-470.

Tversky, Amos. 1967. "A General Theory of Polynomial Conjoint Measurement." *Journal of Mathematical Psychology* 4: 1-20.

Wald, Kenneth D. 1997. *Religion and Politics in the United States*. Third Edition. CQ Press: Washington, D.C.

Welch, Susan. 1985. "The 'More for Less' Paradox: Public Attitudes on Taxing and Spending." *Public Opinion Quarterly* 49: 310-16.

Whitmore, G.A. and Cavadias, G.S. 1974. "Experimental Determination of Community references for Water Quality-Cost Alternatives." *Decision Sciences* 5: 614-31.

Wildavsky, Aaron. 1988. *The New Politics of the Budgetary Process*. Scott, Foresman and Company.

Wilson, L.A. II. 1983. "Preference Revelation and Public Policy: Making Sense of Citizen Survey Data." *Public Administration Review* July/August: 335-342.

Wind, Yoram and Lawrence Spitz. 1976. "An Analytical Approach to Marketing Decisions in health Care Organizations." *Operations Research* 24: 973-90.

Wittink, Dick and Philippe Cattin. 1989. "Commercial Use of Conjoint Analysis: An Update." *Journal of Marketing* 53: 91-96.

-----, D, M. Vriens, and W. Burhenne. 1994. "Commercial Use of Conjoint Analysis in Europe: Results and Critical Reflections." *International Journal of Research in Marketing*, 11, 41-52.

Wlezien, Christopher. 1995. "The Public as Thermostat: Dynamic of Preferences for Spending." *American Journal of Political Science* 39: 981-1000.

Young, Forrest W. 1972. "A Model for Polynomial Conjoint Analysis Algorithms." in *Multidimensional Scaling: Theory and Applications in the Behavioral Sciences* Vol. 1, R. N. Shepard, A. K. Romney, and S. Nerlove, eds. New York: Academic Press: 69-104.

Zaller, John. 1992a. "A Simple Theory of the Survey Response: Answering Questions versus Revealing Preferences." *American Journal of Political Science*, 36, 579-616.

-----. 1992b. *The Nature and Origins of Mass Opinion*. New York: Cambridge University Press.

Zufryden, Fred. 1983. "Course Evaluation and Design Optimization: A Conjoint Analysis Based Application." *Interfaces* 13: 87-94.