To the Graduate Council:

I am submitting herewith a dissertation written by Jeffrey W. Trombly entitled "An Investigation Into the Use of Market Segmentation Analysis in Transportation Energy Planning." I have examined the final copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Civil Engineering.

Frederick J. Wegmann

Frederick J. Wegmann, Major Professor

We have read this dissertation
and recommend its acceptance:

M. L. Brogan

Asoo Chatterjee

Fumiko Samaji

Ernest R. Cadotte

Accepted for the Council:

Lew Minkel

Vice Provost
and Dean of The Graduate School

AN INVESTIGATION INTO THE USE OF MARKET SEGMENTATION ANALYSIS IN TRANSPORTATION ENERGY PLANNING

A Dissertation

Presented for the

Doctor of Philosophy

Degree

The University of Tennessee, Knoxville

Jeffrey W. Trombly

March 1985

DEDICATION

This research is dedicated to the memory of my mother, Agnes Trombly, who passed away during the course of this work.

ACKNOWLEDGMENTS

I am indebted to Professors Fumiko Samejima, Ernest Cadotte, Arun Chatterjee, and Michael Bronzini who served as my committee for this research. I would also like to thank Professor Frederick Wegmann for his service as Chairman of this committee. I appreciate their time, effort, and patience in assisting me in this effort.

I would also like to acknowledge the staff of the Data Analysis Bureau of the New York State Department of Transportation for providing the data base used in this study.

I want to thank the Department of Civil Engineering and the Transportation Center at The University of Tennessee for the generous financial support they provided throughout the course of this research. In addition I am indebted to my associates at the Transportation Center for helping me to prepare this manuscript. Whatever errors remain in this work, however, are solely my responsibility.

ABSTRACT

The purpose of this research is to explore the application of market segmentation analysis in transportation energy planning. The study builds on the concepts of market segmentation developed in the marketing literature to suggest a strategy of segmentation analysis for use in transportation planning.

The results of two statewide telephone surveys conducted in 1979 and 1980 for the New York State Department of Transportation are used as the data base for identifying target segments. Subjects in these surveys were asked to indicate which of 18 energy conservation actions had been implemented over the prior year to conserve gasoline. These responses serve as the basis for segmentation.

Two alternative methods are pursued in identifying target market segments for purposes of energy transportation planning. The first approach consists of the application of conventional multivariate analysis procedures. The second method exploits the principles of latent trait or modern test theory. The results of the conventional analysis suggest that the data collected can be divided into eight segments. The results of the application of latent trait theory identify three market segments.

The results of this study may be used to design future responses to energy shortages in addition to suggesting strategies to be pursued in measuring consumer response.

TABLE OF CONTENTS

CHAPTER		PAGE
1.	INTRODUCTION	1
	Overview and Purpose	1
	Governmental Roles and Responsibilities in Transportation Energy Planning	4
	Summary	7
	Contributions to Theory Development, Empirical Findings, and Methodology	8
	Organization of Report	9
2.	TRANSPORTATION ENERGY AND MARKET SEGMENTATION ANALYSIS	11
	Introduction	11
	Transportation Energy	12
	Energy Crisis of 1973-74 and 1979	19
	Market Segmentation Analysis	30
	Market Segmentation Theory	31
	Market Segmentation in Transportation Planning	38
	Summary and Implications for this Research	42
3.	RESEARCH DESIGN AND DATA SOURCES	44
	Introduction	44
	Background	44
	Survey Design	49
	Survey Results	51
	Research Procedures	65
4.	RESULTS OF CONVENTIONAL MARKET SEGMENTATION ANALYSIS	67
	Introduction	67
	Interaction Correlations	70
	Phi Coefficient	70
	Results for 1979 Survey	72
	Results for 1980 Survey	74
	Summary	77
	Factor Analysis of Interaction Phi Coefficients	77
	Factor Analysis Theory	79
	1979 Factor Analysis Results	84
	Overview	84
	Eigenvalue Configuration	85
	Rotation of Common Factor Loadings	88
	1980 Factor Analysis Results	90
	Overview	90

CHAPTER PAGE

4. (Continued)

Eigenvalue Configuration 90
Rotation of Common Factor Loadings 94
Comparison of 1979 and 1980 Results 94
Estimation of Factor Scores 106
Summary of Results 112
Cluster Analysis of Factor Scores 112
Results of Cluster Analysis 115
Characteristics of Cluster Groups 118
Group Profiles 132
1979 Segments 132
1980 Segments 149
Summary 153

5. RESULTS OF MARKET SEGMENTATION ANALYSIS USING LATENT TRAIT THEORY 155

Introduction 155
Background 158
Theory 159
Scoring Strategies 160
Dimensionality and Local Independence 160
Mathematical Models 161
Parameter Estimation 166
Research Strategy 177
Tetrachoric Method Results 179
Tetrachoric Correlation Matrices 180
1979 Principal Factor Analysis Results 186
1980 Principal Factor Analysis Results 195
Calibration of Item Parameters 200
Item Characteristic Functions 209
Maximum Likelihood Estimates of Individual Parameters 210
Validation of Item Characteristic Functions 214
Test Information Function 215
Logist 5 Method Results 221
Market Segmentation Analysis 228
Group Profiles 251
1979 Segments 251
1980 Segments 254
Summary 256

6. SUMMARY OF RESULTS AND POTENTIAL APPLICATIONS 258

Introduction 258
Results Using Conventional Market Segmentation Analysis 258
Summary of Results 258

CHAPTER PAGE

6. (Continued)

Applications 266
Recommendations 270
Advantages and Disadvantages 270
Results Using Latent Trait Theory 271
Applications 272
Recommendations 274
Advantages and Disadvantages 276
Summary 276

7. SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS 278

LIST OF REFERENCES 284

APPENDICES 291

APPENDIX A. QUESTIONNAIRES 292

APPENDIX B. ITEM VALIDATION FOR NORMAL OGIVE MODEL . . . 298

APPENDIX C. ITEM VALIDATION FOR LOGISTIC MODEL 311

VITA 324

LIST OF TABLES

TABLE		PAGE
2.1	Average Annual Percentage Change for Selected National Aggregate Transportation Energy Related Statistics, 1973-82	14
2.2	New York State Estimated Percentage Total Energy Consumption for Personal Travel by Purpose and Geographic Area, 1980	17
2.3	Annual Average Vehicle Trips and Vehicle Miles per Household and Average Trip Length by Purpose, 1977 National Personal Transportation Study	18
2.4	Summary of Major Events Surrounding 1973-74 and 1979 Energy Crisis in New York State	20
2.5	Impact of 1973-74 and 1979 Energy Crisis in New York State	22
3.1	Summary of Survey Results for Energy Conservation Actions	52
3.2	Summary of Survey Results for Selected Independent Variables	54
3.3	Correlation Matrix of Selected Independent Variables, 1979	58
3.4	Correlation Matrix of Selected Independent Variables, 1980	59
3.5	Results of New York State Surveys	61
3.6	Mean and Variance of Total Summated Score, 1979 and 1980	62
3.7	Correlation of Total Summated Score with Selected Independent Variables, 1979 and 1980	63
4.1	Matrix of Interaction Phi Coefficients, 1979	73
4.2	Action Clusters Suggested by Interpretation of Interaction Phi Coefficient Matrix, 1979	75
4.3	Matrix of Interaction Phi Coefficients, 1980	76
4.4	Action Clusters Suggested by Interpretation of Interaction Phi Coefficient Matrix, 1980	78

TABLE		PAGE
4.5	Squared Multiple Correlations of Each Action with all Other Actions	86
4.6	Eigenvalue Configuration Resulting from Principal Factor Solution Based on Eight Common Factors Developed from Matrix of Phi Coefficients, 1979	87
4.7	Unrotated Factor Loadings Developed from Matrix of Interaction Phi Coefficients, 1979	89
4.8	Rotated Factor Loadings Developed from Matrix of Interaction Phi Coefficients, 1979	91
4.9	Factor Correlations for Rotated Factors, 1979	92
4.10	Correlations of Variables with Factors, 1979	93
4.11	Eigenvalue Configuration Resulting from Principal Factor Solution Based on Eight Common Factors Developed from Matrix of Phi Coefficients, 1980	95
4.12	Unrotated Factor Loadings Developed from Matrix of Interaction Phi Coefficients, 1980	96
4.13	Rotated Factor Loadings Developed from Matrix of Interaction Phi Coefficients, 1980	97
4.14	Factor Correlations for Rotated Factors, 1980	98
4.15	Correlations of Variables with Factors, 1980	99
4.16	Action Items Comprising Factors	100
4.17	Root Mean Square Deviation Calculated from Rotated Loading Matrices, 1979, 1980	104
4.18	Degree of Factorial Similarity Calculated from Rotated Loading Matrices, 1979, 1980	105
4.19	Factor Score Coefficients, 1979	107
4.20	Factor Score Coefficients, 1980	108
4.21	Mean and Standard Deviation of Factor Scores, 1979, 1980	109
4.22	Correlations Between Factor Scores and Composite Variables, 1979	110

TABLE		PAGE
4.23	Correlations Between Factor Scores and Composite Variables, 1980	111
4.24	Cluster Group Size and Average Distance to Cluster Centers for Eight Group Solution, 1979 and 1980	119
4.25	Distance Between Cluster Centers, 1979	120
4.26	Distance Between Cluster Centers, 1980	121
4.27	Cluster Centers, 1979	122
4.28	Cluster Standard Deviations, 1979	123
4.29	Cluster Centers, 1980	124
4.30	Cluster Standard Deviations, 1980	125
4.31	Proportion of Groups Endorsing Each Item, 1979	127
4.32	Proportion of Groups Endorsing Each Item, 1980	128
4.33	Standardized Residuals for 1979 Survey	130
4.34	Standardized Residuals for 1980 Survey	131
4.35	Market Segment Profiles Using Conventional Analysis, 1979	133
4.36	Market Segment Profiles Using Conventional Analysis, 1980	136
4.37	Standardized Residuals Developed from Segment Profiles Using Conventional Analysis, 1979	139
4.38	Standardized Residuals Developed from Segment Profiles Using Conventional Analysis, 1980	142
5.1	Tetrachoric Correlation Coefficients, 1979	181
5.2	Tetrachoric Correlation Coefficients, 1980	182
5.3	Action Clusters Suggested by Interpretation of Interaction Tetrachoric Correlation Coefficients, 1979	183
5.4	Action Clusters Suggested by Interpretation of Interaction Tetrachoric Correlation Coefficients, 1980	184

TABLE		PAGE
5.5	Items Included in Selected Factor Analysis Results	187
5.6	Incremental Proportion of Common Variance Explained for Selected Factor Solutions, 1979	189
5.7	Squared Multiple Correlations of Each Action with all Other Actions	190
5.8	Eigenvalue Configuration Resulting from Principal Factor Solution Based on Eight Common Factors and 18 Items Developed from Matrix of Tetrachoric Correlation Coefficients, 1979	191
5.9	1979 Unrotated Factor Loadings Developed from Matrix of Interaction Tetrachoric Correlation Coefficients	194
5.10	Incremental Proportion of Common Variance Explained for Selected Factor Solutions, 1980	197
5.11	Eigenvalue Configuration Resulting from Principal Factor Solution Based on Eight Common Factors and 18 Items Developed from Matrix of Tetrachoric Correlation Coefficients, 1980	198
5.12	1980 Unrotated Factor Loadings Developed from Matrix of Interaction Tetrachoric Correlation Coefficients . . .	199
5.13	Item Parameters for the Two Parameter Normal Ogive Model Derived from 1979 Data	201
5.14	Item Parameters for the Two Parameter Normal Ogive Model Derived from 1980 Data	202
5.15	Results of Fitting Regression of 1980 Item Parameters on 1979 Item Parameters	205
5.16	Item Parameters Developed from Logist 5, 1979	222
5.17	Item Parameters Developed from Logist 5, 1980	223
5.18	Results of Fitting Regression of Maximum Likelihood Estimates Using the Normal Ogive Model on Maximum Likelihood Estimates Using the Logistic Model	224
5.19	Mean and Standard Deviation of Latent Trait Estimates and Total Scores for Selected Groups, 1979	229

TABLE		PAGE
5.20	Mean and Standard Deviation of Latent Trait Estimates and Total Scores for Selected Groups, 1980	230
5.21	Observed Proportions and Estimated Probabilities Using the Normal Ogive Model for Action Item Endorsement, 1979	232
5.22	Observed Proportions and Estimated Probabilities Using the Normal Ogive Model for Action Item Endorsement, 1980	234
5.23	Correlation of Maximum Likelihood Estimates of Latent Trait with Selected Socioeconomic Variables	237
5.24	Market Segment Profiles Using Latent Trait Theory, 1979	239
5.25	Market Segment Profiles Using Latent Trait Theory, 1980	242
5.26	Standardized Residuals Developed from Segment Profiles Using Latent Trait Theory, 1979	245
5.27	Standardized Residuals Developed from Segment Profiles Using Latent Trait Theory, 1980	248
6.1	Comparison of Segments Developed from Conventional Market Segmentation Analysis Based on Average of Conservation Action Item Endorsement for Eight Common Factors	259
6.2	Characteristics of Overall Segments A to D Resulting from Conventional Segment Analysis	261
6.3	Characteristics of Overall Segments E to H Resulting from Conventional Segment Analysis	262
6.4	Summary of Action Endorsement	267
6.5	Recommended Programs Developed for Conventional Market Segmentation Analysis	269
6.6	General Market Segment Profiles and Recommended Policies and Programs Using Latent Trait Theory	273

LIST OF FIGURES

FIGURE		PAGE
1.1	Map of New York State	3
2.1	Paradigm of Market Segmentation Analysis	34
4.1	Within Group Variation for Selected Cluster Solutions, 1979	116
4.2	Within Group Variation for Selected Cluster Solutions, 1980	117
5.1	Meaning of Item Parameters	165
5.2	Example Item Operating Characteristics	167
5.3	Plot of Item Discrimination Parameters	206
5.4	Plot of Item Difficulty Parameters	207
5.5	Frequency Distribution for 1979 Maximum Likelihood Estimates	212
5.6	Frequency Distribution for 1980 Maximum Likelihood Estimates	213
5.7	1979 Test Information Function	216
5.8	1980 Test Information Function	217
5.9	Selected Item Characteristic Functions Using the Normal Ogive Model, 1979	218
5.10	Selected Item Characteristic Functions Using the Two Parameter Normal Ogive Model, 1980	220
5.11	Selected Item Characteristic Functions Using the Two Parameter Logistic Model, 1979	226
5.12	Selected Item Characteristic Functions Using the Two Parameter Logistic Model, 1980	227
6.1	Summary of Profile Characteristics for Overall Market Segments A, D, F, G	264
6.2	Summary of Profile Characteristics for Overall Market Segments D, C, G	265

CHAPTER 1

INTRODUCTION

Overview and Purpose

Experience gained over the last decade has clearly demonstrated the dependence of the United States transportation sector on petroleum based energy sources. First in 1973 and later in 1979 the nation was faced with rapidly rising fuel prices and spot supply shortages. The policies developed and implemented during this period placed an emphasis on maintaining mobility levels and assuring equitable distribution of scarce energy resources in a framework of limited direct governmental intervention.

Transportation energy policy has been formulated to address the mobility needs of a broad spectrum of consumers. Limited attempts have been made to seek out selected target groups in the general population to assess the ability and willingness of the consumer to participate in energy related conservation programs.

The application of a market segmentation approach can aid in the design of specific programs or projects which are consistent with behavioral characteristics of a particular group as they relate to energy consumption. The overall concept is to implement a range of programs or projects tailored to selected groups so that the aggregate impact results in an improvement in attainment of energy related goals.

The purpose of this research is to investigate the application of market segmentation analysis in transportation energy planning. Data

concerning transportation conservation actions are used as a basis to develop market segments. In two statewide telephone surveys, conducted in New York State (Figure 1.1) during October 1979 and October 1980, over 1,500 adults were asked to indicate which of 18 conservation actions had been implemented over the prior year to conserve gasoline.

Prior analysis of this data conducted by Hartgen and Neveu (1) and Neveu (2) concentrated on estimating the energy savings resulting from these actions. In general, the main concern of this previous analysis centered on the mainfest action item scores. In this research, these scores are hypothesized to represent indicants of an underlying construct. This construct is hypothesized to be related to willingness and ability of a subject to conserve gasoline.

Two alternative analysis procedures are pursued. The first applies the procedures associated with conventional market segmentation analysis. This includes the application of principal factor analysis, cluster analysis, and multidimensional contingency table analysis. The goal of this analysis is to identify groups of individuals in the sample survey data sharing a common response to energy shortages. These groups are then examined in more detail to suggest feasible energy conservation actions which will appeal to segments.

The second analysis method consists of the application of modern test or latent trait theory. In this analysis, detailed functional relationships between the hypothesized latent trait and the probability of an action endorsement are developed. This information is used to

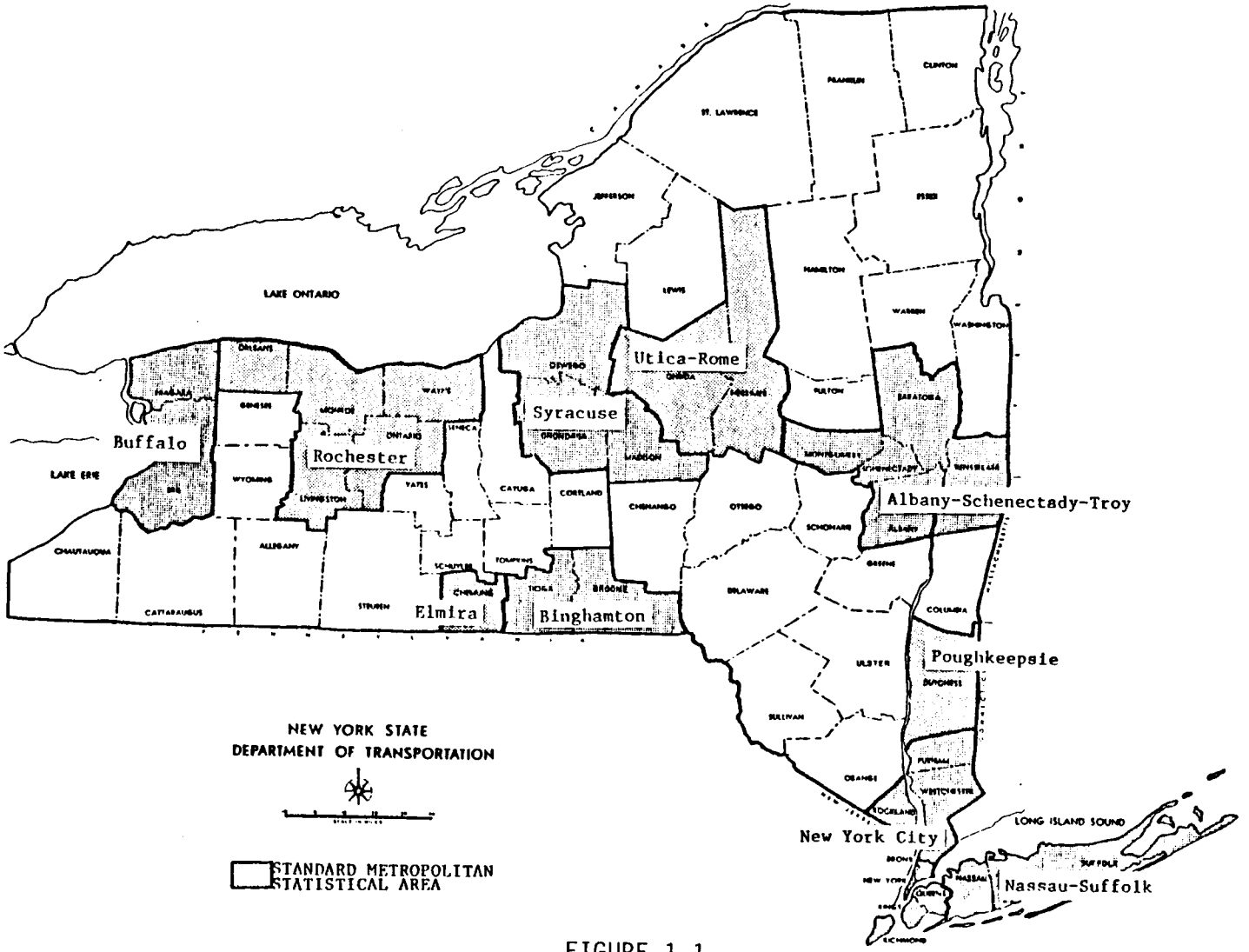


FIGURE 1.1

MAP OF NEW YORK STATE

estimate the position of an individual in the latent space describing the willingness and ability of a subject to conserve gasoline.

Two alternative analysis procedures are applied in order to compare and contrast conventional analysis with latent trait theory. Each has advantages and disadvantages depending on the purpose of the analysis. For example, conventional procedures have been applied more often and are better understood by a wider audience of researchers than latent trait theory. However, latent trait theory holds a great deal of promise because it provides much more information concerning items comprising a measurement instrument. The differences in these analysis approaches are outlined in more detail in later portions of this documentation.

Once market segments have been identified, it is necessary to match each segment with a set of feasible conservation programs. In this research, two scenarios are investigated. First, in 1979 gasoline prices increased rapidly and shortfalls developed. This led to one set of consumer responses. Later, in 1980, prices stabilized and shortages disappeared. This suggests a slightly different consumer response. Market segmentation results serve as a foundation, therefore, to develop policies in future energy shortages consisting of similar background conditions.

Governmental Roles and Responsibilities in Transportation Energy Planning

Over the last decade, transportation energy policymaking has undergone some dramatic shifts in emphasis. This can be attributed, in

part, to the development of price increases and supply shortfalls resulting from the 1973 oil embargo and the 1979 Iranian Revolution. The United States was forced to confront its growing dependence on imported oil which became a major threat to economic, political, and social institutions of the nation. As a result of these events, a strong federal level role was defined.

Of the many legislative initiatives of the period, two are of particular interest to transportation planners. The Energy Policy Conservation Act of 1975 (3) included two major provisions affecting transportation: (1) establishment of automotive fuel economy standards and (2) requirements concerning state energy conservation plans. These plans had to include an assessment of measures to save 5% of the projected 1980 energy demand. Recommended measures included: (1) establishment of a 55 mph speed limit; (2) permissive right turn on red after stopping at signalized intersections; and (3) promotion of carpool and public transit programs.

The 1979 Emergency Energy Conservation Act (4) represented a second legislative initiative of interest to transportation planners. This Act encouraged the development of energy contingency planning and sought to improve the institutional arrangements associated with meeting any challenges presented by future energy emergencies. The 1979 Act authorized the President to require the development of statewide plans to deal with energy shortfalls. Most plans developed in response to this requirement consisted of elements suggested by the legislation. These included: public information programs, minimum fuel purchase

restrictions, employer based travel measures, and speed limit enforcement.

State and local governments have also developed strategies to cope with the emerging energy problems. In most cases, these efforts have followed federal initiatives. Generally speaking, these include proposals to better coordinate current and proposed transportation programs and projects with energy conservation goals. Many states developed cabinet level energy offices. Some local areas have been very active in designating a larger role for transportation energy planning in the decision making process.

The history of the last decade has included a large-scale response on the part of government, the private sector, households, and individuals to rapidly increasing energy prices and supply shortfalls. A major emphasis has been placed on decreasing the vulnerability of society to imported crude oil. In the transportation sector this has meant the development of various legislative and administrative initiatives to help contribute to the reduction of energy requirements. In particular, a strong federal level role was developed which encouraged other levels of government to participate in a transportation energy planning process.

Travel demand is a reflection of a complex set of private decision making. It is important, therefore, to consider the role of the individual consumer in developing transportation energy policy. Experience has shown that consumers respond, in various ways, to price increases

and supply shortfalls. In fact, it can be suggested that the large portion of energy savings realized over the past decade were not an exclusive result of any particular energy policy, but a reflection of a set of measures adopted by consumers to meet individual requirements.

This research is designed to understand how consumers respond to changing energy price and availability by investigating the combinations of actions consumers implement to meet evolving energy trends. This effort will help to establish a better understanding of the socio-economic and travel characteristics of groups exhibiting similar adaptive behavior so that a more effective policy can be designed and implemented.

Summary

This research is designed to investigate several questions. In brief the following statements summarize the fundamental propositions guiding the overall effort:

1. Societal level transportation energy conservation behavior is a reflection of the joint effects of individual and household level decision making; transportation system characteristics; and public policy.
2. Current transportation energy conservation behavior is dependent on current household level constraints; current transportation system constraints; and previous conservation behavior and policies.
3. The transportation energy conservation behavior is a reflection of a latent unidimensional characteristic which can be measured.

4. The position of a household in this latent space is dependent on the characteristics of the household, the transportation system, and transportation policy.
5. Information about consumer energy conservation adaptation can be used to segment the entire travel market.

Contributions to Theory Development, Empirical Findings, and Methodology

The proposed research hypothesizes the existence of an adaptation process which is implemented by individuals and households to cope with changing situational constraints (e.g., household characteristics, transport system characteristics). It is suggested that this process consists, in part, of defining combinations of activity patterns, travel patterns, and investment strategies to meet the activity needs of the individual and the household. This construct is used to segment samples of New York State households so that energy policy and programs can be more effectively designed.

The analysis of the data proposed for use in this study will help to provide additional insight into the nature of specific market segments in the general population. The empirical findings of the study will contribute to a better understanding of the problems and opportunities of market segmentation in addition to identifying a specific framework for describing selected segments. This information can be used in formulating future energy policies and programs tailored to the needs of specific groups.

Finally, the study will help to assess the use of specific market segmentation analysis methods. In particular, the application of

latent trait theory in analysis of the data will help to suggest the problems and opportunities of this approach in the field of transportation planning. Part of the conclusions to the study will include several proposals concerning the use of latent trait theory in future analysis efforts.

Organization of Report

This report consists of seven chapters. The second chapter contains an overview of the concepts of market segmentation analysis. Included in this discussion is an outline of market segmentation as practiced in private sector marketing. In addition, literature concerning the use of market segmentation analysis in the transportation planning field are reviewed.

The third chapter presents an overview of the research design and data sources used in this investigation. As stated earlier, two statewide telephone surveys conducted in New York State serve as the data base for this investigation.

The fourth chapter presents the results of market segmentation analysis using conventional analysis procedures. The results suggest a response to energy price and availability based on the travel requirements of a household and the ability of a household to maintain desired mobility levels.

The fifth chapter presents the results of the use of latent trait theory in market segmentation analysis. These results are contrasted with the earlier efforts using conventional analysis procedures. In

general, latent trait theory represents an interesting, new perspective in developing and interpreting travel related information.

Chapter 6 presents a discussion of the potential applications of this study and suggests a strategy by which overall results may be compared. It is shown that a fairly consistent set of segments can be identified.

Chapter 7 summarizes the results of this research effort concerning the use of market segmentation analysis in transportation energy planning.

CHAPTER 2

TRANSPORTATION ENERGY AND MARKET SEGMENTATION ANALYSIS

Introduction

The purpose of this chapter is twofold. First, it presents an overview of transportation energy supply and demand in the United States. In particular, it is designed to focus on the 1973-82 time frame to review past events and current trends in energy price, supply, and consumption. Literature reviewed includes both aggregate statistics associated with transportation energy and studies of household or consumer level reviews of the impact of the 1973-74 and 1979 shortages.

A second goal of this chapter is to present a review of the concepts of market segmentation analysis. This includes an overview of the concept as developed for use in private sector marketing practice. In addition, examples of the use of market segmentation analysis in transportation planning are reviewed.

The goal of this discussion is to place the current research in the context of both past and current transportation planning efforts. This serves as a foundation to review the research presented in this study and to suggest future research requirements.

Transportation Energy

This section presents an overview of current trends in transportation energy supply and demand. This effort focuses on national level aggregate statistics to summarize the role of transportation energy in meeting the travel needs of the United States. Below are presented summary data concerning the nature of transportation energy supply including fuel types and energy price trends. Also included are data concerning the demand for transportation expressed in vehicle miles of travel (VMT), or the product of miles traveled mutiplied by total vehicles. Most of the information contained in this section was developed from data contained in the Transportation Energy Data Book (5).

Transportation, by all modes, in the United States represents over 26% of the total energy demand. The transportation sector, however, is almost entirely dependent on petroleum as a fuel source and consumes 61% of the total petroleum used by this nation. Nearly 95% of the fuel consumed in transportation is derived from petroleum.

Highway modes including motorcycles, automobiles, buses, and trucks consumed 74% of the total transportation energy in 1981 (5, p. 19). The remaining proportion was consumed by non-highway modes including air, water, pipeline, and rail.

Gasoline meets 64% of the transportation energy needs with distillate fuel oil, liquified petroleum gas, jet fuel, residual fuel oil, natural gas, and electricity representing the remaining major sources of fuel. Highway modes use over 98% of the total gasoline consumed and 67% of the diesel fuel.

Table 2.1 presents information concerning the average annual percentage change in various transportation energy supply and demand variables. This table shows the gradual decrease in domestic petroleum consumption for the period 1978-82. It also suggests that the proportion of imported oil has decreased with domestic production increasing +4.4% for the period 1980-82.

Total transportation energy consumption decreased at an average rate of -0.9% for the period 1973-75 reflecting the impact of the 1973 oil shortage. Consumption increased, however, during the 1975-78 years at an average annual rate of +4.3% before declining in 1978-80 at an average rate of -4.9% as a result of the 1979 shortage. Over the time period 1980-82 total consumption declined at an average annual rate of -1.6%.

Automobiles and trucks combined used 95% of the total gasoline consumed in 1981. The aggregate demand for gasoline continued to decline an average annual percentage of -3.2% for the period 1978-82 after increasing at a rate of +2.2% per year for 1973-78. However, for the period 1973-74 gasoline consumption declined at a rate of -4.3% reflecting the impact of the 1973 Arab Oil Embargo.

It appears that the 1979 energy crisis has a longer term impact on the fuel consumption than the previous 1973 crisis. This can be explained, in part, by the relative price increases for petroleum during these periods. From 1975-78 the average retail price per gallon for leaded regular gasoline increased +3.4% per year in nominal dollars, but it decreased an average of -2.6% in 1982 dollars. This is

TABLE 2.1

AVERAGE ANNUAL PERCENTAGE CHANGE FOR SELECTED NATIONAL AGGREGATE TRANSPORTATION ENERGY RELATED STATISTICS, 1973-82

	1973-1975	1975-1978	1978-1980	1980-1982
Domestic Petroleum Consumption	-2.9	+4.9	-5.0	-5.3
Percentage Imported	+1.4	+5.4	-4.2	-9.2
Percentage Used in Transport	+2.4	-0.4	-0.6	+4.4
Transportation Energy Use[d]	-0.9	+4.3	-4.9	-1.6
Price, Domestic Crude (1982$)	+38.4	+1.5	+27.6	+7.6
Price, Imported Crude (1982$)	a	-4.3	+40.0	-7.3
Price, Unleaded Regular Gasoline (1982$)	a	-2.0	+25.1	-5.1
Price, Leaded Gasoline (1982$)	a	-2.6	+26.6	+7.9
Gasoline Consumption	-0.5	+5.2	-0.9	+2.3
Total Highway Vehicle Travel[b]	+0.8	+5.2	-0.9	+2.3
Total Air Vehicle Travel [b]	-4.3	+5.2	+4.2	-2.3
Total Rail Vehicle Travel[b]	-6.2	+1.4	-0.0	-9.3
Total Urban Transit Bus Travel[c]	+5.9	+2.5	+4.2	a
Total Intercity Bus Travel[c]	-3.9	-1.2	+2.2	a
Total Highway Energy Use[d]	-1.3	+4.8	-3.9	a
Total Air Energy Use[d]	-3.7	+5.5	+2.4	a
Total Rail Energy Use[d]	-4.3	+1.8	+3.2	a

[a]Not available.

[b]Vehicle miles of travel.

[c]Passengers.

[d]Total British Thermal Units.

Source: M. C. Holcomb and S. Koshy, Transportation Energy Data Book: Edition 7, (Oak Ridge, TN: Oak Ridge National Laboratory, 1984).

contrasted with the +13.1% per gallon nominal price increase for the 1978-82 period, or 10% increase in 1982 dollars, for this time span.

Over the past few years, however, the composite price per barrel of crude oil has declined resulting in a moderation of fuel prices. These developments are, in part, a reflection of the conservation actions implemented by the public in response to price increases and supply shortfalls.

In particular, a switch to more fuel efficient automobiles has raised the composite average over the road fuel economy for the nation's fleet of personal motor vehicles. This has resulted in a more efficient use of gasoline by the nation's personal vehicle fleet and maintained desired levels of personal mobility.

The total highway VMT declined -1.75% for the period 1973-74, then increased at an annual rate of +5.3% for 1975-78. VMT declined -1.23% in 1978-79 and -0.54% in 1979-80 reflecting the impact of the 1979 energy crisis. However, for the period 1980-82, VMT again increased at an average annual rate of +2.3%. These readjustments in total travel demand have occurred in spite of higher fuel prices and along with a slow increase in the aggregate gasoline consumption.

The slide in public transportation ridership was halted by recent energy events. In particular, after falling since World War II at a constant rate, ridership increased approximately +3.8% per year for the period 1973-78. In addition, this trend increased to a rate of +4.1% per year for 1978-80. These increases were, in large part, the result of service expansion, but also can be attributed to the mode shift

behavior resulting from energy price increases and supply shortfalls, especially during crisis periods.

A study conducted by the New York State Department of Transportation (NYSDOT) provides some insight into the relationship between energy consumption and household travel (6). Table 2.2 outlines the forecasted 1980 percentage consumption of total automobile fuel consumption by trip purpose. Examination of this data suggests that 48% of total consumption is related to home or work related travel. However, shopping (9.7%), personal business (9.7%), and serving a passenger (9.1%) trips also consume a large portion of transportation energy. In order to maximize effectiveness in conserving automobile fuel, therefore, this table suggests that efforts concentrating on work or home based trips must be pursued as well as those relating to shopping and personal business travel.

Data collected by the National Personal Transportation Study supports this strategy (7). Information from this study, shown in Table 2.3, reports trip rates and vehicle miles for various trip purposes. Of the total vehicle miles per household per year (12,035 veh. mi.), approximately 60% (7,293 veh. mi.) include work, home, or personal business related travel.

Fuel consumption can be reduced by two basic approaches. First, a new technology can be introduced to improve fuel efficiency of vehicles. On the other hand, behavioral changes can be implemented to reduce travel rates, and trip lengths or shift travel modes. It will be seen that most consumers attempt to exploit technological improvement and also implement marginal behavioral changes such as reducing vacation or

TABLE 2.2

NEW YORK STATE ESTIMATED PERCENTAGE TOTAL ENERGY CONSUMPTION FOR PERSONAL TRAVEL BY PURPOSE AND GEOGRAPHIC AREA, 1980

Destination Purpose	Area			
	Rural	Small Urban	Urban	Total
Home	0.409	0.404	0.403	0.405
Work	0.194	0.196	0.170	0.183
Shop	0.096	0.092	0.100	0.097
School	0.010	0.011	0.012	0.011
Soc. Rec.	0.079	0.073	0.099	0.088
Eat Meal	0.020	0.030	0.023	0.024
Per. - Bus.	0.108	0.099	0.088	0.097
Ser. Pass.	0.080	0.091	0.097	0.091
Change Mode	-	0.0004	0.001	0.0004
Ride	0.003	0.005	0.006	0.004
Total	1.000	1.000	1.000	1.000
Share	0.308	0.210	0.482	1.000

Source: N. S. Erlbaum, G. S. Cohen, and D. T. Hartgen, "Automotive Energy Forecasts: Impacts of Carpooling, Trip Chaining, and Auto Ownership," Preliminary Research Report 134, (Albany, NY: New York State Department of Transportation, December 1972), p. 10.

TABLE 2.3

ANNUAL AVERAGE VEHICLE TRIPS AND VEHICLE MILES PER HOUSEHOLD AND AVERAGE TRIP LENGTH BY PURPOSE, 1977 NATIONAL PERSONAL TRANSPORTATION STUDY

Trip Purpose	Trip Rate Yr	Trip Rate Da	Vehicle Mile Yr	Vehicle Mile Da	Average Trip Length (Miles)
Earning a Living					
Home to Work	402	1.1	3,657	10.0	9.2
Work Related	74	0.2	879	2.4	11.9
Family and Personal Business					
Shopping	245	0.7	1,204	3.3	4.9
Medical	19	0.1	205	0.6	10.8
Other	202	0.5	1,348	3.7	6.7
Civic - Religious	92	0.3	566	1.6	6.1
Educational					
Social Recreation					
Visit Friends and Relatives	121	0.3	1,360	3.7	11.2
Pleasure Driving	6	0.0	96	0.3	15.7
Vacations	1	0.0	72	0.2	95.4
Other	150	0.4	1,360	3.7	9.1
Other and Unknown					
	131	0.4	1,288	3.5	9.8

Source: D. Klinger, J. R. Kuzayak, and S. Liss, "Household Travel," 1977 National Personal Transportation Study Report 9, (Washington, D.C.: Federal Highway Administration, July 1982).

recreation travel, and combining or chaining trips to reduce trip rates. In general, changes occur that minimize disruptions, but may not include the most effective set of behavioral modifications.

The overview thus far has pointed out the dependence on petroleum based products. It has been suggested that travel demand adjustments have been made over the last decade because of energy price increases and shortages. It has been shown that weekday home or work related travel represents a large portion of personal transportation energy demand. Finally, it has been suggested that vehicle technology improvements and marginal changes in travel behavior have been used to reduce household transportation fuel consumption.

The following discussion outlines studies of consumer based responses to the 1973 and 1979 energy shortages. Only limited studies of this nature are available, however, they do represent a basis on which to develop a perspective of these periods. In the first part of this section, however, an overview of the events surrounding these shortages is presented with special emphasis on actions taken in New York State to conserve fuel.

Energy Crisis of 1973-74 and 1979

Table 2.4 outlines the major events surrounding the 1973-74 and 1979 energy shortages and associated policy actions implemented in New York State in order to maintain order and contribute to equitable distribution of fuel supplies (8, 9) In addition, Table 2.5 summarizes the major impact of these events on aggregate level energy and travel related variables in New York State.

TABLE 2.4

SUMMARY OF MAJOR EVENTS SURROUNDING 1973-74 AND
1979 ENERGY CRISIS IN NEW YORK STATE

Date	General Situation	Cost of Gasoline Cents per Gallon	Actions in New York State
November 1973 December 1973	Arab Oil Embargo Vehicle Queues at gasoline stations	$0.40	Sunday station closings State allocation system begun
January 1974	Continuing queues, reduced hours at	$0.45	55 MPH speed limit established
February 1974	Continuing queues, reduced hours at stations	$0.45	Daylight savings time, odd/even, one-half tank
April 1974	Arab Oil Embargo ends	$0.50	Resume odd/even
December 1978	Iranian Oil exports stop	$0.70	-
March 1979	Iranian Oil production resumes	$0.80	-
June 1979	Station closings in New York City	$0.87	Odd/even, one-half tank in NYC, state set aside allocation made available for weekends.

TABLE 2.4 (continued)

Date	General Situation	Cost of Gasoline Cents per Gallon	Actions in New York State
July 1979	Continuing lines in NYC	$0.90	Minimum purchase requirements in NYC
September 1979	Reduced lines in NYC	$1.00	Remove odd/even (NYC)

Sources: A. J. Neveu, "The 1973-74 Energy Crisis: Impact on Travel," Preliminary Research Report 131, (Albany, NY: New York State Department of Transportation, December 1977); and D. T. Hartgen et al., "Changes in Travel in Response to the 1979 Energy Crisis, Preliminary Research Report 170, (Albany, NY, New York State Department of Transportation, December 1979).

TABLE 2.5

IMPACT OF 1973-74 AND 1979 ENERGY CRISIS IN NEW YORK STATE

	Percent Change Over Previous Year	
	1973-74	1979
• Maximum Quarterly Gasoline Shortfall	-13%	-11%
• Annual Gasoline Use	-5.1%	-5.3%
• Annual Travel (VMT)	-3.4%	-4.5%
• Nominal Gasoline Price	+30%	+35%
• Transit Ridership	+1.0%	+2.8%

Sources: A. J. Neveu, "The 1973-74 Energy Crisis: Impact on Travel," Preliminary Research Report 131, (Albany, NY: New York State Department of Transportation, December 1977); and D. T. Hartgen et al., "Changes in Travel in Response to the 1979 Energy Crisis, Preliminary Research Report 170, (Albany, NY, New York State Department of Transportation, December 1979).

The outbreak of war in the Middle East in October 1973, and subsequent oil embargo, led to long waiting lines at gas stations throughout the state and the imposition of Sunday station closings for gasoline stations. This policy, along with the implementation of a state level allocation process did little to mitigate against long waiting lines. As a result, between January 1974 and February 1974 a one half tank minimum purchase requirement, and limitation of gasoline sales to alternating days based on the last digit of motor vehicle license plates (odd/even) were placed into effect. In addition, the 55 mile per hour speed limit and daylight savings time were established during this period. In April of 1974, the embargo had ended with gasoline at $0.50 to $0.55 per gallon.

Over the shortage period, the maximum quarterly gasoline shortfall was approximately -13% from the previous year with gasoline use declining -5.1%. The nominal price of gasoline increased +30% over the previous year and statewide travel decreased approximately -3.4%. Transit ridership, however, did experience a +1.0% increase over previous year levels.

The response of households and individuals in New York State was examined in a study by Keck and others (10). The study was conducted in January to April 1974 and included a sample of households from three small cities. Those surveyed indicated that few adjustments in travel behavior were made during the 1973-74 shortage. Most actions that were implemented concentrated on work and shopping travel, in particular driving slower and changing modes.

In a study conducted by Jones and others (11), in the New York City region, it was found that auto travel decreased -16.5% for weekdays and -37% on weekends during the shortage period. The impact on upstate resort areas was dramatic with marked decrease in hotel occupancies and event attendance. This was a result, in large part, of the unavailability of gasoline during weekends because of station closings. Public transportation played a small role in maintaining mobility.

A limited number of additional consumer based studies during the 1973-74 period were completed. Peskin and others (12) in a study of upper income suburban residents of Chicago, Illinois, conducted in July 1974, found that the shortage did not affect work travel, but non-work travel declined -13.0%, primarily due to trip chaining. In addition, this study suggested that energy supply, and not price, was more important in determining travel demand. Finally, transit use did not increase significantly.

In an August 1974 study conducted by Skinner (13), using a small nationwide sample of government workers, it was suggested that gasoline demand is inelastic with respect to price. This study, however, contained an over representation of upper income subjects with few transit alternatives suggesting that the results do not indicate relationships between price and purchase behavior for the average person.

Sacco and Hajj (14) found in an April 1974 study using a before-after survey of households in Columbia, South Carolina, that 60% of the households interviewed made moderate reductions in driving during the shortage. Travel decreased 5% to 10% on weekdays and 10% to 20% on weekends. The major reduction in trip making occurred in shopping and

recreation trips. Once again, supply and not price was found to be a more important factor in determining driving changes.

Stearns (15) in a nationwide study of consumer response at two points in time, November 1973 and February 1974, found that lower income groups experienced modal shifts while upper income subjects reduced discretionary travel. It was concluded that upper income groups had more flexibility to respond than lower income groups.

Corsi and Harvey (16) used data collected in November 1975 to study the impact of the shortage on Milwaukee residents. It was found that only a small proportion of subjects changed travel modes to work. Most who reported a change indicated a switch to carpooling. These individuals tended to be young, middle income, long distance suburban commuters.

Becker and others (17), in a study of Portland, Oregon, used data collected in summer 1974 to investigate changes made during the shortage. Fifty-two percent of those surveyed changed some frequency of travel during the crisis period. Only 16% reverted back to previous behavior after the shortage ended. Upper income groups were more likely to change travel patterns, but lower income groups tended to exhibit a higher rate of retention for changes made than upper income groups.

A study conducted in Dallas, Texas, by the North Central Texas Council of Governments (18) found a reduction of -2.0% to -4.0% in auto travel during the shortage period. Transit ridership increased +5% to +7%. It should be noted that this region did not suffer as severe a shortage in supply as other regions of the nation.

An examination of the results of various studies of household level response to the 1973-74 energy shortage suggests several things. First, only small or incremental changes in non-work related travel were pursued. Next, among higher income groups, demand for gasoline was found to be inelastic with respect to price. In addition, higher income and auto ownership households appeared to have more flexibility to respond to the shortage. As a result, higher income subjects tended to implement more conservation actions than other income groups. Finally, public transportation was not a major source of conservation during this period.

Government response to shortages emphasized actions to maintain order and assure equitable access to available supplies, in addition to speed limit reductions. For the most part, however, consumers saved energy through individual initiatives based on travel needs and modal options available.

In the summer of 1979 a second energy shortage developed as a result of an Iranian oil shutoff combined with misdirected allocation of available supplies. Unlike the 1973-74 shortage, the 1979 shortage had a very severe impact on summer travel and tourism.

Shortages first developed in California in May 1979 as a result of a 7% to 10% shortfall in supply. By June 1979 New York State developed shortages resulting in the imposition of odd/even and one-half tank rationing in New York City. In July 1979 the Governor ordered use of future monthly supply set asides to meet vacation travel needs. Gas stations were ordered open on at least one day during the weekend. By September of 1979 the shortage had ended.

The minimum quarterly shortfall during the shortage was -11% over the previous year. Gasoline sales declined -5.3% and travel was down -4.5% statewide. Fuel prices, now over $1.00 per gallon, had increased +35% over the previous year. Public transit travel was up +2.8%.

Several studies examined the impact on households in New York State during the 1979 shortage.

Hartgen and Neveu (1) reported, in the 1979 statewide survey used in the current study, that consumers made small adjustments to shopping and vacation related travel as a result of the shortage. However, work travel adjustments and car use actions were the most effective in saving fuel. The availability of fuel efficient autos introduced an option not available during the previous shortage and many subjects reported that new fuel efficient autos had been purchased to conserve fuel.

The study also attempted to quantify the savings through actions reported by the sample subjects. Downstate, or New York City residents, concentrated on using transit for work travel which resulted in 31% of the region's energy savings. Upstate residents were shown to save through auto use actions such as driving slower and replacing an auto. Older, low income, and zero auto households were shown not to contribute to savings on a statewide basis. The major savings reported was through purchasing of fuel efficient automobiles.

A similar study was completed using data collected in 1980. These data are also used in the present study. The original analysis of these data by Neveu (2) suggested that one year following the shortage

subjects continued to pursue similar strategies to conserve transportation energy. For the most part, the results paralleled the earlier 1979 study. The conclusions suggested that higher income and auto owning households saved the most energy. It is suggested that consumers respond based on prior experience, economic status and geographic location.

A study by the New York State office of Parks and Recreation (19) found that 40% of those surveyed altered vacation travel through fewer and shorter trips. Sixteen percent of the sample indicated that the fuel shortage prevented them from traveling to recreational areas. Modal shifts for recreation trips were not significant. Over 25% sampled cancelled vacation trips because of the shortage.

The Regional Plan Association of New York City (20) examined the impact of the shortage on that region. Weekday auto travel dropped from -5% to -20% at some reporting stations during the shortage. Weekend travel was found to decline -10% to -30% reflecting a concern about fuel availability. Transit ridership in the region was found to increase +5% to +15%.

A study for New York State conducted by System Design Concepts, Inc. (21) concluded that weekend non-work local travel declined -22% and weekend intercity travel was down -18%. Downstate urban and suburban areas accounted for 77% of the statewide travel decline.

An examination of the role of public transit conducted by New York State (22), using a sample of transit systems throughout the nation, found that transit played a very minor role in conserving fuel. It was found that ridership increases saved less than 5.0% of total fuel

conserved. The cross elasticity of ridership with respect to fuel sales was found to be -0.75 for the entire sample.

A study conducted by the San Diego, California Comprehensive Planning Organization (23) used a sample of households in the San Diego area collected in the summer of 1979 to examine changes in travel as a result of the shortage. Nearly 75% surveyed reported a reduction in travel, with higher income households reducing because of limited supplies and, lower income households because of high price. Less than 20% changed work modes with most choosing to carpool at least once per week. Larger households reported more changes than small households. Finally 25% surveyed considered trading for a more fuel efficient automobile.

Trent and Pollard (24) explored public reaction to increasing fuel prices for the period 1974-79 using a panel of subjects questioned at three points in time (1975, 1976, 1980) in the State of West Virginia. The study findings suggested that the increase in real, not pump, price of gasoline causes a change in travel and fuel use. Next, options cited in the past by a subject are related to current choices. Most subjects expressed a preference for driving reductions, rather than work mode shifts, to conserve fuel.

Finally, in a study conducted by the Washington State Transportation Commission (25) during the summer of 1980 of households nationwide, changes made over the previous two years were examined. Reduction in travel was the highest reported change endorsed by 10% of the subjects. Seventy-seven percent did not change their trip modes.

Nearly 45% had purchased a new auto over the study period with two-thirds indicating that the new one was more fuel efficient.

The studies examined suggest that consumer reaction to the 1979 shortage was similar to the 1973-74 experience. Once again, consumers concentrated on non-work travel. In 1979 vacation and recreation travel in New York State was greatly curtailed. However, unlike 1973-74, during the 1979 period consumers were able to purchase more fuel efficient automobiles to maintain mobility. Transit was shown to play only a minor role in consumer conservation behavior.

Generally speaking, consumers with the highest degree of travel and financial flexibility contributed to the largest savings. Those with economic and geographic or transportation constraints had few options to exercise in fuel conservation.

Market Segmentation Analysis

The purpose of this section is to introduce concepts and procedures associated with market segmentation analysis. Market segmentation is a strategy designed to divide a general product or service market into smaller subgroups, each of which are characterized by different demand elasticities with respect to a set of marketing offers. In other words, market segmentation analysis is designed to exploit the existence of unique demand functions for various combinations of product, price, promotion and place, or market mixes, in order to maximize the achievement of organizational goals. For many years the principles of market segmentation analysis have been applied by

private sector organizations, in large part, in an attempt to enhance corporate profits.

Another goal of this discussion is to present an overview of applications of market segmentation in the field of transportation planning. As will be seen, the term market segmentation has been misused in the classic sense in most of the planning literature. More specifically, there are very few examples of the business or corporate market segmentation strategy as suggested by the definition above in the transportation planning field. Rather, in many cases, market segmentation has been used as a term to describe any division of some travel demand market for purposes of improving forecasting models or in order to identify group differences as they relate to actual or potential demand for selected transportation modes. Many times, no attempt is made to follow through in developing policies or programs to match consumer demand preferences.

Market Segmentation Theory

In the marketing literature a product or service market is described as those individuals, households, or institutions with the ability and willingness to purchase a product or service for end use purposes (26, p. 102). This definition can be contrasted with the more general definition used in the transportation planning literature in which a market is often described in terms of entire populations or geographic areas. Individuals who are unwilling or unable to use a service appear to be included in this latter, rather broad, definition. However, in order for a market to be completely specified, it is

necessary to apply the more complete conceptualization developed in the marketing field.

Engle, Fiorillo, and Cayley (27) suggested that market segmentation is based on three propositions: consumers are different, differences in consumers relate to differences in market demand, segments within an entire population can be identified.

Included among the benefits of market segmentation suggested by Engle, Fiorillo, and Cayley (27) are the following. First, market segmentation leads to a more precise definition of consumer needs. Second, management is better able to design more effective combinations of market mixes to fulfill needs. Third, market segmentation analysis contributes toward a data base which serves as a foundation to meet changing needs. In addition, market segmentation analysis encourages a more focused definition of corporate goals and objectives and contributes to a more precise description of marketing programs. Finally, market segmentation can enhance the effectiveness of organizational objectives.

Corporate use of marketing segmentation is designed, among other things, to maximize consumer satisfaction through the efficient use of organizational resources. In very simple terms the ability of an organization to satisfy consumer demand can be measured in sales while resources can be described as cash, marketable securities, property, plant, equipment and other assets, including personnel. The general purpose of business applications of market segmentation, therefore, is to enhance corporate profitability, market share, and return on investment. In addition, however, segmentation is used to achieve other

qualitative goals such as enhancing corporate image or reinforcing positive attitudes toward past purchases of products or services.

A conceptualization of market segmentation analysis, as suggested by Cravens, Hills, and Woodruff (26), is outlined in Figure 2.1. This paradigm is included as an example of private sector market segmentation analysis. It does, however, suggest a framework for the strategy to be used in this current research.

As shown in Figure 2.1, the first step is to decide on a basis to form market segments. A basis consists of characteristics on which potential buyers differ and which, at least, reflect differences in demand for marketing offers. Cravens, Hills, and Woodruff (26, p. 100). In general terms, bases can be grouped into variables describing purchase or use behavior, customer descriptive characteristics and product characteristics perceived by customers.

Kotler (28) suggests that the selection of a basis for segmentation must meet three criteria: measurability, accessibility, and substantiality. Nicolaidis, Wachs, and Golob (29, p. 23) state that in addition to these three, two additional criteria should be applied in evaluating bases for segmenting travel demand markets. These include the relationship of the basis to travel behavior and the relationship between the basis and potential service or policy proposals.

Measurability refers to the procedure or process involved in procuring the necessary information describing a basis. The data collection effort must be cost-effective in terms of resources required to obtain and process this information.

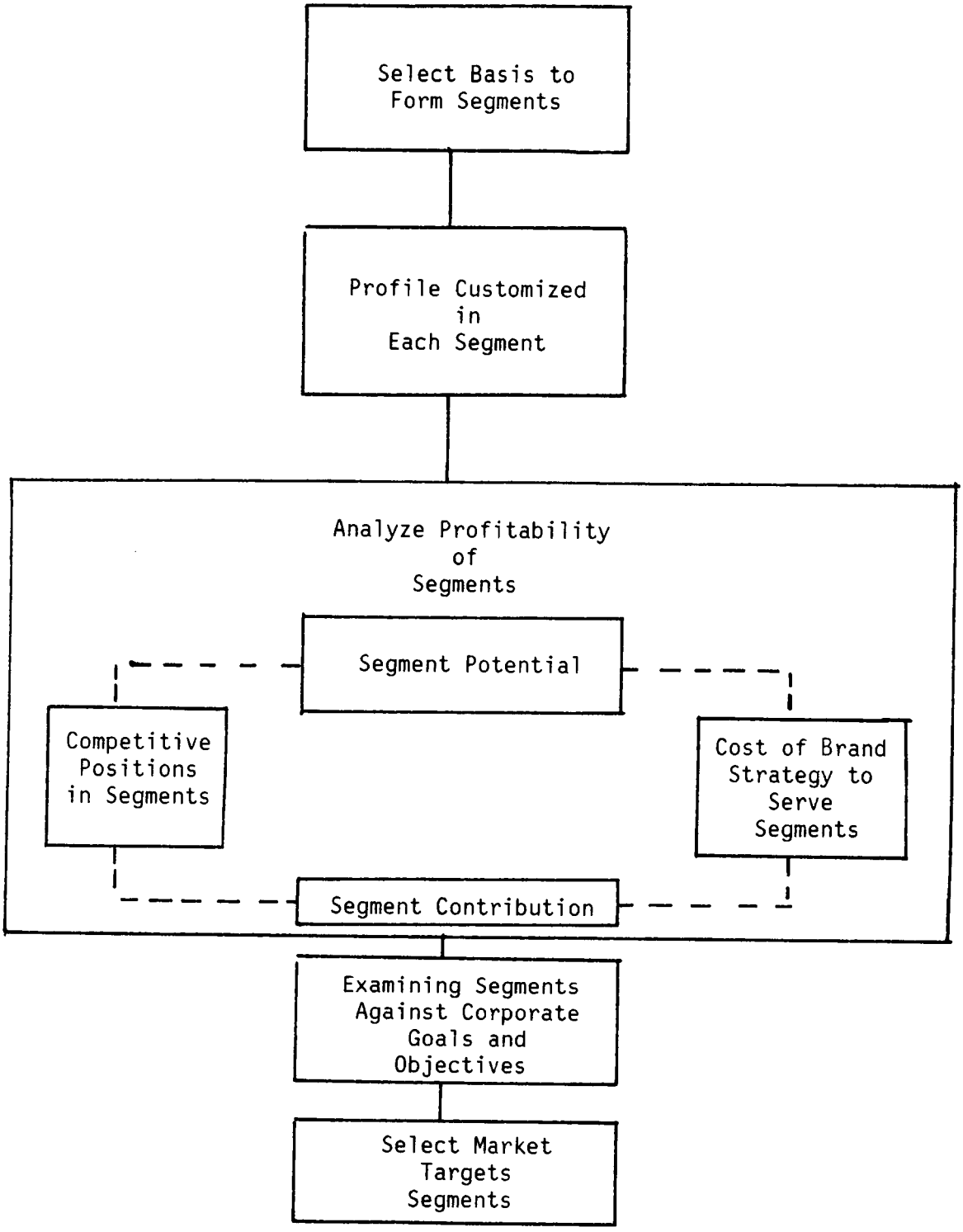


FIGURE 2.1

PARADIGM OF MARKET SEGMENTATION ANALYSIS

Accessibility is related to the statistical robustness of the basis. The market segment identified must be sufficiently different with respect to the behavior under consideration to warrant the pursuit of a market segmentation strategy.

The concept of substantiality refers to the size of market segments identified through the use of a particular basis. Groups identified must be large enough to justify expenditures involved in developing information concerning the segments.

In terms of travel related behavior, it is also necessary for a basis to be a useful indicant of the travel phenomenon under study. Segments should account for a large proportion of the variance for the behavior under investigation.

Finally, a fifth criterion to be used to evaluate potential basis variables includes the relationship between the basis and policy or service options. Basis variables should be useful in examining the relationship between planning actions and market segments.

Once a basis for defining group segments has been developed then it is necessary to apply a procedure which assigns individual consumers to specific groups. This grouping may be done a priori in which discrete levels of basis variables are used to form groups based on the value of these variables for each subject. For example, if age and income are basis variables, a sample of consumers may be cross classified by income and age with the resultant contingency table providing an indication of various market segments. The members of these segments can then be examined in more detail to suggest the relationship between segments and the behavior of interest.

On the other hand, a search procedure using multiple measures can be used. This may involve applying a multivariate statistical procedure such as cluster analysis or factor analysis to place subjects in a hypothesized multidimensional space describing market segments. In addition, multiple group discriminant analysis may be applied to relate observed behavior to a set of independent variables. Differences in group behavior may then be explained in terms of the resultant set of hypothesized discriminant functions containing the original basis measures.

Once market segments have been identified, the next step in market segmentation requires the development of group profiles. This process is designed to search for relationships between the basis variables and other descriptive variables. The profile of any segment attempts to define who the members are that make up a segment; what they are like as people; how they decide to implement the behavior under study; and what outside influences may impact the consumers under study. A market segment profile suggests reasons why segments behave as observed and indicates possible future courses of action, with respect to market offerings, that will appeal to identified segments.

The next step in the market segmentation analysis procedure, as practiced in the private sector, consists of evaluating the profitability of each market segment identified. This includes an assessment of the demand for the product or service under consideration by each segment as well as an indication of the cost of providing the marketing offer. In addition, other offerings from different organizations

should be analyzed to determine the competitive position of the firm with respect to other firms.

After this objective analysis is complete, a more qualitative investigation is undertaken to evaluate the cost of serving a market in light of organizational goals and objectives. This evaluation is designed to allow consideration of such concepts as image of the firm in identification of potential target markets.

Finally, the last step consists of determining particular target market segments to be pursued by the firm. At this point it may be decided to develop strategies to serve the needs of all segments. On the other hand, instead of a full coverage approach, a strategy designed to serve only selected segments may be pursued.

To summarize, in order to be of value to an organization, information describing a market segment must demonstrate that demand is sufficient to warrant expenditure of resources. In addition, the segment identified must be accessible to the marketing offer. Finally, segments must be different from each other in order to justify implementation of the process.

The marketing literature includes a large number of studies using the market segmentation procedure. Among the more useful overviews of these studies are volumes by Engle, Fiorillo, and Cayley (27); Frank, Massey, and Wind (30); and Cravens, Hills, and Woodruff (26). In addition, the Journal of Marketing Research contains many applied studies of the theory of market segmentation in management decisions [McCann (31)], studies on the selection of basis for segmentation [Lessing and

Tollefson (32)], and studies evaluating multivariate analysis procedures [Peters (33)].

The following section presents an overview of the use of market segmentation analysis in the transportation planning literature. For the most part, only portions of the market segmentation analysis procedure presented in this section have been investigated. Market segmentation has been used in various ways including the analysis of consumer travel choice processes; model calibration or simulation; and in the classic sense of developing marketing offers of various public transportation services and policies.

Market Segmentation in Transportation Planning

Numerous applications of the basic concepts of market segmentation have been reported in the transportation planning literature. For the most part, these studies have involved the analysis of users and nonusers of transport systems; the calibration and application of travel demand models; the examination of household travel and activity patterns; and the investigation of mode choice or travel constraints. In addition, urban public transportation planners have applied the principles of market segmentation in systemwide marketing programs. Studies in the area of public transportation planning come closest to fulfilling the goals of market segmentation analysis outlined above. Other research applies the term market segmentation to refer to any division of an overall travel market into groups. In most cases this is done to investigate the relationship between travel demand, often expressed in a trip rate, for various subgroups of the population and to improve the forecasting accuracy of travel simulation systems.

Tye (34) suggests that six types of bases have been used in the application of market segmentation analysis in transportation planning. These include the following: (1) subjective judgments, (2) sociodemographics, (3) relevant choice sets, (4) attributes of choices, (5) use and observed choice, and (6) geography.

Subjective judgments include attitudes, perceptions, beliefs, or behavioral interests toward modes or modal attributes. For example, attitudes toward auto or bus travel can be measured for a population under study. The results can be used to segment the population into groups with positive attitudes toward auto or positive attitudes toward transit.

Sociodemographics includes such variables as age, income, household size, or auto ownership. Analysis may involve simple multidimensional contingency table analysis or the use of multivariate statistical procedures to identify groups.

Relevant choice sets includes using the the availability of modal alternatives as a basis to analyze travel choice. For example, in examining mode choice between auto or bus, the population may be segmented on the basis of availability of selected modes for various trip purposes.

Attributes of choices includes variables related to the characteristics of alternatives. These include both subjective and objective measures of modal characteristics such as travel time, comfort, or convenience. A population under study may be segmented on the basis of measures concerning the attributes of current or proposed modes.

Use and observed choice refers to the division of populations into groups based on current travel choices. This may include segmenting the population based on travel mode to work.

The use of geography to segment travel markets has been an important element in traditional transportation planning. Urban transportation studies often describe analysis areas in terms of geographic zones that serve as a foundation to evaluate current and future travel.

Nicolaidis, Wachs, and Golob (29) found that segmentation of the travel market based on choice constraints performed best in explaining travel behavior. However, it was pointed out that all the bases performed well in some application environments depending on the purpose of the study.

One common use of market segmentation involves investigating users and non-users of particular modes. Tisher and Dobson (35) used a sample of communities in California to collect information about perceptions of work modes. It was found that observed mode choice could be explained by belief about mode characteristics. Costantio and others (36) used beliefs as a basis to segment a study group in order to explain differences in choice for existing and proposed transportation modes. Neveu, Koppleman, and Stopher (37) used attitudes of users toward existing work trip modes and found that the results suggested that mode perceptions are linked to observed use.

A second general area of the use of market segmentation is in the use of disaggregate travel demand models. In order to avoid problems of aggregation bias, Koppleman (38) recommended that the travel market be segmented in various groups based on the independent variables used

in a forecasting model. The results for each group are summed to yield an areawide forecast. Dunbar (39) and Reid (40) also report the use of a market segmentation strategy for use in aggregating forecast results using disaggregate models.

A third general area of interest in the use of market segmentation analysis includes studies of household travel and activity patterns. It has long been recognized that travel is a derived demand. Travel occurs because of the spatial separation between desired activity sets. Out of home activities such as work or shopping are common examples of important activities shared by many households.

Recker and others (41) use travel and activity data to segment a sample population into subgroups based on the type, duration, and sequencing of activity patterns. Pas (42) uses multivariate methods to relate activity patterns to selected socioeconomic variables.

Zahavi and others (43) have constructed an entire forecasting system based on travel time relationships and various activity sets. Hanson (44) reports on the use of age as a stratification variable to segment a sample population into subgroups based on various activities.

In all of these cases, the major goal is to understand travel phenomenon in more detail. Little interest is placed on applying the concept of market segmentation to design marketing offers in the classic sense. However, studies in the field of public transit planning do include some attempt at this comprehensive approach.

Sen and Benjamin (45) report on the results of a study of user requirements applied to define service policies for subgroups.

In a detailed study of users and non-users of public transportation, conducted by Woodruff et al. (46), market segmentation was used to develop services for non-transit users. This information was applied to develop service concepts that were tested in a follow-up survey.

As shown in the overview, market segmentation has been used in transportation planning in various ways. Most applications, however, do not exploit the full concept of market segmentation as suggested by private sector theorists. Some limited examples do exist. For the most part these efforts have been designed to design improved services for urban public transportation systems.

Summary and Implications for this Research

This research is designed to investigate the application of market segmentation analysis in transportation analysis energy planning. As discussed in this chapter, recent energy shortages resulted in an interruption to historic growth trends in gasoline consumption and travel. Limited studies of consumer responses to these shortage periods have indicated a propensity of households to implement conservation actions designed to minimize disruptions to personal travel patterns. The role of the fuel efficient automobile in reducing fuel consumption while maintaining desired travel levels emerged as an important component of the actions pursued by consumers.

Market segmentation analysis is a vital part of private sector marketing. Few examples of the application of the private sector model of market segmentation are contained in the transportation planning

literature. In most cases, the term market segmentation has been applied to any process used to develop subgroups in a total poulation under study.

This research applies the concepts of market segmentation for transportation energy planning. Data collected from two statewide sample surveys collected in New York State are used as a basis for segmentation. Two methods are used to develop market segments. First, conventional multivariate analysis procedures are used. Next, the concepts of latent trait theory are applied. In both cases, the results are used to match programs and policies with target segments in order to more effectively utilize planning resources.

CHAPTER 3

RESEARCH DESIGN AND DATA SOURCES

Introduction

The purpose of this chapter is to present the strategy to be pursued in applying the concepts of market segmentation analysis to energy planning. More specifically, it outlines the data sources to be used and the analysis techniques to be applied in this investigation. The first part presents an overview of the market segmentation strategy. Next, the data sources are presented. This research uses data previously collected by the New York State Department of Transportation for use in transportation energy planning. Data collected through telephone interviews in October 1979 and October 1980 provide the basis for use in market segmentation. These surveys contain over 1,500 adult interviews covering the entire state.

Finally, the third part of this chapter presents an overview of the analysis approaches to be used in this study. Two alternative segmentation analysis methods are described, one based on a set of conventional multivariate procedures and a second based on modern test or latent trait theory.

Background

As stated earlier, New York State experienced the impact of rapid energy price increases and supply shortfalls in the summer of 1979. In

1980 the transportation energy situation was stabilizing resulting in an overall increase in price, but ample fuel supplies. In order to identify the consumer based responses to these developments, a set of statewide telephone interviews were conducted for the New York State Department of Transportation. These surveys included a section where subjects were asked to indicate which of 18 different energy conservation actions had been taken to conserve gasoline. The responses to these questions provide the fundamental information to be used as a basis for market segmentation in this current study.

For purposes of this research a strategy consisting of four major steps comprise the research approach. These include: (1) identification of bases for segmentations, (2) development of market segments, (3) identification of promising segments for conservation actions, and (4) matching promising segments with policies or programs to conserve energy. Each of these are described in more detail below.

Market segmentation analysis will be used in this research to identify groups which share a common response to energy price and supply changes. The basis for identification of groups is provided by a subject's response to a set of 18 energy conservation actions. These actions range from such behaviors as driving more slowly or using public transportation, to buying a new fuel efficient automobile or moving closer to work.

The use of market segmentation is designed to provide insight into the behavioral changes associated with meeting the travel desires of a household within a framework of actual and perceived energy constraints. An individual interacts with other members of a household in defining

activity needs. These needs can be broken down into necessary activities (e.g., work, food shopping) and discretionary activities (e.g., personal business, non-food shopping, recreation). However, the spatial-temporal choices associated with these activities are limited by constraints related to the household (e.g., income, auto ownership, individual roles within the household) and constraints related to the transportation system (e.g., travel times and cost, energy price and supply).

The individual and the household must continually balance activity needs and desires with system constraints to arrive at an optimum combination of travel related decisions. For example, a household characterized by a certain level of auto ownership, household composition, and geographic location may adapt to increasing gasoline prices by reducing the use of the automobile; replacing an auto with a more fuel efficient model; linking work and non-work trips; or moving the location of the household to reduce trip lengths.

Knowledge concerning the various combinations of energy conservation actions undertaken by individuals or households is, therefore, used in this study to develop market segments. This information is then applied in combination with selected socioeconomic variables to relate these actions with characteristics of the household so that programs and policies can be suggested which best meet the needs of these groups.

Two methods are pursued to develop market segments. The first is based on the application of conventional multivariate analysis procedures including factor analysis, cluster analysis and contingency

table analysis. First, matrices of interaction correlation coefficients are factor analyzed using principal factor analysis. Next, a set of factor scores are developed from these results for each subject. The use of factor scores helps to reduce the dimensionality of the problem to a set of factors summarizing the results from the 18 action items. Next, a cluster analysis algorithm is used to identify individuals with similar sets of action responses. The group membership of a subject is then cross-classified with selected independent variables to identify relationships between actions and group clusters. The end product of this effort includes a description of group clusters which help provide an understanding of the nature of energy conservation responses over the study period.

The second method of analysis includes the use of latent trait or modern test theory. The purpose of this analysis is to use the set of action item responses offered by a subject to determine the position of a subject along a hypothesized latent continuum describing a subject's ability and willingness to conserve transportation energy. The subject's socioeconomic characteristics are then related to this psychological construct to provide an understanding of the nature of the relationship between transportation energy conservation and these independent variables.

The major purpose of pursuing two alternative segmentation strategies is based on a desire to investigate the similarities and differences between conventional analysis and modern psychometric methods. The goal of this study is to present an example of where newer methods may fit into transportation analysis. In this manner a certain level of

experience and perspective can be gained through this particular study that can be shared with the larger audience of concerned transportation researchers.

The third step in the analysis strategy applied in this study consists of identifying the characteristics of groups that relate to the effective application of specific conservation policies and programs. These results are derived from interpretation of previous analysis methods. The work is designed to highlight promising segments for conservation actions in future energy price increases and supply shortfalls. Because data to be used in this study has been collected over two years characterized by different background energy price and availability conditions, the types of reactions by groups are expected to differ. For example, the 1979 survey year was characterized by supply shortfalls and rapid price increases throughout the State of New York. The 1980 survey was characterized by higher, but stablizing prices and no supply limitations. Policies to match market segments emerging from these conditions must recognize these background factors in order to be effective.

The last step in the market segmentation consists of matching promising segments with policies and programs to conserve transportation energy. This final portion of the analysis is designed to aid in future policymaking in contending with emergency energy shortages. Because of the nature of the data collected, with respect to background conditions, these recommended policies are designed to meet the policymaking needs arising from a scenario of rapid price increases and shortfalls along

with a later adjustment period characterized by higher prices but steady supplies of gasoline.

Survey Design

The data base used in this research was collected as part of the Crossley Empire State Poll. This poll was conducted four times per year for the years 1979 and 1980 and was designed to allow several organizations, or groups, in the New York State government the opportunity to conduct a shared cost telephone survey. In 1979, wave 79-4 of this survey was used by the New York State Department of Transportation (NYSDOT) to collect data concerning energy conservation behavior (47). In 1980, NYSDOT included a set of similar questions in wave 80-4 of the survey (48). These data sets represent the survey sample base used in this study.

In each wave of the Crossley Empire State Poll, contracting organizations were requested to submit desired survey questions by the 20th day of the month preceding interviewing. Interviewing began on the first day of October in 1979 and the first day of October in 1980. The interviewing, in both cases, continued a maximum of three weeks with results reported on the 15th day of November for the respective years. The 1979 survey resulted in a total sample size of 1,520 adults and the 1980 sample had a sample size of 1,560 adults.

Appendix A contains the set of survey instruments used in this data collection effort. Only the portions relevant to this study are included. The questions included represent data items used in this particular study and do not represent all items collected in these

surveys. However, the questions that have been used in the market segmentation investigation are presented to demonstrate the wording of items. Many of these items have been recoded so that the categories range in value from 0 to 3. This recoding scheme has been done to help simplify the analysis. Code categories are specified in the following section along with the proportion of subjects classified by each code.

The random sample of adults was selected in the following manner. First, New York State was divided into four geographic areas: New York City, Nassau/Suffolk counties, Westchester/Rockland counties, and all upstate counties. Next, each of these regions were divided into smaller units resulting in a total of 14 subregions. The survey was designed to yield responses in proportion to the relationship between the subregion population and the total state population. Within each subregion, a sample of telephone numbers was selected from local telephone directories. The last digit of a number was replaced with a random number so that unlisted telephone numbers would be included in the survey. The survey samples, therefore, are designed to be representative of state adults living in households with telephones. It was estimated that 95% of New York State households had telephones in 1975 (49).

All interviewing was conducted by telephone in the early evening between 5:00 p.m. and 9:00 p.m. during weekdays and weekends. Ten percent of all interviewers' subjects were called back to verify that the survey had been completed.

Telephone numbers were eliminated because of no answer, and no callbacks were made to these numbers. Once a household was contacted,

one adult aged 18 years and over was randomly selected from adults at home to complete the survey. Controls were employed to ensure that a correct proportion of males and females were interviewed. Approximately 10% of households contacted refused to participate in the survey.

Survey Results

The survey instruments were designed to collect information concerning the type of energy conservation actions a subject had taken to conserve gasoline over the previous months. More specifically, interviewers read a list of 18 different actions and asked subjects to indicate whether, since January of the survey year, an action had been taken to reduce fuel consumption because of rising prices or supply shortages. As stated earlier, the background conditions for the 1979 survey were much different from those of the 1980 survey because of energy shortages and price increases resulting from the 1979 Iranian Revolution. The 1980 period was characterized by an increase in prices, but no supply shortfalls.

Table 3.1 presents the results of these surveys for the energy action items. This table outlines the importance of actions requiring changes in shopping related travel. In particular, combining trips or trip chaining, shopping closer to home, and making fewer shopping trips are among the items with the highest proportion of endorsement. In addition, driving slower, shopping on the way home from work and tuning vehicles played a large role. Except for vacation related actions and purchasing fuel efficient autos, remaining actions were endorsed only at a very low proportion.

TABLE 3.1

SUMMARY OF SURVEY RESULTS FOR
ENERGY CONSERVATION ACTIONS

Action	Description	Proportion Sample Endorsing Each Action	
		1979 Survey	1980 Survey
1	Drive more slowly than before	0.418	0.428
2	Carpool to work	0.139	0.133
3	Take bus or subway to work	0.120	0.135
4	Walk or bike to work	0.083	0.108
5	Have car tuned more often	0.372	0.256
6	Move closer to work	0.025	0.031
7	Get a job closer to home	0.052	0.043
8	Shop on way home from work	0.252	0.301
9	Make fewer shopping trips	0.350	0.532
10	Combine shopping and other non-work trips	0.468	0.537
11	Shop closer to home	0.410	0.470
12	Eliminate use of recreation vehicles	0.093	0.117
13	Take vacation closer to home	0.170	0.187
14	Cancel a vacation trip	0.156	0.118
15	Use train, bus or plane for vacation trips	0.159	0.207
16	Replace a car with a more fuel efficient one	0.153	0.173
17	Sell car (don't replace)	0.084	0.051
18	Take bus or subway more often for non-work trips	0.174	0.152

Earlier work by Neveu and Hartgen (1) and Neveu (2) with these data centered on examining cross tabulations of actions with selected independent variables to identify the relationship between group characteristics and action endorsement. In addition, this work involved applying these survey results to estimate energy savings. This work did not examine combinations of actions or investigate the existence of a hypothesized latent trait relating subjects to action endorsement. This earlier work can be examined for details concerning the results. For purposes of this discussion it is sufficient to state that the results indicated that the results of action items endorsement were related to the ability of subjects to respond based on income, auto ownership, location in the state, and household size. It was concluded that high income households had the ability to purchase a fuel efficient automobile to maintain mobility levels. Low to moderate income household postponed this action, but viewed it as an effective solution. In general, subjects attempted to minimize disruptions but not necessarily by using actions which saved the most energy.

Besides data collected concerning the actions taken to conserve transportation energy, selected socioeconomic information was collected. Table 3.2 summarizes the information collected for ten independent variables as well as the coding scheme used. In addition, Table 3.2 presents a comparison of the sample survey results with the 1980 census data collected for New York State, where information is available.

The downstate urban area refers to the greater New York City region including Long Island, Westchester, and Rockland counties. The

TABLE 3.2

SUMMARY OF SURVEY RESULTS FOR SELECTED INDEPENDENT VARIABLES

Variable	Category	Code	Proportion		
			1980 NYS Census	1979 Survey	1980 Survey
Area in State	o Downstate urban	1	0.607	0.601	0.548
	o Upstate urban	2	0.180	0.221	0.262
	o Other	3	0.213	0.178	0.190
Age	o 18-34 years	1	0.357	0.397	0.363
	o 35-64 years	2	0.417	0.482	0.454
	o over 65 years	3	0.156	0.115	0.170
	o Missing		--	0.006	0.013
Sex	o Male	1	0.462	0.495	0.469
	o Female	2	0.538	0.505	0.531
Marital Status	o Married	1	0.522	0.618	0.576
	o Single	2	0.306	0.236	0.249
	o Widowed, divorced or separated	3	0.172	0.141	0.165
	o Missing		--	0.005	0.010
Vehicles Available	o 0 vehicles	0	0.325	0.172	0.195
	o 1 vehicle	1	0.362	0.398	0.416
	o 2 or more vehicles	2	0.313	0.426	0.387
	o Missing		--	0.004	0.002

TABLE 3.2 (continued)

Variable	Category	Code	1980 NYS Census	1979 Survey	1980 Survey
Household Size	o 1-2 persons	1	0.550	0.439	0.503
	o 3-4 persons	2	0.313	0.378	0.343
	o 5 or more persons	3	0.137	0.176	0.137
	o Missing		--	0.007	0.017
Annual Household Income	o Less than $10,000	1	0.304	0.156	0.172
	o $10,000 - $25,000	2	0.402	0.445	0.401
	o over $25,000	3	0.294	0.261	0.249
	o Missing		--	0.138	0.178
Educational Attainment	o Grammar school or some high school	1	a	0.152	0.183
	o High school graduate	2	a	0.374	0.372
	o Some college or college graduate	3	a	0.466	0.570
	o Missing			0.008	0.015
Employment Status	o Employed part-time or full-time	1	a	0.638	0.535
	o Unemployed, retired, student, homemaker	2	a	0.350	0.401
	o Missing			0.012	0.064

TABLE 3.2 (continued)

Variable	Category	Code	1980 NYS Census	1979 Survey	1980 Survey
Housing Type	o Single family	1	a	0.497	0.521
	o Two family	2	a	0.138	0.119
	o Apartment, condo, co-op	3	a	0.338	0.328
	o Missing			0.027	0.032

a - not available

upstate urban area consists of all urban areas outside the New York City region including Albany, Buffalo, Syracuse, and Rochester.

It is clear from these results that fairly representative geographic samples have been collected. However, because the samples only include households with telephones, low income and low auto ownership households are underrepresented. In addition, because telephone numbers that were not answered were not called back, single persons and small households are also understated. Generally speaking, except for income and employment status, only a few records had missing values reported for most independent variables.

As compared to the 1979 survey, the 1980 results tend to be from outside the downstate urban area, are older, tend to be female, are slightly less likely to be married, own fewer vehicles, live in smaller households and are less likely to live in multifamily housing. In addition, the 1980 sample subjects have a lower educational attainment than the 1979 survey, have lower income, and are less likely to be employed.

In order to examine the relationship between various independent variables, matrices of intercorrelations between these variables have been developed. These results are presented in Table 3.3 for the 1979 survey and Table 3.4 for the 1980 data. Consistent results are evident across samples for these data sets.

It appears that upstate urban and other residence have a higher level of auto ownership and are less likely to live in a multifamily dwelling unit. Older subjects tend to own fewer vehicles, live in smaller households, have lower income levels and lower educational

TABLE 3.3

CORRELATION MATRIX OF SELECTED INDEPENDENT VARIABLES, 1979

	Area	Age	Sex	Veh.	Edu.	Marital	HH Size	Income	Employ.	House
Area	1.000									
Age	-0.003	1.000								
Sex	0.004	0.044	1.000							
Vehicle	0.220	-0.208	-0.090	1.000						
Educat.	-0.106	-0.282	-0.118	0.159	1.000					
Marital	-0.096	0.060	0.085	-0.318	-0.051	1.000				
HH Size	0.090	0.247	0.017	0.369	0.062	-0.218	1.000			
Income	-0.124	-0.221	-0.104	0.390	0.390	-0.333	0.177	1.000		
Employ.	0.067	0.293	0.318	-0.135	-0.249	0.031	-0.081	-0.320	1.000	
House	-0.384	-0.042	-0.005	-0.520	0.035	0.286	-0.328	-0.163	-0.052	1.000

TABLE 3.4

CORRELATION MATRIX OF SELECTED INDEPENDENT VARIABLES, 1980

	Area	Age	Sex	Veh.	Edu.	Marital	HH Size	Income	Employ.	House
Area	1.000									
Age	0.050	1.000								
Sex	0.031	-0.006	1.000							
Vehicle	0.264	-0.190	-0.053	1.000						
Educat.	-0.002	-0.240	-0.092	0.223	1.000					
Marital	-0.051	0.073	0.089	-0.389	-0.076	1.000				
HH Size	-0.013	-0.322	0.094	0.324	0.048	-0.280	1.000			
Income	-0.086	-0.317	-0.106	0.438	0.420	-0.323	0.309	1.000		
Employ.	0.035	0.290	0.327	-0.159	-0.293	0.019	-0.034	-0.333	1.000	
House	-0.332	0.004	0.011	-0.546	-0.060	0.308	-0.270	-0.261	-0.005	1.000

attainment levels than younger subjects. In addition, men are more likely to be employed than women. High auto ownership subjects appear to have higher income and household sizes and are more likely to be employed, married, and living in single family households than low auto ownership households. High educational attainment suggests higher income and a greater chance of employment. Married students tend to live in larger households and have higher income than unmarried subjects. Higher household size suggests a higher income and a greater chance of living in a single family house.

Besides reviewing the proportion of subjects endorsing a particular action item, information concerning the total number of actions endorsed by a subject provides an insight into the ability and willingness of a subject to conserve transportation energy. Table 3.5 presents the distribution of the total summated scores for each survey. This score is formed by simply adding all actions endorsed by a subject together. Note that over 16% of the subjects in each sample did not endorse a single action. Table 3.6 shows that on average in 1979, 3.657 actions were endorsed while 3.978 actions were endorsed in 1980.

Examining the correlation of the total score with selected independent variables also suggests some interesting relationships. The results of this effort are presented in Table 3.7. In particular, it appears that subjects from the non-downstate urban area tend to have higher total scores than subjects from downstate areas. A higher score is also found for multiple vehicle households and higher income subjects. In general it appears that as age increases, total score decreases. In addition, employed persons tend to have higher total

TABLE 3.5

RESULTS OF NEW YORK STATE SURVEYS

Total Summated Score	1979 Survey		1980 Survey	
	Total Subjects	Proportion Total Subjects	Total Subjects	Proportion Total Subjects
0	264	0.174	256	0.164
1	189	0.124	154	0.099
2	197	0.130	183	0.117
3	164	0.108	167	0.107
4	154	0.101	187	0.120
5	150	0.099	159	0.102
6	131	0.086	121	0.078
7	94	0.062	98	0.063
8	73	0.048	94	0.060
9	37	0.024	49	0.031
10	23	0.015	39	0.025
11	23	0.015	22	0.014
12	10	0.007	18	0.012
13	6	0.004	6	0.004
14	2	0.001	6	0.004
15	2	0.001	2	0.001
16	0	0.000	0	0.000
17	1	0.001	0	0.000
18	0	0.000	0	0.000

TABLE 3.6

MEAN AND VARIANCE OF TOTAL SUMMATED SCORE, 1979 AND 1980

	Sample Size	Mean of Total Score	Variance of Total Score
1979 Survey	1,520	3.657	3.036
1980 Survey	1,560	3.978	3.163

TABLE 3.7

CORRELATION OF TOTAL SUMMATED SCORE WITH SELECTED INDEPENDENT VARIABLES, 1979 AND 1980

	1979 Survey	1980 Survey
Area	0.204	0.231
Age	-0.221	-0.124
Sex	0.016	-0.023
Vehicle	0.270	0.340
Education	0.087	0.142
Marital	-0.001	-0.134
Household Size	0.145	0.078
Income	0.142	0.133
Employment	-0.132	-0.206
House	-0.154	-0.188

scores than unemployed subjects. Finally, single family housing unit dwellers appear to have higher total scores. In the 1979 survey, educational attainment has a low positive correlation with total score while in 1980 a higher positive correlation is estimated. On the other hand, the 1979 survey appears to have a higher positive correlation with household size than the 1980 data. Marital status appears to be more strongly negatively correlated with total score in 1980 than 1979. This suggests that married subjects are more likely to have endorsed a higher number of actions than unmarried subjects.

This preliminary review of the data sets to be used in this study indicates some general relationships between energy conservation and selected independent variables. In particular, it appears that the types of actions endorsed with the highest rate of frequency are those involving a minimum disruption to travel patterns, such as driving slower or tuning a car, combining non-work travel through trip chaining, and making fewer shopping trips.

Subjects residing in the downstate urban area were more likely to pursue public transit as an alternative to conserve fuel. However, in overall terms, upstate residents were more likely to endorse a larger number of actions. Older subjects had a decreasing chance of endorsing any actions. The number of autos available to a household had a strong relationship with several actions. In particular, the more autos a household owned the more likely it was that the subject drove slower, tuned a car, combined shopping trips, and purchased a more fuel efficient car model. Employed subjects endorsed actions at a higher rate than subjects not working. This was also true for married subjects who also

generally endorsed actions at a higher rate than unmarried subjects. Finally, single family home dwellers generally endorsed actions at a higher rate than those living in multifamily housing.

In terms of data limitations, first it must be recognized that a dicohtomous scoring strategy does not capture the frequency at which an action is undertaken. Ideally, the scoring strategy should have several categories to reflect the intensity to which a subject endorses an action.

In addition, certain independent variables, including income and employment status, have a fairly high degree of missing responses. Any results emerging from this research must recognize this shortcoming.

Finally, this is not a panel survey because the same subjects were not necessarily interviewed in both surveys. However, the same questions were asked in both surveys which will permit a degree of comparison based on socioeconomic characteristics.

This concludes the overview of the data sources to be used in this analysis. These resources provide a unique opportunity to investigate the use of market segmentation in transportation energy planning. A statewide level survey using the same questions relating to conservation behavior in a changing energy environment suggests the possibility of some interesting insights.

Research Procedures

The next chapter presents the results of the market segmentation analysis using conventional multivariate procedures. Following this chapter, the results of a similar type of analysis using modern test or

latent trait theory will be presented. This section presents a brief overview of the steps involved in these analyses. The specific details of these approaches are described at more length in the appropriate chapters.

Conventional market segmentation analysis is completed through first estimation of interaction phi coefficients for each survey year. Next, principal factor analysis is used to identify action clusters. Following this, factor scores are estimated and used to form clusters consisting of subjects with similar responses to energy conservation. Group membership is cross classified with selected independent variables to develop information concerning group profiles. These results help to suggest policies to be used in serving the needs of market segments.

The next analysis procedure consists of applying the concepts of latent trait theory. This involves developing interaction tetrachoric correlation coefficients which are then factor analyzed to suggest the dimensionality of item response variables. These factor analysis results are used to calibrate a set of item characteristic functions which represent the conditional probability of an item response for a fixed value of the latent trait. This is used to estimate the position of subjects in the latent space. Subjects are then grouped among this continuum and assigned to various segments. Group membership is cross classified with selected independent variables to suggest possible policies to meet conservation objectives.

CHAPTER 4

RESULTS OF CONVENTIONAL MARKET SEGMENTATION ANALYSIS

Introduction

The purpose of this chapter is to present the results of a market segmentation analysis using what will be described as a conventional approach. More specifically, in this chapter development of market segments is based on the results of the combination of factor analysis, cluster analysis, and cross-tabulation of items describing transportation energy conservation actions. The next chapter applies the principles of latent trait theory to develop market segments.

As stated earlier, market researchers have suggested that the essence of the segmentation process consists of identifying unique attributes shared by a common group that are related either directly or indirectly to some behavior of interest. In this case the behavior under investigation are the actions taken by individuals to conserve transportation related energy.

A focal point of any segmentation analysis consists of identifying a basis on which to develop group differences. Several types of attributes have been suggested to serve as the bases for group segmentation. These inlcude: purchase or use behavior bases, consumer descriptive characteristics, or product characteristics desired by the consumer Cravens, Hills, and Woodruff (26, p. 178). In the case of this particular research, data concerning behavior related to energy conservation serves as the basis for segmentation.

Once a basis for determining market segments has been defined, it is next necessary to group observations into segments sharing a common set of responses to the variables which serve as the basis for group formation. Many possible approaches can be pursued in order to identify segments. Very often, since multiple measures are collected, it is common to use a multivariate statistical procedure such as factor analysis or cluster analysis to more efficiently reduce the dimensionality of the problem presented by interpreting many variables over a large sample of subjects. This procedure has the additional benefit of adding objectivity to segment formation as well as reducing the complexity of the problem.

In this chapter, market segments will be developed based on a three step process. First, in order to reduce the number of measures used to describe energy conservation actions, correlation matrices consisting of interaction correlations are factor analyzed using principal factor analysis. The results of this effort include the estimation of factor scores for each subject in the data collected in 1979 and 1980.

Next, the factor scores for these subjects are grouped using a cluster analysis algorithm. The results of this analysis serve as the basis for identification of segments.

Finally, once each subject is assigned to a particular segment, group membership is cross-tabulated with each energy conservation action in order to identify the relationship between groups and actions. This process also represents a method to validate cluster analysis results

by helping to check the relationship between factor scores and energy conservation actions.

The development of market segments leads to an analysis of the characteristics of each group as defined by various socioeconomc variables. Once an understanding of the relationship between selected socioeconomic variables and energy conservation actions is established, then energy conservation policies can be developed which can best serve the unique needs of each segment within the framework of conservation goals.

In this chapter, group membership, or segment membership, is cross-classified with several socioeconomic variables to understand the nature of the characteristics describing each segment. The analysis consists of examination of the residual patterns developed through two-way contingency table analysis. These results suggest the relationship between market segments and variables related to household and individual characteristics by indicating the cells in the contingency table that make the highest contribution to a calculated chi-squared statistic. This statistic is designed to test the hypothesis of independence between rows and columns of the table.

In summary, this chapter contains the results of the application of a market segmentation analysis based on conventional development procedures. Data collected from each subject, concerning actions taken to conserve transportation related energy, serve as a basis for developing market segments. Correlation matrices containing interaction correlations for each study year are factor analyzed. Factor scores are estimated for each subject and are then cluster analyzed to identify

segments describing common responses to energy conservation questionnaire items. Socioeconomic characteristics are examined to determine segment differences that may explain the types of actions taken by each segment. This analysis then serves as a framework for suggesting changes that have occurred, with respect to actions taken over the study period, and possible energy conservation policies to be pursued in the future.

In the following chapter latent trait theory is applied to develop market segments. The results of the current chapter are then contrasted with the outcomes using these later procedures.

Interaction Correlations

In the previous chapter preliminary results of the study survey data were presented. The sample proportions endorsing each energy conservation action were presented. It was noted that subjects placed an emphasis on small, frequent actions to conserve energy. In this section the product-moment correlation coefficients between actions within survey years are presented. These results serve two purposes. First, they represent a summarization of the relationship between actions. Next, the variable clusters identified serve as a preliminary basis for evaluation of factor analysis results described in the next section.

Phi Coefficient

Each energy conservation action item was coded either "0" (action not taken) or "1" (action taken) depending on the answer offered by a particular subject. The calculation of a correlation coefficient

between dichotomous items such as these follows a formula that is a special case of the familiar Pearson Product Moment Correlation Coefficient. It is called the phi coefficient. The formula for the phi coefficient is given by the following, Lord and Novick (50, p. 336),:

$$r_{ij} = \frac{P_{ij} - P_i P_j}{\sqrt{P_i Q_i P_j Q_j}} \quad , \tag{4.1}$$

where r_{ij} is the phi coefficient between actions (i) and (j), P_i is the proportion of subjects with a score of "1" for action (i), $Q_i = 1 - P_i$, P_j and Q_j are the similar quantities for items ($_j$), and P_{ij} is the proportion of subjects with a score of "1" for both actions (i) and (j).

The value of the phi coefficient is restricted to $-1 \leq r_{ij} \leq +1$. However, it is not uncommon for the maximum value of this theoretical value to be less than ±1 depending on the proportion of subjects endorsing each item. (Also called the item difficulty.) It has been shown that the maximum value of the phi coefficient is dependent on the difficulty of items that are correlated. In fact, a maximum theoretical value of ±1 is possible only when item difficulties are identical.

The maximum possible value of the phi coefficient, γ_{max}, for two items (i) and (j), where $P_i \geq P_j$, is given by Lord and Novick (50, p. 347) as:

$$\gamma_{max} = \sqrt{P_j(1 - P_i)/P_i(1-P_j)} \tag{4.2}$$

This property has obvious implications when interpreting a set of phi coefficients. That is, it is necessary to remember when reviewing such values that many coefficients, while appearing low in absolute

value, may in fact be the maximum possible value that is attainable based on the difference in difficulties between action items. Ferguson (51), Guilford (52), and Carroll (53) have pointed out that the use of phi coefficients in principal factor analysis results in the emergence of spurious difficulty factors. These factors do not represent any valid construct, but reflect the differences in difficulty levels in items comprising the correlation matrix.

Results for 1979 Survey

Table 4.1 presents the results of calculation of the interaction phi coefficients for data collected in 1979. These coefficients were developed from the "Proc Corr" sub program available from the Statistical Analysis System (SAS).

Note that except for all but one correlation, every phi coefficient is positive. In addition, it can be observed that no correlation exceeds 0.500 in absolute value. In fact, many coefficients are less than 0.200 in absolute value with several correlations close to 0.000. Such results are not surprising, given that the disparity in difficulty levels between items limits the maximum possible value of a correlation to something less than ±1.000 in absolute value.

Table 4.1 has been arranged in such a manner that actions with high correlations can be found clustered together. This results in the placement of higher valued coefficients close to the main diagonal of the matrix with decreasing values found removed from the main diagonal. This rearrangement facilitates interpretation of the relationship between actions by placing highly correlated actions next to each other.

TABLE 4.1

MATRIX OF INTERACTION PHI COEFFICIENTS, 1979

Action	17	16	3	18	15	8	1	5	9	10	11	13	14	12	6	7	4	2
17	1.000																	
16	0.406	1.000																
3	0.028	0.085	1.000															
18	0.031	0.044	0.391	1.000														
15	0.116	0.139	0.249	0.287	1.000													
8	0.110	0.148	0.201	0.168	0.228	1.000												
1	0.105	0.191	0.086	0.146	0.156	0.216	1.000											
5	0.131	0.201	0.030	0.045	0.133	0.271	0.326	1.000										
9	0.107	0.189	0.044	0.062	0.099	0.248	0.281	0.302	1.000									
10	0.151	0.194	0.060	0.073	0.129	0.358	0.274	0.336	0.432	1.000								
11	0.087	0.136	0.076	0.143	0.142	0.287	0.242	0.263	0.390	0.428	1.000							
13	0.155	0.187	0.071	0.101	0.124	0.218	0.164	0.174	0.175	0.233	0.218	1.000						
14	0.073	0.124	0.110	0.151	0.086	0.164	0.139	0.142	0.160	0.193	0.162	0.202	1.000					
12	0.107	0.146	0.028	0.066	0.107	0.131	0.154	0.174	0.163	0.148	0.206	0.198	0.155	1.000				
6	0.165	0.142	0.032	0.038	0.022	0.111	0.061	0.060	0.095	0.112	0.115	0.130	0.071	0.151	1.00			
7	0.090	0.131	0.114	0.107	0.125	0.158	0.066	0.120	0.064	0.095	0.118	0.186	0.144	0.088	0.209	1.000		
4	0.133	0.111	0.146	0.139	0.137	0.150	0.079	0.060	0.010	0.096	0.060	0.118	0.068	0.084	0.105	0.112	1.000	
2	0.064	0.166	0.121	0.078	0.152	0.186	0.117	0.138	0.119	0.190	0.124	0.147	0.052	0.053	-0.004	0.051	0.092	1.000

More careful inspection of Table 4.1 shows that several clusters of items emerge. Table 4.2 outlines clusters of variables presented by the matrix of phi coefficients for the 1979 sample.

Note that action 2, "carpool to work," and action 4, "walk or bike to work," do not appear in the table. This is suggested for two reasons. First, a generally low set of correlation coefficients are indicated. Second, these actions seem to be related to several clusters to about the same intensity.

Several clusters emerge based on the content of the action items. These include: (1) actions related to replacement or purchase of autos; (2) actions related to use of non-auto modes; (3) actions dealing with the use of an auto; (4) shopping related actions; (5) actions related to recreation behavior; (6) actions related to moving behavior.

It is also interesting to observe that the difference in difficulty levels between actions also is reflected by the placement of actions in the table. For example, most actions with similar difficulty levels exhibit generally higher inter-correlations. This points out the problem in understanding the results emerging from a factor analysis investigation using phi coefficients described in the previous section.

Results for 1980 Survey

Table 4.3 presents the interaction phi coefficients developed from the results of the data collected in 1980. This table has also been arranged so that, where possible, actions most highly inter-correlated appear next to each other. Note that these results also reflect the disparity in difficulty levels between actions.

TABLE 4.2

ACTION CLUSTERS SUGGESTED BY INTERPRETATION OF INTERACTION PHI COEFFICIENT MATRIX, 1979

Cluster	Actions in Cluster
1.	16. Buy fuel efficient car
	17. Sell car
2.	3. Bus or subway to work
	18. Bus or subway for non-work
	15. Use train, bus or plane for vacation
3.	1. Drive slower
	5. Tune car
4.	8. Shop on way home from work
	9. Shop less frequently
	10. Combine shop and other
	11. Shop close to home
5.	13. Vacation closer to home
	14. Cancel vacation trip
	12. Eliminate use of recreation vehicles
6.	6. Move closer to work
	7. Job closer to home

TABLE 4.3

MATRIX OF INTERACTION PHI COEFFICIENTS, 1980

Action	4	7	6	2	11	9	10	1	8	5	12	13	14	16	17	3	18	15
4	1.000																	
7	0.171	1.000																
6	0.080	0.324	1.000															
2	0.107	0.047	0.070	1.000														
11	0.108	0.143	0.147	0.156	1.000													
9	0.061	0.135	0.103	0.179	0.412	1.000												
10	0.120	0.114	0.116	0.187	0.386	0.494	1.000											
1	0.075	0.104	0.089	0.228	0.327	0.373	0.360	1.000										
8	0.169	0.144	0.106	0.269	0.265	0.301	0.332	0.258	1.000									
5	0.076	0.115	0.139	0.185	0.243	0.261	0.286	0.366	0.231	1.000								
12	0.099	0.130	0.049	0.110	0.220	0.217	0.230	0.202	0.210	0.235	1.000							
13	0.073	0.118	0.112	0.151	0.218	0.202	0.211	0.231	0.196	0.198	0.287	1.000						
14	0.053	0.070	0.082	0.143	0.233	0.192	0.304	0.186	0.198	0.186	0.201	0.197	1.000					
16	0.103	0.129	0.063	0.155	0.146	0.195	0.228	0.207	0.192	0.198	0.198	0.203	0.174	1.000				
17	0.061	0.081	0.076	0.056	0.116	0.088	0.080	0.119	0.084	0.092	0.135	0.100	0.169	0.173	1.000			
3	0.159	0.083	0.058	0.110	0.067	0.059	0.071	0.097	0.174	0.043	0.090	0.099	0.024	0.022	0.071	1.000		
18	0.095	0.051	0.047	0.081	0.178	0.111	0.085	0.125	0.139	0.096	0.091	0.137	0.139	0.024	0.098	0.329	1.000	
15	0.108	0.110	0.080	0.130	0.159	0.143	0.180	0.184	0.189	0.146	0.169	0.133	0.117	0.122	0.048	0.210	0.304	1.000

Once again, phi coefficients are very low in absolute value. However, several clusters of action items can be observed from these results. Table 4.4 suggests that these consist of the following: (1) actions related to moving; (2) shopping related actions combined with vehicle use and maintenance of autos; (3) use of non-auto modes; (4) recreation or vacation related behavior; and (5) purchase or elimination of vehicles.

The results of the analysis of the 1980 data are very similar to the examination of the 1979 data. Generally speaking, however this matrix is characterized by less well-defined clusters with respect to use of the auto and maintenance than was previously found. The 1980 data seem to suggest one large underlying factor with several smaller, less dominating factors describing the interrelationship between action items.

Summary

Grouping clusters of action items based on phi coefficients suggests the presence of several major groups of variables in each survey. The use of principal factors analysis in the next section represents a more objective approach to clustering action items. The information in this section, however, does provide a basis for examining the factor analysis results.

Factor Analysis of Interaction Phi Coefficients

The next step in the application of conventional market segmentation procedures involves the factor analysis of interaction phi coefficient matrices developed from the data collected in 1979 and 1980.

TABLE 4.4

ACTION CLUSTERS SUGGESTED BY INTERPRETATION OF INTERACTION PHI COEFFICIENT MATRIX, 1980

Cluster	Actions in Cluster
1.	7. Job closer to home
	6. Move closer to work
2.	11. Shop closer to home
	9. Shop less frequently
	10. Combine shop and other
	1. Drive slower
	8. Shop on way home from work
	5. Tune car
3.	3. Bus or subway to work
	18. Bus or subway for non-work
	15. Use train, bus, plane for vacation
4.	12. Eliminate use of recreation vehicles
	13. Vacation closer to home
	14. Cancel vacation trip
5.	16. Buy fuel efficient car
	17. Sell car

Factor analysis is used in this section for two main reasons: (1) to help identify the common factors describing the relationship between action clusters identified earlier and (2) to reduce the dimensionality of the action response data for use in cluster analysis.

Factor analysis is a statistical procedure that is used to investigate the interrelationship between observed variables and unseen constructs or factors. These factors are hypothesized to represent an unseen process or phenomenon which contribute to explanation of some observed behavior. The behavior of interest in this case consists of a set of responses offered by a subject regarding energy conservation actions. It is suggested that the interpretation of these responses is enhanced through developing an understanding of common factors underlying these actions.

Factor analysis theory is a basis for constructing a new set of measures which combine all common variance shared by a set of responses. These new factor scores are then used in further analysis. This results in reducing the number of variables and enhances interpretability.

A brief discussion of the theory behind factor analysis is contained below. Additional details can be found in any number of references including Lawley and Maxwell (54), and Harman (55).

Factor Analysis Theory

Given a set of n measures, denoted $x_1, \ldots, x_n$, the linear factor analysis model maps each of these observable response variables into a new set of variables defined as the product of common factor loadings a_{jr}, $(j = 1, \ldots, n)$, $(r = 1, \ldots, k)$, $(k < n)$, multiplied

by each respective common factor F_r, plus the product of unique factor loadings u_j multiplied by a unique factor Y_j. Mathematically, this is expressed:

$$X_j = a_{j1}F_1 + a_{j2}F_2 + \ldots + a_{jk}F_k + \ldots + u_jY_j \quad . \qquad (4.3)$$

The common factors account for the variance shared by the j variables in the data set, while the unique factor represents the error variance of a measure.

In practice, the principal factor solution is found by identifying the values of the common factor loadings which maximize the common variance of the observed variables. The mathematical solution to this problem is based on transforming the original set of variables into a new set of uncorrelated variables.

Normally, the set of response variables are expressed in standardized form with variance equal to 1 and mean 0. The total variance of a set of n standardized variables is, therefore, equal to the total number of variables or n. In factor analysis each part of the total variance of any variable is hypothesized to consist of a common portion and unique portion. Examining equation (4.3) it is possible to show, Harman (55, p. 18), that the sum of the squares of the common factor loadings are equal to the total common variance. In addition, it can be shown that the sum of the squares of common factor loadings for any factor k is equal to the total variance explained by that common factor (55, p. 18).

The principal factor solution identifies the set of common factors k, ($k < n$), which maximize the values of the common factor loadings.

The first factor represents the linear transformation of the original variables that maximize the overall variance. The second factor, uncorrelated with the first, maximizes the remaining variance. The third uncorrelated with the first and second, maximizes the remaining unexplained variance. The process is continued until all the common variance is accounted for by the factors.

Since the common and unique variance are unknown in practice, it is necessary to provide some estimate of these values for the set of variables. Typically, estimates of the common factor variance (communalities) are placed along the diagonal of the correlation matrix of original variables. This matrix then serves as a basis for defining a new set of common factors to describe the variables.

It can be shown, Tatsuoka (56), that the total variance explained by one factor is equal to the eigenvalue associated with that dimension. A reduced correlation matrix, such as the one described above, will have k such eigenvalues which, in total, correspond to the total variance contributed by all common factors.

Because most such matrices are only sample data, it is unlikely that all eigenvalues will be greater than zero. Therefore, the interpretation of the set of eigenvalues must include the possibility that small positive or negative eigenvalues may appear. These values, however tend to cancel each other out by summing to zero and, therefore, only the dimensions which are included in common factor space remain for evaluation.

In matrix terms, the factor model is given as:

$$Z = AF + U \quad , \tag{4.4}$$

where **Z** is a column vector of order **n** containing observed variables, **A** is a (n x k) matrix of common factor loadings, **F** is a column vector of order k containing common factors and **U** is column vector of unique factors of order n.

The elements of the vector of unique factors have mean 0 and variance μ_j^2. These elements are also independent of every other common factor element and each other unique factor element. The elements of the vector of observed variables are standardized with mean 0 and variance equal to 1.

If the (n x n) positive definite correlation matrix of observed variables **R** is formed, then it can be shown, Lawley and Maxwell (54, p. 5), that such a matrix can be reproduced from the product of common factor loadings **A**, the (k x k) matrix of common factor intercorrelations Φ, and the (n x n) matrix of unique factor loadings μ. In matrix terms this is defined as:

$$R = A\Phi A' + U \quad . \tag{4.5}$$

When communalities are placed along the diagonal of the correlation matrix, then the reduced correlation matrix R is defined as:

$$R = A\Phi A' \quad . \tag{4.6}$$

In this case, the reproduced correlation matrix is described by only the matrix of common factor loadings and the matrix of common factor intercorrelations since the unique portion of the variance has been excluded.

The matrix of common factors can be rotated in common factor space to enhance the interpretability of the solution. This process of attaining simple structure may result in the development of correlated or uncorrelated common factors. In any event, the result is still a valid measure of common factor space in the sense that the total amount of the variance explained remains unchanged.

If an orthogonal transformation is desired, then the intercorrelation matrix of factors is a simple (k x k) identity matrix. In addition, the correlation matrix describing the relationship between factors and observed variables is identical to the matrix of factor loadings. In an oblique rotation, these matrices are not the same, in which case a factor structure matrix **S** is developed to describe the correlation between factors and variables. Such a matrix is found by multiplying the matrix of common factors by factor intercorrelations:

$$\mathbf{S} = \mathbf{A}\Phi \tag{4.7}$$

Besides developing an understanding of the number of common factors in a data set, factor analysis is used to develop a set of new variables or, factor scores, for use in additional analysis. Such scores can have several desirable properties that make them more informative or easier to use than the original response variables. For example, these scores may be uncorrelated with each other while explaining a maximum amount of common variance and reducing the number of variables to be analyzed.

Factor scores are found by developing a (n x n) matrix of factor score coefficients, **B**, from the factor structure matrix, **S**, and the

inverse of the correlation matrix of original variables, R, (55, p. 20):

$$B = R^{-1}S \quad . \tag{4.8}$$

This matrix is used in conjunction with the matrix of standardized original variables to develop the factor score matrix F (55, p. 20):

$$F = ZB \quad . \tag{4.9}$$

Factor scores developed in this manner have mean zero.

This overview of the theory of the classical factor analysis has been brief. However, it is sufficient enough to serve as a basis for interpretation of the factor analysis results described in this chapter. It is shown that factor analysis is designed to identify the number of common factors contained in a matrix of observed interaction correlations with estimates of communalities placed along the diagonal. The results of the factor analysis are used to investigate the dimensionality of the data by indicating clusters of variables. The final output of this process includes estimates of factor scores that are used in later cluster analysis of subjects.

1979 Factor Analysis Results

Overview

The principal factor analysis solution was developed using the BMDP4M program from the library of the Biomedical Programs (BMDP) statistical sofware. Communality estimates were developed from the squared multiple correlation of each variable with all other variables in the

matrix of interaction phi coefficients. Final analysis indicated that the 18 original action items could be explained by eight common factors. The matrix of factor loadings was rotated using the direct Oblimin method which resulted in a set of final factor loadings that exhibited a pattern which enhanced the interpretation of variable clusters. Because the final pattern matrix of factor loadings was developed from an oblique rotation, factors turned out to be intercorrelated, to a certain extent, in common factor space. In addition, a matrix of factor and variable correlations, or structure matrix, was developed to estimate factor score coefficients. This latter matrix is used to develop factor scores for each subject in the data for later use in cluster analysis.

Eigenvalue Configuration

Table 4.5 outlines the squared multiple correlations (SMCs) between actions for both survey years. Many are quite small. Note that in both survey years actions 2, "carpool to work," and 4 "walk or bike to work," are the lowest. This is consistent with the earlier difficulty in identifying cluster membership for these actions.

The SMCs were placed along the diagonal of the corelation matrix and the set of eigenvalues and eigenvectors that maximized the common variance were extracted. After some preliminary analysis, an eight factor solution was found reasonable. This is confirmed from reviewing Table 4.6 containing the final eigenvalues, proportion total variance explained, and the total common variance explained for the 1979 data.

TABLE 4.5

SQUARED MULTIPLE CORRELATIONS OF EACH ACTION WITH ALL OTHER ACTIONS

	Action	Year	
		1979	1980
1.	Drive slower	0.18965	0.28217
2.	Carpool to work	0.09075	0.12430
3.	Bus or subway to work	0.20115	0.15654
4.	Walk or bike to work	0.07717	0.07887
5.	Tune car	0.22636	0.21116
6.	Move closer to work	0.09676	0.12979
7.	Job closer to home	0.10708	0.15405
8.	Shop on way home from work	0.24006	0.23302
9.	Shop less frequently	0.27998	0.34322
10.	Combine shop and other	0.34818	0.34794
11.	Shop close to home	0.28022	0.27892
12.	Eliminate use of recreation vehicle	0.10888	0.17678
13.	Vacation closer to home	0.15872	0.16472
14.	Cancel vacation trip	0.10596	0.13955
15.	Use train, bus, or plane for vacation	0.16261	0.15902
16.	Buy fuel efficient car	0.23658	0.14322
17.	Sell car	0.19341	0.06966
18.	Bus or subway for non-work	0.21889	0.19598

TABLE 4.6

EIGENVALUE CONFIGURATION RESULTING FROM PRINCIPAL FACTOR SOLUTION BASED ON EIGHT COMMON FACTORS DEVELOPED FROM MATRIX OF PHI COEFFICIENTS, 1979

Eigenvalue		Cumulative Proportion of Variance in Data Space	Incremental Contribution of Each Factor
1.	2.9664	0.4848	0.4848
2.	0.9596	0.6416	0.1568
3.	0.7386	0.7623	0.1207
4.	0.4348	0.8334	0.0711
5.	0.2518	0.8745	0.0411
6.	0.2422	0.9141	0.0396
7.	0.1872	0.9447	0.0306
8.	0.1368	0.9671	0.0224
9.	0.0894		
10.	0.0601		
11.	0.0447		
12.	0.0074		
13.	-0.0023		
14.	-0.0110		
15.	-0.0134		
16.	-0.0424		
17.	-0.0572		
18.	-0.0798		

Eight factors explain nearly 97 percent of the total variance. In addition, the values of the remaining eigenvalues virtually cancel each other out. The sum of the remaining positive values is nearly equal to the remaining negative values. In theory, negative eigenvalues should never be found in any factor solution since these values should be equal to a set of squared common factor loadings.

Rotation of Common Factor Loadings

Table 4.7 contains the set of unrotated factor loadings for the eight common factor solution. Note that the sum of squares of loadings on any factor equals the total variance explained by that factor. In addition, the sum of the squares across any row is equal to the final communality estimate for a variable.

In order to develop a better understanding of the relationship between factors and variables, the matrix of factor loadings is rotated. This process does not change the amount of variance explained by the solution, but enhances the interpretation of the results.

In this case, an oblique rotation of the common factors was performed using the criterion developed by Jenrich and Sampson (57). The mathematical form of the objective function for a set of final factor loadings b_{jp}, b_{jq} $(j = 1, \ldots, n)$, $(p = 1, \ldots, k)$, $(q = 1, \ldots, k)$, $(k < n)$, is given for a matrix of factors loadings **A** as:

$$F(A) = \sum_{p<q=1}^{k} \left(\sum_{j=1}^{n} b_{jq}^2 \, b_{jq}^2 - \frac{\Gamma}{n} \sum_{j=1}^{n} b_{jp}^2 \, b_{jq}^2 \right) , \qquad (4.10)$$

Γ in this case was set equal to zero.

TABLE 4.7

UNROTATED FACTOR LOADINGS DEVELOPED FROM MATRIX
OF INTERACTION PHI COEFFICIENTS, 1979

Action	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7	Factor 8
1	0.456	-0.069	-0.095	-0.114	0.190	0.072	-0.095	0.005
2	0.287	0.054	-0.018	-0.160	-0.161	0.126	0.104	0.045
3	0.270	0.515	-0.079	-0.067	-0.044	-0.054	-0.031	-0.050
4	0.229	0.189	0.127	0.033	-0.094	0.060	-0.014	0.103
5	0.503	-0.208	-0.086	-0.106	0.134	0.238	-0.175	-0.037
6	0.245	-0.014	0.264	0.337	-0.044	-0.105	-0.189	0.053
7	0.285	0.130	0.158	0.278	-0.068	0.118	-0.091	-0.148
8	0.527	0.071	-0.116	0.011	-0.189	0.106	-0.044	0.035
9	0.523	-0.267	-0.189	-0.034	0.039	-0.175	-0.021	-0.053
10	0.617	-0.251	-0.192	-0.024	-0.191	-0.098	0.022	-0.021
11	0.543	-0.149	-0.211	0.097	-0.013	-0.155	0.033	0.065
12	0.336	-0.059	0.077	0.178	0.214	0.037	0.104	0.197
13	0.427	-0.005	0.099	0.162	-0.011	0.119	0.217	-0.009
14	0.335	0.052	-0.011	0.147	0.112	0.024	0.166	-0.173
15	0.352	0.313	-0.020	-0.114	-0.003	0.061	-0.017	0.104
16	0.440	-0.060	0.463	-0.247	0.034	-0.057	0.042	-0.079
17	0.326	-0.046	0.471	-0.154	-0.011	-0.115	-0.013	0.030
18	0.313	0.565	-0.149	-0.014	0.144	-0.145	-0.012	-0.020
Total Variance	2.966	0.960	0.739	0.435	0.252	0.242	0.187	0.137

The final results of the oblique rotation are found in Table 4.8 for the 1979 data. Most loadings, in this case, are near zero in absolute value. However, variables are shown to load high on only one (or two) factors which has greatly enhanced the interpretation of the result.

Table 4.9 presents the factor correlations for the rotated factors. For the most part, these values are low indicating that final factors remain independent. In addition, Table 4.10 outlines the matrix of correlations between factors and variables or factor structure matrix. This result is generally consistent with the interpretation of factor loadings described earlier. Most variables are highly correlated with the same factor suggested by the final rotated factor loading matrix.

1980 Factor Analysis Results

Overview

An identical procedure was followed to develop the principal factor solution from the 1980 data matrix as was used for the 1979 data matrix. The results of this effort were very similar to this earlier analysis. The survey results are compared in a later portion of this discussion.

Eigenvalue Configuration

Estimates of SMCs outlined in Table 4.5 provided the basis for communality estimates. It is again interesting to note that Action 4, "walk or bike to work," was again very low. However, Action 17, "sell car" was even lower than Action 4 reflecting a change in the relationship between action endorsement.

TABLE 4.8

ROTATED FACTOR LOADINGS DEVELOPED FROM MATRIX
OF INTERACTION PHI COEFFICIENTS, 1979

Action	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7	Factor 8
1	0.086	0.102	0.056	(0.412)	-0.051	-0.059	0.083	-0.002
2	0.081	0.034	0.078	0.043	-0.090	0.318	-0.012	0.036
3	0.007	(0.571)	0.012	-0.031	0.039	0.063	-0.097	0.031
4	-0.041	0.136	0.071	-0.012	0.133	(0.182)	0.085	-0.040
5	0.035	-0.087	0.012	(0.616)	0.050	0.031	0.018	0.002
6	0.106	-0.024	0.095	-0.051	(0.469)	-0.098	0.102	-0.072
7	-0.065	0.043	-0.001	0.093	(0.407)	0.043	-0.046	0.182
8	0.262	0.116	-0.061	0.149	0.147	(0.283)	-0.016	0.007
9	(0.552)	0.002	0.074	0.125	-0.038	-0.106	-0.013	0.037
10	(0.649)	-0.057	0.041	0.044	0.034	0.155	-0.064	0.038
11	(0.556)	0.064	-0.053	0.006	0.026	0.006	0.131	0.027
12	0.015	-0.020	0.011	0.077	0.015	-0.012	(0.438)	0.063
13	0.073	-0.053	0.067	-0.021	0.066	0.190	(0.292)	(0.288)
14	0.072	0.089	0.030	0.019	0.025	-0.044	0.065	0.356
15	-0.027	(0.356)	0.048	0.116	-0.019	0.163	0.081	-0.051
16	-0.015	-0.001	(0.681)	0.068	-0.035	0.003	-0.041	0.078
17	0.008	-0.012	(0.600)	-0.038	0.051	-0.003	0.022	-0.050
18	0.028	(0.712)	-0.031	-0.021	-0.024	-0.111	0.039	0.053
Total Variance	1.148	1.032	0.869	0.630	0.453	0.347	0.305	0.273

TABLE 4.9

FACTOR CORRELATIONS FOR ROTATED FACTORS, 1979

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7	Factor 8
Factor 1	1.000							
Factor 2	0.148	1.000						
Factor 3	0.268	0.149	1.000					
Factor 4	0.599	0.218	0.330	1.000				
Factor 5	0.130	0.187	0.295	0.077	1.000			
Factor 6	0.204	0.347	0.276	0.281	0.150	1.000		
Factor 7	0.346	0.175	0.341	0.301	0.326	0.097	1.000	
Factor 8	0.312	0.215	0.156	0.321	0.226	0.162	0.268	1.000

TABLE 4.10

CORRELATIONS OF VARIABLES WITH FACTORS, 1979

Action	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7	Factor 8
1	0.372	0.197	0.227	(0.508)	0.045	0.125	0.250	0.189
2	0.193	0.167	0.181	0.215	0.005	(0.371)	0.073	0.123
3	0.070	(0.586)	0.089	0.103	0.132	0.259	0.027	0.141
4	0.062	0.232	0.188	0.091	0.219	(0.258)	0.166	0.066
5	0.414	0.078	0.242	(0.641)	0.100	0.194	0.224	0.216
6	0.151	0.051	0.238	0.055	(0.505)	-0.004	0.275	0.072
7	0.100	0.175	0.163	0.151	(0.446)	0.156	0.152	0.288
8	0.425	0.303	0.193	0.399	0.235	(0.424)	0.196	0.226
9	(0.628)	0.084	0.225	0.456	0.054	0.062	0.229	0.232
10	(0.703)	0.112	0.258	0.472	0.134	0.297	0.214	0.264
11	(0.613)	0.176	0.165	0.387	0.144	0.150	0.334	0.250
12	0.232	0.090	0.197	0.235	0.177	0.064	(0.487)	0.209
13	0.277	0.139	0.257	0.244	0.243	0.275	(0.352)	0.402
14	0.233	0.185	0.142	0.215	0.156	0.083	0.221	(0.424)
15	0.151	(0.441)	0.192	0.245	0.107	0.324	0.181	0.106
16	0.214	0.118	(0.688)	0.294	0.173	0.210	0.216	0.183
17	0.142	0.072	(0.601)	0.156	0.220	0.151	0.219	0.048
18	0.117	(0.682)	0.060	0.137	0.110	0.136	0.152	0.190

Table 4.11 presents the final eigenvalues from an eight factor solution. Once again it appeared that this solution best described the common factors present in the data.

Rotation of Common Factor Loadings

Table 4.12 contains the unrotated factor loadings for the 1980 data. These results are very similar to the earlier analysis using the 1979 data. In this case, however, the first few factors seem to explain a higher proportion of common variance. Generally speaking, these results support the existence of one very large common factor relating the entire data set.

Once again, an oblique rotation was performed using the direct Oblimin method. The final results of this effort are contained in Table 4.13. Except for changes in the factor number describing a cluster of variables, these 1980 factors appear to be comprised of the same variables as the 1979 rotated solution. These results are compared in the following section. Table 4.14 outlines the factor correlations and Table 4.15 presents the correlations of variables with factors for the 1980 solution.

Comparison of 1979 and 1980 Results

Table 4.16 suggests the relationship between action variables and common factors developed from the rotated factor solutions for 1979 and 1980. The first factor contains actions related to shopping behavior. These actions are among the easiest to implement and were earlier identified as a cluster set for both 1979 and 1980.

TABLE 4.11

EIGENVALUE CONFIGURATION RESULTING FROM PRINCIPAL FACTOR SOLUTION BASED ON EIGHT COMMON FACTORS DEVELOPED FROM MATRIX OF PHI COEFFICIENTS, 1980

Eigenvalue		Cumulative Proportion of Total Variance in Data Space	Incremental Contribution of Each Factor
1.	3.2162	0.5143	0.5327
2.	0.8305	0.6471	0.1328
3.	0.6130	0.7451	0.0980
4.	0.4391	0.8153	0.0702
5.	0.3195	0.8664	0.0511
6.	0.2476	0.9060	0.0396
7.	0.2171	0.9407	0.0347
8.	0.1552	0.9655	0.0248
9.	0.0809		
10.	0.0612		
11.	0.0451		
12.	0.0264		
13.	0.0019		
14.	0.0002		
15.	-0.0167		
16.	-0.0507		
17.	-0.0691		
18.	-0.0879		

TABLE 4.12

UNROTATED FACTOR LOADINGS DEVELOPED FROM MATRIX
OF INTERACTION PHI COEFFICIENTS, 1980

Factor	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7	Factor 8
1	0.575	-0.138	-0.124	-0.048	0.025	0.268	-0.116	0.130
2	0.358	0.006	-0.058	0.019	0.281	0.119	0.088	-0.113
3	0.238	0.458	-0.055	-0.064	0.126	-0.047	-0.016	0.096
4	0.229	0.158	0.137	-0.014	0.176	-0.100	0.047	0.079
5	0.482	-0.115	-0.000	0.060	0.026	0.204	-0.138	-0.012
6	0.251	0.059	0.456	-0.132	-0.086	0.123	0.010	-0.107
7	0.296	0.097	0.531	-0.090	-0.024	-0.041	-0.045	0.053
8	0.524	0.031	-0.019	-0.044	0.251	-0.086	0.100	-0.098
9	0.601	-0.236	-0.115	-0.241	-0.113	-0.117	0.037	0.056
10	0.610	-0.217	-0.094	-0.182	-0.002	-0.127	0.030	0.037
11	0.555	-0.088	-0.041	-0.124	-0.172	-0.047	0.047	-0.057
12	0.447	-0.019	0.018	0.323	-0.039	-0.209	-0.220	-0.017
13	0.420	0.013	0.024	0.191	-0.031	-0.027	-0.098	-0.077
14	0.388	-0.025	-0.027	0.210	-0.124	0.033	0.175	-0.162
15	0.355	0.280	-0.061	-0.019	0.018	0.001	-0.086	-0.004
16	0.375	-0.097	0.067	0.212	0.077	0.002	0.078	0.110
17	0.233	0.052	0.073	0.247	-0.123	0.056	0.236	0.180
18	0.319	0.590	-0.209	-0.060	-0.200	0.053	0.014	-0.033
Total Variance	3.216	0.830	0.613	0.439	0.319	0.248	0.217	0.155

TABLE 4.13

ROTATED FACTOR LOADINGS DEVELOPED FROM MATRIX
OF INTERACTION PHI COEFFICIENTS, 1980

Factor	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7	Factor 8
1	0.214	0.073	-0.006	-0.009	0.057	(0.498)	0.086	-0.070
2	-0.008	0.019	-0.010	-0.021	(0.462)	0.107	0.004	0.029
3	-0.037	(0.481)	0.001	-0.004	0.090	-0.026	0.020	-0.164
4	0.026	0.096	0.117	0.035	(0.154)	-0.090	0.065	-0.179
5	0.042	0.001	0.094	0.156	0.101	(0.368)	0.014	0.028
6	-0.008	-0.008	(0.585)	-0.060	0.019	0.042	-0.016	0.091
7	0.027	-0.012	(0.581)	0.072	-0.065	-0.033	0.035	-0.128
8	0.236	0.057	0.028	0.070	(0.417)	-0.068	-0.022	-0.027
9	(0.740)	-0.014	-0.013	-0.029	-0.041	0.041	0.001	-0.021
10	(0.642)	-0.039	-0.019	0.032	0.074	0.030	-0.010	-0.055
11	(0.500)	0.080	0.086	0.039	-0.016	0.021	0.023	0.127
12	0.020	0.009	-0.027	(0.647)	-0.059	-0.001	0.003	-0.021
13	0.010	0.055	0.055	(0.358)	0.059	0.072	0.025	0.080
14	0.085	0.013	0.012	0.141	0.155	-0.043	0.204	(0.291)
15	0.032	(0.359)	0.034	0.123	0.050	0.070	-0.035	-0.024
16	0.043	-0.106	0.005	0.155	0.125	0.078	(0.270)	-0.076
17	-0.010	0.032	0.009	-0.024	-0.044	0.009	(0.492)	0.013
18	0.031	(0.741)	-0.027	-0.025	-0.080	0.008	0.048	0.147
Total Variance	1.327	0.950	0.717	0.651	0.503	0.433	0.376	0.230

TABLE 4.14

FACTOR CORRELATIONS FOR ROTATED FACTORS, 1980

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7	Factor 8
Factor 1	1.000							
Factor 2	0.229	1.000						
Factor 3	0.296	0.199	1.000					
Factor 4	0.514	0.252	0.281	1.000				
Factor 5	0.463	0.303	0.250	0.424	1.000			
Factor 6	0.472	0.062	0.107	0.322	0.302	1.000		
Factor 7	0.322	0.198	0.269	0.470	0.309	0.177	1.000	
Factor 8	0.067	-0.067	-0.101	0.006	-0.128	0.161	0.078	1.000

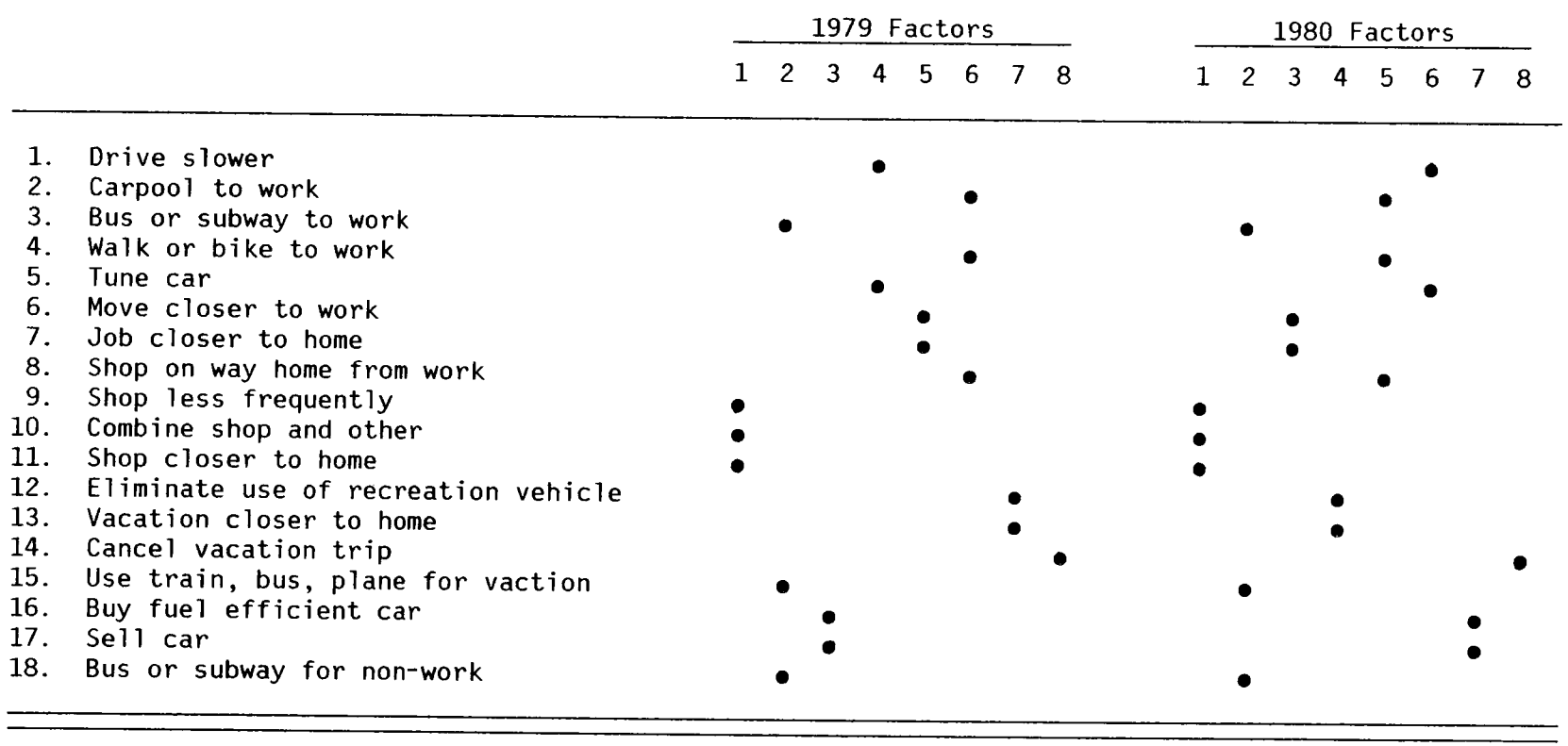
TABLE 4.15

CORRELATIONS OF VARIABLES WITH FACTORS, 1980

Action	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7	Factor 8
1	0.509	0.189	0.168	0.343	0.359	(0.622)	0.264	0.020
2	0.251	0.156	0.111	0.210	(0.483)	0.242	0.158	-0.014
3	0.096	(0.512)	0.126	0.137	0.236	-0.011	0.112	-0.213
4	0.138	0.200	0.218	0.171	(0.255)	-0.019	0.158	-0.224
5	0.377	0.123	0.216	0.372	0.322	(0.485)	0.225	0.070
6	0.162	0.090	(0.560)	0.113	0.129	0.098	0.129	0.034
7	0.190	0.121	(0.610)	0.223	0.137	0.030	0.196	-0.179
8	0.445	0.254	0.222	0.358	(0.556)	0.190	0.221	-0.083
9	(0.718)	0.138	0.192	0.341	0.297	0.363	0.212	0.042
10	(0.685)	0.138	0.197	0.382	0.381	0.350	0.223	-0.013
11	(0.582)	0.214	0.252	0.352	0.275	0.303	0.250	0.153
12	0.319	0.156	0.151	(0.629)	0.224	0.195	0.289	-0.005
13	0.298	0.181	0.191	(0.453)	0.265	0.236	0.259	0.079
14	0.300	0.136	0.139	0.341	0.274	0.174	0.367	(0.285)
15	0.232	(0.418)	0.163	0.268	0.248	0.156	0.140	-0.046
16	0.276	0.045	0.160	0.356	0.297	0.215	(0.384)	-0.048
17	0.131	0.109	0.127	0.198	0.106	0.076	(0.476)	0.055
18	0.172	(0.712)	0.101	0.162	0.140	0.066	0.174	0.117

TABLE 4.16

ACTION ITEMS COMPRISING FACTORS



The second factor is related to non-auto travel comprising behavior related to work, non-work, and vacation mode choice. These were also suggested as a cluster set of variables in previous discussion.

The third factor in 1980 and the fifth factor for the 1979 data consist of moving related behavior. These are examples of the relationship between manifest behavior and opinions which underly the response patterns. It is unlikely that a subject would actually behave in a manner that involved moving for the reason of being close to work, and also finding a job close to home. However, as an opinion these actions are both related to some notion of changing residences to conserve energy. The results indicate, perhaps, that such a subject feels that only such severe actions will result in savings in the face of changing fuel price and availability. It indicates, in part, that these actions are a reflection of an attitude or opinion toward the seriousness of the energy problem.

The fourth factor in the 1980 data and the seventh factor in 1979 appear to consist of actions related to recreation. These factors, however, tend to be closely related to the eighth factor which also includes some action related to this behavior. This latter factor is much more drastic in the sense that it involves a disruptive change, cancelling a vacation trip. This may explain a slightly different perspective by the subjects.

The fifth factor in 1980 and the sixth factor in 1979 suggest actions related to work travel behavior. The interpretation of the relationship of Action 8, "shop on the way home from work," is unclear. This may be related to the overriding concern with work travel, but

also reflects, in part, the notion of opinion based responses. For example, it is generally agreed that carpooling can be disrupted by side trips not related to work. Stopping to shop on the way home from work is inconsistent with this finding. Such results, however, may be explained by the common relationship each action has to work related travel.

Factor 6 for 1980 and Factor 4 for 1979 consist of vehicle related actions. These seem to be consistent reflections of the desire to conserve fuel through proper vehicle utilization.

Finally, Factor 7 for 1980 and Factor 3 for 1979 are related to replacing an auto. These actions comprising this factor are consistent in general content, but may also be a reflection of the opinion, as opposed to behavior, of an individual subject.

The outline of the factor analysis results thus far has relied on a subjective interpretation of the rotated factor results. An additional step to confirm these results includes pair wise comparison of each set of factor loadings through calculation of the root mean square deviation (rms) between factor loadings and calculation of the degree of factor similarity.

Given factor p consisting of a set of a_{jp} factor loadings, (j = 1, . . ., n) from one sample and factor q consisting of a_{jq} factor loadings from another sample, Harman (55, p. 346) states that the root mean square deviation describing the relationship between these factors is:

$$rms_{pq} = \left[\sum_{j=1}^{n} \frac{(a_{jp} - a_{jq})^2}{n} \right]^{\frac{1}{2}} . \qquad (4.11)$$

Besides calculation of the root mean squared deviation, Harman (55, p. 344) suggests that the degree of factorial similarity ρ, ($-1 \leq \rho \leq +1$), can be calculated to compare factors across samples. Harman states that this value is calculated for a set of factor pairs, p, q, for each loading in the first sample, a_{jp}, and the second sample, a_{jq} as follows:

$$\rho_{pq} = \frac{\sum_{j=1}^{n} a_{jp} \cdot a_{jq}}{\sqrt{\sum_{j=1}^{n} a_{jp}^2 \sum_{j=1}^{n} a_{jp}^2}} \qquad (4.12)$$

The results of calculating the root mean square deviation for all pairs of factors for both 1979 and 1980 are contained in Table 4.17. Results of calculating the degree of factorial similarity are contained in Table 4.18 for these data sets.

Highlighted in Table 4.17 are the smallest values of root mean square deviation for each set of factor pairs. These results confirm the interpretation of common factor loadings suggested earlier.

Table 4.18 outlines the values of factor similarity which have the highest positive value for each factor pair. Except for Factor 8 in 1980, these results confirm the earlier interpretation.

In summary, through a series of subjective and objective interpretation, it has been shown that the common factors developed over the survey years reflect similar loadings. These results are used to develop factor scores for each subject that serve as a basis for cluster analysis.

TABLE 4.17

ROOT MEAN SQUARE DEVIATION CALCULATED FROM
ROTATED LOADING MATRICES, 1979, 1980

1979 Factor	Factor 1	2	3	4	5	6	7	8
1	(0.006)	0.111	0.097	0.092	0.072	0.076	0.082	0.069
2	0.125	(0.003)	0.094	0.092	0.074	0.085	0.072	0.070
3	0.113	0.106	0.080	0.071	0.065	0.060	(0.015)	0.067
4	0.077	0.084	0.066	0.051	0.044	(0.010)	0.051	0.052
5	0.094	0.077	(0.005)	0.055	0.049	0.053	0.042	0.041
6	0.084	0.067	0.057	0.042	(0.013)	0.043	0.041	0.040
7	0.085	0.070	0.050	(0.012)	0.046	0.034	0.035	0.024
8	0.077	0.063	0.047	0.027	0.034	0.038	0.026	(0.017)

TABLE 4.18

DEGREE OF FACTORIAL SIMILARITY CALCULATED FROM ROTATED LOADING MATRICES, 1979, 1980

1979 Factor	Factor 1	2	3	4	5	6	7	8
1	(0.961)	0.049	0.066	0.086	0.238	0.154	0.033	0.137
2	0.040	(0.965)	0.027	0.010	0.129	-0.052	0.088	-0.004
3	0.044	-0.052	0.025	0.163	0.153	0.180	(0.861)	-0.116
4	0.312	0.063	0.121	0.282	0.300	(0.843)	0.086	-0.098
5	0.054	0.018	(0.952)	0.106	0.070	-0.076	0.075	-0.082
6	0.120	0.083	0.030	0.251	(0.795)	0.000	-0.030	-0.262
7	0.077	0.001	0.129	(0.825)	-0.014	0.172	0.078	0.199
8	0.179	0.084	0.170	(0.515)	0.209	0.023	0.270	0.405

Estimation of Factor Scores

Now that the common factor space has been identified, factor scores are developed for each subject in the respective surveys. Recall that the factor scores are estimated from the matrix of factor score coefficients premultiplied by the matrix of energy action responses in standardized form. Table 4.19 outlines the matrix of factor score coefficients achieved from the 1979 survey data. Table 4.20 outlines this matrix from the 1980 data.

Table 4.21 summarizes the results of this effort by presenting the mean and standard deviation of each respective factor. Note that each factor score has a mean of zero., and the standard deviations are less than unity.

Composite variables developed from the sum of actions comprising a factor were developed for each subject. These new values were then correlated with factor scores to confirm the relationship between factor scores and action items. The results of this effort are presented in Tables 4.22 and 4.23. As expected, a high value of intercorrelation between factor scores and summated scores resulted from this effort.

Factor 8 and the associated composite variable F8 in Table 4.23 have a slightly lower correlation than Factor 8 and Variable F7. This is a result of the problem associated with identifying the appropriate placement of Action 13 "vacation closer to home." In order to yield a result consistent with the 1979 results it is suggested that this action remain with Factor 7. This is not anticipated to have a large impact on the final result because it may be argued that these results

TABLE 4.19

FACTOR SCORE COEFFICIENTS, 1979

Action	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7	Factor 8
1	0.06999	0.05142	0.03966	(0.23055)	-0.05475	-0.03461	0.07967	0.01768
2	0.01919	0.03816	0.04137	0.03907	-0.05812	(0.23154)	-0.02163	0.01686
3	-0.01665	(0.31378)	0.00727	-0.00455	0.04465	0.12322	-0.07603	0.03265
4	-0.02342	0.07210	0.05332	-0.00809	0.09941	(0.13503)	0.06558	-0.02793
5	0.07247	-0.03167	0.03309	(0.40067)	-0.00206	0.03654	0.03094	0.04063
6	0.03250	-0.01015	0.08089	-0.05260	(0.38513)	-0.08679	0.14838	-0.04660
7	-0.02904	0.03871	0.01696	0.02619	(0.30438)	0.04368	0.00221	0.17278
8	0.10285	0.09487	-0.00341	0.10569	0.10693	(0.25840)	-0.00445	0.02143
9	(0.28185)	-0.01654	0.03476	0.13791	-0.06076	-0.12930	0.02627	0.05223
10	(0.37282)	-0.03345	0.03950	0.11257	0.00474	0.15648	-0.04210	0.05759
11	(0.26133)	0.02952	-0.02057	0.05466	0.02147	-0.02795	0.15466	0.04213
12	0.03001	0.00107	0.03655	0.04472	0.05054	-0.03327	(0.34766)	0.06591
13	0.02917	-0.00215	0.05363	0.01191	0.08656	0.13696	(0.17368)	0.25387
14	0.03052	0.03886	0.00589	0.02939	0.03828	-0.03968	0.06644	(0.28999)
15	-0.01022	(0.18156)	0.04532	0.06889	-0.00209	0.16538	0.06091	-0.03918
16	-0.00116	0.01116	(0.47078)	0.08097	0.01128	0.06210	0.02405	0.06254
17	-0.00417	-0.00039	(0.35215)	-0.01425	0.08996	0.03071	0.07901	-0.07613
18	0.00765	(0.46166)	-0.02056	0.01003	0.00592	-0.04633	0.06443	0.07876

TABLE 4.20

FACTOR SCORE COEFFICIENTS, 1980

Action	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7	Factor 8
1	0.12039	0.03208	-0.00585	0.03252	0.07986	(0.45441)	0.06169	-0.02282
2	0.01250	0.02757	0.00172	0.01103	(0.28447)	0.07355	0.01415	-0.00712
3	-0.01515	(0.26365)	0.02757	0.01516	0.10485	-0.06508	0.01650	-0.22036
4	0.01099	0.06523	0.07847	0.03566	(0.110360	-0.07655	0.05503	-0.18287
5	0.04138	-0.01087	0.05517	0.09948	0.07851	(0.26512)	0.02294	0.05472
6	0.01239	0.00828	(0.37521)	-0.02383	0.01425	0.02445	0.01380	0.07777
7	0.02339	0.02315	(0.43404)	0.06808	-0.01006	-0.06860	0.07261	-0.18174
8	0.10280	0.06873	0.04331	0.07747	(0.32573)	-0.04803	0.01525	-0.09040
9	(0.36346)	-0.01683	0.00742	0.02485	-0.01076	0.07923	-0.00274	0.01187
10	(0.29818)	-0.01252	0.00948	0.07091	0.09753	0.05549	0.00165	-0.06140
11	(0.19740)	0.03710	0.07195	0.05397	0.00040	0.04026	0.03877	0.15230
12	0.04133	0.02082	-0.00168	(0.41921)	-0.00589	0.00227	0.09036	-0.03148
13	0.02559	0.02540	0.03826	(0.18902)	0.04992	0.05060	0.06174	0.06537
14	0.04621	0.00011	0.01575	0.10540	0.08718	-0.00335	0.19307	(0.27482)
15	0.01924	(0.16239)	0.03032	0.06827	0.05707	0.02851	-0.01127	-0.04813
16	0.02506	-0.03619	0.02063	0.11800	0.09617	0.03895	(0.21492)	-0.06252
17	-0.00048	0.01177	0.02776	0.03936	-0.00984	-0.01695	(0.35521)	0.03013
18	0.01915	(0.54413)	0.00029	0.00184	-0.03562	-0.01991	0.05909	0.16943

TABLE 4.21

MEAN AND STANDARD DEVIATION OF FACTOR SCORES, 1979, 1980

Factor	1979		1980	
	Mean	Standard Deviation	Mean	Standard Deviation
1	0.0000	0.8502	0.0000	0.8726
2	0.0000	0.8080	0.0000	0.8029
3	0.0000	0.7974	0.0000	0.7471
4	0.0000	0.7881	0.0000	0.7790
5	0.0000	0.6622	0.0000	0.7297
6	0.0000	0.6352	0.0000	0.7105
7	0.0000	0.6414	0.0000	0.6600
8	0.0000	0.6029	0.0000	0.5187

TABLE 4.22

CORRELATIONS BETWEEN FACTOR SCORES AND
COMPOSITE VARIABLES, 1979

Composite Variable*	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7	Factor 8
F1	(0.9756)							
F2	0.1828	(0.9578)						
F3	0.2554	0.1438	(0.9689)					
F4	0.5676	0.2100	0.3611	(0.8942)				
F5	0.1814	0.1953	0.3097	0.1798	(0.9047)			
F6	0.4497	0.4455	0.3539	0.4921	0.3526	(0.8597)		
F7	0.3892	0.1870	0.3722	0.3917	0.4155	0.3697	(0.8229)	
F8	0.2738	0.2595	0.1777	0.2727	0.2361	0.1301	0.3441	(0.7039)

*Found from sum of item scores comprising a factor.

TABLE 4.23

CORRELATIONS BETWEEN FACTOR SCORES AND
COMPOSITE VARIABLES, 1980

Composite Variable*	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7	Factor 8
F1	(0.9624)							
F2	0.2707	(0.9365)						
F3	0.2489	0.1633	(0.9629)					
F4	0.4379	0.2626	0.2880	(0.8494)				
F5	0.5017	0.3790	0.3655	0.4879	(0.8996)			
F6	0.6196	0.2378	0.3077	0.5532	0.5660	(0.9492)		
F7	0.3257	0.1087	0.2525	0.4873	0.3962	0.2941	(0.8059)	
F8	0.3443	0.1692	0.1858	0.4377	0.3756	0.2448	0.5558	(0.5487)

*Found from sum of item scores comprising a factor.

are sufficiently close to one another as to make a case for placement of Action 13 with Factor 7.

Summary of Results

The results of the principal factor solution for the matrices of interaction phi coefficients indicated that the original 18 actions span a common factor space consisting of eight factors. It appears that a consistent set of variable clusters comprise these factors for the respective years. The factor scores developed exhibit a high intercorrelation with composite variables developed from the sum of the values of the actions comprising these factors.

The next section uses these results to develop a set of groups comprised of subjects that share similar action response sets. This section applies a cluster analysis algorithm to the factor scores to develop an understanding of the nature of these groups. This information is then used to investigate the relationship between group characteristics and the actions undertaken by these groups.

Cluster Analysis of Factor Scores

Cluster analysis can be broadly defined as strategy designed to assist a researcher in the problems presented by classifying observations into groups. Typically, a set of measures is collected which describe the phenomenon of interest in an investigation. The researcher may then wish to partition the observations into a set of smaller groups in order to understand the relationship between these groups described by the attributes measured in the study.

In this particular research, cluster analysis is used to identify subjects sharing common energy conservation behavior as measured by the set of action items describing various transportation related behaviors. The measure used to develop these sets of groups consist of the factor scores developed from earlier results. The use of factor scores results in a reduction in the number of attributes describing a subject and leads to a simplification of the analysis. This includes an associated decrease in the amount of computer time and memory required to find group clusters which is important when analyzing such large data sets.

There has been a great deal of literature developed describing the theory and method of cluster analysis of subjects. In general, the process of cluster analysis involves a three step process: (1) selection of a measure of inter object similarity; (2) selection of a method to form groups; and (3) deciding on the number of groups.

For purposes of this analysis, an algorithm available from the BMDP catalog of programs was applied. This program, BMDPKM, uses a measure of Euclidean distance between cluster centers and subjects to partition the data of interest. Output includes a measure of the within cluster variation which, when compared with the total variation for the entire sample, gives some indication of the number of clusters in the data.

The first step in a cluster analysis includes agreement on some measure of similarity between entities comprising a cluster. In this research, each subject is described by a set of eight factor scores. If the center of a cluster j (j = 1, . . ., g) is designated as (c_{j1} c_{j2}, . . . c_{jp}) for a set of n (n = 1, . . ., p) measures then the

distance between this quantity and the set of x_{ip} measures on any subject i (i = 1, . . ., N) or (x_{i1}, x_{i2}, . . ., x_{ip}) is given by:

$$\left[\sum_{n=1}^{p} (x_{ip} - c_{jp})^2\right]^{\frac{1}{2}} \qquad (4.13)$$

The BMDPKM algorithm requires that the user first specify the number of clusters hypothesized to exist in the data. Using this as a basis, the algorithm then splits the overall sample into the requested number of partitions through the following procedure. First, the entire sample is divided into two clusters by calculating the variance of each variable in the sample. Next, the sample is partitioned into two groups based on the variable with the highest variance. Those variables less than the mid-point of the range of the value of this variable are assigned to one group. The remaining subjects are assigned to a second group. This process is repeated until the requested number of groups have been found.

Next, each subject is assigned to the cluster whose center is closest to that subject. Cluster centers are recomputed and distances between subjects and centers are recomputed. If no subjects change groups, then the process is stopped and the results are reported.

The output of this process includes the mean, minimum and maximum, and variance of each variable for each cluster; the distance of cluster centers to each subject; histograms displaying the distances for subjects in a cluster and not in a cluster.

One criteria commonly used to decide on the number of clusters is to find the solution which minimizes the within cluster variation (58,

p. 69). This is equivalent to minimizing the trace of the matrix of the pooled within-group dispersion matrix. The value of this quantity is used in the present research to assess the results of several applications of the BMDPKM algorithm with the current data.

Results of Cluster Analysis

In order to identify the acceptable number of group clusters present in the data sets, the BMDPKM algorithm was applied several times to each set of factor scores. This resulted in the specification of 5, 8, 10, 15 and 20 groups for each year. For each iteration, the within group variation for each factor was developed and summed to find the total within group variation for the sample.

The results of this effort were interpreted by developing a plot of within group variation and group size for each sample. These relationships are presented in Figure 4.1 for the 1979 data and Figure 4.2 for the 1980 survey.

These figures indicate that the within group variation decreases as the number of groups is increased. However, it appears that an inflextion point occurs at the eight group solution for each survey. This suggests that the additional gain in within group variation increases at a decreasing rate for solutions behind eight groups. More specifically, it requires the addition of 12 more groups beyond the eight group solution to reduce the within group variation an amount that is similar in size to the drop occurring between the five and eight groups.

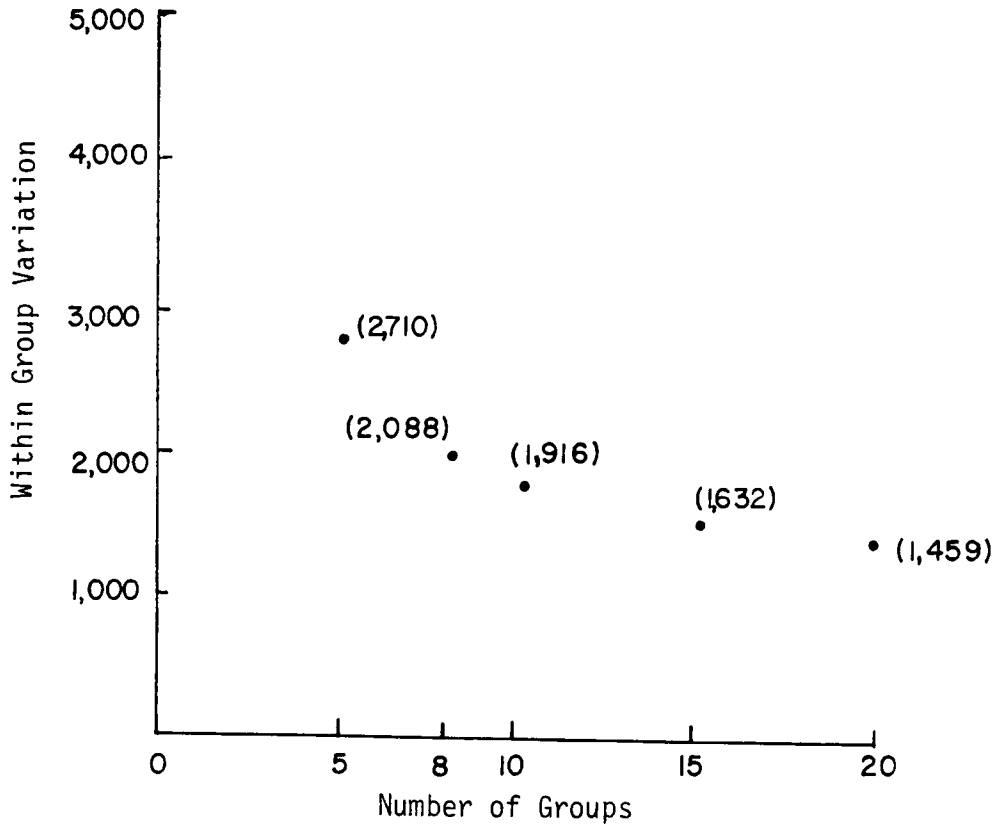


FIGURE 4.1

WITHIN GROUP VARIATION FOR SELECTED CLUSTER SOLUTIONS, 1979

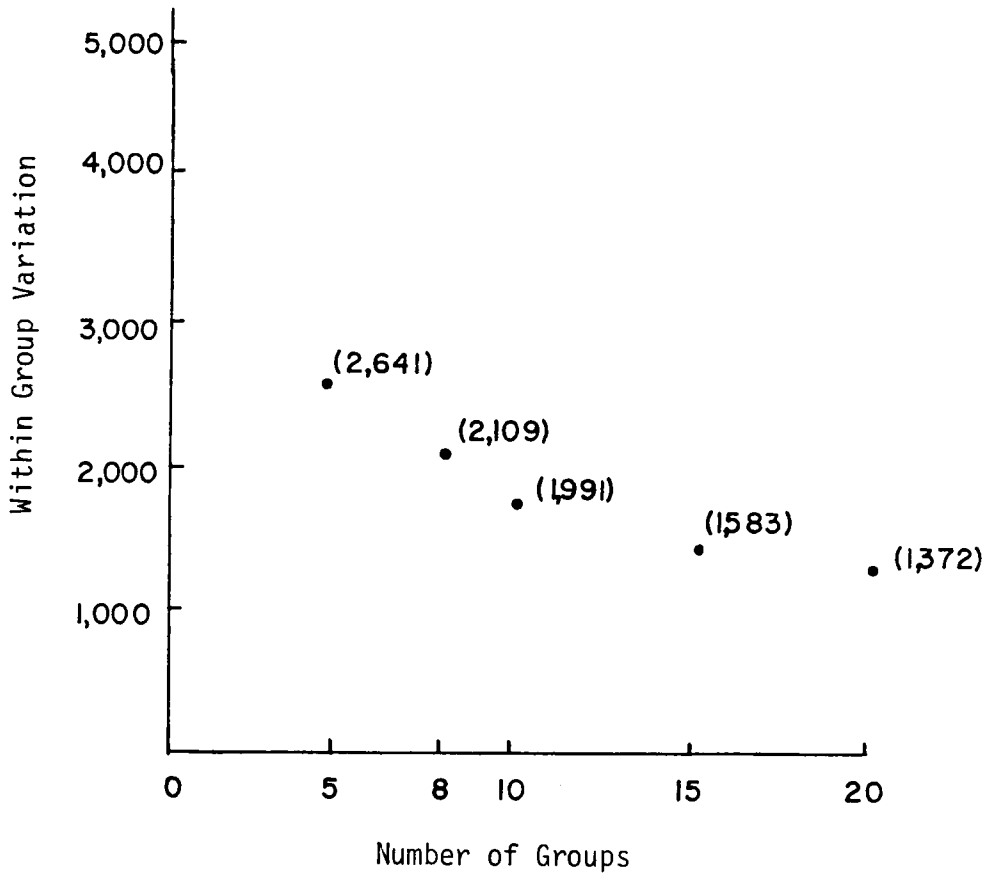


FIGURE 4.2

WITHIN GROUP VARIATION FOR SELECTED CLUSTER SOLUTIONS, 1980

Based on this observation, it was decided that the eight group solutions best described the number of clusters in the survey data. This interpretation, while somewhat subjective, represents a reasonable application of the principles of cluster analysis that results in a manageable number of groups for further analysis.

Characteristics of Cluster Groups

Table 4.24 presents the number of subjects in each of the eight groups for the various years as well as the average Euclidean distance between group centers and subjects assigned to a group. Group sizes range from 37 to 518 in the case of the 1979 results and from 78 to 530 in the 1980 data. Most cases are within 1.5 units of the cluster center with several lying an average of less than 1.0 units.

Tables 4.25 and 4.26 present the Euclidean distances between various cluster centers. Many cluster centers are over 3.0 units apart with only several less than 1.5 units. This indicates, for the most part, cluster centers are distributed over a fairly broad range in variable space.

Tables 4.27, 4.28, 4.29 and 4.30 outline the various group cluster centers and standard deviations resulting from the eight group solution. Because these means are factor scores developed earlier, it is somewhat difficult to interpret these results as presented. High positive values for cluster means suggest that a group has undertaken the conservation actions comprising a particular factor. On the other hand, low negative values for cluster means indicates a low endorsement rate for the actions comprising a factor.

TABLE 4.24

CLUSTER GROUP SIZE AND AVERAGE DISTANCE TO CLUSTER CENTERS FOR EIGHT GROUP SOLUTION, 1979 AND 1980

	1979		1980	
Cluster	Group Size	Average Distance to Center	Group Size	Average Distance to Center
1	62	1.5491	78	2.1941
2	139	1.2371	111	1.1834
3	70	1.8527	103	1.7146
4	124	1.3364	105	1.4882
5	37	2.3523	155	1.1319
6	315	0.9171	157	1.1002
7	518	0.5478	530	0.5719
8	225	1.2901	321	0.9537

TABLE 4.25

DISTANCE BETWEEN CLUSTER CENTERS, 1979

Cluster	1	2	3	4	5	6	7
2	3.50299						
3	2.51512	3.01577					
4	1.86902	2.46006	3.29146				
5	3.06244	4.63537	3.74154	3.81838			
6	3.26817	1.79536	3.93322	1.87023	4.66174		
7	4.25120	1.98652	4.70000	2.60554	5.30781	1.37377	
8	2.21827	2.23007	2.92076	1.82360	3.98260	1.55306	2.84842

TABLE 4.26

DISTANCE BETWEEN CLUSTER CENTERS, 1980

Cluster	1	2	3	4	5	6	7
2	4.09319						
3	3.05001	2.53982					
4	3.01580	3.37087	2.15072				
5	3.73358	1.77107	3.04470	2.96045			
6	3.13076	2.63636	2.12865	1.60700	1.92991		
7	4.81217	2.11090	4.08221	3.81606	1.47166	2.96678	
8	3.78907	2.10415	2.92438	2.38360	1.43195	1.47767	1.79495

TABLE 4.27

CLUSTER CENTERS, 1979

Cluster	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7	Factor 8
1	1.0776	0.1399	1.8092	1.1439	0.4155	0.8710	0.9652	0.7913
2	-0.3207	1.2820	-0.3120	-0.2235	-0.0528	0.0343	-0.1050	0.0312
3	0.7417	2.3071	0.7929	1.0404	0.6608	1.4109	0.7197	1.0376
4	0.2107	-0.2585	1.3789	0.4612	-0.0029	0.1836	0.1024	0.0565
5	1.0241	0.4290	1.6720	0.5069	3.1481	0.1200	1.8642	0.6231
6	0.1835	-0.3966	-0.3456	-0.0080	-0.2078	-0.1903	-0.1553	-0.1324
7	-0.8661	-0.3973	-0.4132	-0.7686	-0.2160	-0.3785	-0.4272	-0.4381
8	0.9908	-0.0057	-0.1343	0.8315	-0.0134	0.2796	0.3644	0.4414

TABLE 4.28

CLUSTER STANDARD DEVIATIONS, 1979

Cluster	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7	Factor 8
1	0.4280	0.5530	0.6403	0.4651	0.5105	0.6095	0.6714	0.6661
2	0.5707	0.5307	0.3019	0.4755	0.3447	0.5674	0.3748	0.4120
3	0.6192	0.5487	0.9115	0.5794	0.6586	0.6676	0.7231	0.6972
4	0.6287	0.3176	0.5727	0.5289	0.4015	0.5037	0.4183	0.4503
5	0.6932	1.0241	1.1476	0.7506	0.9094	0.5532	0.8173	0.8771
6	0.4445	0.2243	0.0990	0.4260	0.3082	0.4411	0.3385	0.3298
7	0.2090	0.2671	0.2710	0.2463	0.2304	0.2408	0.2037	0.1995
8	0.4269	0.4940	0.2163	0.4483	0.4479	0.5783	0.5525	0.5513

TABLE 4.29

CLUSTER CENTERS, 1980

Cluster	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7	Factor 8
1	1.0111	0.6724	2.8333	1.0718	0.8607	0.4680	0.9576	-0.4018
2	-0.2708	1.4246	-0.1668	-0.3019	-0.3032	-0.3606	-0.1063	0.2942
3	0.9857	1.8714	0.1272	0.9342	0.8824	0.6305	0.7113	0.2577
4	1.0141	-0.1386	0.0854	1.5631	0.6941	0.7835	1.1151	0.3176
5	-0.1961	0.0149	0.0267	-0.1778	0.2123	-0.4696	-0.1696	-0.6055
6	0.8342	-0.0609	0.0131	0.3841	0.9825	0.6989	0.1676	-0.0953
7	-0.9059	-0.4775	-0.3440	-0.5942	-0.6045	-0.5858	-0.4084	0.0035
8	0.3830	-0.4004	-0.1507	-0.0842	-0.2003	0.4064	-0.1118	0.1438

TABLE 4.30

CLUSTER STANDARD DEVIATIONS, 1980

Cluster	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7	Factor 8
1	0.4965	1.0002	0.9540	0.8400	0.6618	0.6690	0.8145	0.9177
2	0.6221	0.4433	0.2488	0.3284	0.4356	0.5169	0.3679	0.4973
3	0.4168	0.4875	0.1930	0.8529	0.7431	0.5393	0.8395	0.7490
4	0.4153	0.3925	0.1509	0.6272	0.6633	0.5368	0.8127	0.5802
5	0.4863	0.4015	0.5602	0.3286	0.3930	0.4179	0.4036	0.4194
6	0.3892	0.3750	0.2195	0.4336	0.4760	0.4637	0.3687	0.4938
7	0.3913	0.1523	0.0777	0.1626	0.2040	0.2489	0.2294	0.2372
8	0.4188	0.1857	0.2875	0.3700	0.2973	0.5220	0.3691	0.3358

The results can be better understood by cross-tabulating group membership with response scores of each of the 18 different actions. The end product of this effort is displayed in Tables 4.31 and 4.32. Besides cross-tabulating the actions by groups, a chi-squared test of independence was conducted for each 2 x 8 contingency table. In all, this resulted in a total of 36 tables covering both 1979 and 1980 data sets.

The chi-squared test is conducted to determine if group membership is independent from response score (e.g., 0, 1 where: 0 "did not take action" and 1 "did take action"). Each 2 x 8 contingency table contains a total of 16 cells. In order to develop a chi-squared test statistic, it is first necessary to find each expected cell frequency e_{ij} $(i = 0, 1)$, $(j = 1, \ldots, 8)$. These are found by multiplying the row (r) total corresponding to the ith response score by the column (c) total found in the jth group and dividing by the entire sample size N. If f_{ij} represents the observed cell frequency corresponding to a particular cell, then the χ^2 statistic is given by:

$$\chi^2 = \sum_{i=0}^{1} \sum_{j=1}^{g} \frac{(f_{ij} - e_{ij})^2}{e_{ij}} \qquad (4.14)$$

This value is then compared to a theoretical value of χ^2_α associated with a particular confidence level (α) and $(r - 1)(c - 1)$ or 7 degrees of freedom. If the calculated χ^2 is greater than χ^2_α then the null hypothesis is rejected and dependence between group membership and an action is implied.

TABLE 4.31

PROPORTION OF GROUPS ENDORSING EACH ITEM, 1979

		Factor	Group 1	2	3	4	5	6	7	8	Total Sample
1.	Drive slower	4	0.710	0.475	0.786	0.581	0.676	0.416	0.124	0.702	0.418
2.	Carpool to work	6	0.419	0.101	0.514	0.169	0.135	0.095	0.044	0.244	0.139
3.	Bus or subway to work	2	0.032	0.482	0.857	0.056	0.189	0.019	0.039	0.051	0.120
4.	Walk or bike to work	6	0.210	0.115	0.371	0.040	0.351	0.048	0.033	0.082	0.083
5.	Tune car	4	0.823	0.180	0.671	0.589	0.595	0.400	0.054	0.761	0.372
6.	Move closer to work	5	0.000	0.000	0.000	0.000	0.919	0.006	0.004	0.000	0.025
7.	Job closer to home	5	0.129	0.036	0.271	0.048	0.432	0.010	0.010	0.067	0.052
8.	Shop on way home from work	6	0.597	0.230	0.814	0.242	0.541	0.175	0.017	0.561	0.252
9.	Shop less frequently	1	0.758	0.201	0.529	0.460	0.676	0.400	0.035	0.761	0.350
10.	Combine shop and other	1	0.968	0.273	0.786	0.613	0.865	0.622	0.041	0.914	0.468
11.	Shop closer to home	1	0.806	0.345	0.743	0.395	0.811	0.521	0.050	0.800	0.418
12.	Eliminate use of recreation vehicles	7	0.339	0.029	0.271	0.073	0.459	0.048	0.012	0.200	0.093
13.	Vacation closer to home	7	0.677	0.101	0.500	0.177	0.541	0.089	0.021	0.377	0.170
14.	Cancel vacation trip	8	0.371	0.165	0.486	0.161	0.324	0.105	0.031	0.298	0.156
15.	Use train, bus, plane for vacation	2	0.323	0.324	0.800	0.153	0.243	0.035	0.050	0.220	0.159
16.	Buy fuel efficient car	3	0.839	0.029	0.443	0.855	0.541	0.000	0.027	0.024	0.153
17.	Sell car	3	0.581	0.022	0.171	0.419	0.432	0.000	0.015	0.000	0.084
18.	Bus or subway for non-work	2	0.129	0.791	0.886	0.000	0.270	0.010	0.015	0.118	0.152
	Group Size		(62)	(139)	(70)	(124)	(37)	(315)	(518)	(255)	1520
	Proportion of Sample		(0.041)	(0.091)	(0.046)	(0.082)	(0.024)	(0.207)	(0.314)	(0.168)	
	Average Number of Actions		(8.710)	(3.899)	(9.900)	(5.032)	(9.000)	(2.997)	(0.622)	(6.118)	

TABLE 4.32

PROPORTION OF GROUPS ENDORSING EACH ITEM, 1980

		Factor	Group 1	2	3	4	5	6	7	8	Total Sample
1.	Drive slower	6	0.769	0.288	0.835	0.779	0.187	0.835	0.111	0.682	0.428
2.	Carpool to work	5	0.269	0.027	0.340	0.212	0.084	0.538	0.032	0.037	0.133
3.	Bus or subway to work	2	0.295	0.270	0.573	0.019	0.381	0.152	0.019	0.012	0.135
4.	Walk or bike to work	5	0.321	0.117	0.233	0.163	0.297	0.146	0.028	0.016	0.108
5.	Tune car	6	0.603	0.090	0.583	0.654	0.135	0.475	0.030	0.318	0.256
6.	Move closer to work	3	0.538	0.009	0.000	0.000	0.006	0.006	0.000	0.012	0.031
7.	Job closer to home	3	0.718	0.000	0.000	0.000	0.065	0.000	0.000	0.003	0.043
8.	Shop on way home from work	5	0.679	0.243	0.631	0.577	0.561	0.816	0.015	0.125	0.301
9.	Shop less frequently	1	0.846	0.423	0.893	0.885	0.394	0.867	0.140	0.813	0.532
10.	Combine shop and other	1	0.859	0.324	0.922	0.952	0.606	0.905	0.130	0.729	0.537
11.	Shop closer to home	1	0.885	0.468	0.845	0.827	0.271	0.741	0.140	0.642	0.470
12.	Eliminate use of recreation vehicle	4	0.308	0.018	0.350	0.750	0.032	0.108	0.002	0.059	0.117
13.	Vacation closer to home	4	0.513	0.135	0.427	0.538	0.084	0.316	0.032	0.174	0.187
14.	Cancel vacation trip	8	0.269	0.036	0.379	0.481	0.006	0.171	0.026	0.087	0.118
15.	Use train, bus, plane for vacation	2	0.462	0.333	0.641	0.250	0.245	0.259	0.065	0.140	0.270
16.	Buy fuel efficient car	7	0.462	0.036	0.291	0.596	0.148	0.310	0.036	0.146	0.173
17.	Sell car	7	0.141	0.027	0.126	0.269	0.032	0.025	0.011	0.028	0.051
18.	Bus or subway for non-work	2	0.282	1.000	0.961	0.038	0.000	0.006	0.000	0.000	0.152
	Group Size		(78)	(111)	(103)	(104)	(155)	(158)	(530)	(321)	1560
	Proportion of Sample		(0.050)	(0.071)	(0.066)	(0.064)	(0.099)	(0.101)	(0.340)	(0.206)	
	Average Number of Actions		(9.218)	(3.847)	(9.029)	(7.990)	(3.535)	(6.677)	(0.760)	(4.025)	

In order to find cells that contribute the highest to the calculated χ^2 and thus indicate sources of dependence, it is possible to examine the chi-squared calculated value, or residual, associated with a particular cell. However, it is more meaningful to standardize these residual values to a common unit mean and variance so that the impact of very large or small values do not distort this interpretation. Such standardized residuals can be compared to values of the unit normal distribution to better understand the significance of these contributions. Values close to zero indicate that no difference exists between expected and observed values under the null hypothesis. Large positive values indicate that a much higher than expected observed value has been found. A very low negative value indicates that the observed cell frequency is much lower than expected.

The standardized residual a_{ij} is found for any cell by calculating the following value (59):

$$a_{ij} = \frac{(f_{ij} - e_{ij})}{\sqrt{e_{ij}}} \qquad (4.16)$$

It should be noted that the sum of the squares of the set of adjusted residuals is equal to the calculated chi-squared test statistic.

Tables 4.33 and 4.34 contain the adjusted residuals corresponding to the row in each cell that contained the number of subjects that undertook a particular action (i.e., i = 1). Also contained in these tables are the overall chi-squared statistic calculated for the entire set of contingency tables. Since each is much larger than the theoretical chi-squared value, at the 0.05 significance level and 7 degrees

TABLE 4.33

STANDARDIZED RESIDUALS FOR 1979 SURVEY

		Group								
		1	2	3	4	5	6	7	8	x^2
1.	Drive slower	3.5	1.0	4.8	2.8	2.4	-0.1	-10.4	7.0	355.045
2.	Carpool to work	5.9	-1.2	8.4	0.9	-0.1	-2.1	-5.8	3.6	184.225
3.	Bus or subway to work	-2.0	12.3	17.8	-2.0	1.2	-5.2	-5.3	-3.2	619.209
4.	Walk or bike to work	3.5	1.3	8.4	-1.6	5.7	-2.2	-4.0	-0.0	151.937
5.	Tune car	5.8	-3.7	4.1	3.9	2.2	0.8	-11.9	10.2	525.467
6.	Move closer to work	-1.2	-1.9	-1.3	-1.8	34.4	-2.1	-3.0	-2.5	1243.636
7.	Job closer to home	2.7	-0.8	8.1	-0.2	10.2	-3.3	-4.2	1.0	218.810
8.	Shop on way home from work	5.4	-0.5	9.4	-0.2	3.5	-2.7	-10.6	9.8	463.577
9.	Shop less frequently	5.4	-3.0	2.5	2.1	3.3	1.5	-12.1	11.1	511.374
10.	Combine shop and other	5.8	-3.4	3.9	2.4	3.5	4.0	-14.2	10.4	759.351
11.	Shop closer to home	4.9	-1.2	4.4	-0.3	3.8	3.1	-12.8	9.7	552.989
12.	Eliminate use of recreation vehicle	6.3	-2.5	4.9	-0.8	7.3	-2.7	-6.1	5.6	219.228
13.	Vacation closer to home	9.7	-2.0	6.7	0.2	5.5	-3.5	-8.2	6.5	354.866
14.	Cancel vacation trip	4.3	0.3	7.0	0.2	2.6	-2.3	-7.2	5.7	194.659
15.	Use train, bus, plane for vacation	3.2	4.9	13.4	-0.2	1.3	-5.5	-6.2	2.4	346.464
16.	Buy fuel efficient car	13.8	-3.7	6.2	20.0	6.0	-6.9	-7.3	-5.3	952.936
17.	Sell car	13.5	-2.5	2.5	12.9	7.3	-5.1	-5.4	-4.6	538.878
18.	Bus or subway for non-work	-0.5	19.3	15.7	-4.3	1.8	-6.5	-8.0	-1.4	886.694

TABLE 4.34

STANDARDIZED RESIDUALS FOR 1980 SURVEY

		Group								
		1	2	3	4	5	6	7	8	x^2
1.	Drive slower	4.6	-2.3	6.3	5.5	-4.6	7.8	-13.1	7.0	698.123
2.	Carpool to work	3.3	-3.1	5.7	2.2	-1.7	13.9	-6.4	-4.7	366.583
3.	Bus or subway to work	3.8	3.9	12.1	-3.2	8.3	0.6	-7.3	-6.0	397.802
4.	Walk or bike to work	5.7	0.3	3.9	1.7	7.2	1.5	-5.6	-5.0	180.200
5.	Tune car	6.1	-3.5	6.6	8.0	-3.0	5.4	-10.3	2.2	409.364
6.	Move closer to work	25.3	-1.3	-1.8	-1.8	-1.8	-1.8	-4.1	-1.9	695.113
7.	Job closer to home	22.8	-2.2	-2.1	-2.1	1.3	-2.6	-4.8	-3.4	923.893
8.	Shop on way home from work	6.1	-1.1	6.1	5.1	5.9	11.8	-12.0	-5.8	649.088
9.	Shop less frequently	3.8	-1.6	5.0	4.9	-2.4	5.8	-12.4	6.9	654.883
10.	Combine shop and other	3.9	-3.1	5.3	5.8	1.2	6.3	-12.8	4.7	675.602
11.	Shop closer to home	5.3	-0.0	5.5	5.3	-3.6	5.0	-11.1	4.5	506.377
12.	Eliminate use of recreation vehicle	4.9	-3.0	6.9	18.9	-3.1	-0.3	-7.7	-3.0	585.962
13.	Vacation closer to home	6.7	-1.3	5.7	8.3	-3.0	3.8	-8.2	-0.5	292.833
14.	Cancel vacation trip	3.9	-2.5	7.7	10.8	-4.0	1.9	-6.1	-1.6	291.564
15.	Use train, bus, plane for vacation	4.9	2.9	9.7	1.0	1.0	1.4	-7.2	-2.6	239.414
16.	Buy fuel efficient car	6.1	-3.5	2.9	10.4	-0.7	4.1	-7.6	-1.1	292.759
17.	Sell car	3.5	-1.1	3.4	9.9	-1.0	-1.4	-4.0	-1.8	153.792
18.	Bus or subway for non-work	2.9	22.9	21.1	-3.0	-4.9	-4.7	-9.0	-7.0	1370.004

of freedom, of 14.067, the null hypothesis of independence was rejected in every case.

Group Profiles

The application of the k-means clustering algorithm results in an eight group solution. The eight groups in each survey year represent individuals in each sample sharing a similar response to changing energy price and availability. In each case, one large group consisting of subjects endorsing few actions was identified. Several other groups have endorsed several actions and a smaller number have endorsed a large number of actions.

Profiles of each group were developed through cross-tabulation of group membership and selected socioeconomic variables. Table 4.35 contains the results of this effort for the 1979 results. Table 4.36 presents the results of this analysis for the 1980 survey data. Also shown in Tables 4.37 and 4.38 are the standardized residuals associated with each contingency table. The group profile data is used in association with the cluster analysis results to describe market segments.

1979 Segments

Group 1. The results for this group indicate a high level of participation in energy conservation. Generally speaking, actions requiring various degrees of disruption in daily travel and activity behavior have been undertaken by this group. This includes such actions as combining trips to purchasing a fuel efficient vehicle. With few exceptions, the group has indicated a willingness to implement

TABLE 4.35

MARKET SEGMENT PROFILES USING CONVENTIONAL ANALYSIS, 1979

	Segment								
	1	2	3	4	5	6	7	8	Total Sample
Income									
Under $10,000	0.09	0.21	0.17	0.14	0.15	0.16	0.25	0.11	0.18
$10,000 - $25,000	0.54	0.55	0.51	0.51	0.52	0.51	0.49	0.54	0.52
Over $25,000	0.37	0.24	0.32	0.35	0.33	0.33	0.26	0.35	0.30
Household Size									
1-2	0.21	0.50	0.46	0.34	0.38	0.36	0.35	0.34	0.38
3-4	0.50	0.68	0.41	0.46	0.38	0.36	0.35	0.44	0.38
5+	0.29	0.18	0.13	0.20	0.24	0.22	0.11	0.22	0.18
Employment Status									
Employed	0.66	0.73	0.84	0.65	0.76	0.61	0.59	0.62	0.64
Unemployed	0.34	0.27	0.16	0.35	0.26	0.39	0.41	0.38	0.36
Marital Status									
Married	0.83	0.55	0.44	0.68	0.51	0.68	0.54	0.74	0.62
Single	0.06	0.33	0.50	0.20	0.25	0.18	0.26	0.18	0.24
Widow, Dv, Sep	0.07	0.12	0.06	0.12	0.24	0.14	0.20	0.08	0.14

TABLE 4.35 (continued)

	Segment								
	1	2	3	4	5	6	7	8	Total Sample
Area of State									
Downstate Urban	0.29	0.84	0.80	0.38	0.35	0.55	0.75	0.41	0.60
Upstate Urban	0.29	0.10	0.19	0.35	0.22	0.25	0.16	0.30	0.22
Other	0.42	0.06	0.01	0.27	0.43	0.20	0.09	0.29	0.18
Number of Vehicles									
0	0.00	0.27	0.20	0.02	0.08	0.04	0.37	0.01	0.17
1	0.48	0.51	0.49	0.32	0.46	0.47	0.32	0.44	0.40
2+	0.52	0.22	0.31	0.47	0.46	0.49	0.31	0.55	0.43
Age									
18-34	0.60	0.48	0.69	0.40	0.60	0.35	0.33	0.40	0.40
35-64	0.39	0.41	0.30	0.51	0.40	0.54	0.48	0.54	0.49
65+	0.02	0.11	0.11	0.10	0.00	0.11	0.18	0.06	0.12
Sex									
Male	0.47	0.61	0.46	0.54	0.46	0.49	0.50	0.53	0.49
Female	0.53	0.39	0.54	0.46	0.54	0.51	0.50	0.47	0.51

TABLE 4.35 (continued)

	Segment								
	1	2	3	4	5	6	7	8	Total Sample
Educational Status									
Grammar or Some High School	0.07	0.18	0.07	0.11	0.09	0.11	0.19	0.14	0.15
High School Graduate	0.50	0.26	0.35	0.36	0.43	0.37	0.38	0.41	0.38
Some College or College Graduate	0.44	0.56	0.58	0.53	0.38	0.49	0.43	0.46	0.47
Housing Type									
Single Family	0.69	0.29	0.21	0.75	0.57	0.65	0.37	0.78	0.52
Two Family	0.08	0.17	0.19	0.10	0.16	0.10	0.15	0.18	0.14
Apt. Condo, Co-Op	0.18	0.54	0.60	0.60	0.32	0.25	0.49	0.12	0.34

TABLE 4.36

MARKET SEGMENT PROFILES USING CONVENTIONAL ANALYSIS, 1980

	Segment								
	1	2	3	4	5	6	7	8	Total Sample
Income									
Under $10,000	0.17	0.34	0.18	0.11	0.10	0.05	0.30	0.21	0.21
$10,000 - $25,000	0.56	0.41	0.56	0.56	0.46	0.64	0.45	0.47	0.49
Over $25,000	0.27	0.27	0.26	0.33	0.44	0.31	0.25	0.32	0.30
Household Size									
1-2	0.50	0.55	0.55	0.40	0.45	0.44	0.59	0.51	0.51
3-4	0.36	0.32	0.30	0.40	0.38	0.40	0.33	0.35	0.35
5+	0.14	0.13	0.15	0.20	0.17	0.17	0.11	0.15	0.14
Employment Status									
Employed	0.86	0.60	0.68	0.63	0.79	0.76	0.47	0.43	0.57
Unemployed	0.24	0.40	0.32	0.37	0.21	0.24	0.53	0.57	0.43
Marital Status									
Married	0.53	0.37	0.56	0.74	0.51	0.69	0.51	0.72	0.58
Single	0.29	0.43	0.34	0.18	0.35	0.24	0.25	0.15	0.25
Widow, Dv, Sep	0.17	0.20	0.10	0.08	0.14	0.01	0.23	0.13	0.17

TABLE 4.36 (continued)

	Segment								
	1	2	3	4	5	6	7	8	Total Sample
Area of State									
Downstate Urban	0.47	0.87	0.69	0.26	0.67	0.37	0.75	0.46	0.60
Upstate Urban	0.39	0.08	0.19	0.41	0.14	0.33	0.13	0.26	0.21
Other	0.18	0.05	0.17	0.33	0.19	0.30	0.12	0.27	0.19
Number of Vehicles									
0	0.09	0.35	0.14	0.00	0.10	0.01	0.42	0.02	0.20
1	0.37	0.43	0.62	0.28	0.54	0.42	0.31	0.51	0.42
2+	0.54	0.22	0.25	0.72	0.36	0.58	0.27	0.47	0.39
Age									
18-34	0.51	0.46	0.36	0.33	0.36	0.39	0.32	0.29	0.36
35-64	0.45	0.41	0.49	0.58	0.44	0.60	0.41	0.47	0.47
65+	0.04	0.13	0.15	0.09	0.05	0.01	0.27	0.24	0.18
Sex									
Male	0.40	0.47	0.48	0.60	0.47	0.44	0.46	0.46	0.47
Female	0.60	0.53	0.52	0.40	0.53	0.56	0.54	0.54	0.53

TABLE 4.36 (continued)

	Segment								
	1	2	3	4	5	6	7	8	Total Sample
Educational Status									
Grammar or Some High School	0.16	0.18	0.14	0.13	0.11	0.05	0.28	0.19	0.19
High School Graduate	0.37	0.35	0.39	0.46	0.36	0.39	0.36	0.40	0.38
Some College or College Graduate	0.47	0.47	0.48	0.42	0.53	0.57	0.37	0.41	0.44
Housing Type									
Single Family	0.52	0.33	0.46	0.82	0.48	0.73	0.43	0.67	0.54
Two Family	0.21	0.15	0.14	0.11	0.11	0.11	0.12	0.12	0.12
Apt. Condo, Co-Op	0.27	0.52	0.41	0.13	0.38	0.11	0.46	0.21	0.34

TABLE 4.37

STANDARDIZED RESIDUALS DEVELOPED FROM SEGMENT PROFILES
USING CONVENTIONAL ANALYSIS, 1979

	Segment							
	1	2	3	4	5	6	7	8
Income								
Under $10,000	-1.7	0.7	-0.2	-1.1	-0.4	-1.0	3.5	-2.4
$10,000 - $25,000	0.3	0.5	-0.1	-0.1	0.1	-0.1	-0.8	0.5
Over $25,000	1.5	0.8	-0.2	0.4	-0.2	0.6	-1.9	0.5
Household Size								
1-2	-2.8	1.0	0.2	-1.7	0.6	-0.5	3.5	-2.4
3-4	1.5	-1.2	0.4	1.4	0.0	-0.7	-1.2	1.4
5+	2.1	-0.1	-0.9	0.6	0.9	1.8	-3.6	1.7
Employment Status								
Employed	0.2	1.4	2.1	0.2	0.9	-0.7	-1.3	-0.1
Unemployed	-0.3	-1.8	-2.8	-0.3	-1.2	0.9	1.7	0.1
Marital Status								
Married	-2.0	-1.0	-1.9	-0.8	-0.8	1.4	-2.4	2.4
Single	2.0	2.3	4.5	-0.8	0.1	-2.0	-1.0	-1.8
Widow, Dv, Sep	-1.6	0.8	-1.9	-0.6	1.6	-0.2	3.6	-2.7

Table 4.37 (continued)

	Segment							
	1	2	3	4	5	6	7	8
Area of State								
Downstate Urban	-3.2	3.6	-2.2	-3.2	-2.0	-1.1	4.3	-4.1
Upstate Urban	1.2	-2.8	-1.1	3.0	-0.1	1.0	-2.9	3.0
Other	4.5	-3.5	-2.7	2.5	3.7	0.9	-4.7	4.2
Number of Vehicles								
0	-5.6	-3.4	0.6	-6.1	-1.4	-7.4	3.7	-9.6
1	-1.0	2.4	1.1	-1.5	0.6	1.8	2.9	1.0
2+	2.7	-4.2	-1.5	3.6	0.4	1.7	-4.4	2.8
Age								
18-34	2.5	1.5	-3.8	0.1	1.9	-1.4	-2.4	0.0
35-64	1.1	-1.2	-2.2	0.4	-0.7	1.4	-0.1	1.2
65+	-2.3	-0.2	-0.2	-0.6	-2.1	0.2	4.6	-2.5
Sex								
Male	-0.3	1.9	-4.0	0.7	-0.2	-0.1	0.1	-1.5
Female	0.3	-2.0	0.5	-0.7	0.3	0.2	0.1	1.4

Table 4.37 (continued)

	Segment							
	1	2	3	4	5	6	7	8
Educational Status								
Grammer or Some High School	-1.8	0.9	-1.7	-1.1	0.6	-0.5	1.9	0.5
High School Graduate	1.6	-2.2	0.4	-0.4	0.5	-0.1	0.2	0.9
Some College or College Graduate	0.4	1.5	1.3	1.0	-0.8	0.4	-1.3	-0.6
Housing Type								
Single Family	2.4	-3.7	-3.6	3.6	0.0	-3.1	-4.7	4.1
Two Family	-1.2	-0.8	1.1	-1.3	0.4	-1.8	0.3	1.8
Apt, Condo, Co-Op	-2.2	4.0	3.7	-3.6	-0.2	-2.7	5.5	-6.2

TABLE 4.38

STANDARDIZED RESIDUALS DEVELOPED FROM SEGMENT PROFILES
USING CONVENTIONAL ANALYSIS, 1980

	Segment							
	1	2	3	4	5	6	7	8
Income								
Under $10,000	0.7	2.4	-0.6	-2.0	-2.6	-3.9	3.9	-0.1
$10,000 - $25,000	0.1	-1.1	1.0	1.0	-0.5	1.3	-0.9	-0.3
Over $25,000	-0.5	-0.6	-0.7	0.4	2.8	1.6	-2.1	0.4
Household Size								
1-2	-0.1	0.6	-0.5	-1.6	-1.1	-1.3	1.6	-0.1
3-4	0.2	-0.5	-0.8	0.8	0.7	1.0	-0.7	-0.1
5+	0.5	-0.3	0.2	1.7	1.0	1.0	-2.1	0.3
Employment Status								
Employed	3.2	0.4	1.4	0.7	3.4	2.9	-3.2	-3.2
Unemployed	-3.7	-0.5	-1.6	-0.9	-3.9	-3.4	3.6	3.7
Marital Status								
Married	-0.5	-2.9	-0.2	2.1	-1.2	1.8	-2.2	3.3
Single	0.7	3.8	1.8	1.4	2.5	-1.0	-0.1	-3.7
Widow, Dv, Sep	0.1	0.8	-1.7	-2.2	-0.7	-2.1	4.2	-1.4

Table 4.38 (continued)

	Segment							
	1	2	3	4	5	6	7	8
Area of State								
Downstate Urban	-1.8	3.8	-0.6	-4.4	1.1	-3.7	4.6	-3.1
Upstate Urban	3.3	-3.0	-4.0	4.5	-1.9	3.2	-4.1	2.1
Other	-0.3	-3.5	-0.6	3.1	0.0	3.2	-3.7	3.3
Number of Vehicles								
0	-2.4	3.2	-1.7	-8.1	-2.9	-8.7	9.5	-10.8
1	0.6	0.3	2.9	-2.3	2.2	0.0	-4.0	-2.6
2+	2.0	-3.3	-2.4	4.6	-0.6	3.4	-4.9	2.2
Age								
18-34	-2.2	1.8	-0.1	0.4	-1.7	0.7	1.4	-1.9
35-64	-0.2	-0.9	0.3	1.6	0.8	2.3	-2.1	0.1
65+	-2.8	-1.1	-0.7	-2.0	-3.7	-4.8	5.3	2.6
Sex								
Male	-0.9	0.0	0.1	1.9	0.0	-0.5	-0.2	-0.1
Female	0.9	0.0	-0.1	-1.8	0.0	0.4	0.2	0.1

Table 4.38 (continued)

	Segment							
	1	2	3	4	5	6	7	8
Educational Status								
Grammer or Some High School	-0.5	-0.1	-1.2	-1.4	-2.3	-4.1	4.8	0.2
High School Graduate	-0.1	-0.5	0.2	1.3	-0.3	0.1	-0.8	0.6
Some College or College Graduate	0.4	0.5	0.6	-0.3	1.8	2.5	-2.4	-0.7
Housing Type								
Single Family	-0.2	-3.0	-1.1	-3.9	-1.0	3.2	-3.4	3.0
Two Family	2.2	0.8	0.4	-1.9	0.7	-0.4	-0.7	0.0
Apt, Condo, Co-Op	-1.1	3.3	1.2	-3.8	0.8	3.7	4.7	3.8

actions at a rate much higher than the overall sample results. In brief, this group is characterized by a very high participation rate for a wide range of conservation actions.

The group profile suggests that the subjects are young, highly educated, married, employed, living in single family homes with moderate to large household sizes located in upstate urban or other areas. These characteristics suggest a high level of personal and household travel demands and interactions. These patterns may explain, in part, the need for a wide range of actions to balance constraints faced by limited fuel supplies. It appears that this segment requires a broad set of conservation options and is willing to adopt several measures simultaneously to meet household needs.

Group 2. This group appears to concentrate efforts in non-work travel and using non-auto modes for work and vacation travel. Purchasing a new fuel efficient car is not a high priority. In general, most actions are implemented at a rate much lower than the overall sample results indicate.

The group profile suggests that the subjects comprising this segment are young, highly educated, single persons, employed, living in small households with low to moderate incomes. This group is generally comprised of households living in the downstate urban area and live in multi-family housing.

The low levels of income and auto ownership suggest that a fuel efficient car is not a viable option for this group. In addition, the urbanized nature of the location of these subjects makes the use of public transportation an immediate conservation option. It appears that

non-auto based travel options and changes in shopping trip behavior appeal most highly to this segment.

Group 3. This group, in some ways, is very similar to Group 1. However, the emphasis on non-auto travel modes is suggestive of Group 2. Unlike Group 2, however, this group has appeared to have purchased fuel efficient cars to meet travel needs. In addition, Group 3, has implemented non-auto actions at a much higher level than Group 1. Note in particular the use of bus or subway for work and non-work travel. It appears, in brief, that Group 3 falls somewhere in between Groups 1 and 2 with respect to energy conservation behavior.

The profile data concerning this group indicates that it is generally comprised of young, single, college educated, employed persons who live in small sized households on moderate to high incomes. Like Group 2, these subjects live in multi-family housing in the downstate urban area. The main difference between Group 3 and Group 2 appears to be reflected in the higher incomes. On the other hand, Group 3 members are different from Group 1 members with respect to geographic location, vehicle ownership, household size, and educational level.

Group 3 appears to rely not only on vehicle use actions but also on non-auto modes to meet travel needs. Such a group may best be able to take advantage of policies combining such actions.

Group 4. Group 4 appears to rely on purchasing fuel efficient vehicles while making efficient use of the auto through maintenance and shopping travel adjustments. No indication is given that the use of non-auto modes has been pursued to any great extent.

These subjects are older than the previous group members, college educated, married, employed, live in large household in single-family homes in the upstate urban or other areas. The major difference suggested between Group 4 and Group 1 is centered in the slightly older age of the Group 4 members. This may explain why this group emphasis is on the use of the auto and generally avoids other actions. It appears that any policies designed for this group must reflect this general pattern.

Group 5. The interesting observation developing from this group is suggested by the very high proportion of subjects endorsing the moving related actions. Note in particular the percentage moving closer to work. In general, this group appears to have implemented a series of actions at a higher rate than the overall sample proportion. Special emphasis seems to be placed on shopping travel by this group as well as vacation and auto maintenance and use.

These subjects are generally young, single, employed, and live in two-family housing in non-urban areas. They have a large household size and have moderate to high incomes. Finally, this group seems to have a low educational attainment level when compared to other groups.

The higher incomes of this segment perhaps make moving closer to work possible as well as purchasing or replacing an auto. However, it appears that this group is willing to implement a series of actions and this must be considered in any future action related policies.

Group 6. Group 6 does not exhibit a great willingness to implement actions that are not easy to pursue. The emphasis in this group

appears to center on shopping related actions and maintenance and use of the auto.

This group is comprised of middle-aged subjects who are college educated and married but are less likely to be employed. These subjects enjoy a moderate to high income and have a high auto ownership rate. In addition, it is suggested that these subjects live in the upstate or other areas in single family homes. The group is similar in profile to Group 4, but has not purchased fuel efficient cars to the same degree. This difference may be explained by the slightly higher educational attainment by Group 4 and the greater likelihood that Group 6 members are located in the downstate urban area.

Group 6 appears to be willing to pursue a limited number of actions. Any policies designed for this segment must include an emphasis on actions that disrupt overall travel only minimally and encourage efficient use of the automobile.

Group 7. Group 7 consists of a large segment of subjects who did not take any actions over the study year to conserve gasoline. This may be a reflection of the low travel needs of this group. More specifically, the segment includes older persons with low incomes who are single and live in small households in multi-family dwellings in the downstate urban area with only low auto ownership levels.

These results reflect, in part, the strong likelihood that this group has only limited travel needs and, perhaps, only limited travel options. It is, therefore, unlikely that the current travel patterns can change. It is interesting to note that this single groups is by

far the largest segment that many individuals and households are not willing or able to adapt to changing energy price and supply situations.

Group 8. The last group is in some ways similar to Group 6 with respect to actions implemented. However, unlike Group 6 this group has emphasized a wider range of actions at a higher proportion of frequency.

These subjects tend to be young to middle-aged, highly educated, married, live in mid to large households, have moderate to high incomes, have high auto ownership rates, and live in single-family housing in the upstate urban or other areas. Compared to Group 6, Group 8 members tend to come from slightly larger households, are more likely to be married, are older, have more autos, and are more likely to live in the upstate area in single family homes.

Group 8 appears to consist of individuals willing to participate in actions requiring a minimum amount of disruption, but at a very high rate. Future policy decisions must recognize this willingness and encourage actions that support these desires.

1980 Segments

Group 1. Group 1 consists of a set of individuals who have exhibited a willingness to undertake a variety of conservation actions. A high proportion have implemented actions related to changes in shop travel as well as making efforts to improve the maintenance and use of the automobile. Interestingly, a high proportion indicated that some moving related action had been implemented. Many subjects reported the purchase of a fuel efficient auto. In addition, subjects report changes in vacation travel designed to conserve fuel.

These subjects tend to be young, have a high educational attainment, have moderate incomes, are employed, live in upstate urban or other areas, own several autos, and live in single family dwellings. These profile characteristics are reflected by the need to implement many actions to maintain the desired life-style. A broad range of options seems to best appear to this segment.

Group 2. Group 2 places a great deal of emphasis on actions related to the use of non-auto travel modes. Note that the entire group reports using a bus or subway for non-work travel. In addition, many have adjusted shop travel and used non-auto modes for work travel.

This group is generally young in age, single, lives in small to moderate sized households in downstate urban areas with a low auto ownership rate. It appears that these subjects are best served by transit related options that provide service that meets basic travel desires.

Group 3. This segment also places a great emphasis on non-auto travel modes. However, unlike Group 2, this segment also appears to participate in a broader range of actions at a high level. This indicates that the group members have a willingness to consider a wide spectrum of actions to maintain travel.

Group 3 subjects are likely to be young or middle-aged, married, have an auto, live in small to moderate sized households in the downstate urban area. This segment appears to be more likely to be married than Group 2 and chances are that Group 3 subjects also live in larger households. These differences may explain the higher rate of action implementation suggested by the Group 3 results. Policies for this

group should be centered on non-auto modes, but must recognize the willingness of the group to implement a range of actions.

Group 4. This segment has implemented a set of actions related to changes in travel patterns, improved use and maintenance of the auto, and has been willing to replace an auto with a fuel efficient model. Little emphasis is placed on non-auto modes by this group.

The segment profile suggests that these subjects are generally middle-aged, highly educated, married, live in moderate to large households, have moderate to high incomes, high auto ownership rates, live in single family homes in the upstate urban or other areas. These characteristics, in part, suggest a limited desire or opportunity to use public transportation and may explain the need to make best use of the automobile. A broad range of auto use and travel pattern adjustments appears to be the type of actions that best appeal to this group.

Group 5. This segment appears to consist of subjects willing to undertake only a limited number of actions that involve minimum travel related adjustments. The group is generally characterized by a low rate of participation in actions beyond shopping travel adjustments.

This group tends to be young or middle-aged, highly educated, married, live in small households, have one or more autos, have high income levels, and live in the downstate urban area in multi-family housing.

The high income and auto ownership levels combined with low household sizes suggest a flexibility in the adaptation of the household to the energy price and supply situation actions that assist in providing

non-auto work trip modes, and combining trips seems to appeal to this group.

Group 6. This segment has implemented a number of actions related to shopping travel and improved vehicle use and maintenance. However, it is interesting to note the high proportion of subjects using carpools to work.

The profile of this group suggests that subjects are middle-aged, live in single-family houses, have large auto ownership patterns, are college educated, live in small to moderate sized households, are married and live in upstate urban or other areas.

The auto ownership and education levels may suggest reasons for the endorsement of trip adjustment actions and auto use improvements. The carpool action may be explained by the education level or trip distances traversed in upstate or other areas to reach work places. In summary, it appears that improved use of the auto appeals to this particular segment.

Group 7. Group 7 reflects the large number of subjects who implemented one or fewer actions. Except for the shop trip relation actions, this group indicates an almost complete unwillingness to take any conservation actions.

These subjects live in the downstate urban area, are middle-aged or elderly, have few autos, and live in small-sized households. It appears that these characteristics are reflected in an unwillingness, ability, or desire to implement energy conservation actions. It may

be suggested that such a group make fewer travel related demands and thus has no need to conserve fuel.

Group 8. The final segment consists of subjects who concentrate on shopping travel adjustments as well as improved use of the auto.

These subjects tend to be married, live in small households, have moderate to high incomes, have several autos, are highly educated, middle-aged, and live in the upstate or other urban area.

This segment, in some ways, is similar to Group 6. However, unlike Group 6, this segment does not participate in actions to a high level in general. It appears that actions related to trip planning and auto use will best appear to this group.

Summary

The application of conventional market segmentation analysis resulted in the identification of eight segments in both survey years. One large segment consisted of individuals that did not endorse any conservation actions. These subjects were found to have very low auto ownership levels and resided in the downstate urban area in multifamily housing. In addition, low education attainment and income were characteristics of this group.

Remaining groups consisted of subjects endorsing conservation actions at a very high rate to a somewhat moderate rate. High auto ownership, educational levels, upstate urban or other locations, and young to middle aged subjects were found to be more likely to endorse a series of conservation actions. Most groups concentrated on non-work and vacation related travel to conserve gasoline.

The conventional analysis approach is understood and practiced by a wide range of practitioners. In addition, it is based on a set of commonly accepted multivariate analysis procedures. On the other hand, it is difficult to generalize the results of one sample to the next.

Chapter 6 outlines the possible conservation actions which will meet the needs of the segments identified in this chapter. However, the following chapter presents the results of the application of latent trait theory in market segmentation so that further comparison can be made between the use of conventional and modern test methods.

CHAPTER 5

RESULTS OF MARKET SEGMENTATION ANALYSIS USING LATENT TRAIT THEORY

Introduction

The previous chapter contained the results of a market segmentation analysis based on the application of conventional multivariate statistical procedures. The results of this effort suggested that market segments can be formed based on such variables as auto ownership, geographic location, household size, and marital status. In most cases, such an approach would be adequate to suggest strategies to implement energy conservation policies and programs. However, in this research an alternative method to identify market segment is explored.

This chapter presents a different strategy to identify market segments. While the previous chapter centered on the observed responses offered by each subject to the list of 18 conservation actions, this current chapter suggests that these observable item scores represent an indicant of an unseen psychological construct related to transportation energy conservation attitudes and behavior. It is hypothesized that such a construct is a reflection of a combination of situational variables (i.e. household size, auto ownership) and psychological variables (i.e. attitudes, opinions, beliefs) which influence the choice process a subject implements to decide a course of action to pursue in conserving transportation related energy.

Psychometric research methods are designed to develop measurement procedures and theories to describe an unseen psychological contruct. Typically, a 'paper and pencil' type format is employed to solicit responses from subjects for a set of questionnaire items designed to measure the construct of interest. The scores assigned to each item are used as a reflection of the relationship between a subject and the construct.

In most cases, analysis of the results of such questionnaires or tests is based on the principles of classical test theory. Typically, a sum of individual items scores would be developed to form a total observed test score. This total test score then serves as an estimate of the subject's true score for the construct under study.

In the case of this particular study, the total observed test score for a subject can range from 0 to 18 using a item score of "0" for did not take action and "1" for did take action. However, many possible combinations of actions yield the same total test score even through the characteristics or content of items comprising this score vary. Lord (60, p. 5) suggests that classical test theory is built on a set of tautalogies which cannot be proven or disproven by the data collected.

The results of an analysis using classical test theory are exclusive to the test and cannot be generalized beyond the test. For example, measures of reliability, validity, and the standard error of estimation are dependent on the characteristics of the subjects and the items comprising a test (61).

Because of these problems, and in order to explore new analysis theories and methods, this research applies the principles of modern test theory or latent trait theory in developing a measure of the hypothesized psychological construct understudy. In theoretical terms, latent trait theory can be defined as a collection of postulates concerning the relationship between data and hypothesized constructs or traits. In practical terms, latent trait theory consists of identifying the relationship between items comprising a guestionnaire and the characteristics of study subjects.

Unlike classical test theory, latent trait theory is called "population free" since the standard error of estimation of the construct, as well as parameters used in mathematical models used to describe individual traits and test items are not dependent on the characteristics of the group sampled or the items comprising a test. In addition, while both theories make use of some hypothetical construct in modern test theory, the ability under study is common to all tests used to measure a phenomenon. The true score concept in classical test theory is used for a specific test only. As will be seen, latent trait theory is much more scientific than classical test theory since it is based on rigorous statistical proofs and not a set of theoretical assumptions.

The first part of this chapter consists of an introduction to the principles of latent trait theory. This includes a description of the strategy to be used in this study to investigate the existence of various market segments. Following this theoretical introduction, the results of applying the principles of latent trait theory will be presented. This analysis suggests that market segments can be formed

based on the estimate of the latent trait under study and provides some interesting information about the characteristics of the action items comprising each questionnaire. The final part of this chapter summarizes the results of this analysis.

Background

Psychometric research methods are designed to develop measurement procedures and theories that are used to describe an unseen psychological construct. Typically, the researcher developes a set of questionnaire items designed to measure a phenomenon of interest and administers the set of such questions to a group of subjects under study. Since the construct is unseen, the scores to the questionnaire items represent the sole indicant of a subject's relationship to the phenomenon.

In this study, the construct of interest consists of a phenomenon describing a subject's overall opinions, beliefs, and attitudes toward transportation energy conservation. This construct is measures through the responses offered by each subject to questionnaire items concerning actions taken to conserve gasoline. The principles of latent trait, or modern test theory, are used to describe a subject's position along a unidimensional continuum reflecting the unseen construct. A subject's position in this space indicates the degree to which a respondent is able or willing to conserve transportation energy and serves as a basis for segmenting the entire sample under study.

Theory

Latent trait theory suggests that the item score assigned to any item g (g = 1, . . ., n) from a questionnaire or test consisting of n items is a reflection of an underlying continuous process called the response tendency or item variable denoted Xg (62, p. 359). This tendency for any item can be described as a linear combination of common factors, Θ_s (s = 1, . . ., k), premultiplied by a set of factor loadings, $_{gs}$, plus a unique factor Ug:

$$X_g = \sum_{s=1}^{k} {}_{gs}\,\Theta_s + Ug. \qquad (5.1)$$

In terms of latent trait theory, the set of common factors is taken to represent the dimensions of the latent space. In most cases, since questionnaires are designed to include items measuring a single construct it is likely that a single common factor is obtained on s = 1. Therefore, the response tendency is described by a linear combination of one common factor or latent trait plus a unique factor (63, p. 116).

The dimensionality of a set of item response tendencies is established by applying the principal factor solution to the matrix of inter item polychoric correlation coefficients. This correlation coefficient reflects the underlying relationship that exists between two item response tendencies and is obtainable from the two-by-two contingency table of observed item scores.

Scoring Strategies

Item responses may be classified in several ways depending on the item scoring strategy and the nature of the item. It is possible to assign item scores on a continuous level, but most often some set of discrete item score categories are used to indicate the level of response tendency for a particular item g. The discrete level includes the nominal level, where item scores are not ordered, and the graded level where item scores, x_g, may be rank ordered. In the graded case, non-negative integers, ($x_g = 0, 1, 2, \ldots, m_g$) are used to categorize responses. If $m_g = 1$, then the result is described as the dichotomous response level (64, p. 32). In this case $x_g = 0$, $x_g = 1$ are the item score categories.

Dimensionality and Local Independence

In latent trait theory the conditional probability of an item response score, denoted $P_{x_g}(\Theta)$ for the graded case, takes an important role. If the latent space is unidimensional this conditional probability is assigned to each position along a unidimensional continuum (i.e. $-\infty < \Theta < \infty$). The conditional probability of an item score given a fixed level of Θ is called the operating characteristic of an item response. If an item is dichotomous then the operating characteristic associated with item score $x_g = 1$ is the item characteristic function denoted $P_g(\Theta)$, and or any item the operating characteristic associated with item score $x_g = 0$ is denoted $Q_g(\Theta)$.

Latent trait theory is based on the assumption of local independence. This suggests that if the dimensions of latent space have been

completely identified, then the distributions of item response categories at a fixed point of the latent space are independent of each other (63, p. 4). This means that for any subject, the response to a particular item is dependent only on the subject's position in latent space and is not related to the responses offered to other items.

As an example, recalling the action items used in this study, if prob (x_1, = 1 Θ) is the probability with which a subject assigned a score of 1 to the question "drive slower?" where 0 means "did not take action" and 1 means the subject "did take action," and if prob (x_2 = 1 Θ) is the probability with which a subject assigned a score of 1 to the question "carpool to work?," the assumption of local independences suggests:

$$\text{Prob}\ (x_1 = 1\ \Theta,\ x_2 = 1\ \Theta) = \text{Prob}\ (x_1 = 1\ \Theta) \cdot \text{Prob}\ (x_2 = 1\ \Theta). \tag{5.2}$$

Mathematical Models

Several mathematical models have been suggested to describe the operating characteristics of an item response category. Perhaps the easiest way to describe these models is the case where there are only two graded response categories (i.e. $x_g = 0, 1$) and a unidimensional homogeneous thinking process, or where the latent space is one dimensional and the reasoning process applied for the set of item categories is similar in quality (64, p. 19).

Let $Q_g\ (\Theta)$ and $P_g\ (\Theta)$ be item operating characteristic functions for item scores $x_g = 0, 1$. In addition, let a_g be a parameter called item discrimination of the rate at which the conditional probability of

an item response varies with the latent trait Θ; let b_g be another parameter called the difficulty parameter which positions the item response operating characteristic along the latent continuum; let c_g be the guessing parameter designed to account for the probability of a random guess as the basis for an item score assignment. The random guessing parameter is important in mental tests of ability, but it plays no role in items dealing with attitudes or opinions since it is unlikely that a subject would base a response on a guess in these cases. Finally, let D be a scaling parameter.

Models describing the relationship between item scores and the latent trait may take any number of functional forms. The most common types include the normal ogive and logistic models using one, two or three parameters. Selected models are outlined below. In each case, the item response characteristic function, $P_g(\Theta)$, is presented. This refers to the conditional probability of item score $x_g = 1$ and is used as a convenient basis for introducing the model forms used in this study. Recall that such an item score means that a subject did endorse a transportation energy conservation action.

The one parameter logistic model or Rasch Model for the item response characteristic function is given by:

$$P_g(\Theta) = \frac{1}{1 + e^{-D(\Theta - b_g)}}. \qquad (5.3)$$

This model assumes that all discrimination parameters are constant and equivalent. It has been used, in part, because it requires that only one parameter be estimated.

The two parameter normal ogive model includes the discrimination parameter to account for the different rate of change between item responses for each item and the latent trait. It is given as:

$$P_g(\Theta) = \int_{-\infty}^{a_g(\Theta - b_g)} \frac{1}{\sqrt{2\Pi}} e^{-t^2/2} dt \quad . \tag{5.4}$$

The two parameter normal ogive model is similar in shape to the logistic model and if an appropriate scaling factor, D = 1.7, is selected for the logistic model the difference between the results for the two forms is less than 0.01 over the range of Θ (59, p. 14). The Z-parameter logit model is given as:

$$P_g(\Theta) = \frac{1}{1 + e^{-Da_g(\Theta - b_g)}}. \tag{5.5}$$

The three parameter forms of each of these models includes a guessing parameter which some have found useful in mental testing of ability. For example, if a subject is asked to select one of four choices in an item designed to measure verbal ability, it is theoretically possible that a subject has a one in four chance in selecting the correct response. For completeness, the three parameter logistic and normal ogive models are given below:

$$P_g(\Theta) = c_g + \frac{1 - c_g}{1 + e^{-Da_g(\Theta - b_g)}}, \tag{5.6}$$

$$P_g(\Theta) = c_g + (1 - c_g) \int_{-\infty}^{a_g(\Theta - b_g)} e^{-t^2/2} \, dt \quad . \qquad (5.7)$$

Several points concerning these models should be noted. First, each is a monotonically increasing function of Θ with range $-\infty < \Theta < +\infty$ and domain $0.0 < Pg(\Theta) < 1.0$. These models, thus, represent the conditional probability of an item response given ability. If a subject's value of Θ has been determined, this can be used to suggest the probability of a particular response. No other information is necessary. More importantly, the concept of latent trait is not based on any particular set of items or group. Theoretically, the value of the latent trait assigned to any subject is "population free" and applies to any additional set of items or any other group, except for scaling differences and sampling fluctuation.

This result suggests that once a set of items has been calibrated that it is possible to use them to estimate the value of the latent trait for any subject without administering the entire set of test items. Rather, the relationship between item responses can be used to focus on a subject's position along the latent continuum.

Figure 5.1 presents a typical item characteristic function using the normal ogive model with no random guessing and $a_g = 1.0$, $b_g = 0.0$. Note that the discrimination parameter is proportional to the scope of the inflextion point, and in the case of the normal ogive model, this slope is equal to $a_g/\sqrt{2\Pi}$ or $0.3989\ a_g$. In the logistic model this slope is given as $0.4250\ a_g$. In addition, the point at which $P_g(\Theta) = 0.5$ corresponds to the value $\Theta = b_g$ when there is no

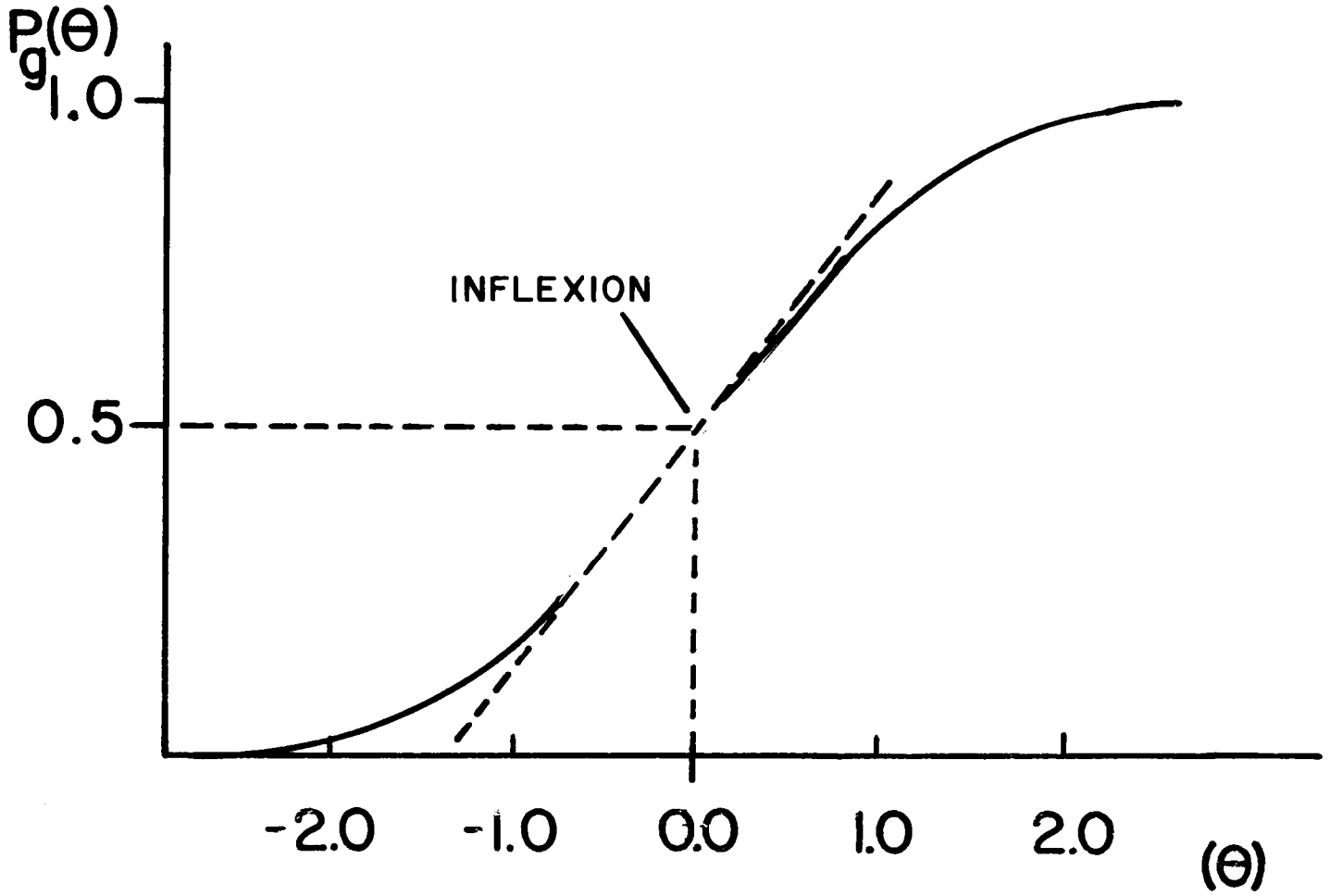


FIGURE 5.1

MEANING OF ITEM PARAMETERS

random guessing. The discrimination parameter is always $0 < a_g < +\infty$ while the difficulty parameter has the range $-\infty < b_g < +\infty$.

Samejima (64, p. 23) points out that any graded item where there are more than two response categories can be easily rescored such that any category less than some value of x_g is 0 and categories greater than or equal to x_g are assigned to category 1. This dichotomization results in m_g set of item response characteristic functions where $x_g = 0, 1, \ldots m_g$. These can be used to derive the operating characteristics for each response category. Figure 5.2 presents the results of this effort for an example where $m_g = 3$ using the two parameter normal ogive model. Note that with the exception of extreme values (i.e. $x_g = 0$ or $x_g = 3$) the operating characteristic functions are neither strictly increasing nor decreasing in value along the latent continuum.

Parameter Estimation

In practical terms, the use of latent trait theory requires the estimation of a number of parameters related to the operating characteristic functions and the individual parameter describing the position of a subject in latent space. Several approaches to estimate these parameters have been suggested including maximum likelihood, Baysian approaches and marginal maximum likelihood (65). In addition, a method combining the results of item calibration using principal factor analysis to estimate item parameters and maximum likelihood estimates of individual parameters has been suggested (63, p. 115). Because this research emphasizes this latter approach, the following discussion outlines the theory behind parameter estimation utilizing this method.

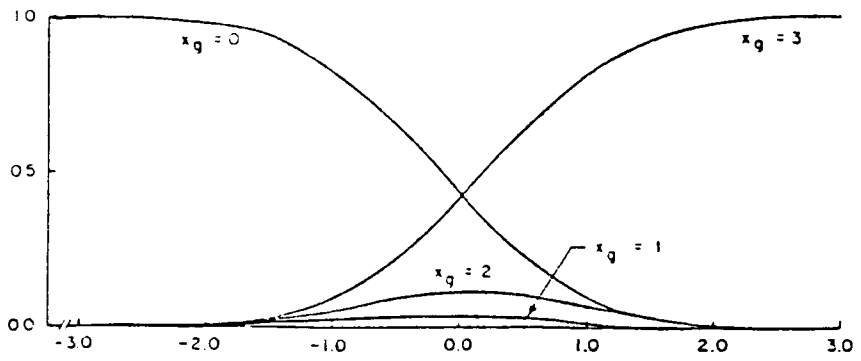


FIGURE 5.2

EXAMPLE ITEM OPERATING CHARACTERISTICS

However, some analysis has been included concerning the use of joint maximum likelihood estimation.

In the use of tetrachoric correlation coefficients and factor analysis for two parameter estimation, the two parameter normal ogive model serves as the functional form for the item characteristic functions. On the other hand, in the use of both joint maximum likelihood estimation of both item and individual parameters the two parameter logistic model served as the functional form. This analysis relies on the use of the Logist 5, version 1.0 computer program, developed by Wingersky, Barton, and Lord (66), that became available for use during this research. Computer programs written by Samejima and McCarter (67, 68) were used to apply the normal ogive model.

For purposes of this discussion, the approach using the normal ogive model will be called the tetrachoric method while the alternative is labeled the Logist 5 approach.

The tetrachoric method. For purpose of this analysis the use of the tetrachoric correlation is required since it is these quantities which outlines the continuous underlying relationship between response tendencies. The phi coefficient, on the other hand measures the relationship between observed item scores. The range of a tetrachoric correlation coefficient is not constrained by the difficulty levels of items like the phi coefficients so these values can range from +1.0 to -1.0 and in general are larger than corresponding phi coefficients.

Recall earlier that the observed item score category assigned to a particular item represents a reflection of some unseen continuous

process called the response tendency or item variable. Let X_g represent the item variable for Item g, X_h be the item variable for item h and γ_g be the critical value of response tendency x_g and γ_h be this value for x_h. Suppose that when $X_g \geq \gamma_g$ iten g is scored $x_g = 1$ and when $X_g < \gamma_g$ the item is scored $x_g = 0$. In addition, a similar process is assumed for item h or when $X_h \geq \gamma_h$ then $x_h = 1$ and when $X_h < \gamma_h$ then $x_h = 0$. One way to describe the correlation between these item variables is to use the observable joint relationship between item scores by calculating a phi coefficient. However, when the underlying process is of real interest, then a tetrachoric correlation should be used. No simple formula exists to develop this coefficient; however, several computer algorithms are available to calcualte this quantity.

It was suggested earlier that the item variable X_g can be described in factor analysis terms as a linear combination of a common and unique factor. If it is assumed that the distribution of Θ is normal with mean zero and unit variance, and no guessing is assumed in assigning item score categories, this factor analysis based model of the item variable can be used to develop parameters for the two parameter normal ogive model (62, p. 370)

Birnbaum (62, p. 373) shows that if Λ_g is the factor loading for item variable X_g on common factor Θ this loading can be used to estimate item parameters for the two parameter normal ogive item response characteristic function. The item discrimination parameter a_g is defined:

$$a_g = \frac{\Lambda g}{\sqrt{1 - g\Lambda^2}} . \qquad (5.8)$$

The item difficulty parameter b_g is defined using γ_g or the normal deviate corresponding to the proportion of subjects with item score $x_g = 1$ as follows:

$$bg = \frac{\gamma_g}{\Lambda_g} . \tag{5.9}$$

In summary, item calibration for the two-parameter normal ogive model consists of factor analyzing the matrix of tetrachoric correlation coefficients. The factor loadings corresponding to the common factor shared by these item variables are then used to estimate the item discrimination and item difficulty parameters. In practical terms the factor analysis results also help to establish the dimensionality of the set of items. If the latent root for the first common factor is much larger than all the remaining eigenvalues, then the latent space can be considered to be unidimensional (60, p. 21).

Once item parameters have been estimated, through the principle of local independence, the position of a subject in the latent space can be explored. Recall that under the assumption of local independence the conditional joint distribution of a set of item responses is equal to the product of the separate conditional joint distribution of each individual item.

For simplicity, assume that a set of n dichotomous items have been collected with item categories $x_g = 0$, $x_g = 1$. If u_g is defined as a special case of x_g when item g is a dichotomous item then $u_g = 0$ and $u_g = 1$ are used to designate item scores. Let the response patterns or the collection of item score be $V = (u_1, u_2, u_3)$.

The conditional joint distribution of item scores for this response pattern is called the response pattern operating characteristic $P_v(\theta)$, and because of the principle of local independence this can be written:

$$P_v(\theta) = \prod_{g=1}^{n} P_g(\theta)^{u_g} Q_g(\theta)^{1-u_g} \tag{5.10}$$

Since θ is the unknown parameter to be estimated this function can be considered as a likelihood function for estimating $\hat{\theta}$. The maximum likelihood equation can be formed from this function and used to find $\hat{\theta}$. Since the value of $\hat{\theta}$ that maximizes equation is also the same as the value of $\hat{\theta}$ which maximize the natural logarithm of this quantity, the product notation can be replaced by a summation sign and the likelihood equation expressed as:

$$\frac{\partial}{\partial \phi} \ln P_v(\theta) = \frac{\partial}{\partial \phi} \sum_{g=1}^{n} (\ln P_g(\theta)^{u_g} Q_g(\theta)^{1-u_g}) = 0. \tag{5.11}$$

Samejima (64, p. 9) shows that $\hat{\theta}$ exists and is unique for any response pattern and introduces the use of the basic function $A_{u_g}(\theta)$ for each response category to facilitate the numerical solution to this problem:

$$A_{u_g}(\theta) \begin{cases} \frac{\partial}{\partial \phi} \ln Q_g(\theta) = 0 & u_g = 0 \\ \frac{\partial}{\partial \theta} \ln P_g(\theta) = 0 & u_g = 1. \end{cases} \tag{5.12}$$

The likelihood equation thus becomes:

$$\sum_{u_g \varepsilon V}^{n} A_{u_g} = 0 \qquad (5.13)$$

A numerical search algorithm is applied to find the value of $\hat{\Theta}$ which satisfies this equation.

A problem is presented, however, in the case where the response pattern for a subject consists of all zeros or ones. In this case, the likelihood function has no local maximum. The response pattern operating characteristics in these cases are either strictly increasing or decreasing functions of the latent trait. In the case where all zeros are recorded in a response pattern, the value of the operating characteristic increases asympotically toward 1.0 as the value of Θ approaches $-\infty$. In the case of all ones in a response pattern the opposite occurs. That is, the operating characteristic increases toward 1.0 asympotically toward 1.0 as Θ nears $+\infty$.

The implications of this result suggests, in part, that the items do not sufficiently span the range of the value of the latent trait in the group under study. It suggests that the questions do not provide enough information about the subject to provide an estimation of the latent trait beyond a certain range.

Unfortunately, this becomes a problem in this particular study. It results from the fact that many subjects in both survey years did not endorse any actions. The value of the latent trait for these subjects cannot be estimated and have been arbitrarily assigned a high negative value.

In spite of this limitation, however, the study does suggest that some consistencies emerge from the group comprised of subjects endorsing no conservation actions. Future studies of this type will have to take such possibilities into account if latent trait theory is to be fully explored.

Another topic of interest includes the validation of item response characteristic models. In this research model validation consists of grouping the estimated individual parameters for each sample into 21 groups of approximately size 70 and calculating the proportion of group members endorsing a particular action Π_g or not endorsing an action $(1 - \Pi_g)$. These ratios are then plotted over theoretical item characteristic functions to compare the theoretical result with the observed proportions. In this manner the fit of each item characteristic function can be examined.

So far, the discussion has explained how item and individual parameters have been estimated for the respective surveys. In addition, however, an understanding of the accuracy of these individual parameters can be attained by the use of various information functions.

The concept of information is important to latent trait theory since it represents a quantification of the accuracy of a set of test items to measure the construct or latent trait of interest. The amount of information provided by a single item is defined as the conditional expectation of the item response function, $I_{u_g}(\Theta)$, for a fixed level of Θ (64, p. 37). This is given by:

$$I_g(\Theta) = \sum_{u_g=0}^{1} I_{u_g}(\Theta)\, P_g(\Theta)^{u_g}\, Q_g(\Theta)^{1-u_g}. \tag{5.14}$$

In more specific terms $I_{u_g}(\Theta)$ is described as the negative of the first derivative with respect to Θ of the basic function:

$$I_{u_g}(\Theta) = \frac{-\partial}{\partial\Theta} A_{u_g}(\Theta). \tag{5.15}$$

It can be shown mathematically (62, p 460; 64, p. 37) that, for the dichotomous case, the item information function is given as:

$$I_g(\Theta) = \left\{ \frac{\frac{\partial}{\partial\Theta} P_g(\Theta)}{P_g(\Theta)\, Q_g(\Theta)} \right\}^2 \tag{5.16}$$

Of additional interest is the concept of the amount of information contained in an entire test or the test information function. Let $I_v(\Theta)$ represent the response pattern information function, Samejima (64, p. 37) shows that this function is given as:

$$I_v(\Theta) = \sum_{u_g \in V} I_{ug}(\Theta) . \tag{5.17}$$

The test information function $I(\Theta)$ is then given as the conditional expectation of the response pattern information function:

$$I(\Theta) = \sum_{v} I_v(\Theta)\, P_v(\Theta) . \tag{5.18}$$

It can be shown, however, that $I(\Theta)$ can be rewritten as the sum of the item information functions for a fixed level of Θ (62, p. 40):

$$I(\theta) = \sum_{g=1}^{n} I_g(\theta) \quad . \tag{5.19}$$

The value of the test information function lies in the relationship between the test information function and the accuracy of the maximum likelihood estimate $\hat{\theta}$. In the case of a unidimensional latent trait, the asymptotic standard error of the maximum likelihood estimate $\hat{\theta}$ given θ, is $I(\theta)^{-\frac{1}{2}}$ (62, p. 457). This knowledge can be used to stablish confidence intervals for the estimates $\hat{\theta}$ and allows an enhanced understanding of the variability in the accuracy of these estimates over the range of the latent trait.

The concept of information can be exploited in test construction. For example, a set of items which have been calibrated may be selected in such a manner that allows a highly accurate estimate of the latent trait in a position along the continuum. This approach permits the researcher to design tests so that the error of classifying a subject above or below a certain range is minimized.

This research includes the development of test information functions for the survey years. These results describe the range of the latent trait which provide the most accurate estimates of the latent trait. These information functions are also used to construct confidence intervals for the estimates of the latent trait.

Logist 5 approach. A limited amount of analysis using the Logist 5 computer program developed by Wingersky, Barton, and Lord (66) has been included in this research. This includes the estimation of item

parameters following the two parameter logistic model and the estimation of the latent trait for each subject.

Given a set of n items and N subjects and letting $P_g(\Theta_a)$ be the probability of item score $u_g = 1$ for subject "a" at level Θ_a and $Q_g(\Theta_a)$ be the probability of item score $u_g = 0$ for subject "a" at level Θ_a, the Logist 5 Fortran Program is designed to maximize the function:

$$\text{Max } L(\Theta, a_g, b_g, c_g) = \prod_{a=1}^{N} \prod_{g=1}^{n} [P_g(\Theta_a)]^{u_g} [Q_g(\Theta_a)]^{1-u_g} \quad (5.20)$$

The program operates in a four step fashion each comprised of a number of stages. Essentially, individual parameters and item parameters are estimated in a joint fashion designed to maximize the likelihood function of observations. The user may input various estimation options which influence the results. Only subjects which have endorsed at least one action are included in the estimation process.

The output includes item parameters and estimates of the latent trait for each subject. LOGIST 5 is designed for the three parameter logistic model for purposes of this research. However, the guessing parameter has been held equal to 0.0. Any subject not endorsing any action has been assigned a high negative value of the latent trait. The discussion of these results consists of comparing the parameter estimates from these results and the tetrachoric method.

The analysis using the Logist 5 method is limited. However it is included to investigate the implications of the two approaches to parameter estimation. As the discussion of results progresses, the comparisons and contrasts of these approaches will be developed.

Research Strategy

To summarize, the major portion of the research documented in this chapter consists of the results of the application of the tetrachoric method to item and individual parameter estimation. This includes plots of the item characteristic functions, test information functions and the relatively frequency ratios described previously. The goal of this process is to help identify useful bases for segmenting the surveys into groups for energy conservation policy, and program development. In addition, another goal is to suggest new approaches to analysis of transportation planning questions using latent trait theory. The following steps are included in the application of the tetrachoric method:

1. Obtaining tetrachoric correlation coefficients using a computer program developed by Samejima and McCarter (67).
2. Principal factor analysis of the matrix of interaction tetrachoric correlation coefficients using the factor analysis program "factor" from the BMDP system.
3. Estimation of item parameters for the two parameter normal ogive model using the program developed by Samajima and McCarter (68).
4. Estimation of individual parameters θ using a program developed by Samejima (69).
5. Development of test information function using program developed by Samejima (70).
6. Model validation through calculating frequency ratios or the proportion of subjects in selected groups, defined

by similar values of Θ, selecting particular item score categories, and comparing these proportions with theoretical probabilities described by item response operating characteristics.

7. Development of market segments by grouping subjects with similar estimates of the latent trait and examining group profiles by cross tabulation of group membership with selected independent variables.
8. Development of selected policies to meet needs of segments and interpretation of results.

In addition, the following steps are taken to apply the Logist 5 method:

1. Estimation of item parameters and individual parameters for two parameter logistic model using Logist 5 computer program (66).
2. Comparison of item parameters and individual parameters with previous results using the normal ogive model.

The remaining portion of this chapter presents the results of the application of latent trait theory for market segmentation. The next section concentrates on the use of the tetrachoric method. Following this the results of item and individual parameters using the Logist 5 approach are presented. This entire analysis suggests that latent trait theory has some real potential for application in the field of transportation energy planning.

Tetrachoric Method Results

The first step in the application of latent trait theory using the tetrachoric method described earlier involves the estimation of tetrachoric correlation coefficients. These quantities represent the unobservable relationship between item variables or response tendencies as reflected in the observable item scores. In order to use this approach in the development of item parameters for the two parameter normal ogive model the value of the latent trait, θ, is assumed to be normally distributed in the sample population. In addition, it is assumed that no guessing is permitted in assigning item scores by a subject (62, p. 375). These assumption can be safely applied in the case.

The formula for calculating tetrachoric correlation coefficients is quite complex, but fortunately a computer program written by Samejima and McCarter (67) can be used to calculate these quantities. In simple terms, the algorithm is designed to use the observed contingency table containing the distribution of item scores for each item pair to find the estimate of the tetrachoric correlation. This involves fitting a polynomial of a specified degree that best describes the unseen bivariate relationship between items.

With this brief background, the next section outlines the results of the estimation of the tetrachoric correlation coefficients for each survey sample year. Note that each correlation matrix has been rearranged in a manner to aid in identifying clusters of variables.

Tetrachoric Correlation Matrices

Table 5.1 presents the matrix of tetrachoric correlation coefficients for the 1979 survey and Table 5.2 presents corresponding results for the 1980 survey. Note that the coefficients in this table are generally larger in value than the associated values in the previous tables of phi coefficients. Notice that all coefficients are positive suggesting a single common factor between items. Note, however, that several coefficients have been assigned a zero value. This is a result of truncation by the computer program which occurs in cases of extremely small estimates of the tetrachoric correlation coefficients for item pairs.

In an effort to help clarify clusters of variables in data sets, an attempt has been made to suggest possible action clusters from the tetrachoric correlation coefficients. This represents a preliminary attempt to assess the dimensionality of the hypothesized latent trait. If a single strong underlying common factor can be established, then the latent trait can be described as unidimensional. Before factor analysis results are presented, however, a visual inspection of the data will provide some insight into the anticipated results. Tables 5.3 and 5.4 suggest possible variable clusters presented by the matrices of interaction tetrachoric correlation coefficients.

1979 survey results. It appears that at least three major clusters of variables emerge from the matrix of tetrachoric correlation coefficients developed from the 1979 survey sample. One large cluster consisting of actions related to shopping trip behavior and auto use

TABLE 5.1

TETRACHORIC CORRELATION COEFFICIENTS, 1979

Action	17	16	6	7	13	12	11	9	10	8	5	1	15	18	3	14	4	2
17	1.000																	
16	0.705	1.000																
6	0.458	0.406	1.000															
7	0.250	0.321	0.545	1.000														
13	0.355	0.334	0.376	0.429	1.000													
12	0.262	0.313	0.429	0.240	0.403	1.000												
11	0.184	0.249	0.373	0.295	0.389	0.432	1.000											
9	0.229	0.344	0.291	0.152	0.312	0.335	0.583	1.000										
10	0.335	0.362	0.380	0.240	0.423	0.317	0.626	0.642	1.000									
8	0.236	0.276	0.331	0.373	0.386	0.276	0.464	0.403	0.579	1.000								
5	0.281	0.365	0.176	0.296	0.310	0.359	0.406	0.466	0.511	0.438	1.000							
1	0.227	0.350	0.182	0.156	0.294	0.324	0.373	0.434	0.418	0.357	0.495	1.000						
15	0.259	0.275	0.000	0.308	0.243	0.237	0.259	0.180	0.237	0.406	0.243	0.284	1.000					
18	0.047	0.079	0.103	0.270	0.202	0.146	0.263	0.109	0.131	0.310	0.072	0.270	0.510	1.000				
3	0.030	0.181	0.075	0.293	0.147	0.030	0.144	0.073	0.111	0.383	0.003	0.165	0.469	0.659	1.000			
14	0.168	0.248	0.219	0.349	0.378	0.329	0.296	0.291	0.359	0.302	0.259	0.255	0.172	0.298	0.231	1.000		
4	0.321	0.251	0.328	0.302	0.263	0.210	0.122	0.000	0.208	0.318	0.121	0.166	0.301	0.307	0.330	0.156	1.000	
2	0.148	0.328	0.000	0.129	0.290	0.118	0.233	0.222	0.364	0.347	0.258	0.219	0.302	0.161	0.256	0.102	0.215	1.000

TABLE 5.2

TETRACHORIC CORRELATION COEFFICIENTS, 1980

Action	4	7	6	11	9	10	8	1	5	2	12	13	14	16	17	3	18	5
4	1.000																	
7	0.423	1.000																
6	0.242	0.686	1.000															
11	0.269	0.404	0.483	1.000														
9	0.116	0.405	0.331	0.605	1.000													
10	0.251	0.332	0.387	0.571	0.700	1.000												
8	0.334	0.366	0.301	0.423	0.487	0.536	1.000											
1	0.148	0.278	0.262	0.493	0.559	0.542	0.409	1.000										
5	0.153	0.294	0.383	0.401	0.440	0.482	0.379	0.580	1.000									
2	0.236	0.124	0.208	0.302	0.359	0.377	0.486	0.435	0.347	1.000								
12	0.225	0.339	0.146	0.451	0.462	0.495	0.401	0.401	0.442	0.238	1.000							
13	0.152	0.302	0.313	0.388	0.368	0.386	0.341	0.404	0.349	0.297	0.523	1.000						
14	0.116	0.195	0.246	0.473	0.403	0.433	0.379	0.367	0.359	0.301	0.409.	0.383	1.000					
16	0.220	0.328	0.183	0.265	0.362	0.429	0.339	0.369	0.352	0.305	0.387	0.371	0.346	1.000				
17	0.167	0.257	0.258	0.303	0.229	0.206	0.207	0.303	0.228	0.145	0.338	0.249	0.405	0.405	1.000			
3	0.334	0.226	0.171	0.122	0.106	0.131	0.326	0.183	0.073	0.231	0.196	0.200	0.000	0.000	0.188	1.000		
18	0.206	0.135	0.131	0.333	0.208	0.157	0.257	0.230	0.180	0.169	0.195	0.265	0.288	0.000	0.250	0.575	1.000	
15	0.224	0.284	0.229	0.276	0.252	0.319	0.325	0.317	0.258	0.255	0.333	0.246	0.237	0.229	0.113	0.392	0.525	1.000

TABLE 5.3

ACTION CLUSTERS SUGGESTED BY INTERPRETATION OF INTERACTION TETRACHORIC CORRELATION COEFFICIENTS, 1979

Cluster	Item	Actions in Cluster
1	1	Drive slower
	8	Shop on way home from work
	5	Tune car
	9	Shop less frequently
	10	Combine shop and other trips
	11	Shop closer to home
2	3	Bus or subway for work
	15	Use train, bus, or plane for vacation
	18	Bus or subway for non-work
3	6	Move closer to work
	7	Job closer to home
	12	Eliminate use of recreation vehicles
	13	Vacation closer to home
	16	Buy fuel efficient car
	17	Sell car

TABLE 5.4

ACTION CLUSTERS SUGGESTED BY INTERPRETATION OF INTERACTION TETRACHORIC CORRELATION COEFFICIENTS, 1980

Cluster	Item	Actions
1	1	Drive slower
	2	Carpool to work
	5	Tune car
	8	Shop on way home from work
	9	Shop less frequently
	10	Combine shop and other trips
	11	Shop closer to home
	12	Eliminate use of recreation vehicles
	13	Vacation closer to home
	14	Cancel vacation trip
	16	Buy fuel efficient car
2	3	Bus or subway for work
	15	Use train, bus, or plane for vacation
	18	Bus or subway for non-work
3	4	Walk or bike to work
	6	Move closer to work
	7	Job closer to home

and maintenance is suggested. In addition, a second cluster including the actions related to non-auto modes appears to emerge. Finally a smaller cluster of the actions related to replacement or elimination of the automobile, moving actions, and vacation travel also is suggested. Note that three actions including Item 2 "carpool to work," Item 4 "walk or bike to work" and Item 14 "cancel vacation trip" are not included by any cluster. Rather, it appears that these actions are correlated with several item clusters.

1980 survey results. The results of the 1980 analysis suggests a similar grouping of items as those emerging from the 1979 results. However, it appears that the largest cluster of items includes a few more actions than those outlined in the 1979 data. In particular it appears that the vacation related actions are more closely correlated with a large cluster of variables including shopping actions and auto use and maintenance. Once again, however, a cluster of actions relating to use of non-auto modes is suggested. Unlike the 1979 results, no actions are left out of the interpretation of clusters. Recall earlier that several items were not assigned to a particular cluster since it was unclear as to what cluster these items could be assigned.

Now that a subjective analysis of action items has been introduced, it is necessary to continue with any objective evaluation using principal factor analysis. If a large dominate principal factor is present in the data sets, then it will emerge from the results of the principal factor solution. The importance of factor analysis in the use of latent trait theory has been introduced in an earlier section.

Essentially factor analysis is used to help identify the dimensionality of the latent space and to calibrate item parameters in the two parameter normal ogive model.

1979 Principal Factor Analysis Results

In order to explore the dimensionality of the 1979 matrix of tetrachoric correlations the "factor" program available in the BMDP system of statistical programs has been used. The major goal of this analysis is to identify the relationship between response tendencies and the hypothesized latent trait. In evaluating the results it is necessary to concentrate on the configuration of eigenvalues emerging from the results of the computer analysis. More specifically, if the first eigenvalue is much larger than the second latent root and this latter eigenvalue is approximately the same size as the remaining roots then the latent space may be considered unidimensional.

Several principal factor analysis results were developed to confirm the dimensionality of the matrix of tetrachoric correlation coefficients. This included several combinations of the number of factors extracted and the number of actions to be included in the analysis. Table 5.5 summarizes the specific details associated with each solution.

The results were investigated in a series of steps based on the analysis of previous evaluations. For example, the first solution included all action items and required the extraction of eight common factors from the data matrix. This result led to the investigation of the dimensionality based on eight common factors and 15 action items.

TABLE 5.5

ITEMS INCLUDED IN SELECTED FACTOR ANALYSIS RESULTS

Action	Description	Variables Used in Factor Analysis					
1	Drive slower	•	•	•	•	•	•
2	Carpool to work	•	•	•	•	•	•
3	Bus or subway to work	•		•			•
4	Walk or bike to work	•	•	•	•	•	
5	Tune car	•	•	•	•	•	•
6	Move closer to work	•					•
7	Job closer to work	•	•	•			•
8	Shop on way home from work	•	•	•	•	•	
9	Shop less frequently	•	•	•	•	•	
10	Combine shop and other	•	•	•	•	•	•
11	Shop closer to home	•	•	•	•	•	•
12	Eliminate use of rec veh	•	•	•	•	•	
13	Vacation closer to home	•	•	•	•	•	•
14	Cancel vacation trip	•	•	•	•	•	•
15	Use train bus plane or vac	•					
16	Buy fuel efficient car	•	•	•	•	•	•
17	Sell car	•	•				•
18	Bus or subway for non-work	•	•				•
	Total items	18	15	14	12	11	12
	Common factors	8	8	8	6	6	6

This continued until only 11 items served as a basis for analysis using five common factors.

In selecting a suitable solution it was decided that a maximum number of items be included while still attaining a unidimensinal characteristic. In order to understand the implications on dimensionality of various solutions, the incremental contribution of each factor for the various solution were to examined. Table 5.6 contains these results for the 1979 data set.

In all cases, the first common factor explains a large proportion of the total variance relative to the remaining values. In addition, it appears that two secondary factors emerge in each solution. However, the contribution of these factors is much less than the first common factor.

These results surggest a large common factor in the data indicating a unidimensional latent space. Therefore, the principal factor solution based on eight common factors and using all 18 items was selected for further analysis.

Table 5.7 presents the squared multiple correlations for each item with all remaining items that were used as communality estimates. Note that these values are consistently higher than those used earlier in the factor analysis of phi coefficients.

Table 5.8 presents the eigenvalues emerging from the solution. Notice that the value of the last eigenvalue and the ninth latent root are approximately equal, in absolute value, suggesting, as stated in the previous chapter, that the eight common factor solution spans the common factor space. Recall that the eigenvalues are expected to be

TABLE 5.6

INCREMENTAL PROPORTION OF COMMON VARIANCE EXPLAINED FOR SELECTED FACTOR SOLUTIONS, 1979

Common Factors	Total Items in Analysis[a]					
	18	15	14	12	11	12
1	0.312	0.329	0.333	0.360	0.384	0.317
2	0.096	0.076	0.083	0.057	0.047	0.107
3	0.072	0.067	0.048	0.042	0.045	0.082
4	0.051	0.039	0.035	0.033	0.034	0.044
5	0.026	0.028	0.025	0.019	0.019	0.034
6	0.026	0.024	0.023	0.017	0.016	0.027
7	0.019	0.020	0.019			
8	0.017	0.018	0.014			

[a]Note: These items correspond to items described in Table 5.5.

TABLE 5.7

SQUARED MULTIPLE CORRELATIONS OF EACH ACTION WITH ALL OTHER ACTIONS

	Action	Year	
		1979	1980
1.	Drive slower	0.38506	0.54777
2.	Carpool to work	0.32612	0.38298
3.	Bus or subway to work	0.56493	0.51924
4.	Walk or bike to work	0.32364	0.33516
5.	Tune car	0.48841	0.49126
6.	Move closer to work	0.65123	0.67861
7.	Job closer to work	0.54626	0.66923
8.	Shop on way from work	0.52212	0.48040
9.	Shop less frequently	0.53782	0.63500
10.	Combine shop and other	0.66202	0.65475
11.	Shop close to home	0.54138	0.58388
12.	Eliminate use of rec veh	0.43262	0.57433
13.	Vacation closer to home	0.39739	0.42358
14.	Cancel vacation trip	0.33336	0.44540
15.	Use train, bus, or plane or vac	0.52010	0.43725
16.	Buy fuel efficient car	0.62980	0.45488
17.	Sell car	0.61353	0.39103
18.	Bus or subway for non-work	0.56517	0.59443

TABLE 5.8

EIGENVALUE CONFIGURATION RESULTING FROM PRINCIPAL FACTOR SOLUTION BASED ON EIGHT COMMON FACTORS AND 18 ITEMS DEVELOPED FROM MATRIX OF TETRACHORIC CORRELATION COEFFICIENTS, 1979

	Eigenvalue	Cumulative Proportion of Total Variance	Incremental Contribution of Each Factor
1.	5.612386	0.311799	0.311799
2.	1.719118	0.407306	0.095507
3.	1.303316	0.478712	0.072406
4.	0.920682	0.530861	0.051149
5.	0.474762	0.557237	0.026376
6.	0.462401	0.582926	0.025689
7.	0.339270	0.601774	0.018848
8.	0.312881	0.619156	0.017382
9.	0.162404		
10.	0.151883		
11.	0.071372		
12.	0.030184		
13.	0.008939		
14.	-0.020612		

TABLE 5.8 (continued)

Eigenvalue	Cumulative Proportion of Total Variance	Incremental Contribution of Each Factor
15. -0.046587		
16. -0.068527		
17. -0.128394		
18. -0.174099		

positive or zero in value, but because of sampling error negative values may emerge. However, the sum of remaining eigenvalues following number 8 is approximately equal suggesting a satisfactory solution. Table 5.8 shows the incremental proportion of total variance each factor contributes. This result demonstrates the importance of the first common factor and presents the sharp decline in variance explained between the first and second factor as well as the gradual decline in contribution offered by remaining values.

Turning to the matrix of factor loadings for this solution, shown in Table 5.9, the importance of the first common factor is clearly evident. Recall that the factor loadings for the set of orthogonal factors are equivalent to the correlation of items with factors. Therefore, the loadings shown for the first factor are equal to the correlation between item variables or response tendencies and the latent trait.

It should be noted, however, that some large loadings are present for several other factors. In particular, Factor 2 contains such loadings for Item 3 "bus or subway to work," Item 5 "use train, bus or plane for vacation," and Item 18 "bus or subway for non-work." In addition, Factor 3 contains a high loading on Item 6 "move closer to work." Finally, Factor 4 has a large common factor loading for Item 16 "buy fuel efficient car."

The large loadings for these items on Factor 3 and Factor 4 can be considered as a portion of the item's unique variance since these action items are the primary items loading on these factors. That is

TABLE 5.9

1979 UNROTATED FACTOR LOADINGS DEVELOPED FROM MATRIX
OF INTERACTION TETRACHORIC CORRELATION COEFFICIENTS

Action	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7	Factor 8
1	0.549	-0.044	-0.228	0.102	0.129	-0.143	-0.018	-0.139
2	0.404	0.119	-1.165	0.262	-0.116	0.218	0.117	0.111
3	0.386	0.710	0.005	0.026	-0.137	-0.032	-0.140	0.017
4	0.414	0.263	0.283	0.057	-0.082	0.098	0.253	-0.162
5	0.607	-0.241	-0.281	0.123	0.311	0.155	-0.105	-0.267
6	0.585	-0.272	0.572	-0.403	-0.206	-0.086	-0.048	-0.154
7	0.565	0.148	0.362	-0.271	0.225	0.299	-0.236	0.074
8	0.681	0.134	-.156	-0.062	-0.139	0.223	0.026	-0.106
9	0.610	-0.301	-0.369	-0.058	-0.130	-0.162	-0.169	0.069
10	0.736	-0.252	-0.299	-0.047	-0.276	0.120	0.044	0.032
11	0.651	-0.133	-0.261	-0.233	-0.081	-0.095	0.017	0.048
12	0.540	-0.172	0.044	-0.143	0.240	-0.228	0.254	-0.009
13	0.614	-0.051	0.090	-0.085	0.133	0.104	0.203	0.269
14	0.477	0.058	-0.023	-0.148	0.158	-0.062	-0.022	0.198
15	0.491	0.431	-0.073	0.194	0.081	-0.011	0.058	-0.053
16	0.636	-0.190	0.329	0.532	-0.010	-0.115	-0.171	0.131
17	0.522	-0.213	0.401	0.375	-0.059	-0.086	0.039	-0.045
18	0.426	0.695	-0.060	-0.109	0.019	-0.277	-0.029	-0.026
Total Variance	5.612	1.719	1.303	0.921	0.475	0.462	0.339	0.313

to say that these factors may not be true clusters of variables, but consist only of the unique portion of a single item.

Factor 2, however, contains high loadings for two items (greater than 0.500). This result rules out the possibility of labeling these results a unique portion of the variance. In a strict sense, the course of action suggests a multidimensional latent space. However, for purposes of this analysis, it was decided to proceed with the analysis using only one common factor or a unidimensional latent space. This course is not completely unjustified given the dominance of the first common factor and the desire to include all possible action items. In addition, the various solutions attempted suggested a secondary factor in spite of the inclusion of as few as 11 items. Therefore, the factor loadings on the first common factor using all items and eight factors were used to calibrate the normal ogive item parameters.

1980 Principal Factor Analysis Results

An identical strategy to that applied in identifying the dimensionality of the 1979 data set was applied to the 1980 data. The results of this analysis also indicated a large common factor with several secondary factors.

Once again, esimtates of the square multiple correlation between items were used as communality estimates. These were outlined in Table 5.7 for the eight common factor solution. Note in this table that Item 17 "sell car" exhibits a decrease in the size of squared multiple correlation over the 1979 data.

Several factor solutions were developed in a manner similar to the 1979 analysis. These approaches are presented in Table 5.10. Once again the incremental contribution of each factor is outlined in Table 5.10. The dominance of one common factor is evident from these results. Therefore, the eight common factor solution using all 18 items was selected for further analysis.

Table 5.11 presents the eigenvalues developed from the eight common factor solution. The first common factor explains a large proportion of the variance with the remaining latent roots contributing a decreasing amount. The larger size of the first latent root in this solution over the 1979 solution suggests a stronger common factor in response tendencies. This confirms the subjective item clustering performed earlier. Recall that in 1980, the number of items included in this result was larger for the first cluster than in 1979.

Table 5.12 presents the matrix of common factor loadings for the common factor solution. Once again a secondary factor of non-auto mode actions emerges. In addition, however, Item 6 "move closer to work" and Item 7 "job closer to home" have high loadings (greater than 0.500) on the Factor 3. Factor 4 includes a high loading (greater than 0.500) on Item 17 "sell car." This was expected, in part, because of the small value of the squared multiple correlation or communality estimate described earlier.

Once again, however, it was decided to proceed with the analysis using the factor loadings on the first common factor for the eight factor solution. This appears to be a reasonable approach that is required, in part, to maintain the entire number of items. In brief,

TABLE 5.10

INCREMENTAL PROPORTION OF COMMON VARIANCE EXPLAINED FOR SELECTED FACTOR SOLUTIONS, 1980

Common Factors	Total Items in Analysis[a]					
	18	15	14	12	11	12
1	0.339	0.366	0.376	0.409	0.437	0.347
2	0.081	0.054	0.070	0.046	0.044	0.104
3	0.062	0.046	0.043	0.038	0.090	0.083
4	0.042	0.039	0.036	0.035	0.030	0.043
5	0.035	0.036	0.028	0.020	0.022	0.031
6	0.027	0.024	0.023	0.017	0.005	0.017
7	0.020	0.022	0.017			
8	0.017	0.012	0.012			

[a]Note: These items correspond to items described in Table 5.5 (p. 187).

TABLE 5.11

EIGENVALUE CONFIGURATION RESULTING FROM PRINCIPAL FACTOR SOLUTION BASED ON EIGHT COMMON FACTORS AND 18 ITEMS DEVELOPED FROM MATRIX OF TETRACHORIC CORRELATION COEFFICIENTS, 1980

Eigenvalue	Cumulative Proportion of Total Variance	Incremental Contribution of Each Factor
1. 6.109996	0.339444	0.339444
2. 1.450173	0.420009	0.080565
3. 1.113642	0.481878	0.061869
4. 0.756039	0.523881	0.042003
5. 0.630624	0.558915	0.035034
6. 0.477617	0.585449	0.026534
7. 0.357786	0.605326	0.019877
8. 0.306226	0.622339	0.017013
9. 0.196539		
10. 0.118840		
11. 0.092792		
12. 0.053482		
13. 0.014754		
14. -0.002928		
15. -0.033711		
16. -0.114189		
17. -0.146412		
18. -0.194418		

TABLE 5.12

1980 UNROTATED FACTOR LOADINGS DEVELOPED FROM MATRIX
OF INTERACTION TETRACHORIC CORRELATION COEFFICIENTS

Action	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7	Factor 8
1	0.693	-0.139	-0.201	-0.129	0.048	-0.167	0.302	0.169
2	0.516	-0.001	-0.169	-0.243	-0.195	-0.328	-0.050	-0.109
3	0.346	0.699	0.008	-0.085	-0.178	-0.012	0.038	0.123
4	0.375	0.231	0.213	-0.018	-0.308	-0.009	-0.115	-0.005
5	0.619	-0.175	-0.056	-0.066	0.040	-0.085	0.281	-0.076
6	0.564	-0.010	0.677	-0.015	0.238	-0.154	0.077	-0.132
7	0.587	0.044	0.614	0.059	-0.101	0.153	0.010	0.069
8	0.664	0.055	-0.041	-0.201	-0.204	-0.078	-0.209	-0.059
9	0.731	-0.247	-0.081	-0.220	0.182	0.152	-0.148	0.273
10	0.743	-0.221	-0.065	-0.204	-0.022	0.078	-0.112	0.042
11	0.705	-0.081	0.040	-0.018	0.275	0.067	-0.110	0.004
12	0.662	-0.091	-0.194	0.195	-0.193	0.414	0.109	-0.104
13	0.578	-0.023	-0.073	0.108	-0.035	0.093	0.110	-0.153
14	0.587	-0.134	-0.205	0.260	0.178	-0.096	0.110	-0.226
15	0.483	0.351	-0.069	-0.076	-0.012	0.074	0.081	-0.078
16	0.527	-0.024	-0.045	0.201	-0.268	-0.074	-0.001	0.028
17	0.460	0.021	-0.018	0.604	-0.028	-0.218	-0.042	0.225
18	0.439	0.755	-0.232	0.075	0.318	0.021	-0.012	-0.012
Total Variance	6.110	1.450	1.114	0.756	0.631	0.478	0.358	0.306

the generally small values of factor loadings on the secondary factors for most items should not interfere, to a great extent, with the final result.

Calibration of Item Parameters

The next step in the investigation consists of applying the results of the principal factor solution in item calibration. A computer algorithm written by Samejima and McCarter (68) was used, in association with factor loadings and the proportion of subjects assigning an item score of "1" to each action, to calculate item discrimination and difficulty parameters for the two parameter normal ogive model. The formulas to estimate these values were described earlier in equations 5.8 (p. 169) and 5.9 (p. 170). These parameters are presented in Table 5.13 for the 1979 data and Table 5.14 for the 1980 results.

For the most part in comparing the results for each year, it appears that item parameters are very similar. In fact, if no sampling fluctuation is assumed and there is no change in the meaning of the content of any item, it is expected that the parameters should differ only by differences in unit and origin selected for the measurement of the latent trait (62, p. 379).

Normally, the values of the latent trait Θ assigned to each subject will be rescaled with mean 0 and unit variance. If it is assumed that the item parameters in this case are the same, except for scale differences, then it is possible to suggest the nature of this linear relationship by understanding the distribution of the value of the latent trait in each group.

TABLE 5.13

ITEM PARAMETERS FOR THE TWO PARAMETER
NORMAL OGIVE MODEL DERIVED
FROM 1979 DATA

Item g	Discrimination a_g	Difficulty b_g
1	0.6568	0.3751
2	0.4416	2.6800
3	0.4184	3.0474
4	0.4548	3.3476
5	0.7638	0.5364
6	0.7213	3.3504
7	0.6848	2.8780
8	0.9300	0.9814
9	0.7698	0.6317
10	1.0872	0.1099
11	0.8576	0.3500
12	0.6416	2.4444
13	0.7779	1.5557
14	0.5427	2.1203
15	0.5636	2.0320
16	0.8242	1.6067
17	0.6120	2.6467
18	0.4709	2.4132

TABLE 5.14

ITEM PARAMETERS FOR THE TWO PARAMETER
NORMAL OGIVE MODEL DERIVED
FROM 1980 DATA

Item g	Discrimination a_g	Difficulty b_g
1	0.9612	0.2611
2	0.6024	2.1527
3	0.3688	3.1846
4	0.4045	3.3038
5	0.7881	1.0605
6	0.6830	3.2987
7	0.7251	2.9258
8	0.8880	0.7870
9	1.0713	-0.1100
10	1.1101	-0.1234
11	0.9941	0.1072
12	0.8833	1.8003
13	0.7083	1.5410
14	0.7251	2.0192
15	0.5516	1.6909
16	0.6201	1.7876
17	0.5181	3.5623
18	0.4886	2.3422

If σ_{79} is the standard deviation of the distribution of the value of the latent trait for the 1979 data and μ_{79} is the mean of this distribution, and if σ_{80} and μ_{80} are the corresponding values for the 1980 data, then it can be shown that the linear relation between item parameters is given by the following (62, p. 379):

$$a_{g80} = \frac{\sigma_{80}}{\sigma_{79}} (a g_{79}), \tag{5.21}$$

$$b_{g80} = \frac{1}{\sigma_{80}} (\mu_{79} - \mu_{80}) + \frac{\sigma_{79}}{\sigma_{80}} (b_{g79}) \tag{5.22}$$

where a_{g79}, a_{g80} are discrimination parameters for the 1979 and 1980 data and b_{g79}, b_{g80} are the difficulty parameters for 1979 and 1980 results.

The important points to note are that the parameter estimates are invariate from one group to another except for a linear transformation relating scale differences, and that the intercept of the relationship between discrimination parameters passes through the origin while the intercept for the relationship between the difficulty parameters does not.

The relationship between item parameters may be better understood by fitting a linear regression equation describing the 1980 item parameters as a function of the 1979 parameters. Because of the presence of sampling and measurement error in the estimated item parameters used in fitting the regression equation, a method consisting of the use of instrumental variables must be used to estimate the regression parameters.

This procedure, as outlined by Kmenta (71, p. 309), is based on the introduction of a new instrumental variable which is uncorrelated with the error of measurement in the dependent and independent variable and the error term in the resultant regression equation. In addition, these new variables must be correlated with the independent variable.

In this case, values of the independent variables, or 1979 item parameters, are rank ordered from lowest to highest value and divided into three groups. The first group, consisting of six items, is assigned an instrumental variable, Z_g, of -1, the second group of six items is assigned a value of 0, and the final 6 A value of +1. In addition, let $Y^*_g = (Y_{\lessgtr g} -)\bar{Y})$ and $X^*_g = (Y_{\lessgtr g} -)X)$, where Y_g is the observed parameter value for the 1980 data for Item g (g = 1, . . ., 18) and $\bar{Y}$ is the mean of all these values and X^*_g, $\bar{X}$ are the associated values for the independent variables defined by the item parameters from the 1979 data.

Kmenta (71, p. 309) shows that consistent estimations of the scope, β_1, and intercept, β_0, are given by the following:

$$\beta_1 = \frac{\sum_{g=1}^{n} Y^*_g \; Z_g}{\sum_{g=1}^{n} X^*_g \; Z_g} \quad , \tag{5.23}$$

$$\beta_0 = \bar{Y}^* - \beta_1 \; \bar{X}^*. \tag{5.24}$$

Using this method, the linear relationship between 1979 item parameters and 1980 item parameters have been estimated. The results of this effort are summarized in Table 5.15. In addition, Figures 5.3 and 5.4

TABLE 5.15

RESULTS OF FITTING REGRESSION OF 1980 ITEM PARAMETERS ON 1979 ITEM PARAMETERS

	a_g Item Discrimination	b_g Item Discrimination
Intercept	0.0784	-0.2661
Standard Deviation	(0.1559)	(0.0361)
t	0.5216	-7.3712
Score	0.9560	1.0989
Standard Deviation	(0.1503)	(0.0927)
t	6.1321	11.8554

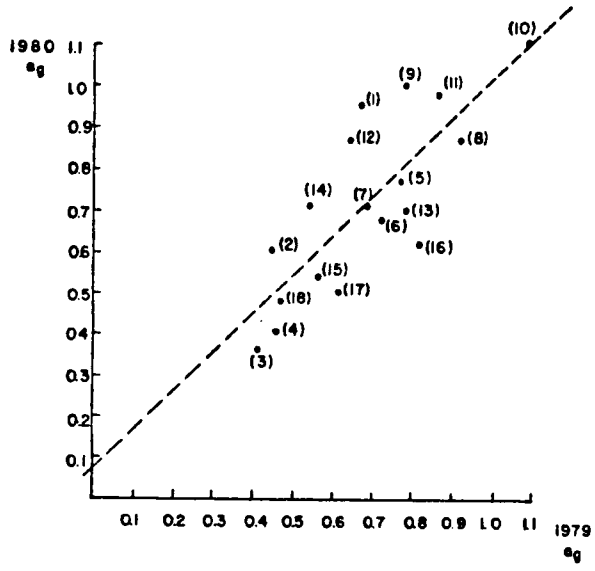


FIGURE 5.3

PLOT OF ITEM DISCRIMINATION PARAMETERS

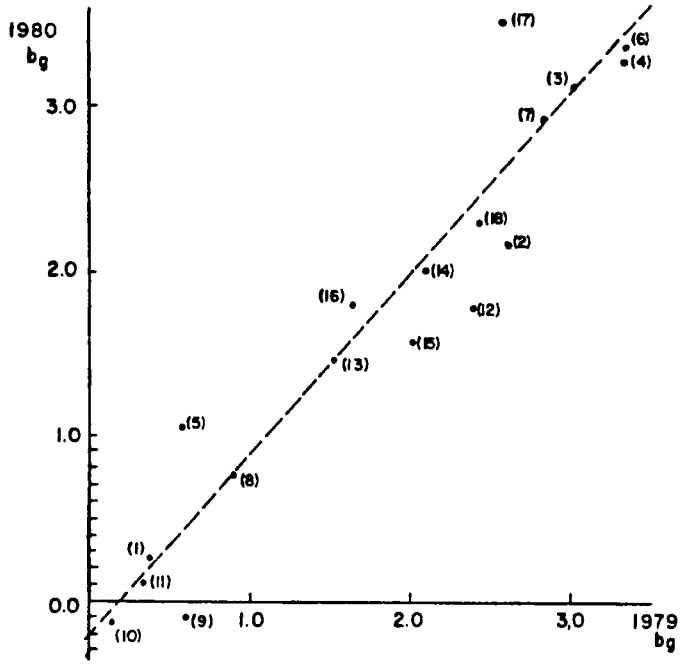


FIGURE 5.4

PLOT OF ITEM DIFFICULTY PARAMETERS

contain a scattergram of the parameters along with a plot of the respective regression equations.

The results suggest that the item parameters estimated from the 1980 data, as a group, tend to have higher discrimination and lower difficulty when compared to the 1979 item parameters. The regression line estimated for the difficulty parameters indicates that the mean value of the latent trait estimated using the 1980 data is higher than the 1979 data and must be adjusted downward to find a comparable 1979 estimate. This is suggested by reference to equation 5.22 where it is shown that the intercept term for the regression equation includes the difference between group means.

In addition, the results suggest a slightly higher group variance for the 1979 group as opposed to the 1980 results. This is indicated by the slope estimates for each model. Recall that in the case of the relationship between item discrimination parameters that this is given as the ratio of group standard deviations as shown in equation 5.21. This value will be less than one only if $\sigma_{79} < \sigma_{80}$. A similar conclusion can be drawn from examining equation 5.22. In this case the estimated scope is consistent with the expected result.

Finally, the results of the parameter estimation for the intercept terms show that in the case of the discrimination parameter this estimated parameter is not significantly different from zero. This is also expected, based on the relationship outlined in equation 5.21. Note, that as expected, the intercept term is significant in the case of the item difficulty parameters.

These results suggest that item parameters estimated for the two samples are fairly consistent, except for scale differences and sampling fluctuation. It is important to remember the relationships between parameters when examining the results of latent trait estimates. In general, the values of the latent trait estimates are expected to be slightly higher in the 1980 sample data. This difference in part, reflects the relationship described above and is not entirely related to a shift in the value of the latent trait between groups.

Item Characteristic Functions

Item parameters developed for each survey provide the basis for development of item characteristic functions for the two parameter normal ogive model. Recall that this function describes the conditional probability of an item response $x_g = 1$ for a given level of the latent trait. It can also be regarded as the regression of item score on the latent trait. In this case, an item score $x_g = 1$ is assigned when a subject indicates that a particular action had been taken to conserve transportation energy.

Using a computer program written by Samejima (72) and the estimated item parameters, the set of 18 item characteristic functions for each sample have been developed. The results have been used to develop a set of figures contained in Appendix B. These figures serve as a basis for the following discussion which presents a brief overview of the results.

In general, items which have the highest values of discrimination also have low values of difficulty. In particular, actions related to shopping travel seem to exhibit this tendency, while many actions requiring large changes in behavior also have high difficulty levels and low discrimination. In summary, in the context of these actions items, the results appear to confirm the pattern of very high difficulty, low discrimination items.

A fairly consistent set of results emerge for the various items over the time period studied. A review of the figures in Appendix B supports the observation that item parameters emerge in a manner consistent with action item content.

The results of the calibration of item characteristic functions serve as a basis for estimating the value of the hypothesized latent trait for each subject. These results are outlined below.

Maximum Likelihood Estimates of Individual Parameters

Applying computer programs written by Samejima (69) the set of individual item parameters for each subject have been developed. This program is designed to find the maximum likelihood estimates for each pattern by using the response pattern offered and the item parameters calibrated earlier. As explained in a previous section, the analysis method makes use of a numerical search procedure that finds the value of the latent trait which maximizes the likelihood equation.

Immediately, it must be noted that a total of 264 subjects in the 1979 survey and 256 subjects in the 1980 survey did not endorse any of the 18 actions. As stated earlier, this means that the estimate of the

value of the latent trait for these subjects cannot be found. Therefore each of these subjects has been assigned a high negative value of -99.9.999 for the value of the latent trait. The remaining values for other subjects are standardized estimates with zero mean and unit variance. If the distribution of the latent trait is normal in each group, it is expected that 68% of the subjects will have estimates of ±1.0 from the mean of zero. In theory 95% of the total should lie within ±2.0 from the mean.

Figures 5.5 and 5.6 present the frequency distributions of the estimated values of the latent trait for each survey year. Note that because of the presence of the high negative values, this distribution is skewed to the positive side in each case. In theory these distributions should be centered over 0.0, however it is not the case in this result because of the high negative values, which have not been included in the analysis.

This result is a problem which emerges in the use of latent trait theory. Specifically, it suggests that insufficient information exists to describe subjects not endorsing any action item. This is clearly reasonable and is a reflection of the limitations present in using only 18 actions to measure the latent trait. This can be reduced if additional actions are added which subjects with lower values of the latent trait may endorse. However, in the case of this data it is not possible to find these estimates. However, a large number of subjects do remain which can provide an meaningful basis for segmentation.

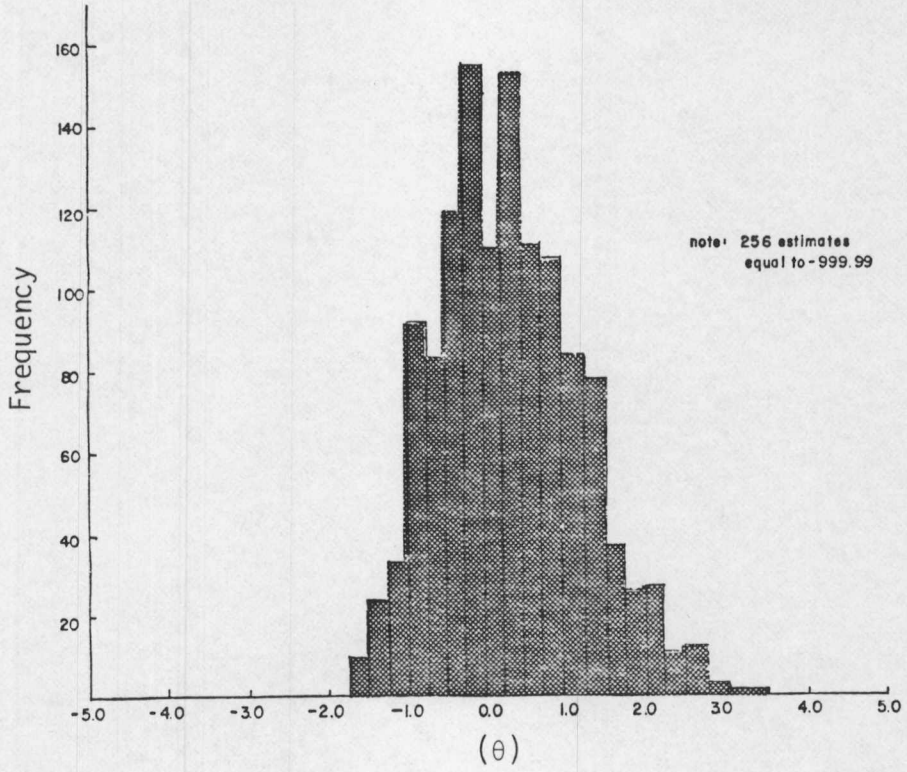


FIGURE 5.5

FREQUENCY DISTRIBUTION FOR 1979
MAXIMUM LIKELIHOOD ESTIMATES

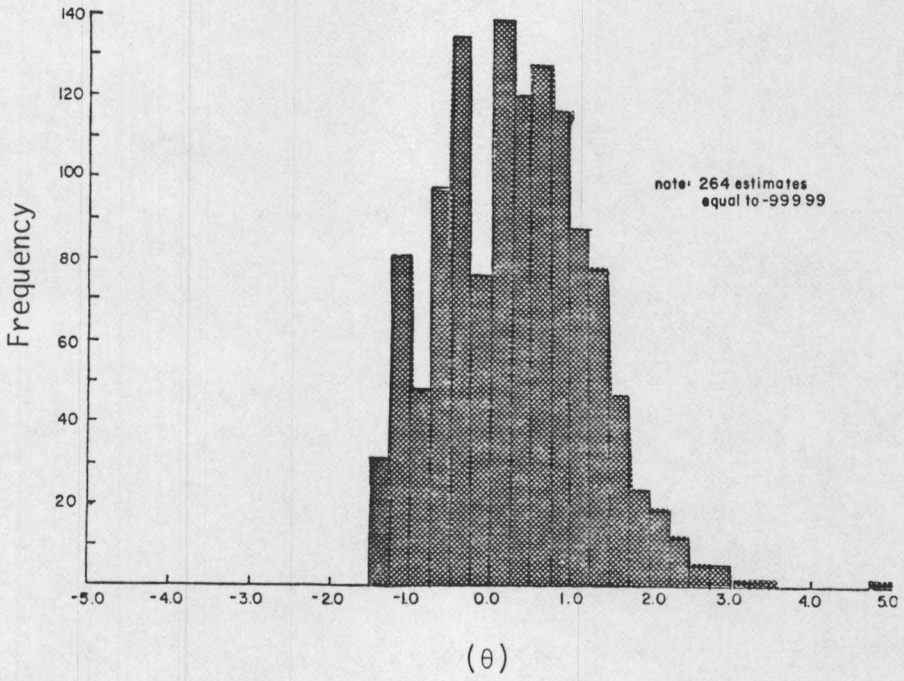


FIGURE 5.6

FREQUENCY DISTRIBUTION FOR 1980
MAXIMUM LIKELIHOOD ESTIMATES

Validation of Item Characteristic Functions

Once maximum likelihood estimates of the latent trait were made available this information was used to validate the item characteristic functions developed earlier. This involved grouping estimates of the latent trait into groups of approximately the same size along the latent trait continuum. This process resulted in a total of 20 groups in addition to the group comprised of subjects assigned high negative values. Next, the proportion of subjects in each group assigning an item score $x_g = 1$ and $x_g = 0$ were estimated. These frequency ratios were then plotted along with the item characteristic function for each item over the two survey years. This resulted in a set of 36 figures each containing three plots. The purpose of this effort was to examine differences between observed proportions and predicted probabilities for each action item. The results of this validation process are presented in Appendix B.

Generally speaking, the estimated item characteristic functions appear to capture the relationship between the latent trait and item responses. Recall that except for sampling fluctuation most items should exhibit the largest variation around the point where $P_g(\Theta) = 0.5$. However, some items do seem to have the highest variation at extreme values particularly at the lower values of the latent trait. In conclusion, the two parameter normal ogive model describes the relationship between item responses and the value of the latent trait in a reasonable fashion.

Test Information Function

As stated earlier, the test information function provides an indication of the accuracy of the maximum likelihood estimates of the latent trait for each subject. Using a computer program written by Samejima (70), test information functions for each survey year have been prepared. Recall that $[I(\Phi)]^{-\frac{1}{2}}$ is equal to the standard error of the estimate which can be used to construct confidence intervals for the individual parameter estimates.

Figures 5.7 and 5.8 contain the test information functions for each sample year. The important point to note is that the most accurate estimates are contained in the range $0<\theta<2$ for each survey. This suggests that at extreme values of the latent trait the estimates are subject to a higher degree of error. This has important implications in designing future questionnaires. More specifically, it is necessary to include action items which will contain higher levels of information at lower and higher positions along the latent continuum in order to develop more accurate estimates in these ranges.

Generally speaking, the validation procedure suggests that most calibrated item characteristic functions describe the relationship between the latent construct and the observed proportion of item endorsement, for each of the 20 groups fairly well. In several items, however, some discrepancies between the theoretical model and the observed results are noted.

Figure 5.9 presents the two items from the 1979 data set exhibiting a high degree of difference between the theoretical item characteristic function and observed proportions of subjects with item scores

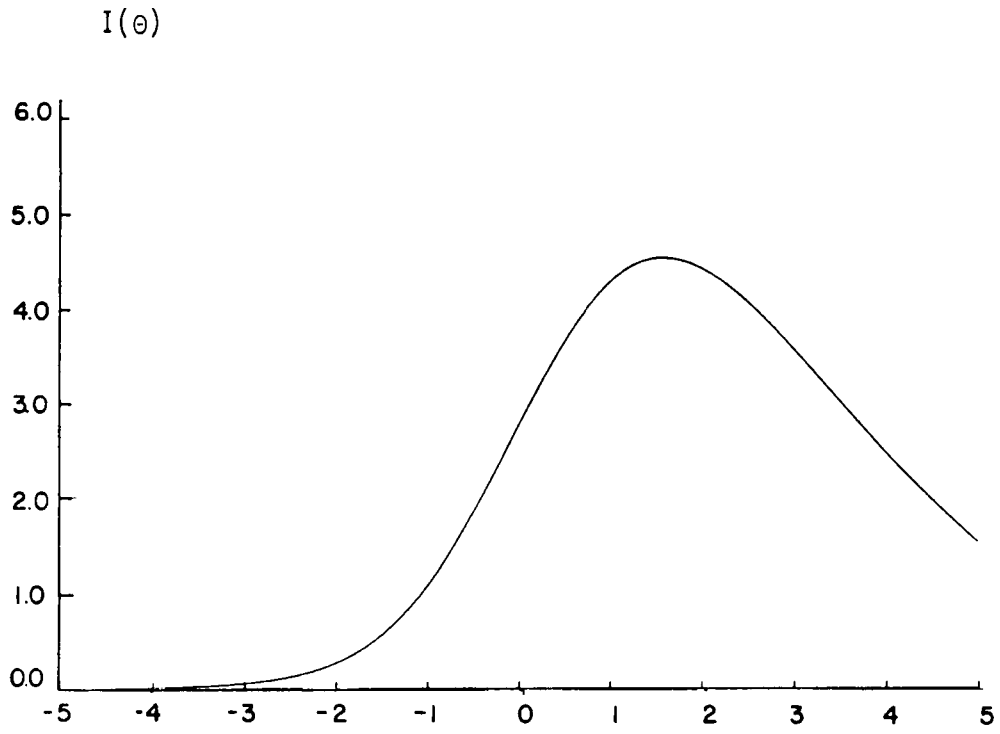


FIGURE 5.7

1979 TEST INFORMATION FUNCTION

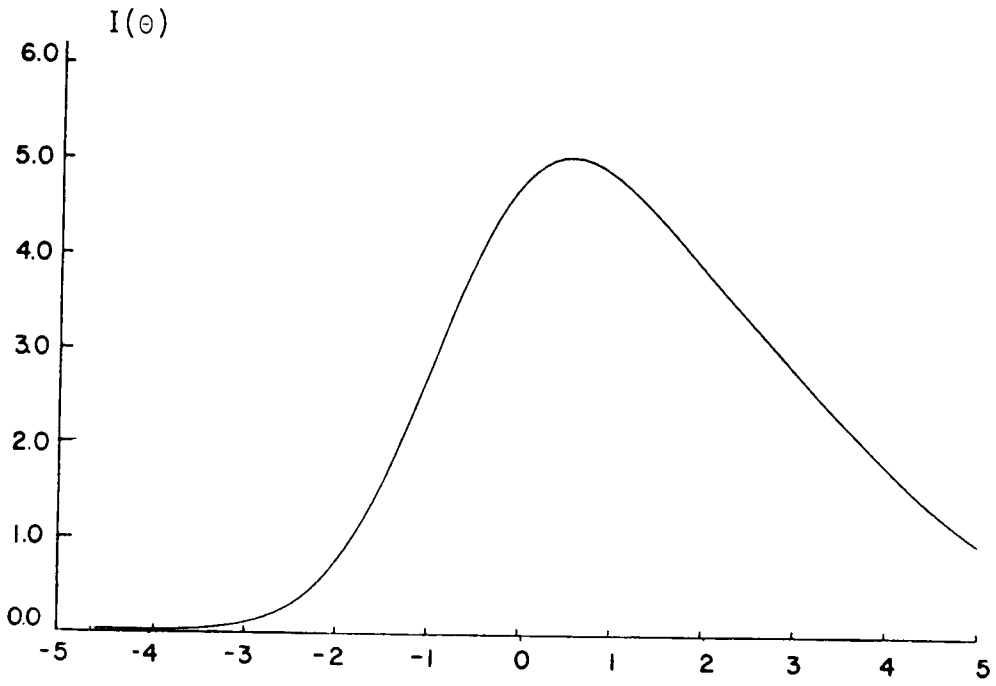
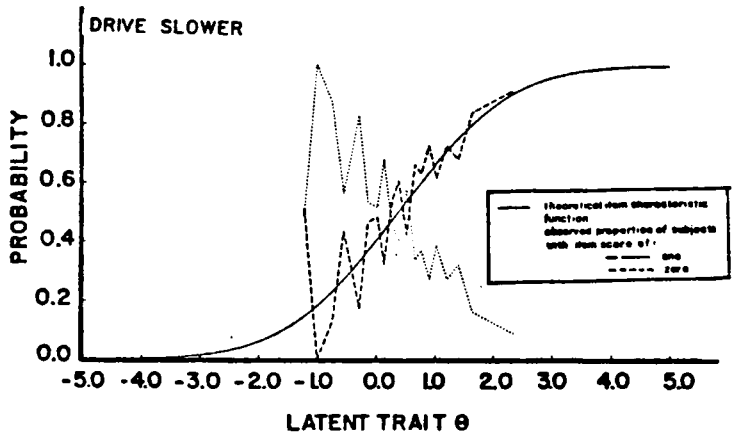


FIGURE 5.8

1980 TEST INFORMATION FUNCTION

(a)



(b)

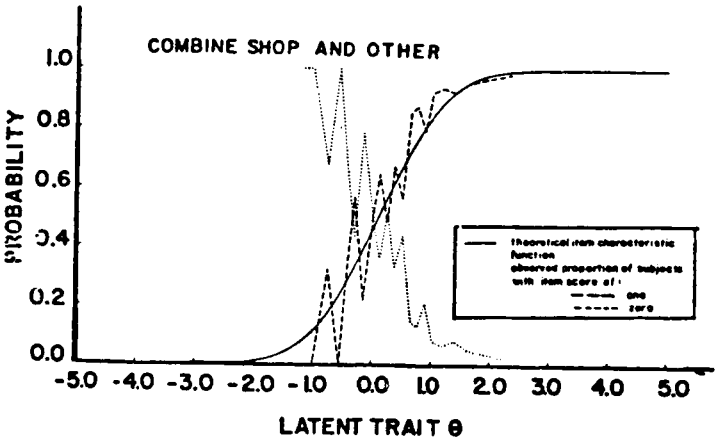


FIGURE 5.9

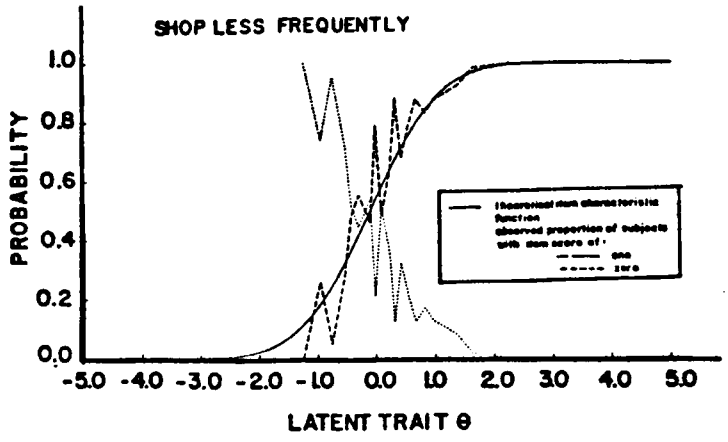
SELECTED ITEM CHARACTERISTIC FUNCTIONS USING THE NORMAL OGIVE MODEL, 1979

$x_g = 0$ or $x_g = 1$. The first item shown is the action "drive slower" while the second item shown is the action "combine shop and other non-work trips." In both cases, the difference between the theoretical probability and the observed proportion is highest at the lower values of the latent trait. This may be the result of the very low difficulty levels of these items. Recall that the proportion of individuals endorsing the action "combine shop and other" was 0.468 with 0.418 endorsing "drive slower." In other words, perhaps because these actions were endorsed at a high rate, relative to other actions, and because the accuracy of the latent trait estimates is lowest at the ranges where discrepancies are high and theoretical results is enlarged. If more accurate estimates of the latent trait were available for the lower end of the continuum then the differences between the observed and theoretical results may be reduced.

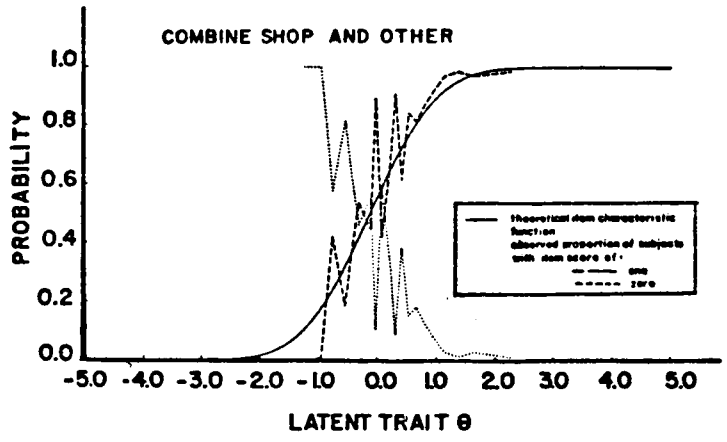
Figure 5.10 outlines three items from the 1980 results exhibiting a high variability between observed and theoretical results. Once again, the discrepancy is highest at the lower ranges of the latent trait. In this case the items include "shop less frequently," "combine shop and other non-work trips," and "shop closer to home."

Once again, the reason for these discrepancies may be explained by the low amount of information provided by the items for low values of the latent trait. This suggests a larger variability in the accuracy of the latent trait estimates for these ranges which contributes to the observed discrepancies for these low ranges.

(a)



(b)



(c)

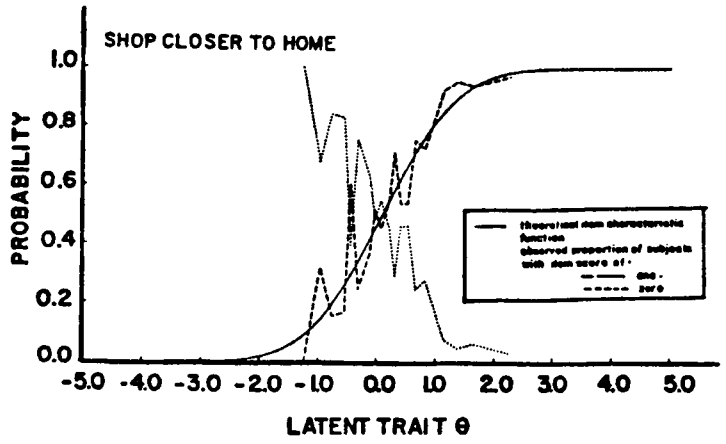


FIGURE 5.10

SELECTED ITEM CHARACTERISTIC FUNCTIONS USING THE TWO PARAMETER NORMAL OGIVE MODEL, 1980

Logist 5 Method Results

Table 5.16 presents the item parameter for the two parameter logist model estimated from the Logist 5 program for the 1979 data. Table 5.17 presents these results for the 1980 data. These parameters appear to be similar to the earlier parameters developed for the two parameter normal ogive model.

Individual estimates of the latent trait have also been developed for each subject in the sample survey data. In order to compare these estimates with those developed using the normal ogive model a linear regression of the normal ogive results on the logistic results has been developed. The parameters resulting from this analysis are presented in Table 5.18.

The results of this effort suggest that a strong positive relationship exists between the maximum likelihood estimates developed through these procedures. It is interesting to note that the maximum likelihood estimates developed from the Logist 5 approach are generally smaller in value than associated likelihood estimates developed through the tetrachoric method. In addition, the parameters for the regression models are almost identical over the survey year, suggesting a systematic difference between the estimation procedures.

It appears that the maximum likelihood estimates using the Logist 5 method yield results consistent with those developed from the tetrachoric method. It is unlikely that the estimates using the former approach will result in drastically different placement of a subject in latent space over the latter approach.

TABLE 5.16

ITEM PARAMETERS DEVELOPED
FROM LOGIST 5, 1979

Item g	Item Discrimination a_g	Item Difficulty b_g
1	0.5171	-0.0311
2	0.4235	2.4521
3	0.3542	3.1664
4	0.3847	3.6226
5	0.6853	0.2216
6	0.8115	3.1579
7	0.7075	2.7555
8	1.0194	0.7358
9	0.7945	0.3155
10	1.0877	-0.2305
11	0.8369	0.0182
12	0.6926	2.1744
13	0.7566	1.3820
14	0.5297	1.8876
15	0.5180	1.8865
16	0.7060	1.5705
17	0.6083	2.5115
18	0.3839	2.4906

TABLE 5.17

ITEM PARAMETERS DEVELOPED
FROM LOGIST 5, 1980

Item g	Item Discrimination a_g	Item Difficulty b_g
1	0.8806	-0.0559
2	0.6271	1.8840
3	0.2813	3.5941
4	0.3367	3.5437
5	0.7771	0.8237
6	0.8577	2.8507
7	0.8723	2.5716
8	0.8577	0.5489
9	0.9508	-0.5230
10	1.0632	-0.5187
11	0.8099	-0.2579
12	1.0336	1.5379
13	0.7134	1.3121
14	0.7779	1.7756
15	0.4719	1.5683
16	0.5937	1.5924
17	0.5762	3.2104
18	0.3979	2.4269

TABLE 5.18

RESULTS OF FITTING REGRESSION OF MAXIMUM LIKELIHOOD ESTIMATES USING THE NORMAL OGIVE MODEL ON MAXIMUM LIKELIHOOD ESTIMATES USING THE LOGISTIC MODEL

	1979 Survey	1980 Survey
Intercept	0.304	0.309
Standard Error	0.002	0.002
t	165.67	142.74
Slope	0.903	0.895
Standard Error	0.002	0.002
t	499.58	413.09
Sample Size[a]	1256	1304

[a]Cases with high negative values (i.e., -999.99) for latent trait estimates have been removed.

Appendix C presents the item response characteristic functions developed for the two parameter logistic model. Also contained in Appendix C are the frequency ratios developed for each item score category using 21 groups along the latent continuum. The validation process suggests a reasonable fit of the two parameter logist model for this data set.

Once again, the results of the item validation process suggest that most theoretical item characteristic functions compare favorably with observed proportions. However, the 1979 results for items "drive slower" and "combine shop and other non-work trips" exhibit a high degree of variability for low ranges of the latent trait. In 1980 similar results are observed for items "shop closer to home," "combine shop and other non-work trips" and "shop closer to home." The 1979 results are contained in Figure 5.11 and the 1980 results in Figure 5.12.

The use of the Logist 5 method does not include an evaluation of the dimensionality of the latent space. It is necessary to explore the data using principal factor analysis. While the implementation of the Logist 5 is less complex than the tetrachoric method, in the sense that the program is self contained, care must be exercised in evaluating the results. In particular, the dimension of the latent space must be confirmed through factor analysis in order to maximize the accuracy of results.

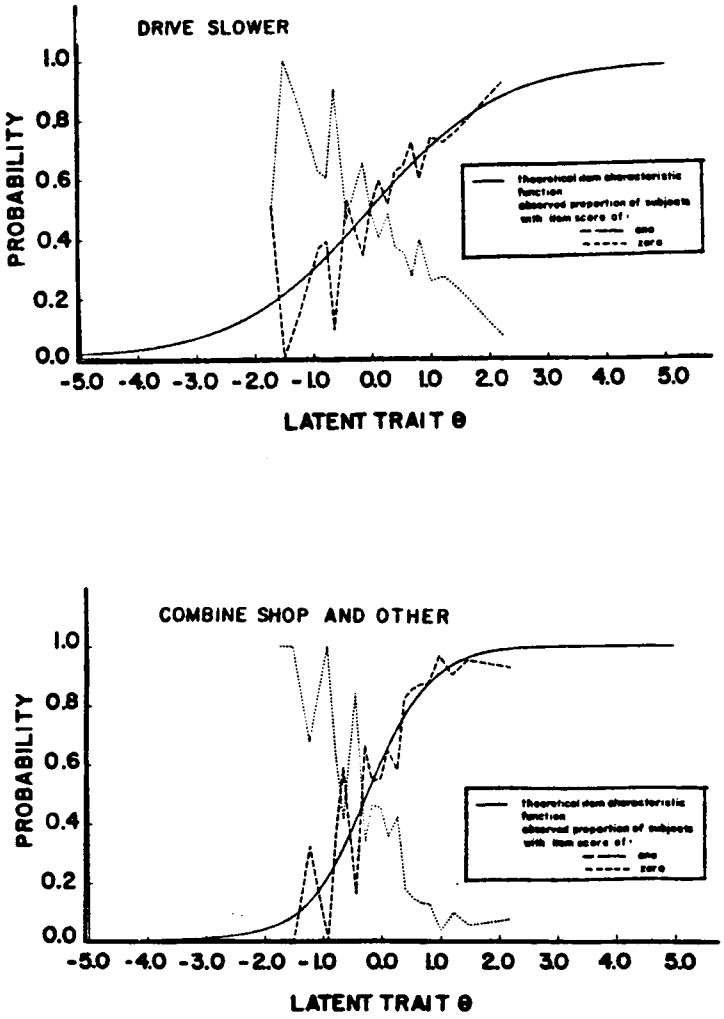
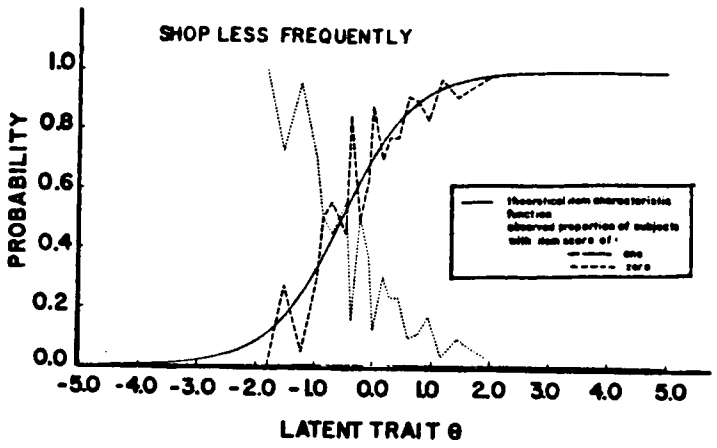


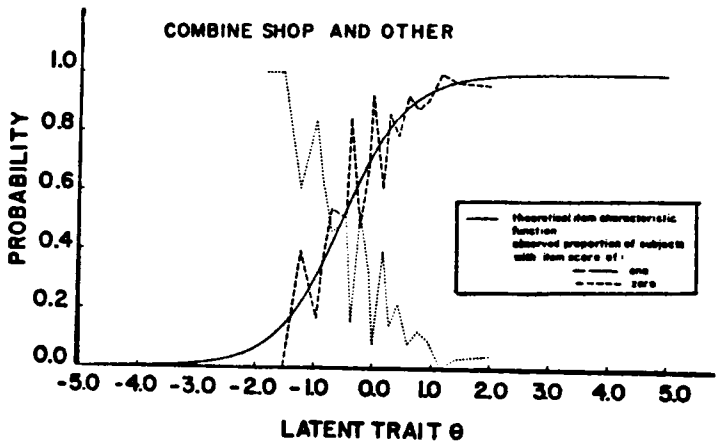
FIGURE 5.11

SELECTED ITEM CHARACTERISTIC FUNCTIONS USING THE TWO PARAMETER LOGISTIC MODEL, 1979

(a)



(b)



(c)

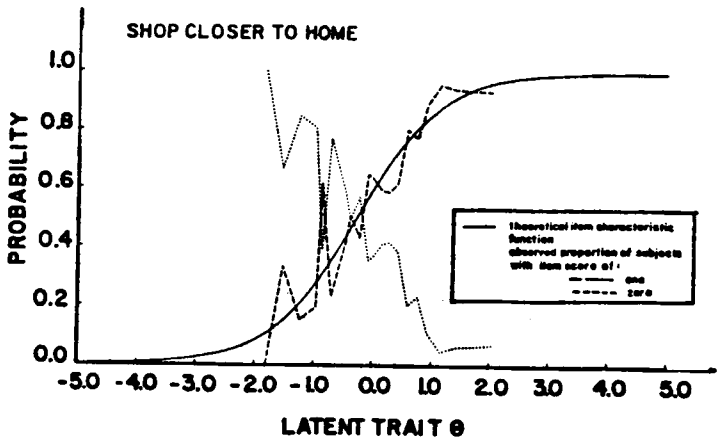


FIGURE 5.12

SELECTED ITEM CHARACTERISTIC FUNCTIONS USING THE TWO PARAMETER LOGISTIC MODEL, 1980

Market Segmentation Analysis

The maximum likelihood estimates of the hypothesized latent trait serve as a basis for segmenting sample survey data. In simple terms, these estimates present a measure of the willingness and ability of each subject to implement various transportation energy conservation actions. It has been shown that these estimates can be used to understand in a probabilistic sense, the relationship between a subject and an energy conservation action.

In order to apply these estimates in market segmentation analysis, the entire sample for each year has been divided into six groups. This arbitrary classification is done in part, to simplify further analysis and is also based on a desire to have fairly similar group sizes. Recall that over 250 subjects in each sample have been assigned a very low negative value for the latent trait since these subjects did not endorse a single action. Table 5.19 outlines group sizes and the range of the value of the latent trait included for each group for the 1979 survey data. Also included in this table is data concerning the mean and standard deviation for each group's estimated latent trait position along the unidimensional latent continuum. Also shown in this table is the estimated mean total summarized score with associated standard deviation. This is formed from the sum of all actions endorsed divided by the group size. Table 5.20 presents similar information for the 1980 survey sample.

It is interesting to observe that, as expected, the mean of the total summated group score increased along with the increase in the

TABLE 5.19

MEAN AND STANDARD DEVIATION OF LATENT TRAIT ESTIMATES AND TOTAL SCORES FOR SELECTED GROUPS, 1979

Group	Group Size	Range of Latent Trait	Latent Trait		Total Score	
			Mean	Std	Mean	Std
1	264	<-1.40	-999.999	0	0	0
2	270	-1.40<θ<-0.47	-0.907	0.269	1.30	0.459
3	244	-0.47<θ<0.08	-0.204	0.167	2.56	0.567
4	249	0.08<θ<0.059	0.321	0.146	4.06	0.663
5	248	0.59<θ<1.111	0.829	0.142	5.75	0.681
6	245	>1.111	1.637	0.507	8.75	1.850
	1520					

TABLE 5.20

MEAN AND STANDARD DEVIATION OF LATENT TRAIT ESTIMATES AND TOTAL SCORES FOR SELECTED GROUPS, 1980

Group	Group Size	Range of Latent Trait	Latent Trait		Total Score	
			Mean	Std	Mean	Std
1	256	<-1.54	-999.999	0	0	0
2	267	-1.54<θ<-0.46	-0.882	0.282	1.446	0.541
3	254	-0.46<θ<0.05	-0.209	0.168	2.787	0.598
4	257	-0.05<θ<0.48	0.261	0.128	4.361	0.584
5	260	0.48<θ<0.88	0.746	0.169	6.004	0.854
6	266	>0.88	1.620	0.471	9.132	1.800
	1560					

value of the latent trait. The group with the highest mean value of the latent trait also exhibits the highest mean summated score. This relationship of total score on the ability, or test characteristic function, follows a similar relationship to the item characteristic function on regression of item score on ability (62, p. 386)

A problem with using the total summated score as a measure is reflected in the observation that the total test score provides no information about the items comprising this score. In other words many combinations yield an identical total score yet one may contain actions easily implemented while another does not. The latent trait estimate, however can be related to the probability of an action endorsement function.

The group value for the latent trait estimate has been used in conjunction with calibrated item characteristic functions to determine theoretical probabilities for action item endorsement for each action. In addition, the observed proportion of action item endorsement has been calculated for each group. The results of this effort are presented in Table 5.21 for the 1979 data and Table 5.22 for the 1980 data. Each set of groups exhibit the same general pattern of action endorsement. As the mean value of the latent trait increases, so do action endorsement probabilities. As stated earlier, Group 1 subjects did not endorse a single action. Group 2 members appear to demonstrate only a low proportion of action endorsement. Group 3 subjects start to exhibit larger probabilities of action endorsement especially for action items with lowest difficulty parameters. This includes shopping related actions and measures relating to vehicle use and maintenance.

TABLE 5.21

OBSERVED PROPORTIONS AND ESTIMATED PROABILITIES USING THE NORMAL OGIVE MODEL FOR ACTION ITEM ENDORSEMENT, 1979

	Observed 1	Estimated 1	Observed 2	Estimated 2	Observed 3	Estimated 3	Observed 4	Estimated 4	Observed 5	Estimated 5	Observed 6	Estimated 6
Drive Slower	0.0000	0.0000	0.2815	0.1999	0.3566	0.3519	0.4699	0.4859	0.6573	0.6173	0.7878	0.7964
Carpool to Work	0.0000	0.0000	0.0704	0.0566	0.0861	0.1014	0.1486	0.1488	0.1935	0.2069	0.3551	0.3226
Bus Subway to Work	0.0000	0.0000	0.10000	0.0490	0.0984	0.0869	0.1044	0.1270	0.1532	0.1768	0.2735	0.2776
Walk or Bike to Work	0.0000	0.0000	0.0667	0.0265	0.0369	0.0531	0.0884	0.0843	0.0968	0.1260	0.2163	0.2183
Tune Car	0.0000	0.0000	0.1259	0.1352	0.3115	0.2859	0.4378	0.4347	0.6331	0.5885	0.7755	0.7998
Move Closer to Work	0.0000	0.0000	0.0037	0.0011	0.0082	0.0052	0.012	0.0144	0.0081	0.0345	0.1224	0.1083
Job Closer to Home	0.0000	0.0000	0.011	0.0048	0.0164	0.0174	0.0321	0.0400	0.0561	0.0803	0.2041	0.1978
Shop on Way Home from Work	0.0000	0.0000	0.0185	0.0395	0.1148	0.1352	0.2651	0.2696	0.4194	0.4438	0.7347	0.7290

TABLE 5.21 (continued)

	Observed 1	Estimated 1	Observed 2	Estimated 2	Observed 3	Estimated 3	Observed 4	Estimated 4	Observed 5	Estimated 5	Observed 6	Estimated 6
Shop Less Frequently	0.0000	0.0000	0.0889	0.1181	0.2172	0.2601	0.4297	0.4055	0.6411	0.5605	0.7741	0.7806
Combine Shop and Other	0.0000	0.0000	0.0741	0.1345	0.4057	0.3665	0.5904	0.5908	0.8589	0.7829	0.9469	0.9516
Shop Closer to Home	0.0000	0.0000	0.1185	0.1405	0.3279	0.3174	0.5301	0.4901	0.6895	0.6595	0.8490	0.8652
Eliminate Use of Rec. Vehicle.	0.0000	0.0000	0.0185	0.0158	0.0287	0.0465	0.0924	0.0866	0.1492	0.1500	0.2857	0.3023
Vacation Closer to Home	0.0000	0.0000	0.0222	0.0277	0.0861	0.0855	0.1767	0.1684	0.2419	0.2860	0.5184	0.5253
Cancel Vacation Trip	0.0000	0.0000	0.0593	0.0502	0.1184	0.1036	0.1647	0.1644	0.2056	0.2418	0.4122	0.3966
Use Train Bus, Plane for Va.	0.0000	0.0000	0.1037	0.0488	0.0984	0.1038	0.1245	0.1675	0.2258	0.2489	0.4204	0.4120
Buy Fuel Efficient Car	0.0000	0.0000	0.0222	0.0191	0.0697	0.0677	0.1486	0.1445	0.2500	0.2606	0.4531	0.5097
Sell Car	0.0000	0.0000	0.0259	0.0148	0.0287	0.0454	0.0643	0.0773	0.1089	0.1330	0.2857	0.2683
Bus Subway for Non-Work	0.0000	0.0000	0.0889	0.0590	0.1557	0.1089	0.1847	0.1623	0.1613	0.2279	0.3388	0.3574

TABLE 5.22

OBSERVED PROPORTIONS AND ESTIMATED PROABILITIES USING THE NORMAL OGIVE MODEL FOR ACTION ITEM ENDORSEMENT, 1980

	Observed 1	Estimated 1	Observed 2	Estimated 2	Observed 3	Estimated 3	Observed 4	Estimated 4	Observed 5	Estimated 5	Observed 6	Estimated 6
Drive Slower	0.0000	0.0000	0.0000	0.1360	0.1461	0.3256	0.2741	0.5001	0.5486	0.6794	0.7154	0.9042
Carpool to Work	0.0000	0.0000	0.0637	0.0338	0.0315	0.0774	0.0973	0.1273	0.2115	0.1984	0.3872	0.3740
Bus Subway to Work	0.0000	0.0000	0.1161	0.0669	0.1220	0.1054	0.1401	0.1405	0.1654	0.1842	0.2632	0.2819
Walk or Bike to Work	0.0000	0.0000	0.0787	0.0452	0.0669	0.0776	0.1440	0.1092	0.0923	0.1504	0.2594	0.2478
Tune Car	0.0000	0.0000	0.0674	0.0629	0.1102	0.1585	0.2685	0.2644	0.4269	0.4021	0.6504	0.6702
Move Closer to Work	0.0000	0.0000	0.0000	0.0021	0.0000	0.0083	0.0117	0.0190	0.0462	0.0406	0.1278	0.1257
Job Closer to Home	0.0000	0.0000	0.0000	0.0029	0.0118	0.0115	0.0311	0.0267	0.0308	0.0570	0.1805	0.1718
Shop on Way Home from Work	0.0000	0.0000	0.0524	0.0692	0.2126	0.1881	0.2957	0.3203	0.4769	0.4855	0.7556	0.7701

TABLE 5.22 (continued)

Group	Observed 1	Estimated 1	Observed 2	Estimated 2	Observed 3	Estimated 3	Observed 4	Estimated 4	Observed 5	Estimated 5	Observed 6	Estimated 6
Shop Less Frequently	0.0000	0.0000	0.1461	0.2042	0.5798	0.4576	0.6732	0.6546	0.8423	0.8204	0.9511	0.9680
Combine Shop and Other	0.0000	0.0000	0.1461	0.1999	0.5669	0.4620	0.6576	0.6653	0.8654	0.8328	0.9774	0.9735
Shop Closer to Home	0.0000	0.0000	0.1610	0.1627	0.4291	0.3765	0.5642	0.5608	0.7077	0.7373	0.9474	0.9336
Eliminate Use of Rec. Vehicle.	0.0000	0.0000	0.037	0.0089	0.0118	0.0379	0.0584	0.0870	0.1923	0.1759	0.4248	0.4366
Vacation Closer to Home	0.0000	0.0000	0.0712	0.0431	0.0591	0.1075	0.1790	0.1823	0.2769	0.2867	0.5226	0.5522
Cancel Vacation Trip	0.0000	0.0000	0.0300	0.0177	0.0315	0.0531	0.0817	0.1012	0.1615	0.1780	0.3947	0.3860
Use Train Bus Plain for Vac.	0.0000	0.0000	0.1423	0.0779	0.0945	0.1473	0.2451	0.2152	0.2885	0.3011	0.4624	0.4843
Buy Fuel Efficient Car	0.0000	0.0000	0.0749	0.0489	0.0748	0.1078	0.1868	0.1719	0.2115	0.2592	0.4812	0.4585
Sell Car	0.0000	0.0000	0.0262	0.0107	0.0118	0.0253	0.0428	0.0436	0.0808	0.0723	0.1391	0.1571
Bus Subway for Non-Work	0.0000	0.0000	0.1199	0.0576	0.1063	0.1063	0.1362	0.1546	0.2115	0.2177	0.3308	0.3120

Group 4 continues this trend with the addition of vacation travel related action items. As expected, Group 5 members demonstrate an increased probability to endorse additional action items at a higher rate than previous groups. Finally, Group 6 is the most positive group with very high probabilities of endorsement for many actions.

Obviously, this grouping strategy is arbitrary. It does, however, represent an attempt to summarize the relationship between the hypothesized latent trait and various energy conservation actions in a useful manner. Grouping subjects in this manner helps to focus the analysis in a simple manner so that results can be more meaningfully interpreted. More detailed review of the socioeconomic characteristics of these groups is contained below. A hint of the results of this effort is presented in Table 5.23 where these variables where correlated with the estimates of the latent trait for each subject.

The values of these correlation coefficients are not large in absolute value, but the magnitude and sign of these values do present some interesting insights. First of all it appears as auto ownership levels increase so does the estimated latent trait value. On the other hand a negative relationship appears to exist between age and the latent trait, marital status and the latent trait, employment status and the latent trait, and housing unit type and the latent trait. In other words, older subjects generally have lower values on the latent trait than younger subjects; married subjects have a higher value than widowed, divorced, or separated subjects for the latent trait estimate; subjects with jobs appear to have a generally higher latent trait

TABLE 5.23

CORRELATION OF MAXIMUM LIKELIHOOD ESTIMATES
OF LATENT TRAIT WITH SELECTED
SOCIOECONOMIC VARIABLES

	1979 Survey	1980 Survey
Area in State	0.1763	0.2076
Age of Subject	-0.1754	-0.1114
Sex of Subject	-0.0957	-0.0622
Education Level of Subject	0.0900	0.1759
Marital Status of Subject	-0.1781	-0.1946
Employment Status of Subject	-0.1171	-0.1614
Number of Vehicles in Household	0.4144	0.4187
Size of Household	0.1679	0.0999
Annual Income of Household	0.1592	0.1870
Type of Dwelling Unit	-0.2674	-0.2587

estimate than unemployed or retired subjects; and subjects living in single family housing seem to have higher latent trait estimates.

It also appears that subjects in upstate non-urban areas have a higher positive relationship with the latent trait estimates than downstate urban subjects. Household income appears to be generally positively related to the latent trait estimate.

An unusual pattern seems to emerge from the coefficients calculated for household size and educational level of subject. More specifically, the 1979 results suggest a positive relationship between household size and latent trait estimates. However, in 1980 this relationship is unclear. A similar occurrence appears for the educational attainment relationship with the latent trait except the 1979 data shows an unclear relationship and the 1980 data suggest a positive relationship.

The relationship between these coefficients seemed to suggest a coding error. However, a review of the data indicated that the information had been properly recorded. Therefore, it is difficult to determine the reason for this reversal. Of course, it may have to do with the sample characteristics but this pattern is not manifested in any other variable.

In order to examine in more detail the characteristics of the six groups described earlier. Tables 5.24 and 5.25 have been prepared. These tables contain the proportion of each group classified by various levels of selected independent variables. In addition, Tables 5.26 and 5.27 present the standardized residuals calculated for the chi-squared test of independence for these tables. These tables, in association

TABLE 5.24

MARKET SEGMENT PROFILES USING LATENT TRAIT THEORY, 1979

	Segment						
	1	2	3	4	5	6	Total Sample
Income							
Under $10,000	0.35	0.20	0.17	0.16	0.14	0.11	0.18
$10,000 - $25,000	0.45	0.51	0.54	0.56	0.51	0.52	0.52
Over $25,000	0.24	0.29	0.29	0.28	0.35	0.37	0.30
Household Size							
1-2	0.61	0.45	0.48	0.39	0.34	0.37	0.44
3-4	0.32	0.39	0.34	0.37	0.44	0.44	0.38
5+	0.07	0.16	0.18	0.24	0.32	0.23	0.18
Employment Status							
Employed	0.52	0.68	0.64	0.64	0.63	0.74	0.64
Unemployed	0.48	0.32	0.36	0.36	0.37	0.24	0.36
Marital Status							
Married	0.46	0.63	0.67	0.61	0.71	0.65	0.62
Single	0.27	0.24	0.19	0.26	0.20	0.25	0.24
Widow, Dv, Sep	0.27	0.13	0.14	0.12	0.09	0.09	0.14

TABLE 5.24 (continued)

	Segment						
	1	2	3	4	5	6	Total Sample
Area of State							
Downstate Urban	0.80	0.71	0.62	0.53	0.44	0.48	0.60
Upstate Urban	0.11	0.19	0.24	0.25	0.28	0.25	0.22
Other	0.09	0.10	0.13	0.22	0.28	0.27	0.18
Number of Vehicles							
0	0.59	0.16	0.09	0.07	0.05	0.07	0.17
1	0.24	0.42	0.47	0.47	0.39	0.43	0.40
2+	0.18	0.43	0.44	0.46	0.57	0.51	0.43
Age							
18-34	0.27	0.39	0.37	0.44	0.39	0.55	0.40
35-64	0.48	0.49	0.50	0.48	0.53	0.43	0.49
65+	0.24	0.12	0.13	0.09	0.09	0.02	0.12
Sex							
Male	0.39	0.61	0.57	0.49	0.47	0.43	0.49
Female	0.61	0.39	0.43	0.51	0.53	0.57	0.51

TABLE 5.24 (continued)

	Segment						
	1	2	3	4	5	6	Total Sample
Educational Status							
Grammar or Some High School	0.24	0.13	0.17	0.12	0.14	0.11	0.15
High School Graduate	0.34	0.43	0.34	0.36	0.39	0.40	0.38
Some College or College Graduate	0.42	0.44	0.49	0.52	0.47	0.48	0.47
Housing Type							
Single Family	0.26	0.49	0.50	0.64	0.67	0.56	0.52
Two Family	0.14	0.14	0.10	0.17	0.14	0.14	0.14
Apt. Condo, Co-Op	0.62	0.37	0.35	0.26	0. 16	0.30	0.34

TABLE 5.25

MARKET SEGMENT PROFILES USING LATENT TRAIT THEORY, 1980

	Segment						
	1	2	3	4	5	6	Total Sample
Income							
Under $10,000	0.42	0.23	0.20	0.18	0.14	0.11	0.21
$10,000 - $25,000	0.39	0.45	0.49	0.47	0.52	0.59	0.49
Over $25,000	0.19	0.32	0.31	0.35	0.34	0.30	0.30
Household Size							
1-2	0.65	0.51	0.45	0.52	0.51	0.45	0.51
3-4	0.24	0.39	0.40	0.34	0.36	0.36	0.35
5+	0.11	0.11	0.16	0.14	0.13	0.19	0.14
Employment Status							
Employed	0.39	0.85	0.54	0.56	0.60	0.76	0.57
Unemployed	0.61	0.42	0.46	0.44	0.40	0.24	0.43
Marital Status							
Married	0.40	0.59	0.60	0.59	0.64	0.64	0.58
Single	0.29	0.25	0.22	0.28	0.23	0.23	0.25
Widow, Dv, Sep	0.31	0.16	0.15	0.13	0.43	0.13	0.17

TABLE 5.25 (continued)

	Segment						
	1	2	3	4	5	6	Total Sample
Area of State							
Downstate Urban	0.83	0.73	0.62	0.53	0.41	0.41	0.60
Upstate Urban	0.11	0.13	0.19	0.23	0.35	0.35	0.21
Other	0.06	0.15	0.19	0.24	0.24	0.24	0.19
Number of Vehicles							
0	0.64	0.27	0.11	0.06	0.06	0.06	0.20
1	0.22	0.36	0.54	0.50	0.40	0.40	0.42
2+	0.14	0.37	0.36	0.44	0.54	0.54	0.39
Age							
18-34	0.31	0.34	0.36	0.37	0.37	0.38	0.40
35-64	0.37	0.47	0.46	0.47	0.51	0.54	0.49
65+	0.33	0.19	0.18	0.16	0.12	0.08	0.12
Sex							
Male	0.40	0.52	0.43	0.49	0.50	0.47	0.47
Female	0.60	0.48	0.57	0.51	0.50	0.53	0.53

TABLE 5.25 (continued)

	Segment						
	1	2	3	4	5	6	
Educational Status							
Grammar or Some High School	0.36	0.19	0.18	0.12	0.13	0.11	0.19
High School Graduate	0.33	0.37	0.41	0.39	0.38	0.38	0.38
Some college or College Graduate	0. 31	0.45	0.40	0.46	0.49	0.50	0.44
Housing Type							
Single Family	0.28	0.51	0.57	0.59	0.64	0.64	0.54
Two Family	0.11	0.11	0.14	0.13	0.12	0.18	0.12
Apt. Condo, Co-Op	0.61	0.38	0.30	0.27	0.24	0.24	0.34

TABLE 5.26

STANDARDIZED RESIDUALS DEVELOPED FROM SEGMENT PROFILES USING LATENT TRAIT THEORY, 1979

	Segment					
	1	2	3	4	5	6
Income						
Under $10,000	6.3	0.6	-0.3	-1.0	-2.1	-3.2
$10,000 - $25,000	-1.5	-0.1	0.6	1.0	0.1	-0.1
Over $25,000	-2.9	0.4	0.1	1.2	1.0	0.0
Household Size						
1-2	4.1	0.2	0.9	-1.1	-2.4	-1.7
3-4	-1.6	-0.6	-1.0	0.4	1.4	1.5
5+	-4.0	-0.1	0.0	2.4	1.8	0.6
Employment Status						
Employed	-2.5	0.7	-0.1	0.0	0.2	-2.7
Unemployed	-3.5	0.1	-0.1	0.0	-0.1	-2.0
Marital Status						
Married	-3.3	0.2	1.0	-0.1	1.8	0.6
Single	1.1	-0.5	-1.4	0.8	-1.3	0.5
Widow, Dv, Sep	5.4	2.3	-0.3	-0.7	-2.09	-2.0

TABLE 5.26 (continued)

	Segment					
	1	2	3	4	5	6
Area of State						
Downstate Urban	4.2	-1.0	0.5	-1.5	-3.3	-2.5
Upstate Urban	-3.7	-3.0	0.74	1.1	2.1	1.1
Other	-3.7	-0.5	-1.7	1.6	3.7	3.4
Number of Vehicles						
0	16.0	0.4	-3.1	-3.8	-4.9	-4.2
1	-4.1	0.4	1.8	1.7	-0.4	0.7
2+	-6.2	0.0	0.2	0.8	3.5	0.2
Age						
18-34	-3.2	-0.2	-0.7	0.9	-0.4	3.7
35-64	0.0	0.2	0.2	-0.2	1.0	-1.3
65+	6.0	0.0	0.8	-1.3	-1.4	-4.2
Sex						
Male	-2.4	2.7	1.8	-0.1	-0.5	-1.5
Female	2.4	-2.7	-1.7	0.1	0.5	1.5

TABLE 5.26 (continued)

	Segment					
	1	2	3	4	5	6
Educational Status						
Grammar or Some High School	3.7	-1.0	0.8	-1.4	-0.5	-1.7
High School	-1.0	1.5	-1.1	0.4	0.4	0.6
Some College or College Graduate	-1.2	-0.8	0.5	1.2	-0.1	0.4
Housing Type						
Single Family	-6.0	-0.6	-0.3	2.8	3.3	1.0
Two Family	0.0	0.0	0.4	-1.7	1.3	0.0
Apt., Condo, Co-Op	7.3	0.7	0.2	-2.3	-4.9	-1.2

TABLE 5.27

STANDARDIZED RESIDUALS DEVELOPED FROM SEGMENT PROFILES USING LATENT TRAIT THEORY, 1980

	Segment					
	1	2	3	4	5	6
Income						
Under $10,000	4.5	0.7	-0.5	-0.9	-1.6	-3.2
$10,000 - $25,000	-1.9	-0.7	0.1	-0.3	0.6	-0.1
Over $25,000	-1.6	-0.4	-0.7	1.2	0.0	0.0
Household Size						
1-2	3.0	-0.1	-1.4	0.1	-0.1	-1.7
3-4	-2.8	1.0	1.3	-0.2	0.4	1.5
5+	-1.2	-1.4	0.7	0.1	-0.5	0.6
Employment Status						
Employed	-3.7	0.1	-0.6	-0.2	0.5	-2.7
Unemployed	-4.2	-0.1	0. 7	0.3	-0.6	-2.0
Marital Status						
Married	-3.9	0.2	0.9	0.3	1.2	0.6
Single	1.2	0.0	-1.0	0.9	-0.5	0.5
Widow, Dv, Sep	5.7	-0.3	-0.5	-1.6	-1.6	-2.0

TABLE 5.27 (continued)

	Segment					
	1	2	3	4	5	6
Area of State						
Downstate Urban	4.9	2.8	0.5	-1.5	-2.7	-2.5
Upstate Urban	-3.5	-3.0	-0.8	0.7	1.8	1.1
Other	-4.9	-1.7	-0.1	1.9	2.9	3.4
Number of Vehicles						
0	16.2	2.8	-3.2	-4.9	-5.6	-4.2
1	-4.9	-1.4	2.9	2.1	1.7	0.7
2+	-6.4	-0.5	-0.8	0.4	2.2	0.2
Age						
18-34	-1.3	-0.4	0.0	0.5	0.6	0.7
35-64	-2.3	0.0	-0.2	-0.1	0.9	1.6
65+	5.7	0.5	0.3	-0.5	-2.2	-3.6
Sex						
Male	-1.7	1.3	-0.9	0.5	0.7	0.1
Female	1.5	-1.3	0.9	-0.5	-0.6	0.0

TABLE 27 (continued)

	Segment					
	1	2	3	4	5	6
Educational Status						
Grammar or Some High School	5.3	0.0	0.0	-1.6	-2.1	-3.0
High School	-1.2	-0.2	0.9	0.4	0.1	1.2
Some College or College Graduate	-3.2	0.2	-0.8	0.7	1.2	1.6
Housing Type						
Single Family	-5.6	-0.7	0.6	1.2	2.1	2.3
Two Family	-0.6	-0.5	0.6	0.5	0.0	0.0
Apt., Condo, Co-Op	7.4	1.1	-1.1	-1.8	-2.6	-2.9

with earlier data describing the item characteristic function, provide a basis for describing the results of this market segmentation effort.

Group Profiles

1979 Segments

Group 1. As stated earlier, this group has the lowest latent trait value, each member having been assigned a low negative value for the estimate of the latent trait.

It appears that these subjects tend to live in the downstate urban area; are generally old; tend to be female; own few automobiles; tend to be either single, widowed, divorced or separated; have small household sizes; have low incomes; and live in apartments, condominiums or co-ops.

The characteristics of these subjects suggest limited travel demands and opportunities. This suggests that only small transportation energy needs are required in an already limited set of travel behaviors. Therefore, formulation of transportation energy planning policies and programs for this group is very difficult. It appears that any efforts will be met with limited response. However, so little savings may result because of the limited travel demands of this group that efforts may be better placed in other areas.

Group 2. The second group consists of subjects with low estimated latent trait values. On average, this group only endorsed actions. Not surprisingly, these actions included items with lowest difficulty parameters.

The members of this group tend to display characteristics similar to those of the previous group. These include downstate urban location, low educational attainment, low income, small household size. However, these group members tend to have a younger age distribution, higher auto ownership rates, and are more likely to be employed than Group 1 subjects.

Once again, it is difficult to suggest specific conservation actions. Certainly actions related to auto use and maintenance and shopping travel will appeal to this group to some extent. However, it appears that so far over one third of the entire sample is either unwilling or unable to endorse a large proportion of conservation actions.

Group 3. This group and Group 4 appears to represent transition groups between socioeconomic variables and the latent trait summarized by the correlation coefficients described earlier.

Group 3 members tend to reside in downstate urban or upstate urban areas, are middle aged, have high auto ownership levels, high educational attainment, are married, have low to moderate incomes, and tend to live in multiple family dwelling units.

These subjects will best respond to actions requiring only limited changes in behavior. Therefore, it is suggested that shopping related travel policies and vehicle use and maintenance policies be pursued to encourage conservation among this group.

Group 4. This group is very similar to the previous group both in terms of actions taken and socioeconomic characteristics. Besides

shopping related actions and vehicle use and maintenance actions, however, this group appears to include a willingness to pursue acations related to vacation travel.

Group 4 members tend to be young, reside in the upstate urban or other areas, own several autos, are highly educated, tend to be married, have moderate to large household sizes, be middle income, and live in single family dwelling units.

This group, as a transition group, appears to best respond to low difficulty items and some moderate difficulty items such as vacation or recreation actions.

Group 5. Group 5 members have among the highest estimates for the hypothesized latent trait. This group includes subjects with an increasing willingness and ability to conserve transportation related energy. As expected, these subjects are characterized by an upstate urban or other place of residence, mid to young age, high auto ownership, high education, mid to large household size, and high income. In addition, these subjects tend to be married and live in single family homes.

Besides all previously mentioned actions suggested for earlier groups, this segment appears to express a willingness to pursue actions related to use of non-auto modes and replacement of automobiles. This group appears to have high travel needs, based on socioeconomic profile, but also has flexibility in meeting these needs in the face of contained energy supplies.

Group 6. This group clearly represents a collection of subjects with the highest willingess and ability to conserve transportation energy. On the average, each member has endorsed over nine actions.

Continuing with the pattern described thus far, the members of this group tend to live outside the downstate urban area, are young, highly educated, own several vehicles, are married, live in large households, have high incomes, and live in single family homes.

It appears that this group tends to have an interest in pursuing a full range of conservation actions. In addition, it appears that concentrating efforts in this group may have a large impact on fuel conservation since the demographic characteristics suggest a large travel requirement for these households.

1980 Segments

Group 1. This group, once again, includes subjects not endorsing a single conservation action. These subjects tend to be middle aged or older, reside in the downstate urban area, tend not to own an automobile, have small household sizes, low incomes, are not umployed, live in multifamily dwelling units, and have low educational attainment levels.

Like the similar group in the 1979 data, this group does not appear to represent a collection of subjects in a position to take advantage of any set of conservation actions. Efforts are better spent concentrating on other segments.

Group 2. The second group consists of individuals which demonstrate a low estimated value for the hypothesized latent trait. As expected, only an average number of action items were endorsed by this group. As a group, subjects tend to live in the downstate urban area, are middle aged to elderly, own no automobiles, are married, live in mid-sized households, and tend to live in multifamily housing.

The lower difficulty items appeal to this group to some extent. For example car use and maintenance, and shopping related actions appear to have been important actions to this group. In summary, however, this group will not respond in a positive manner to many policies and programs.

Group 3. Once again, this group appears to be a transition group which starts to represent subjects with an increasing willingness and ability to conserve transportation energy. Besides all the previous action items suggested in Group 2, this group also starts to endorse items related to the use of non-auto modes.

The socioeconomic characteristics also represent the trends urban area, are young to middle aged, own one auto, are highly educated, are married, live in small to medium sized households, in single family homes.

This group responds, in a limited fashion to actions involving only low difficulty levels. Therefore, actions related to auto use and maintenance, shopping travel, and the use of non-auto modes may appeal to this group in an effective manner.

Group 4. This group also represents a transition segment. The results suggest that this group appears to endorse several actions including those related to vacation travel.

The socioeconomic characteristics of this group are consistent with earlier analysis trends. Finally, it appears that these subjects represent a segment that may respond to similar policies designed for Group 3, except at a higher rate.

Group 5. This group consists of subjects with an increasing propensity to endorse many items. These subjects include a number of upstate and other residents, are young to middle aged, have a high auto ownership level, high education, mid to large household sizes, and high income. A full range of conservation actions appeal to this group.

Group 6. The final group has the highest willingness to conserve gasoline. The data indicates a continuing trend of high action endorsement. This group confirms the relationship between selected socioeconomic variables and conservation opinions and behavior.

Summary

The application of latent trait theory in market segmentation analysis resulted in the development of a set of estimates describing the position of each subject along the latent continuum. In addition, item response characteristic functions were estimated to relate a subject's position in latent space to the probability of item endorsement. These models provided insight into the nature of item content not fully captured by simply tabulating the proportion of subjects endorsing an

action item. In particular, the item and test information functions provide a deeper understanding of the accuracy of individual parameter estimates over the range of the latent continuum. The results of the latent trait theory approach found that the willingness and ability of a subject to conserve gasoline increased if subjects had access to several vehicles, lived in the upstate urban or other areas, had high incomes, were employed, lived in single family housing, and were young and the middle aged.

The latent trait theory approach has not been applied in market segmentation analysis and is not well understood by a large group of researchers. However, the use of modern test theory provides a great deal of detail about the item comprising a measuring instrument and the accuracy of individual parameter estimates. In addition, through the use of the test information function, items may be selected in such a manner that accurate estimates of the latent trait may be developed over the range of interest.

CHAPTER 6

SUMMARY OF RESULTS AND POTENTIAL APPLICATIONS

Introduction

The purpose of this chapter is to compare the results of the application of market segmentation analysis using conventional analysis procedures and the principles of latent trait theory. In addition, this discussion outlines advantages and disadvantages of each method. Finally, the potential applications of the concepts of market segmentation analysis in future energy planning and policymaking are reviewed. This includes the implications of pursuing a strategy based on conventional analysis methods or latent trait theory in achieving energy conservation goals and objectives. The major emphasis in this entire discussion is on summarizing the results of this overall research effort and suggesting the contribution this study makes to improving the transportation energy planning process.

Results Using Conventional Market Segmentation Analysis

Summary of Results

In order to better understand the results of the application of the conventional market segmentation, a series of summary tables have been prepared. Table 6.1 presents an overview of the results across survey years.

In this table an attempt has been made to identify similar market segments over the survey years through calculating the average of the

TABLE 6.1

COMPARISON OF SEGMENTS DEVELOPED FROM CONVENTIONAL MARKET SEGMENTATION ANALYSIS
BASED ON AVERAGE OF CONSERVATION ACTION ITEM
ENDORSEMENT FOR EIGHT COMMON FACTORS

Overall Segment	A		B		C		D		E		F		G		H	
Year	1979	1980	1979	1980	1979	1980	1979	1980	1979	1980	1979	1980	1979	1980	1979	1980
Segment	(5)	(1)	(2)	(2)	(3)	(3)	(1)	(4)	(4)	(5)	(8)	(6)	(7)	(7)	(6)	(8)
Shop less frequently; combine shop and other; shop closer to home	0000	0000	00	00	000	0000	0000	0000	00	00	0000	0000	0	0	00	00
Bus or subway to work; use train, bus, plane for vacation; bus or subway for non-work	0	00	000	000	0000	000	0	0	0	0	0	0	0	0	0	0
Buy fuel efficient car, sell car	00	00	0	0	0	0	0000	00	00	0	0	0	0	0	0	0
Drive slower, tune car	000	000	0	0	000	0	000	000	0	000	000	0	0	0	00	000
Move closer to home; job closer to work	000	000	0	0	0	00	0	00	0	0	0	0	0	0	0	0
Carpool to work; walk or bike to work; shop on way home from work	00	000	0	0	000	00	00	00	0	00	00	000	0	0	0	0
Eliminate use of recreation vehicles; vacation closer to home	000	00	0	0	00	00	000	00	0	00	00	0	0	0	0	0
Cancel vacation trip	00	00	0	0	00	00	00	0	0	0	00	0	0	0	0	0

0000 1.0 to 0.75 percent
000 0.75 to 0.50 percent
00 0.25 to 0.50 percent
0 0.00 to 0.25 percent

proportion of subjects in each segment endorsing the actions comprising a previously defined common factor. For example, one factor consists of three shopping related actions. In order to arrive at the overall average level of endorsement for a segment, the proportion of each segment endorsing individual actions are summed and divided by three. The information provided through this process showed that a fairly consistent set of overall market segments can be developed. These have been called segments A to H.

An examination of the results for overall Group E shows that it is difficult to match the combinations of action clusters for this segment over the survey years. In particular, the factor related to auto replacement has been endorsed at a much higher rate for the 1979 results than for the 1980 results. In addition, a similar result is observed for the factor related to "driving slower" or "tuning car."

It is difficult to suggest that these segments comprise a consistent pair both in terms of item endorsement and group characteristics. Additional analysis is required to better understand this result. Further discussion, therefore, concentrates on the seven remaining segments.

In addition to examining the average action endorsement proportions, a set of tables describing the characteristics of eight segment pairs have been developed and are presented in Tables 6.2 and 6.3. Using this information, two figures have been prepared which summarize the relationship between segment characteristics and energy conservation actions. Only seven of the eight groups are included because of the limitations discussed above with overall segment E.

TABLE 6.2

CHARACTERISTICS OF OVERALL SEGMENTS, A TO D RESULTING
FROM CONVENTIONAL SEGMENT ANALYSIS

Variable	Overall Segment 1979 Segment 1980 Segment	A (5) (1)	B (2) (2)	C (3) (3)	D (1) (4)
Income		Moderate	Low to Moderate	Moderate to High	Moderate
Household Size		Moderate to High	Low to Moderate	Low to Moderate	Moderate to Large
Employment Status		Employed	Employed	Employed	Employed
Marital Status		Married	*Unclear	*Unclear	Married
Area of State		Upstate Urban or Other	Downstate Urban	Downstate Urban	Upstate Urban or Other
Number of Vehicles		High	Low to Moderate	Moderate	High
Age		Young	Young to Middle Aged	Young to Middle Aged	Young to Middle Aged
Sex		*Unclear	*Unclear	*Unclear	*Unclear
Educational Status		High School or College	High School or College	College	High School or College
Housing		Single Family	Multifamily	Multifamily	Single Family

TABLE 6.3

CHARACTERISTICS OF OVERALL SEGMENTS E TO H RESULTING FROM CONVENTIONAL SEGMENT ANALYSIS

Variable	Overall Segment 1979 Segment 1980 Segment	E (4) (5)	F (8) (6)	G (7) (7)	H (6) (8)
Income		Moderate to High	Moderate	Low to Moderate	Moderate to High
Household Size		Low to Moderate	Low to Moderate	Low to Moderate	Low to Moderate
Employment Status		Employed	Employed	*Unclear	*Unclear
Marital Status		*Unclear	Married	*Unclear	Married
Area of State		*Unclear	Upstate Urban or Other	Downstate Urban	Upstate Urban and Other
Number of Vehicles		Moderate to High	High	Low to Moderate	High
Age		Young to Middle Aged	Young to Middle Aged	Young to Middle Aged	Young to Middle Aged
Sex		*Unclear	*Unclear	*Unclear	*Unclear
Educational Status		College	High School or College	Grammar or High School	High School or College
Housing		*Unclear	Single Family	Multifamily	Single Family

Figure 6.1 presents the results for overall segments A, F, D, H and Figure 6.2 presents the results for overall segments C, B, G. In general terms, these figures suggest that the segments can be described first in terms of area location with Figure 6.1 presenting the results for segments from the upstate urban or other areas and Figure 6.2 containing groups from the downstate urban area.

In Figure 6.1, the four upstate urban or other area groups are all characterized by high vehicle ownership, single family housing, and high school or college education. Next, age becomes an important element in defining a segment. After this, groups are described by income levels and, finally, household size.

Figure 6.2 outlines the relationship between group characteristics and actions for downstate urban area segments. As shown in this presentation, all groups are characterized by young to middle age, multi-family housing, and low to moderate household size. Next, income and ownership levels are used to define segments. Finally, the three segments are characterized by different levels of educational attainment.

To summarize, the results suggest the importance of geographic location in understanding the nature of energy conservation response. Downstate urban subjects tend to endorse a lower number of actions and concentrate on transit options to maintain mobility. The upstate urban or other area subjects are more willing to endorse a higher number of actions than the downstate urban area subjects. Income, auto ownership, and education are also related to action endorsement. In the

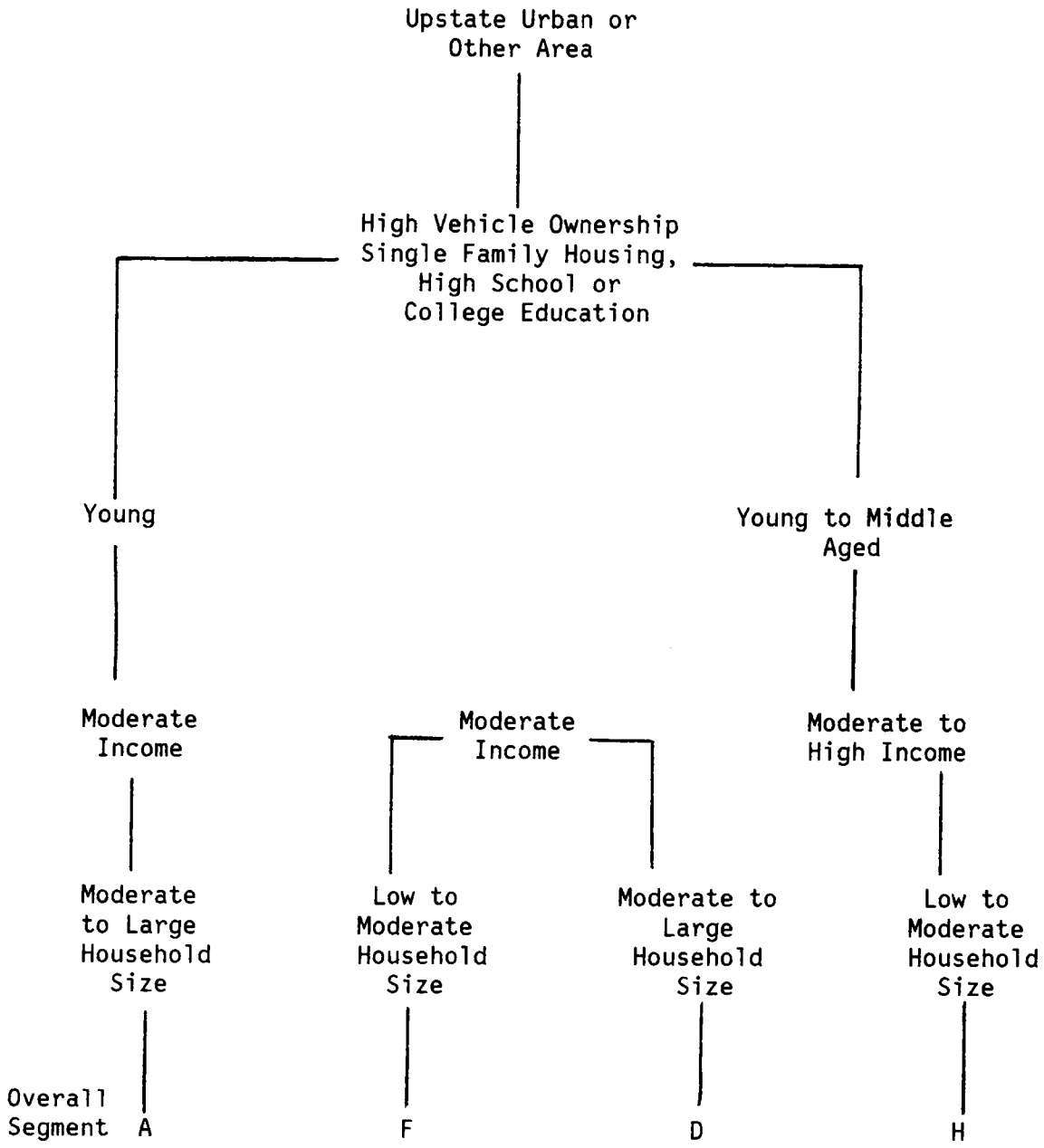


FIGURE 6.1

SUMMARY OF PROFILE CHARACTERISTICS FOR OVERALL MARKET SEGMENTS A, D, F, G

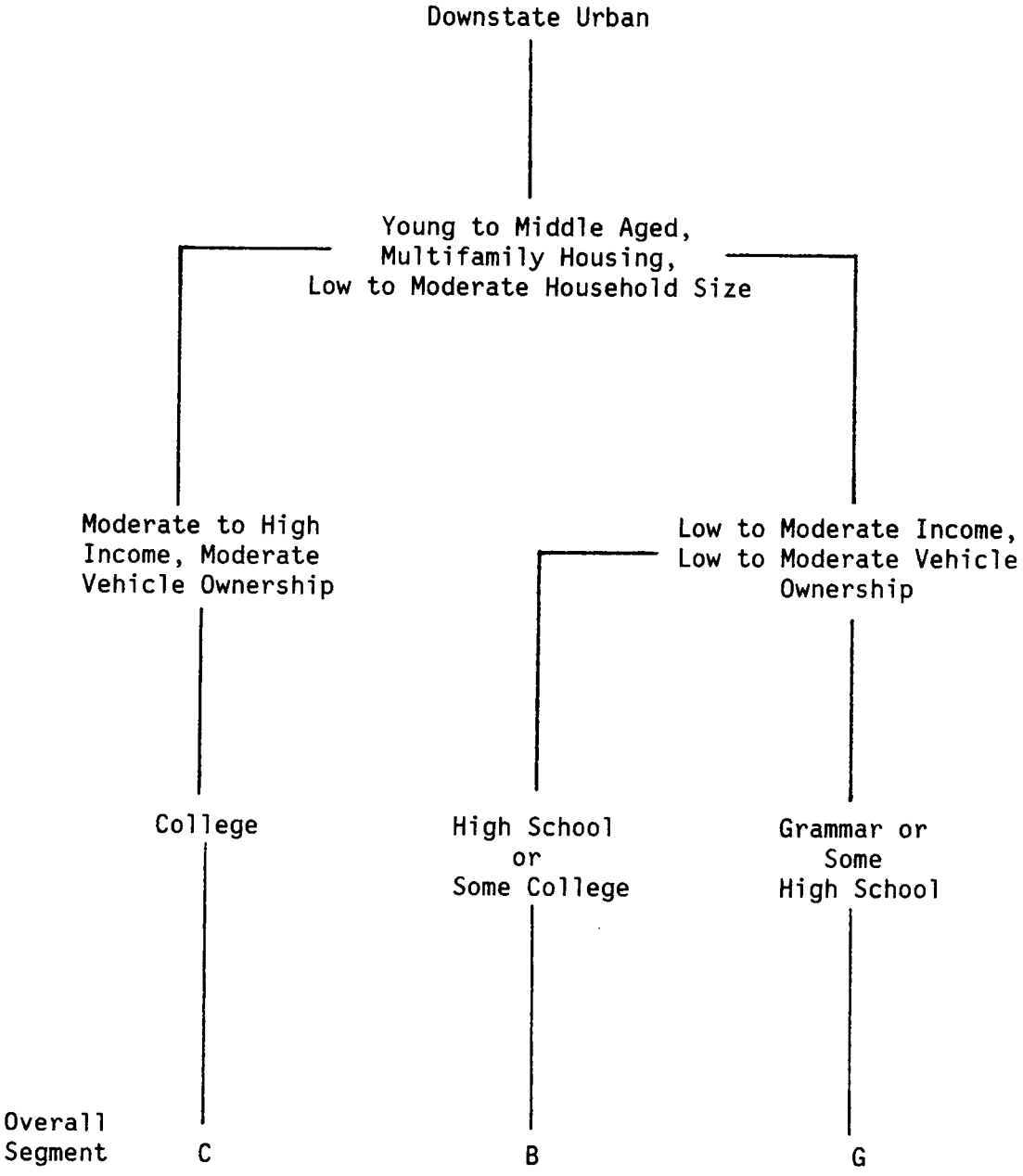


FIGURE 6.2

SUMMARY OF PROFILE CHARACTERISTICS FOR OVERALL MARKET SEGMENTS D, C, G

upstate area, age, income, and household size are important in understanding action endorsement.

Table 6.4 summarizes the nature of conservation actions endorsed by the eight groups emerging from the conventional analysis. Note that for the most part, all groups have undertaken actions related to shopping travel. In addition, auto use and maintenance are also important actions undertaken by most groups. Of course, overall segment G consists of a large number of subjects not endorsing any conservation actions.

Applications

Now that a fairly consistent set of segments have been defined, it is necessary to describe a set of feasible conservation policies and programs to meet the needs of these groups. Based on the analysis results, two recommended sets of program areas emerge: informational campaigns and improved service provisions such as expansion of public transportation or carpool matching.

The set of informational campaigns are designed to help affirm certain behavior and raise the consciousness of potential participants in such conservation actions. This includes providing newspaper, radio, or television messages concerning the advantages of planning daily travel so that trips can be consolidated or eliminated. Other messages concerning auto use and maintenance can be described and implemented to serve the needs of target segments. Fuel availability literature can be distributed through the media to ease concern about the availability of gasoline in vacation or resort areas. One

TABLE 6.4

SUMMARY OF ACTION ENDORSEMENT

Conservation Emphasis	Segment A	B	C	D	E	F	G[a]	H
• Shop Travel	•		•	•	•	•		•
• Non-Auto Mode Choice		•	•					
• Auto Use and Maintenance	•	•	•	•	•	•		•
• Vacation and Leisure	•		•	•				
• Non-Auto Work Travel	•	•	•	•		•		
• Vehicle Replacement	•			•	•			
• Moving	•							

[a]Overall, segment G consists of subjects with low endorsement proportions for all actions.

additional example concerns moving behavior and may consist of campaigns targeted at segments able and willing to change residence. As a result of this research, several target markets have been identified which may be influenced by these programs. In the event of future energy shortages, this information may be used to reach segments in a cost effective manner.

The second general program area consists of improvement of selective transportation related services. These include expanding bus, train, or plane service for vacation travel or improving local transit service. In addition, carpool matching and promotional programs can be targeted at selected segments in order to use energy planning funds in an effective manner.

Table 6.5 outlines a series of recommended programs for target market segments. Note that most groups are targeted to some sort of informational campaign.

These results of the application of conventional analysis can be applied in the event of future energy shortages. Specifically, the results concerning target markets can be used to implement a set of informational strategies and service improvement programs to effectively meet the needs of various groups. For example, the target markets identified in this study can provide an indication of the types of planning programs or policies to implement. In addition, the results can be used to match programs and specific segments. In this manner, informational campaigns and service improvements can be concentrated in areas with the greatest need.

TABLE 6.5

RECOMMENDED PROGRAMS DEVELOPED FOR CONVENTIONAL MARKET SEGMENTATION ANALYSIS

Recommended Programs	Segment A	B	C	D	E	F	G[a]	H
Informational Campaigns								
• Trip Planning	•		•	•	•	•		•
• Fuel Availability	•		•	•				
• Driving Techniques and Maintenance	•	•	•	•	•	•		•
• Housing Availability	•							
Services								
• Carpool Matching	•		•	•		•		
• Improved Public Transit		•	•					
• Improved Plane, Bus, Train Service for Vacations		•	•					

[a]Overall, segment G consists of subjects with low endorsement proportions for all actions.

Recommendations

In addition to suggesting potential conservation strategies, these results can also be used as a data base for increased understanding of the stability of market segments over time. A similar analysis may now be pursued through the use of another statewide telephone survey to conduct further analysis of this nature. If the results of this new study are consistent with these results, in terms of the number of clusters, action endorsement, and characteristics, then additional information concerning the dynamics of conservation opinion and behavior can be better understood. Unlike the background conditions under which the data used in this current research were collected, the present energy situation is much more stable. Awareness of energy shortages is on the decline and it would be useful to understand the relationship between current responses and past responses.

These results represent an example of the use of market segmentation that can be reviewed by analysts interested in undertaking similar studies. For example, researchers in other states may wish to pursue a similar analysis to identify target segments, using data collected in the same time period. Results can be compared to identify any consistencies that emerge. This will contribute to an improved understanding of the transferability of these conclusions and may aid in the design and implementation of conservation strategies in other geographic areas.

Advantages and Disadvantages

The procedures applied in the conventional analysis approach are understood by a wide group of researchers. The computer programs to

implement the analysis are available at most research centers. The results of the analysis suggest that groups can be compared across surveys. This, however, is based on a certain level of subjective analysis. As will be seen in the following section, the principals of latent trait theory are more scientific and comparisons across survey years can be more accurately replicated by other researchers using a similar process. Perhaps the greatest drawback of the use of the conventional analysis approach is that a certain level of subjective interpretation must be applied which will not assure the replicability of study conclusions by other researchers.

Results Using Latent Trait Theory

The use of latent trait theory to identify market segments has resulted in the estimation of a set of individual parameters to describe the relationship between action items responses and the hypothesized latent trait. It has been shown that the sample data consists of three broad categories of subjects. The first consists of subjects with very low willingness and ability to endorse action items. The second segment consists of a transition group with an increasing willingness to participate in gasoline conservation. The third general category consists of subjects with a very high propensity to conserve fuel through a series of conservation actions.

The results are fairly consistent across sample surveys. Older, unmarried, subjects living in multifamily housing with no autos, residing in the downstate urban area exhibit an unwillingness to participate in conservation actions. Higher income, upstate adults living in

single family homes with high auto ownership levels have a generally higher value of the latent trait and, thus, are more likely to undertake a series of conservation actions.

Unlike the previous analysis, it is difficult in this case to identify discrete categories of market segments. This is a reflection of the continuous nature of the estimate of the latent trait as well as the relationship between item characteristic functions and these individual item parameters. However, the use of latent trait theory does yield meaningful information about the characteristics of subjects and action item endorsement. Rather than identifying unique sets of groups, this process suggests important indicants of conservation attitudes and behavior to serve as a focus for policy and program development.

Table 6.6 has been prepared to summarize the results of the application of latent trait theory for market segmentation analysis. The first column describes the most salient socioeconomic characteristics for the segment profile. The second column describes the energy conservation actions endorsed by a particular segment. The final column outlines the recommended policies and programs designed to meet the needs of these target segments. These programs and policies have been described in the previous section.

Applications

The results of this analysis may now be used to target specific segments for future conservation policymaking. Because these groups

TABLE 6.6

GENERAL MARKET SEGMENT PROFILES AND RECOMMENDED POLICIES
AND PROGRAMS USING LATENT TRAIT THEORY

Overall Segment	Segment Profile	Energy Conservation Actions	Recommended Programs and Policies
A	Downstate urban area, old, unmarried unemployed, zero vehicles, low income, multi-family housing.	No actions	
B	Upstate Urban or other areas, moderate vehicle ownership, moderate income	Non-work and vacation travel, auto use, limited non-auto use	Trip planning, car literature, and fuel efficiency, fuel availability information, transit promotion
C	Upstate urban or other area, married, employed, several vehicles, high income, single family housing	Full range of actions including work and non-work travel, auto use and vacation travel, and moving	Trip planning car literature, and fuel efficiency, carpooling, transit promotion and capacity, fuel availability information, housing availability

have been identified, the planning and program funds can be concentrated on these segments to maximize the effectiveness of resource allocation.

In addition to applying these results to future energy planning and programming, this analysis can be used in designing and measuring the relationship between conservation opinion and behavior with socioeconomic characteristics in future studies. For example, the calibrated item characteristic functions can be used to design a form of adaptive testing for market segmentation analysis. In adaptive testing, each item that a subject is administered depends on the answer to a previous item. A subject may first be asked to respond to an item with a high difficulty parameter. If the subject endorses the item, a more difficult one is administered. If the subject does not endorse the item a less difficult item is administered. The process is continued until a stable estimate of the individual latent trait parameter is developed. In the use of adaptive testing, therefore, the full set of items may not be administered to a subject thus saving time and money to conduct an investigation.

The calibrated item parameters may be used to compare samples from different geographic locations with some confidence that the estimates of the latent trait are valid across survey samples.

Recommendations

The overall results of this current research effort can be used to design a future investigation. One of the major drawbacks of the current effort is the limited range of the latent trait measured by the

questionnaire action items. Recall that a large segment did not endorse any action and, therefore, the latent trait individual parameter could not be estimated. Additional action items must be developed to measure the value of the latent trait for these subjects.

Another recommendation emerging from this study for future investigations concerns the scoring of individual action items. It would be more desirable to use a polychotomous graded response score. A five category scoring strategy, for example, would give a better measure of the intensity to which a subject agrees with or endorses a conservation action. This would require rewording items, perhaps, to reflect some level of agreement. This would be important in such action items as selling or buying a car or moving because these are discrete behaviors.

The use of latent trait theory would be an interesting and effective method to measure changes in energy conservation behavior and opinions for a long-term panel of subjects. If such a study is contemplated in the future, procedures of latent trait theory should be given strong consideration for a number of reasons. First, the individual latent trait estimates can be compared on a common scale so that the shift between a subject's position in latent space can be traced over time. In addition, subjects can be compared within time periods and across time periods. The item characteristic functions developed in such a study can be used in future studies to improve the accuracy of latent trait estimates over a range of interest. For example, a study may be concerned with identifying subjects with very high values of the latent trait. A questionnaire consisting of items with required

levels of information may then be developed and administered to a study population.

Advantages and Disadvantages

Latent trait theory consists of a set of analysis methods which are for the most part, not well understood by members of the transportation planning community. In addition, the computer programs and other automated algorithms are not widely available.

The use of latent trait theory, however, does permit the consistent comparison of subjects over time and space. In addition, it allows the development of detailed information about test items and the accuracy of individual estimates of the latent trait.

Summary

The results of the use of conventional market segmentation analysis procedures and the principles of latent trait theory can be used for future energy program and policymaking. The market segments identified can be applied to design responses to future energy shortages that will serve the needs of target segments in a cost-effective manner. The results of both analysis approaches can serve as a guide for other similar studies in other geographic settings.

The segments described through the conventional analysis appear to be consistent across surveys in terms of responses to action items and profile characteristics. The same conclusion can be drawn from the latent trait theory results.

Analysis using latent trait theory offers the researcher an opportunity to evaluate groups on a common scale. In addition, the results can be compared with other studies using the same constructs. However, there is a great deal of work involved in understanding the principles of modern test theory, which may inhibit its widespread dissemination in the transportation planning field.

Using the conventional analysis approaches provides the researcher the opportunity to more easily communicate findings to a wide audience. However, this sacrifices the scientific rigor of latent trait theory.

In general, this research has provided a framework for further investigation into the use of market segmentation analysis. It has also assisted in describing a consistent set of market segments in the general population. Finally, this research has introduced the concept of modern test theory which can serve as a starting point for further investigation into the use of these procedures in transportation planning and analysis.

CHAPTER 7

SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

The purpose of this research was to investigate the application of market segmentation analysis in the field of transportation energy planning. Two statewide telephone questionnaires, collected in 1979 and 1980, served as the data base. The survey population consisted of New York State adults. During this period, the nation experienced a severe gasoline shortfall and rapidly increasing gasoline prices. Survey subjects indicated a willingness and ability to undertake a series of actions related to non-work and vacation related travel. In addition, a certain segment of the survey sample reported the purchase of a more fuel efficient automobile in order to maintain desired level of mobility. Interestingly, few subjects pursued actions related to work travel or transit in spite of the great emphasis placed on these actions by governmental agencies.

The theory of market segmentation analysis was introduced. The process consists of identification of a suitable basis, identification of groups or segments within the data, identification of promising segments for conservation actions, and matching promising segments with policies or programs to conserve energy. Some of the advantages of market segmentation are that it provides a focus for organizational goals and objects, contributes to efficient use of resources, establishes a data base for future analysis, and enhances effectiveness of organizational objectives.

Many examples of the use of market segmentation analysis are contained in private sector marketing. The transportation planning literature presents many incomplete applications of the concepts of market segmentation. For the most part these include attempts to better define specific travel markets or improve travel demand forecasting models. Selected examples of the complete use of market segmentation analysis in public transit planning, however, are difficult to find.

The 1973 Arab Oil Embargo and the 1979 Iranian Revolution resulted in gasoline shortfalls and rapid price increases. The dependence of the transportation sector on imported petroleum resulted in a widespread effort to introduce energy demand into transportation planning theory and practice. Among the most effective policies implemented at this time included the development of more fuel efficient vehicles. This allowed consumers to maintain desired travel levels in the face of rising gasoline prices. Studies of consumer behavior during these periods are travel limited. However, most studies that were completed suggested that consumers implemented a series of small, frequent actions to conserve fuel based on the ability of a household to respond to activity requirements.

Subjects' responses to a set of 18 energy conservation actions served as the basis for segmenting the survey sample data. It was hypothesized that these item responses reflected an underlying psychological construct related to a subject's willingness and ability to conserve gasoline. Two alternative methods were pursued to identify market segments. The first consisted of application of principal factor analysis, cluster analysis, and multidimensional contingency

table analysis. This was called the conventional analysis method. A second approach, using the principles of latent trait theory, consisted of calibrating a set of item response functions to describe the position of a subject along a unidimensional latent continuum related to energy conservation behavior and opinion, attitudes and beliefs.

The following conclusions summarize the results of the conventional analysis approach:

1. The results of the application of conventional market segmentation analysis suggested that the data sets collected in 1979 and 1980 can be divided into eight groups.
2. The eight group solution was found to yield consistent results both in terms of actions taken or endorsed, and segment profiles describing these groups over the two-year survey period.
3. One large group consisting of subjects not endorsing any conservation action was reported. The subjects comprising this group had very low auto ownership levels and tended to live in the downstate urban area in multifamily housing. In addition, these subjects were generally older and had lower incomes and smaller household sizes than members of other segments.
4. Market segments consisting of subjects endorsing a large number of conservation actions tended to be characterized by high auto ownership levels, high incomes, upstate urban or other residence, and high educational levels.

5. The major emphasis on conserving gasoline was placed on actions related to shopping travel, auto use and maintenance, and vacation travel.
6. Among the recommended programs to implement included those policies aimed to promoting certain behavior, such as dissemination of information about trip planning, fuel efficiency and use, fuel availability, and service expansion programs such as improving transit service and carpool matching programs.

The following summarize the results of the application of latent trait theory in market segmentation analysis:

1. A unidimensional latent continuum describing willingness and ability to conserve gasoline was established.
2. The results of the application of latent trait theory suggested that the sample survey data can be divided into three segments.
3. The three group solution was found to consist of one segment with a very low willingness and ability to conserve gasoline and a group with a very high willingness and ability to conserve gasoline.
4. Among the recommended program policies designed to meet the conservation needs of these three groups were informational programs and service expansion programs.
5. A comparison of the results of theoretical item response characteristic functions, using the two parameter normal ogive model and the two parameter logistic model, with

observed frequency ratios of item scores found that the theoretical results fit the observed data quite well.

6. A comparison of the estimates of individual parameters, describing the position of a subject in latent space, developed through the tetrachoric method and the Logist 5 approach indicated a strong positive relationship between the estimates developed through these methods.
7. The item characteristic function parameters, developed for the two parameter normal ogive model over the survey years, were found to be linearly related reflecting scale deficiencies.
8. The use of latent trait theory suggested that upper income, multiple vehicle, large households in the up-state urban and other area had the highest willingness to conserve gasoline. Low income, low ownership levels, and downstate urban location were found to be related to a low willingness to conserve gasoline.
9. The actions with the highest probability of endorsement, for a fixed value of the latent space, were found to be shopping actions, auto use and maintenance, and vacation actions.
10. The use of the test information function can contribute toward developing questionnaires with high levels of accuracy for specific values of the latent trait.

It is difficult to recommend one market segmentation analysis approach over the other. The use of conventional analysis results in a

set of information which can be applied in future energy planning. In addition, the statistical analysis methods contained in this approach are available to many transportation researchers through various computer program packages.

The application of latent trait theory on the other hand, suggests that tests or questionnaires can be designed and implemented which contain a more detailed understanding between the relationship of items with a subject's position in latent space. In addition, because of the "population free" nature of this approach, results can be better compared across different survey samples.

LIST OF REFERENCES

LIST OF REFERENCES

1. Hartgen, D. T. and Neveu, F. J. "The 1979 Energy Crisis: Who Conserved How Much?," Preliminary Research Report 173. Albany, NY: New York State Department of Transportation, April 1980.

2. Neveu, F. J. "The Sensitivity of Work and Non-Work Travel to Energy Shortages or Price Rises," Preliminary Research Report 201. Albany, NY: New York State Department of Transportation, July 1981.

3. U.S. Congress, Senate. Energy Policy and Conservation Act of 1975, Pub. L. 94-163, 94th Cong. 2d Sess., 1975, S. 622.

4. U.S. Congress, Senate. Emergency Energy Conservation Act of 1979, Pub. L. 96-102, 96th Cong. 1st. Sess., 1979, S. 1030.

5. Holcomb, M. C. and Koshy, S. Transportation Energy Data Book: Edition 7. Oak Ridge, TN: Oak Ridge National Laboratory, June 1984.

6. Erlbaum, N. S. Cohen, G. S. and Hartgen, D. T. "Automotive Energy Forecasts: Impact of Carpooling, Trip Chaining, and Auto Ownership," Preliminary Research Report 134. Albany, NY: New York State Department of Transportation, December 1977.

7. Klinger, D., Kuzayak, J. R., and Liss, Susan. "Household Travel," 1977 National Personal Transportation Study Report 9. Washington, D.C.: Federal Highway Administration, July 1982.

8. Neveu, A. J. "The 1973-74 Energy Crisis: Impact on Travel," Preliminary Research Report 131. Albany, NY: New York State Department of Transportation, December 1977.

9. Hartgen, D. T. et al. "Changes in Travel in Response to the 1979 Energy Crisis," Preliminary Research Report 170. Albany, NY: New York State Department of Transportation, December 1979.

10. Keck, C. et al. "Changes in Individual Travel Behavior During the Energy Crisis 1973-74," Preliminary Research Report 67. Albany, NY: New York State Department of Transportation, August 1974.

11. Jones, H. G. et al. An Analysis of the 1973-74 Energy Shortage in the New York City Region. Stony Brook, NY: Bookhaven National Laboratory, June 1975.

12. Peskin, R. L. et al. The Immediate Impact of Gasoline Shortages on Urban Travel Behavior. Washington, D. C.: Federal Highway Administration, April 1975.

13. Skinner, L. E. The Effect of Energy Constraints on Travel Patterns: Gasoline Purchase Study. Washington, D.C.: Federal Highway Administration, July 1975.

14. Sacco, J. F. and Hajj, H. M. "Impact of the Energy Shortages on Travel Patterns and Attitudes," Transportation Research Record 561. Washington, D.C.: Transportation Research Board, 1976.

15. Stearns, M. D. The Behavioral Impacts of the Energy Shortage: Shifts in Trip Making Characteristics. Washington, D.C.: U.S. Department of Transportation, 1979.

16. Corsi, T. M. and Harvey, M. F. Changing Journey to Work Patterns in Response to Motor Fuel Shortages and Higher Prices. College Park, MD: University of Maryland, 1979.

17. Becker, B. W. et al. "Behavior of Car Owners During the Gasoline Shortage," Traffic Quarterly, 30 (July 1976): 100-110.

18. North Central Texas Council of Governments. A Metropolitan Plan for National Energy Emergencies. Dallas, TX: North Central Texas Council of Governments, August 1977.

19. New York State Office of Parks and Recreation. "Fuel Shortages and Recreation Travel: Summer 1979," Technical Report No. 2. Albany, NY: Office of Parks and Recreation, 1979.

20. Regional Plan Association. The Last Gas Crisis: How It Affected Travel. New York, NY: New York State Power Authority, November 1979.

21. System Design Concepts, Inc. New York State Transportation Energy Contingency Planning. Albany, NY: New York State Department of Transportation, January 1981.

22. Boyle, D. K. "Transit Use and Energy Crisis: Experience and Possibilities," Preliminary Research Report 190. Albany, NY: New York State Department of Transportation, August 1981.

23. Comprehensive Planning Organization. Transportation Energy Opinion Survey. San Diego, CA: Comprehensive Planning Organization of the San Diego Region, September 1979.

24. Trent, R. B. and Pollard, C. R. Individual Responses to Rising Gasoline Prices. Morgantown, WV: Department of Sociology and Anthropology, West Virginia University, August 1981.

25. Iverson, E. E. et al. Response of Washington State Residents to Higher Transportation Costs and Energy Shortages. Olympia, WA: Washington State Transportation Commission, August 1984.

26. Cravens, D. W., Hills, G. E., and Woodruff, R. B. Marketing Decision Making. Homewood, FL: Richard D. Irwin, Inc., 1980

27. Engle, J. F., Fiorillo, H. F., and Cayley, M. A. (eds.). Market Segmentation Concepts and Application. New York, NY: Holt, Rinehart and Winston, Inc., 1972.

28. Kotler, P. Marketing Management. Englewood Cliffs, NJ: Prentice-Hall, Inc., 1972.

29. Nicolaidis, G. C., Wachs, M., and Golob, T. F. "Evaluation of Alternative Market Segmentations for Transportation Planning," Transportation Research Record 649. Washington, D.C.: Transportation Research Board, 1976, pp. 23-31.

30. Frank, R. E., Massey, W. E., and Wind, Y. Market Segmentation. Englewood Cliffs, NJ: Prentice-Hall, Inc., 1972.

31. McCann, J. M. "Market Segmentation Response to Marketing Decision Variables," Journal of Marketing Research, 10 (November 1974): 399-405.

32. Lessing, V. P. and Tollefson, J. O. "Market Segmentation and Numerical Taxonomy," Journal of Marketing Research. 8 (November 1971): 480-485.

33. Peters, W. H. "Using MCA to Segment New Car Markets," Journal of Marketing Research. 7 (August 1970): 360-365.

34. Tye, W. B. "Consumer Segmentation," in Behavioral Travel Modelling, pp. 200-210. Edited by D. A. Hensher and P. K. Stopher. London: Croom Helm, 1979.

35. Tisher, M. L. and Dobson, R. "An Empirical Analysis of Behavioral Intentions of Single Occupant Auto Drivers to Switch to High Occupancy Vehicles," Transportation Research, 13A (1979): 143-58.

36. Costantio, D. P. et al. "Investigation of Modal Choice for Dual Mode Transit," Special Report 170. Washington, D.C.: Transportation Research Board, 1976, pp. 67-70.

37. Neveu, A. J., Koppleman, F. S. and Shopher, P. R. "Perceptions of Comfort, Convenience, and Reliability for the Work Trip," Transportation Research Record 723. Washington, D.C.: Transportation Research Board, 1979, pp. 59-63.

38. Koppleman, F. S. "Guidelines for Aggregate Travel Predictions Using Disaggregate Choice Models," Transportation Research Record 610. Washington, D.C.: Transportation Research Board, 1976, pp. 19-24.

39. Dunbar, F. C. "Policy Contingent Travel Forecasting With Market Segmentation," Transportation Research Record 637. Washington, D.C.: Transportation Research Board, 1977, pp. 27-32.

40. Reid, F. A. "Minimizing Error in Aggregate Predictions from Disaggregate Models," Transportation Research Record 675. Washington, D.C.: Transportation Research Board, 1979, pp. 59-65.

41. Recker, W. W. et al. An Empirical Analysis of Household Activity Patterns. Irvine, CA: University of California, 1980.

42. Pas, E. I. "Methodology for Analyzing Daily Travel and Activity Patterns." Presented at 61st Annual Meeting, Transportation Research Board.

43. Zahavi, Y. et al. The UMDT/Urban Interactions. Washington, D.C.: U.S. Department of Transportation, 1981.

44. Hanson, P. "The Activity Patterns of Elderly Households," Geographisica Annaler, 109 (1978): 124.

45. Sen, L. and Benjamin, J. Travel Behavior and Market Segmentation of Low and Middle Income Residents of Richmond, Virginia. Washington, D.C.: U.S. Department of Transportation, December 1979.

46. Woodruff, R. B. et al. "Market Opportunity Analysis for Short-Range Public Transportation Planning Method and Demonstration," National Cooperative Highway Research Program Report 212. Washington, D.C.: Transportation Research Board, September 1981.

47. Crossley Surveys, Inc. Travel Behavior and Fuel Constraints. New York, NY: Crossley Surveys, Inc., October 1979.

48. Crossley Surveys, Inc. Travel Behavior and Fuel Constraints. New York, NY: Crossley Surveys, Inc., November 1980.

49. ATT Comptrollers Accounting Division. Telephone Availability by Household Characteristics. New York, NY: American Telephone and Telegraph, April 1975.

50. Lord, F. M. and Novick, M. R. Statistical Theories of Metal Test Scores. Reading, MA: Addison-Wesley Publishing Co., Inc., 1968.

51. Ferguson, G. A. "The Factorial Interpretation of Test Difficulty," Psychometrika, 6 (October 1981): 323-329.

52. Guilford, J. P. "The Difficulty of a Test and Its Factor Composition," Psychometrika, 6 (April 1981): 67-77.

53. Carroll, J. B. "The Nature of Data, or How to Choose a Correlation Coefficient," Psychometrika, 26 (December 1961): 347-372.

54. Lawley, D. N. and Maxwell, A. E. Factor Analysis as a Statistical Method. New York: Elsevier Publishing Company, Inc. 1971.

55. Harman, H. H. Modern Factor Analysis. Chicago, IL: University of Chicago Press, 1976.

56. Tatsuoka, M. M. Multivariate Analysis. New York, NY: John Wiley and Sons, 1971.

57. Jenrich, R. I. and Sampson, P. F. "Rotation for Simple Loadings," Psychometrika, 31 (September 1966): 313-323.

58. Lorr, M. Cluster Analysis for Social Scientists. San Francisco, CA: Jossey-Bass Publishers, 1983.

59. Haberman, S. J. "The Analysis of Residuals in Cross-Tabulated Tables," Biometrics, 29 (March 1973): 205-220.

60. Lord, F. M. Applications of Item Response Theory to Practical Testing Problems. Hillsdale, NJ: Lawrence Ellbaum Assoc., Inc., 1980.

61. Samejima, F., "The Use of the Information Function in Tailored Testing," Applied Psychological Measurement, 1 (Spring 1977): 233-247.

62. Birnbaum, A. "Some Latent Trait Models and Their Use in Inferring in Examinee's Ability," in Statistical Theories of Mental Test Scores, pp. 358-378. Edited by F. M. Lord and M. R. Novick. Reading, MA: Addison-West Publishing Company, Inc., 1968.

63. Samejima, F., "Normal Ogive Model on the Continuous Response Level in the Multidimensional Latent Space," Psychometrika, 29 (March 1974): 111-121.

64. Samejima, F. "Estimation of Latent Ability Using a Response Pattern of Graded Scores," Psychometrika Monograph, 17 (December 1969).

65. Lord, F. M. "Technical Problems Arising in Parameter Estimation." Paper Presented at the American Educational Research Association Annual Meeting, March 1984, New Orleans, LA.

66. Wingersky, M. S., Barton, M. A., and Lord, F. M. Logist User's Guide. Princeton, NJ: Educational Testing Services, 1982.

67. Samejima, F. and McCarter, C. "Polychoric Correlation Coefficients Computed from Response Patterns or Contingency Table." (Computer Program)

68. Samejima, F. and McCarter, C. "Item Parameter Computation from Proportion Correct and Correlation Between Item Variable and Theta." (Computer Program)

69. Samejima, A. F. "MLE Theta." (Computer Program)

70. Samejima, F. "Item Information Functions and Test Information Functions and Their Square Roots." (Computer Program)

71. Kmenta, J. Elements of Econometrics. New York, NY: MacMillian Publishing Co., Inc., 1980.

72. Samejima, F. "Operating Characteristics in Normal Ogive Model With and Without Quessing." (Computer Program)

APPENDICES

APPENDIX A

QUESTIONNAIRES

CROSSLEY SURVEYS, INC.
109 Third Avenue
New York, New York 10022

EMPIRE STATE POLL
Wave 79-4

#16070
October, 1979

Introduction
Hello, my name is _____ of Crossley Incorporated. I am talking to people in this area about some important issues that concern us all.

I would like to speak with the youngest person at home who is over 18 years of age. Are you that person?

Yes
ASK Q'S. A,B,C BELOW

No
Not Youngest At Home
Would you please ask the youngest person to come to the phone?

Interviewer repeat the above introduction with the proper person and then ask Q's. A,B,C below

A. Please tell me your approximate age? Are you ...? (READ LIST. CHECK ONE)

18 - 34 years	5-1
35 - 64 years	-2
65 years or older . .	-3

B. And you are a ...

Man . . .	6-1
Woman . .	-2

C. The town or place where you live is _____ (Town) and the county is _____ (County) New York -- is that correct? (Fill in correct location above)

PLEASE CONTINUE WITH QUESTION 1 ON THE NEXT PAGE

Mail Point_____ P.S.U.#_____ 7-

Telephone # of Household_____ 8-

Respondent's Name_____ 9-

Interviewer's Name_____ Date_____ 10-

Time Started_____ Time Ended_____

Validated by_____ Date_____

7a. I am going to read a list of things you could do as a result of higher gasoline prices or limited supplies of gasoline. At any time since January, 1979 did you ...? (READ LIST BELOW)

b. If the price of gasoline increased to $1.50 a gallon, would you ...? (READ LIST)

c. If you have 20% less gasoline available for use in your car or cars, would you ...? (READ LIST)

	(Q.7a) Changes made since January 1979	(Q.7b) Changes would make if gasoline cost $1.50	(Q.7c) Changes would make if 20% less gasoline available
Drive more slowly	17-1	19-1	21-1
Carpool to work	-2	-2	-2
Take a bus or subway to work	-3	-3	-3
Walk or bicycle to work	-4	-4	-4
Have car tuned	-5	-5	-5
Move closer to work	-6	-6	-6
Get a job closer to home	-7	-7	-7
Shop on the way home from work	-8	-8	-8
Shop less frequently	-9	-9	-9
Combine shopping and other trips	-0	-0	-0
Shop closer to home	18-1	20-1	22-1
Eliminate use of snowmobiles, power boats, or other recreational vehicles	-2	-2	-2
Take a vacation closer to home	-3	-3	-3
Cancel a vacation trip	-4	-4	-4
Use a train, bus or airplane for vacation trips	-5	-5	-5
Buy a more fuel efficient car	-6	-6	-6
Sell a car	-7	-7	-7
Take bus or subway more often for non-work trips	-8	-8	-8

8. How many cars, vans, or light trucks do members of your household own?

None	23-0
One	-1
Two	-2
Three	-3
Four or more	-4

9. How far do you live from your place of work? (Miles, one way)

0 - 0.5 miles	24-1
0.6 - 1.0	-2
1.1 - 2.0	-3
2.1 - 5.0	-4
5.1 - 10.0	-5
10.1 - 15.0	-6
15.1 - 25.0	-7
25.1 - 50.0	-8
Over 50.0	-9

Finally, I would like to ask you a few more questions which will be used only to combine your answers with others from your area.

a. What is the last grade of school you have completed? (READ LIST)

Grammer school (1-8)	70-1
Some high school (9-11)	-2
High school graduate (12)	-3
Some college (1-3 years)	-4
College graduate (4 & more)	-5

b. Are you married, single, widowed or divorced?

Married	71-1
Single	-2
Widowed, divorced or separated	-3

c. Including yourself, how many people live in this household? (READ LIST)

One person	72-1
Two persons	-2
Three or four persons	-3
Five or more	-4

d. Do you consider yourself to be? (READ LIST)

White	73-1
Black	-2
Hispanic	-3
Other	-4

e. Are you currently employed ...? (READ LIST)

Full-time outside of home	74-1
Part-time outside of home	-2
Unemployed seeking work	-3
Homemaker	-4
Retired	-5
Student	-6

f. Is your total annual family income before taxes over $15,000?

No			Yes		75-
Is it $10,000 or higher?	Yes	-2	Is it $25,000 or higher?	Yes	-4
	No	-1		No	-3

g. Do you own or rent your home?

Own	76-1
Rent	-2

h. Do you live in an apartment building, two family or single family house?

Single family house	77-1
Two family house	-2
Apartment, condominium, co-op	-3

i. Finally, are you registered to vote

Yes	78-1
No	-2

j. Are you registered as a ... (READ LIST)

Democrat	79-1
Republican	-2
Other party	-3

RETURN TO PAGE ONE AND COMPLETE THE INFORMATION IN BOX AT BOTTOM OF PAGE. 80- 1

Transportation Survey

#16183

7a. I am going to read a list of things you could do as a result of higher gasoline prices or limited supplies of gasoline. At any time since January, 1980 did you ...? (READ LIST BELOW)

b. If the price of gasoline increased to $2.00 a gallon, would you ...? (READ LIST)

	(Q.7a) Actions Taken Since January 1980	(Q.7b) Actions You Would Take If Gasoline Cost $2.00
Drive more slowly than before	17-1	19-1
Carpool to work	-2	-2
Take a bus or subway to work	-3	-3
Walk or bicycle to work	-4	-4
Have car tuned more often	-5	-5
Move closer to work	-6	-6
Get a job closer to home	-7	-7
Shop on the way home from work	-8	-8
Make fewer shopping trips	-9	-9
Combine shopping and other nonwork trips	-0	-0
Shop closer to home	18-1	20-1
Eliminate use of snowmobiles, power boats, or other recreational vehicles	-2	-2
Take a vacation closer to home	-3	-3
Cancel a vacation trip	-4	-4
Use a train, bus or airplane for vacation trips rather than auto	-5	-5
Replace a car with a more fuel efficient one	-6	-6
Sell a car (don't replace)	-7	-7
Take bus or subway more often for non-work trips	-8	-8

8. How many cars, vans, or light trucks do members of your household own?

None	21-0
One	-1
Two	-2
Three	-3
Four or more	-4

9. How far do you live from your place of work? (Miles, one way)

10. How far do you live from your usual place of shopping? (Miles, one way)

	Q.9	Q.10
0 - 0.5 miles	22-1	23-1
0.6 - 1.0	-2	-2
1.1 - 2.0	-3	-3
2.1 - 5.0	-4	-4
5.1 - 10.0	-5	-5
10.1 - 15.0	-6	-6
15.1 - 25.0	-7	-7
25.1 - 50.0	-8	-8
Over 50.0	-9	-9

Finally, I would like to ask you a few more questions which will be used only to combine your answers with others from your area.

IF SAME WOMEN INTERVIEWED ON BLUE QUESTIONNAIRE, SKIP TO Q.h

a. What is the last grade of school you have completed? (READ LIST)

Grammer school (1-8)	24-1
Some high school (9-11)	-2
High school graduate (12)	-3
Some college (1-3 years)	-4
College graduate (4 & more)	-5

b. Are you married, single, widowed or divorced?

Married	25-1
Single	-2
Widowed, divorced or separated	-3

c. Please tell me your approximate age? Are you (READ LIST)

15-17	26-0
18-34	-1
35-64	-2
65 or older	-3

d. Including yourself, how many people live in this household? (READ LIST)

One person	27-1
Two persons	-2
Three or four persons	-3
Five or more	-4

e. Do you consider yourself to be? (READ LIST)

White	28-1
Black	-2
Hispanic	-3
Other	-4

f. Are you currently employed ...? (READ LIST)

Full-time outside of home	29-1
Part-time outside of home	-2
Unemployed seeking work	-3
Homemaker	-4
Retired	-5
Student	-6

g. Is your total annual family income before taxes over $15,000?

	No			Yes		30-
Is it $10,000 or higher?	Yes	-2	Is it $25,000 or higher?	Yes	-4	
	No	-1		No	-3	

h. Do you live in an apartment building, two family or single family house?

Single family house	31-1
Two family house	-2
Apartment, condominium, co-op	-3

i. Sex:

Male	32-1
Female	-2

j. County:_____ 33/34/35-

k. Do you own or rent your home?

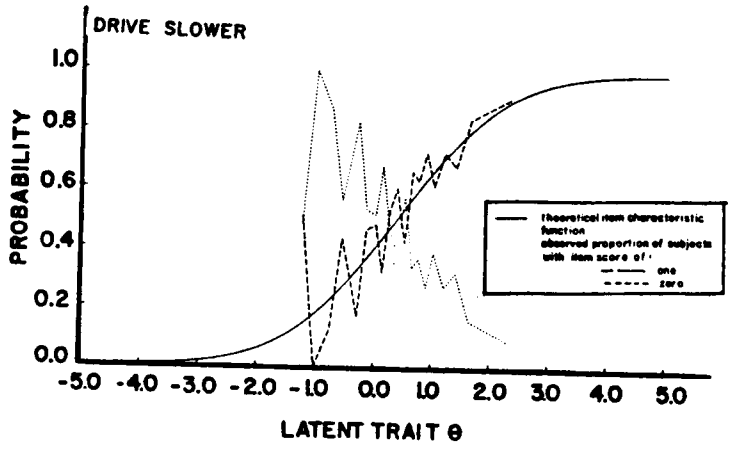
Own	63 -1	GO TO SCREENER
Rent	-2	PART C Q.2

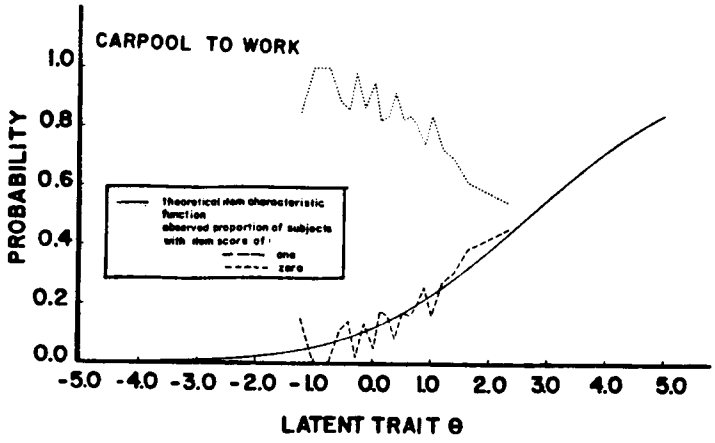
ENTER RESPONDENT'S NAME AND TELEPHONE NUMBER AND TERMINATE

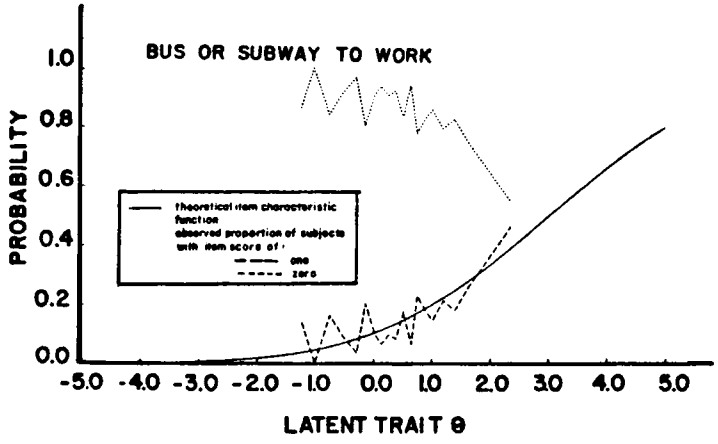
RESPONDENT'S NAME:_____ TELEPHONE #( )_____

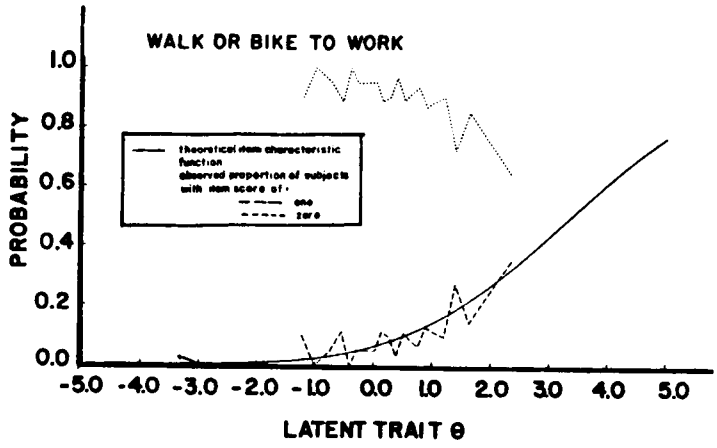
APPENDIX B

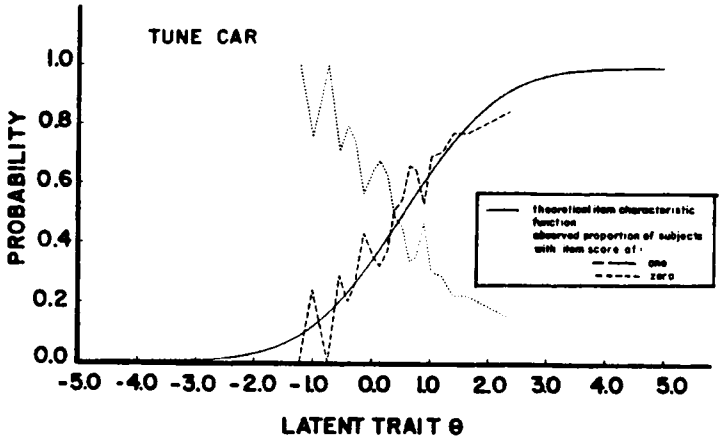
ITEM VALIDATION FOR NORMAL OGIVE MODEL

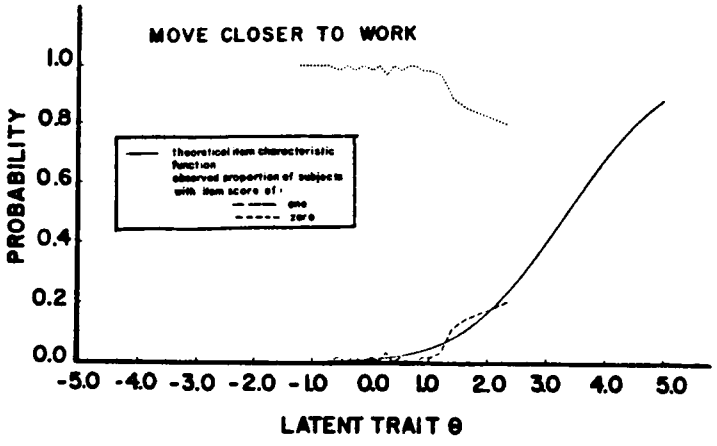

DRIVE SLOWER
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ

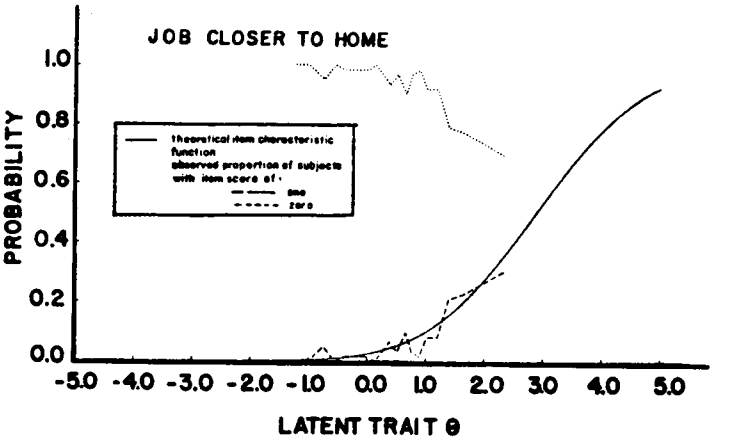

CARPOOL TO WORK
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ

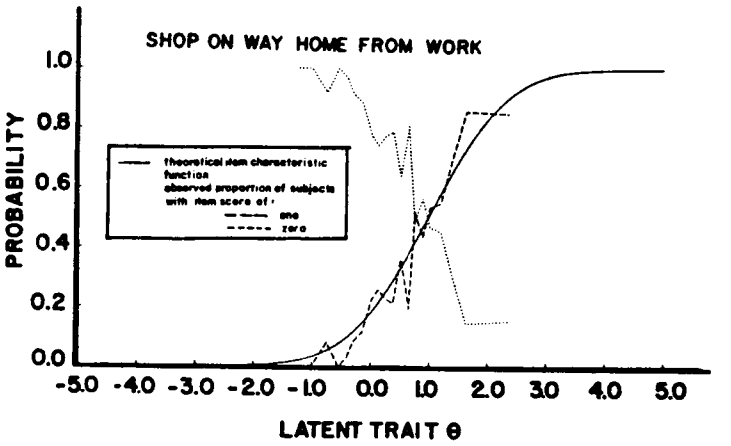

BUS OR SUBWAY TO WORK
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ

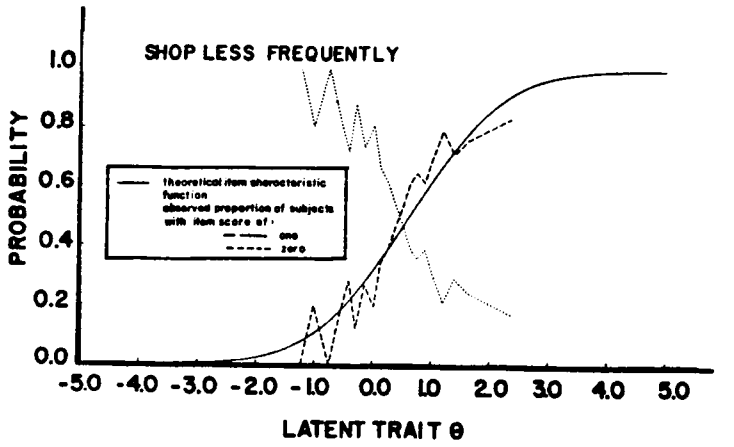

WALK OR BIKE TO WORK
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ

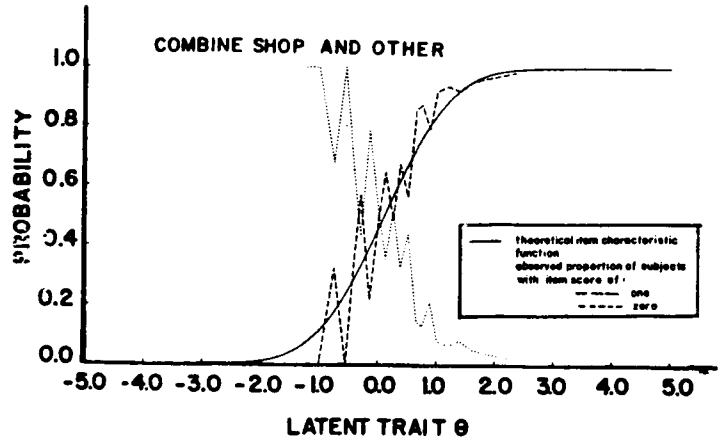

TUNE CAR
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ

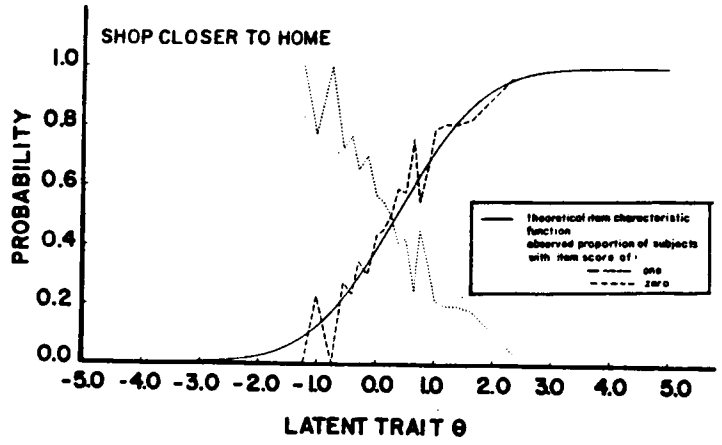

MOVE CLOSER TO WORK
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ

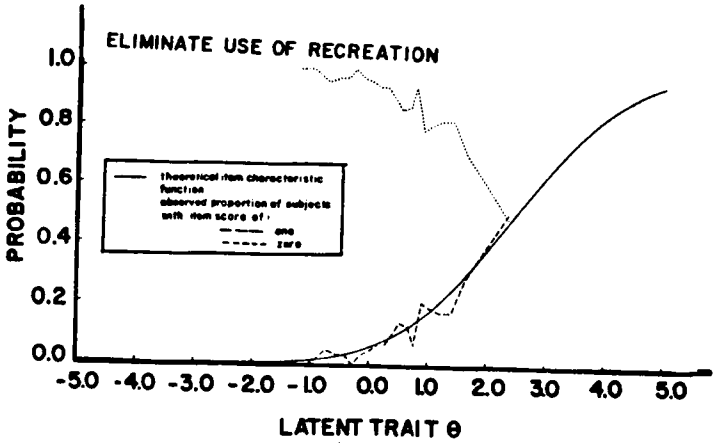

JOB CLOSER TO HOME
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ

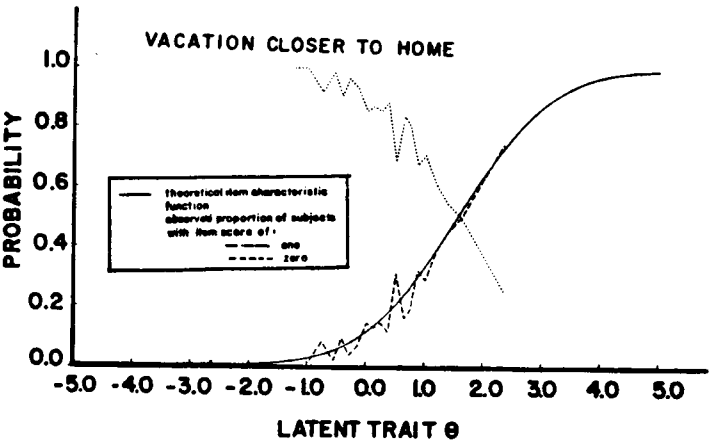

SHOP ON WAY HOME FROM WORK
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ

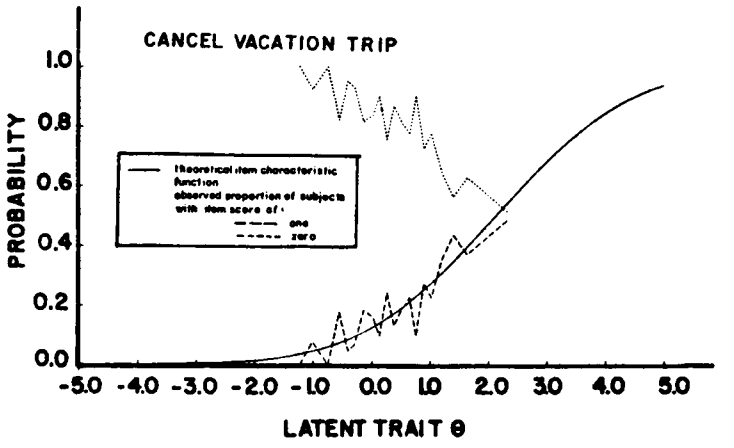

SHOP LESS FREQUENTLY
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ

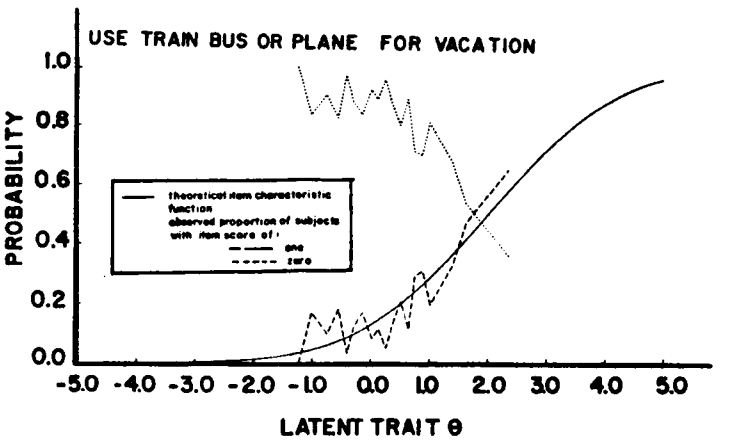

COMBINE SHOP AND OTHER
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ

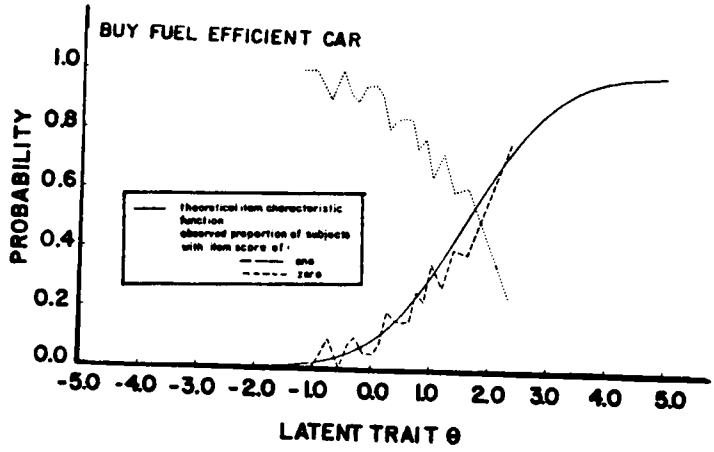

SHOP CLOSER TO HOME
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ

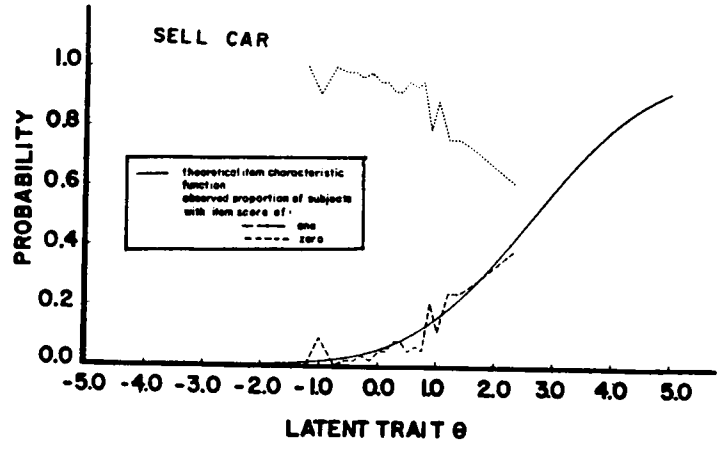

ELIMINATE USE OF RECREATION
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ

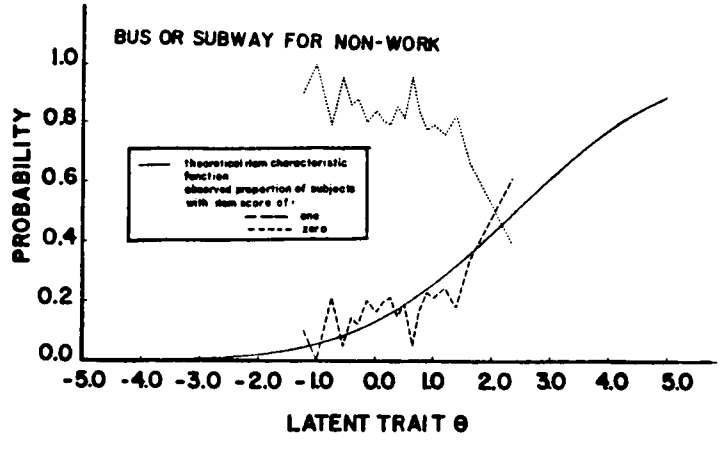

VACATION CLOSER TO HOME
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ

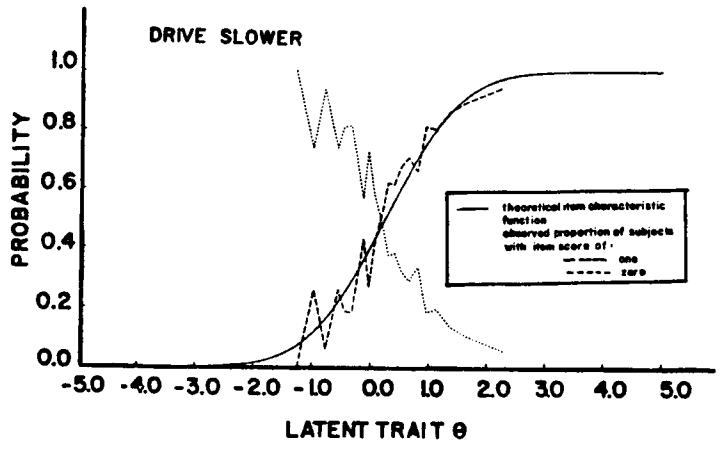

CANCEL VACATION TRIP
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ

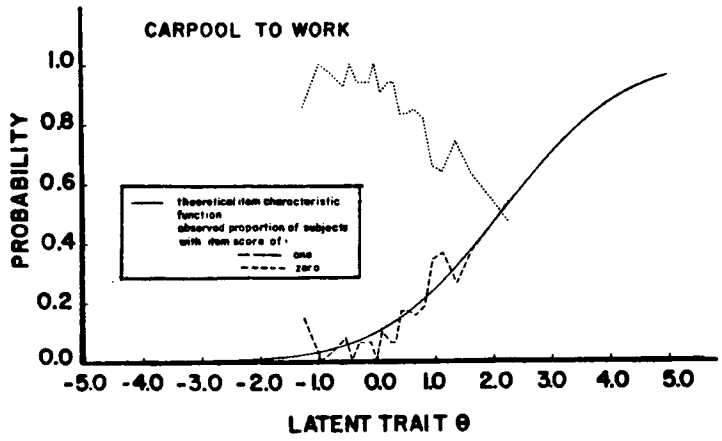

USE TRAIN BUS OR PLANE FOR VACATION
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ

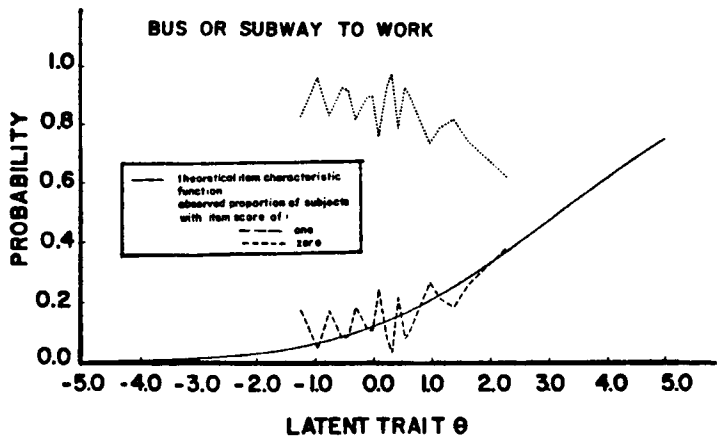

BUY FUEL EFFICIENT CAR
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ

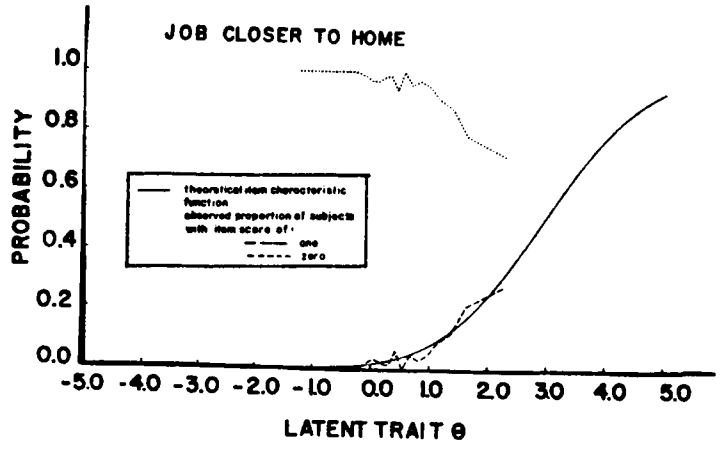

SELL CAR
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ

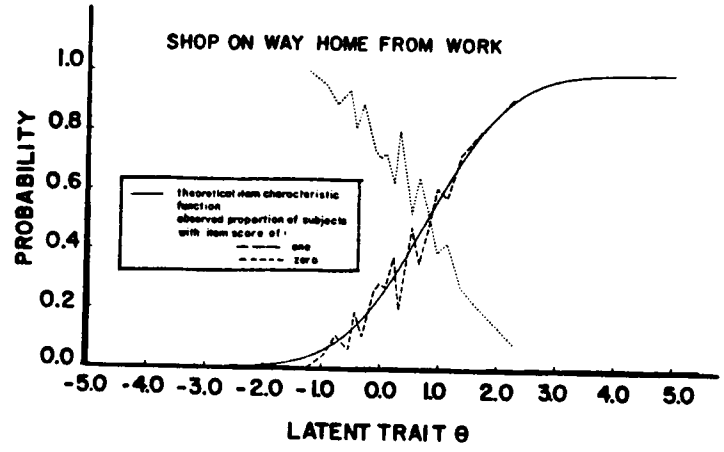

BUS OR SUBWAY FOR NON-WORK
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ

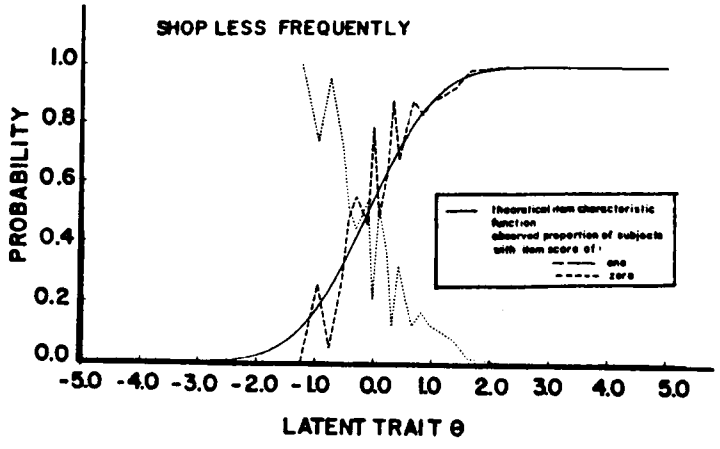

DRIVE SLOWER
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ

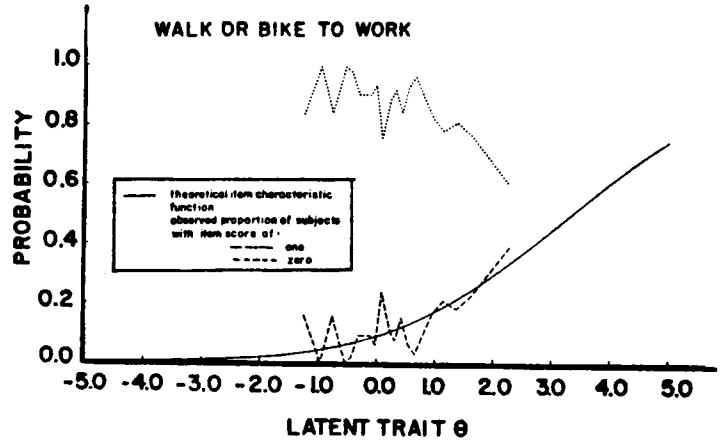

CARPOOL TO WORK
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ

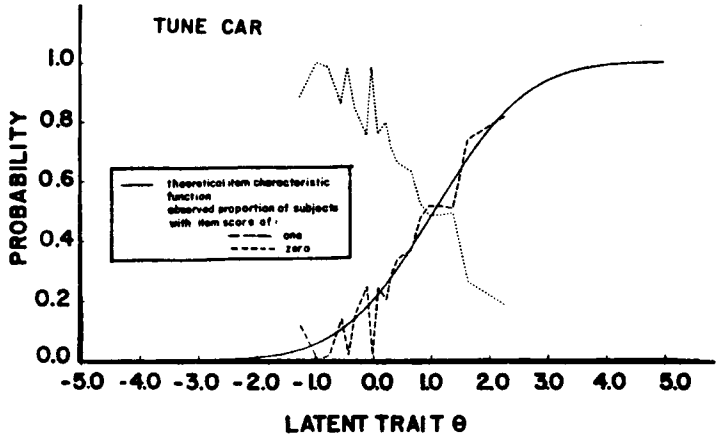

BUS OR SUBWAY TO WORK
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ

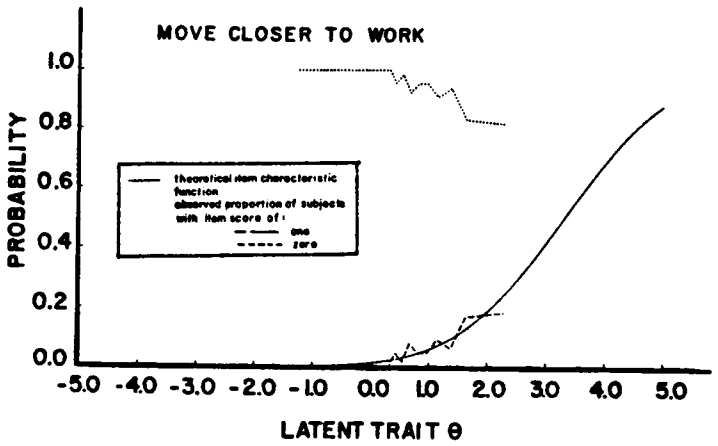

JOB CLOSER TO HOME
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ

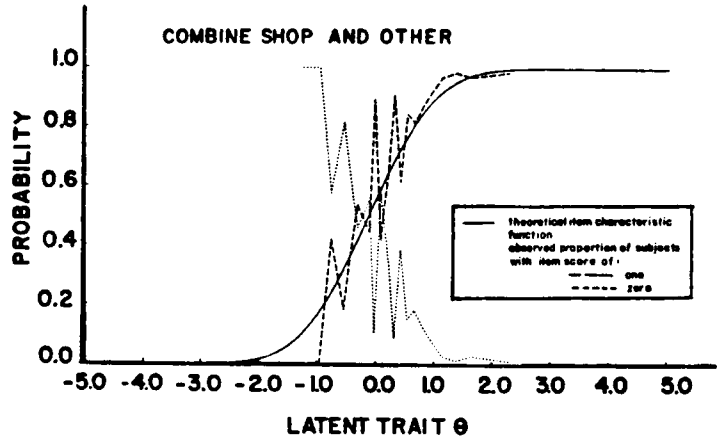

SHOP ON WAY HOME FROM WORK
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ

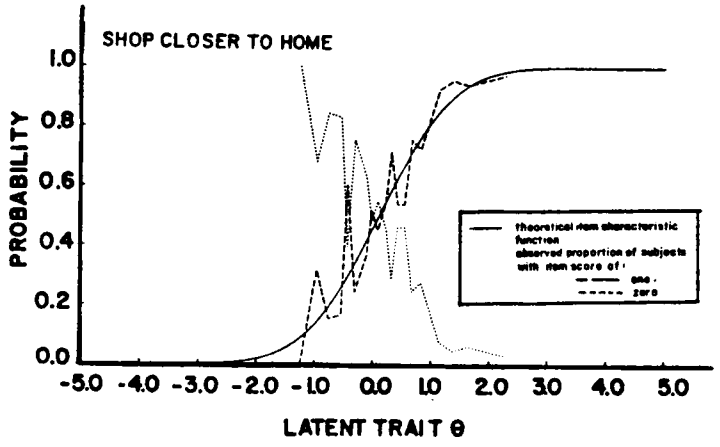

SHOP LESS FREQUENTLY
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ

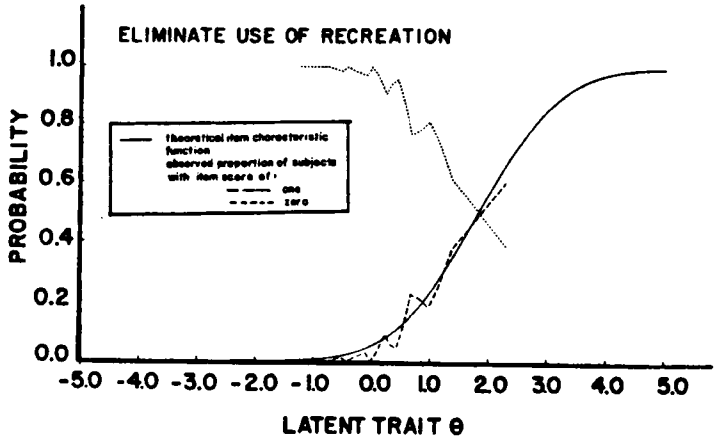

WALK OR BIKE TO WORK
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ

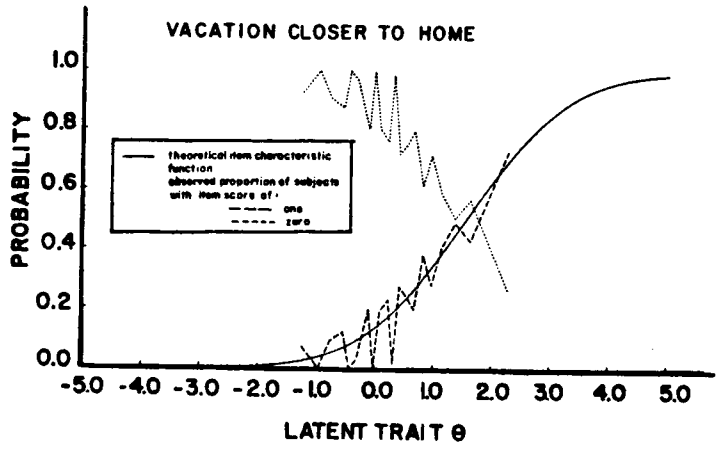

TUNE CAR
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ

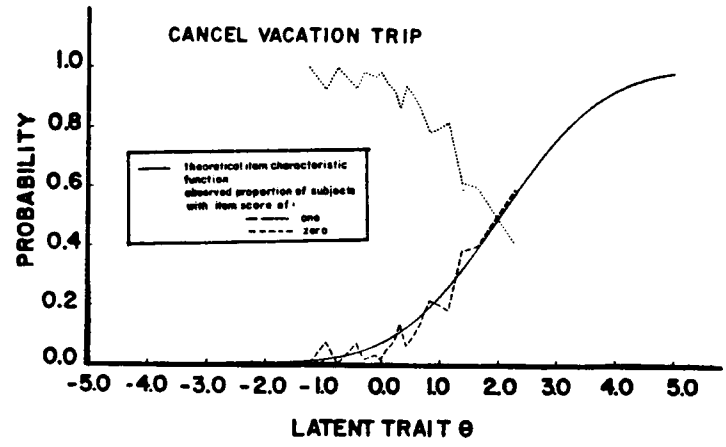

MOVE CLOSER TO WORK
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ

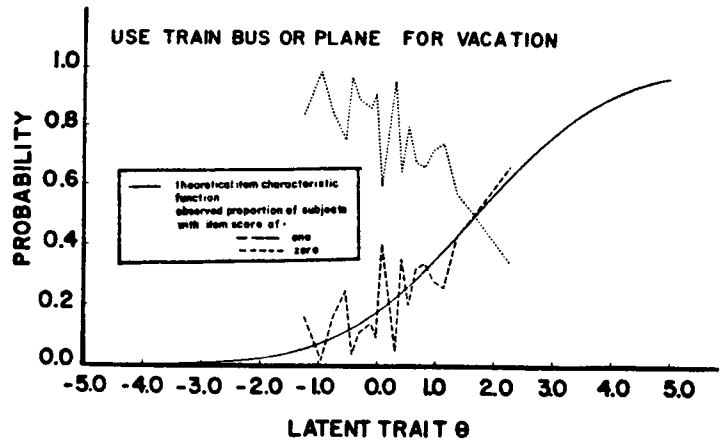

COMBINE SHOP AND OTHER
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ

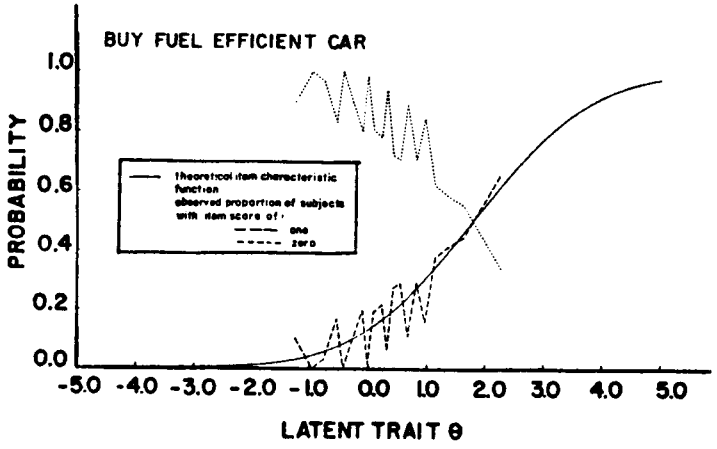

SHOP CLOSER TO HOME
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ

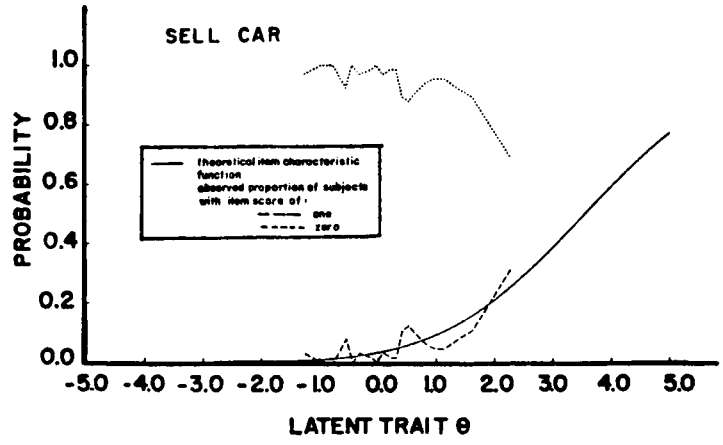

ELIMINATE USE OF RECREATION
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ

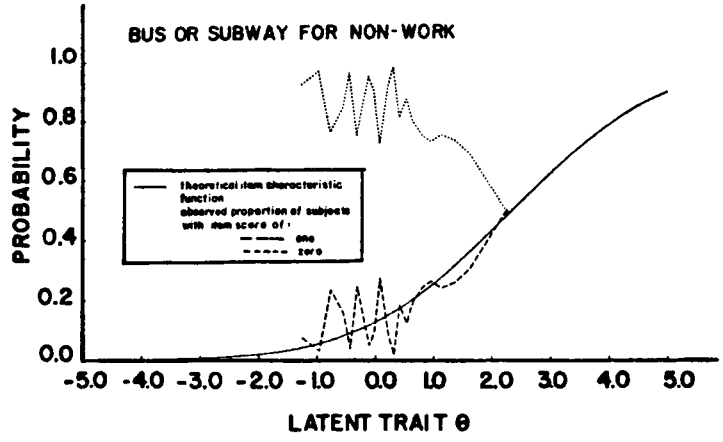

VACATION CLOSER TO HOME
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ


CANCEL VACATION TRIP
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ


USE TRAIN BUS OR PLANE FOR VACATION
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ

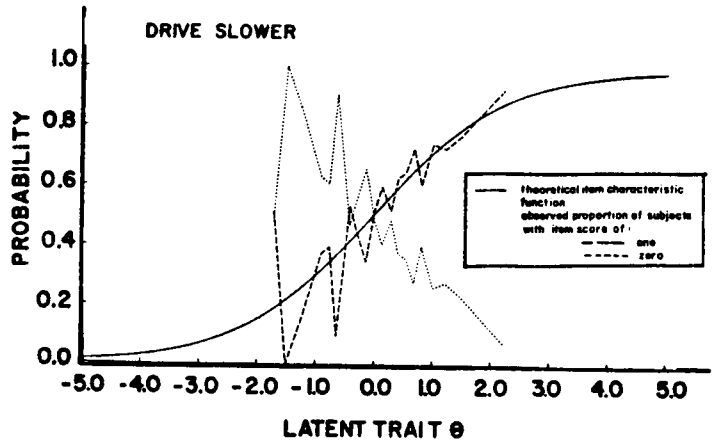

BUY FUEL EFFICIENT CAR
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ

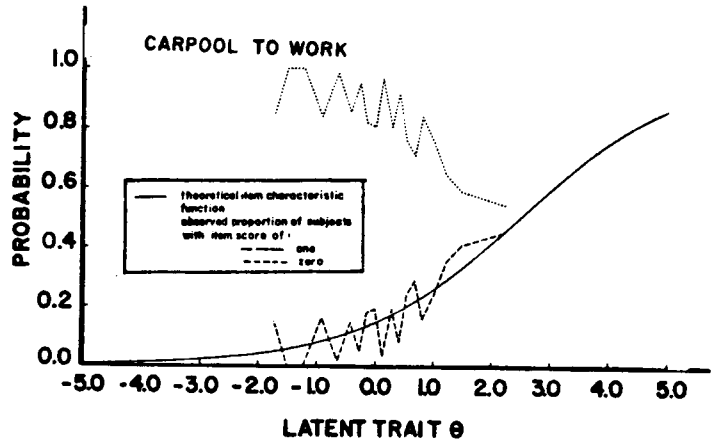

SELL CAR
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ

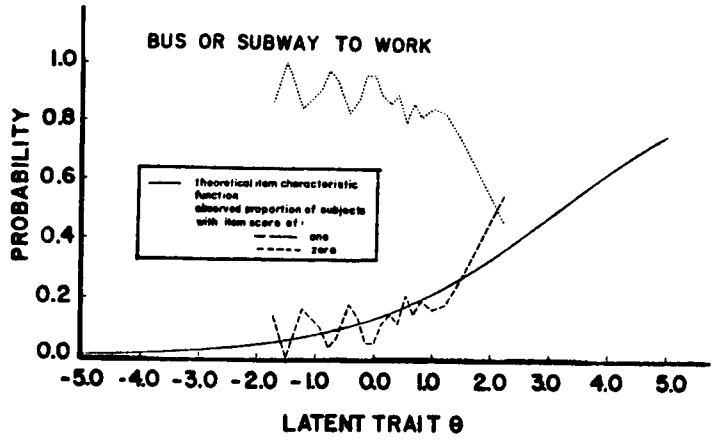

BUS OR SUBWAY FOR NON-WORK
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ

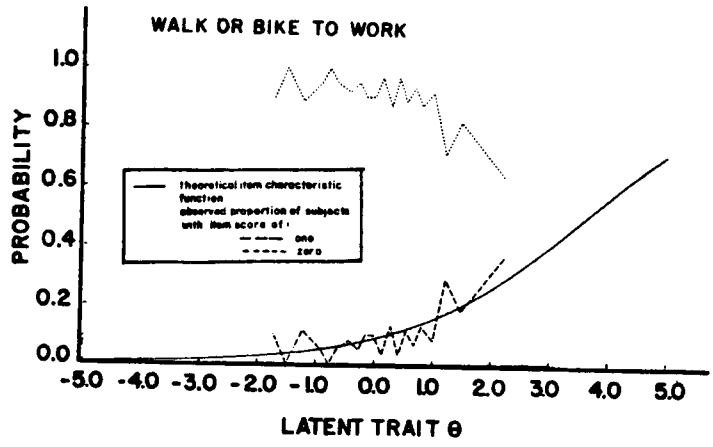
APPENDIX C

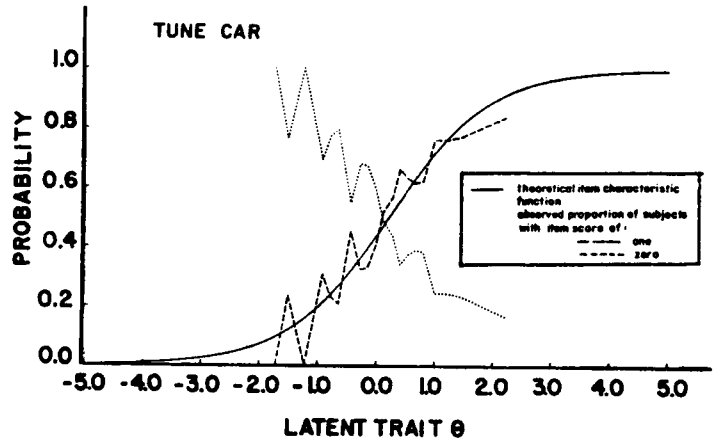
ITEM VALIDATION FOR LOGISTIC MODEL

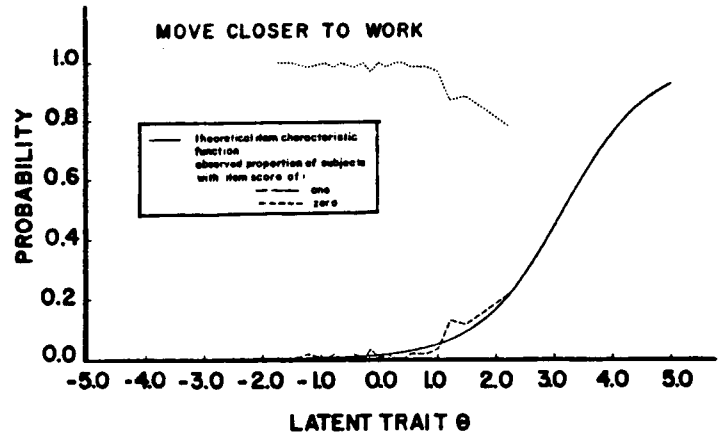

DRIVE SLOWER
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ

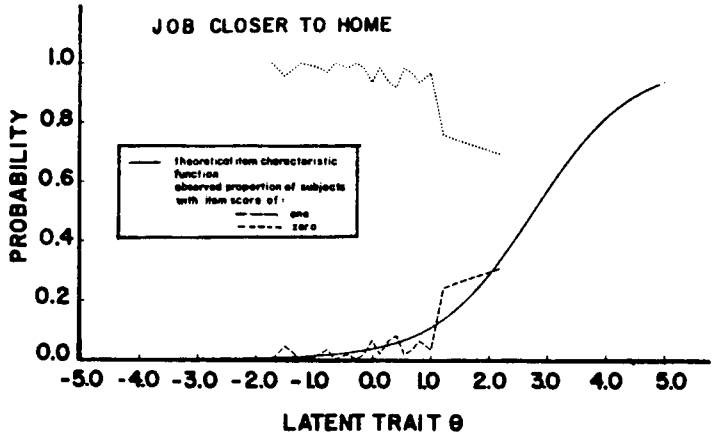

CARPOOL TO WORK
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ

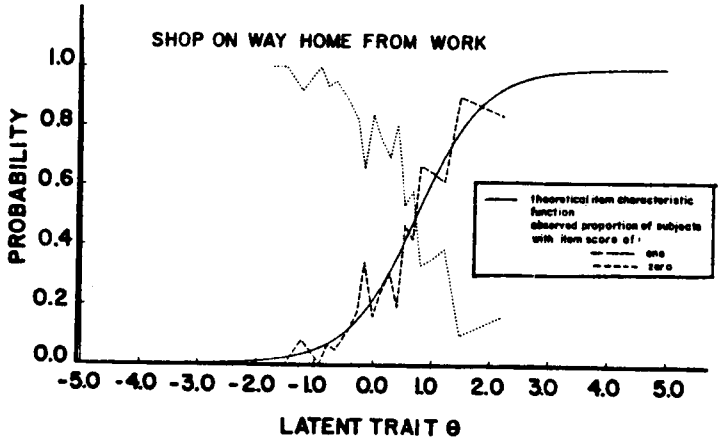

BUS OR SUBWAY TO WORK
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ

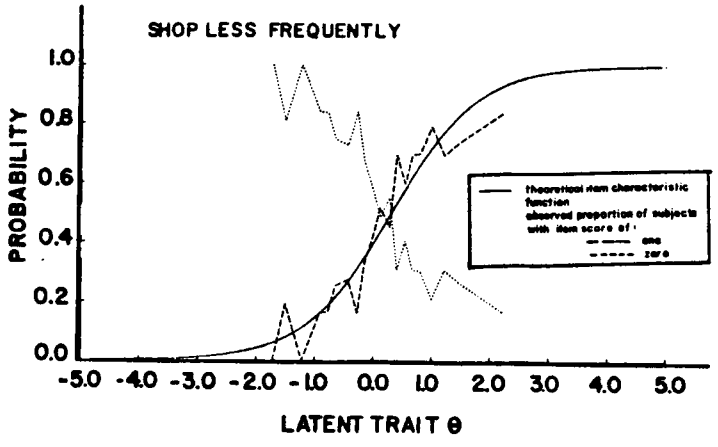

WALK OR BIKE TO WORK
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ

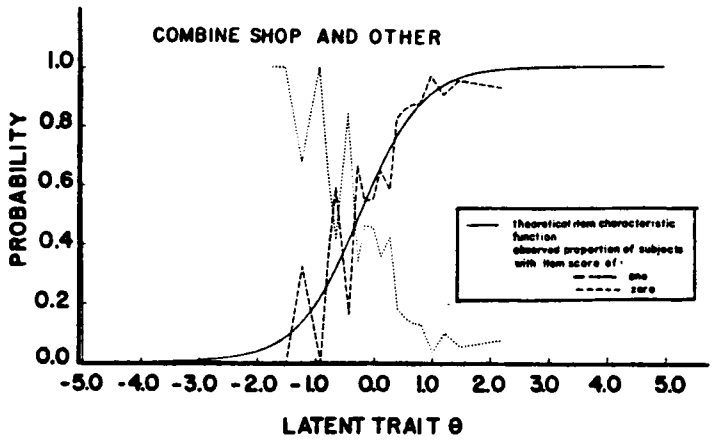

TUNE CAR
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ

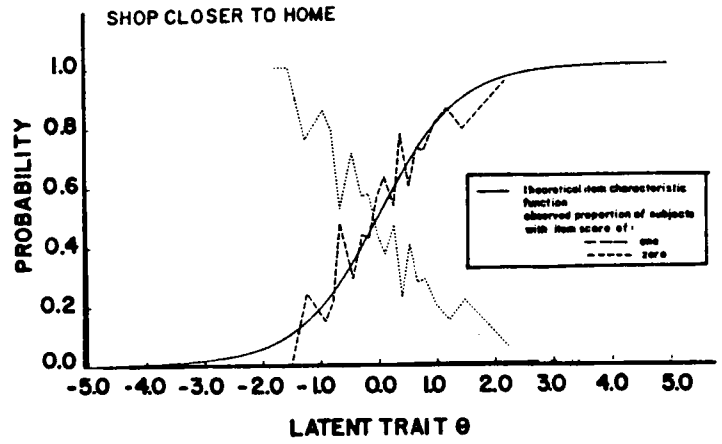

MOVE CLOSER TO WORK
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ

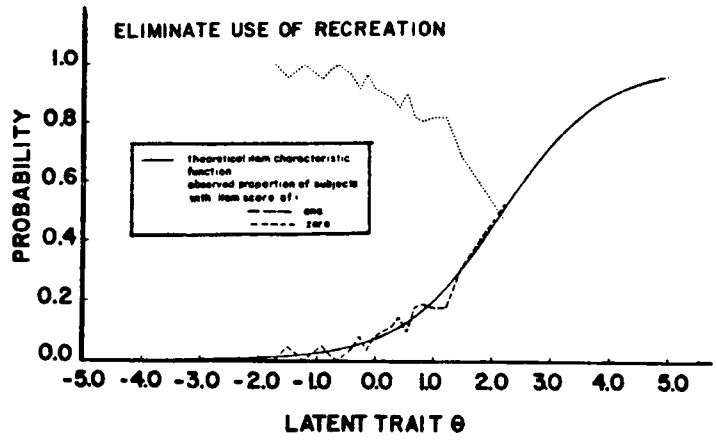

JOB CLOSER TO HOME
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ

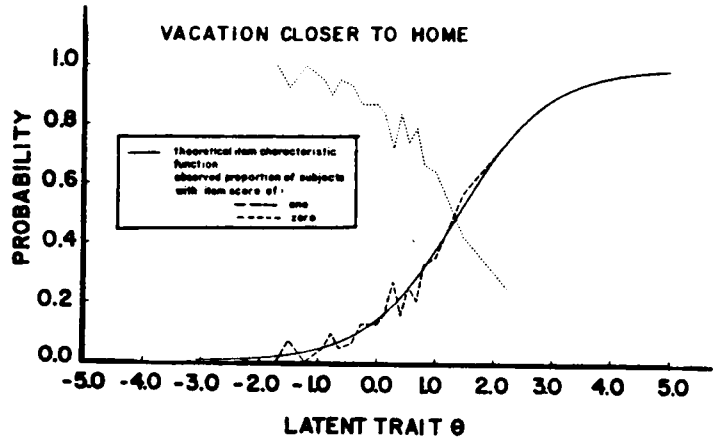

SHOP ON WAY HOME FROM WORK
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ

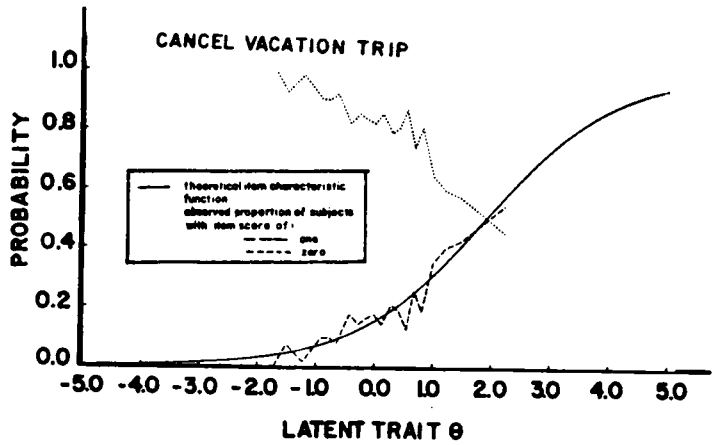

SHOP LESS FREQUENTLY
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ

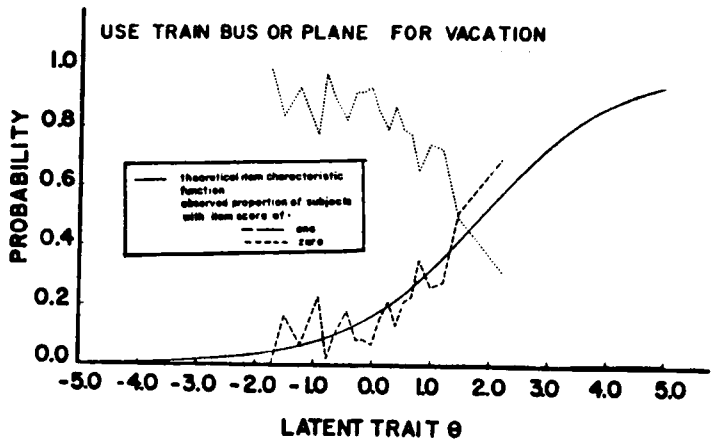

COMBINE SHOP AND OTHER
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ

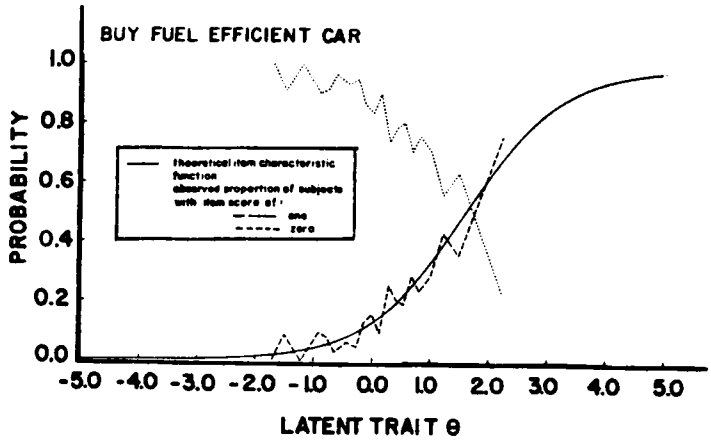

SHOP CLOSER TO HOME
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ

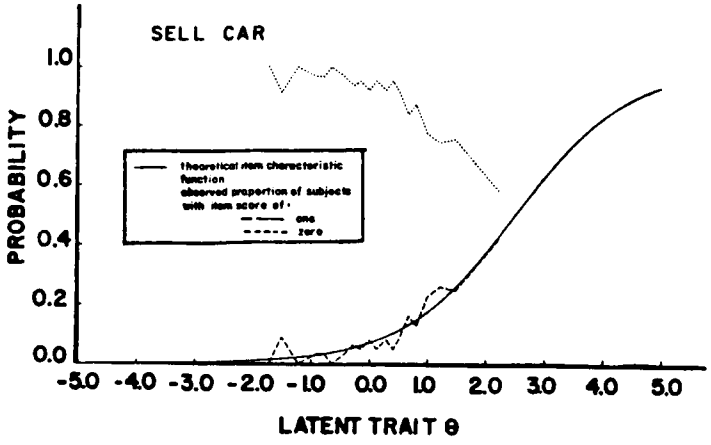

ELIMINATE USE OF RECREATION
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ

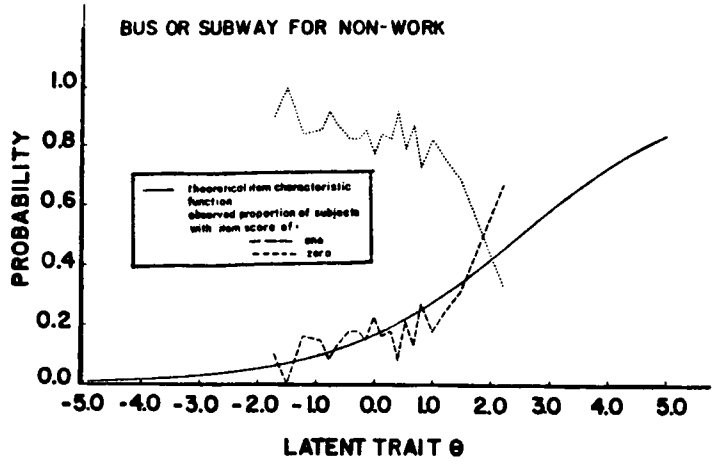

VACATION CLOSER TO HOME
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ

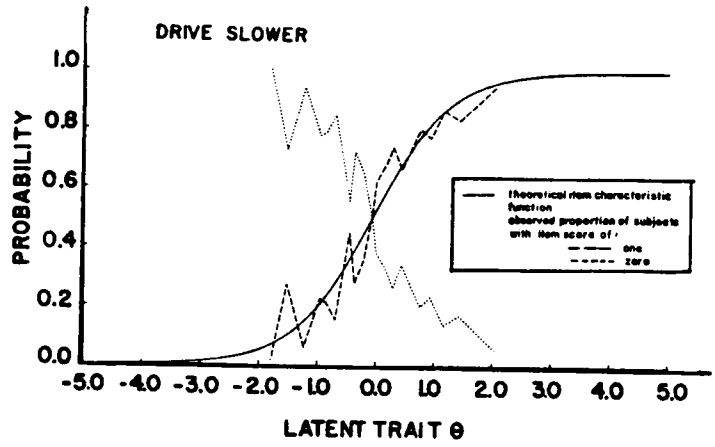

CANCEL VACATION TRIP
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ

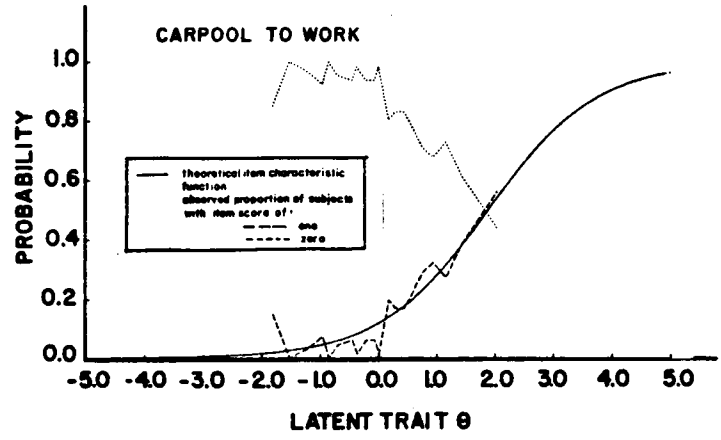

USE TRAIN BUS OR PLANE FOR VACATION
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ

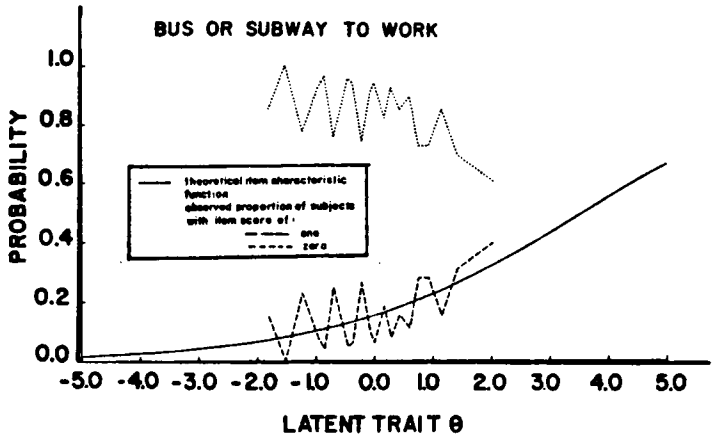

BUY FUEL EFFICIENT CAR
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ

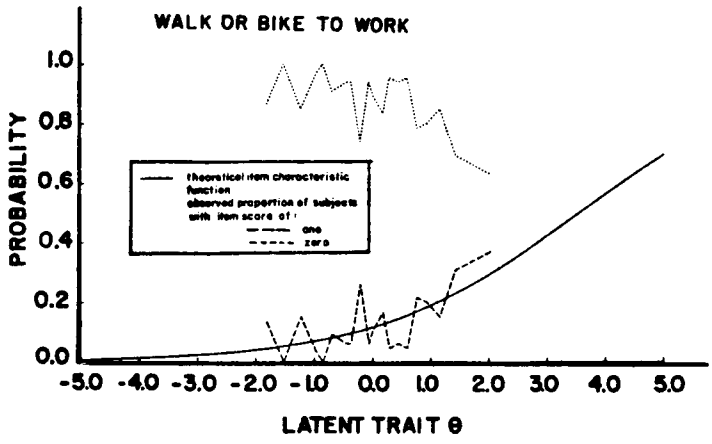

SELL CAR
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ

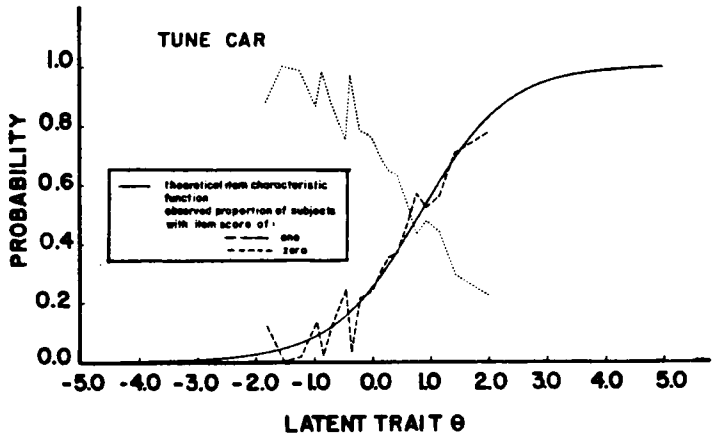

BUS OR SUBWAY FOR NON-WORK
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ

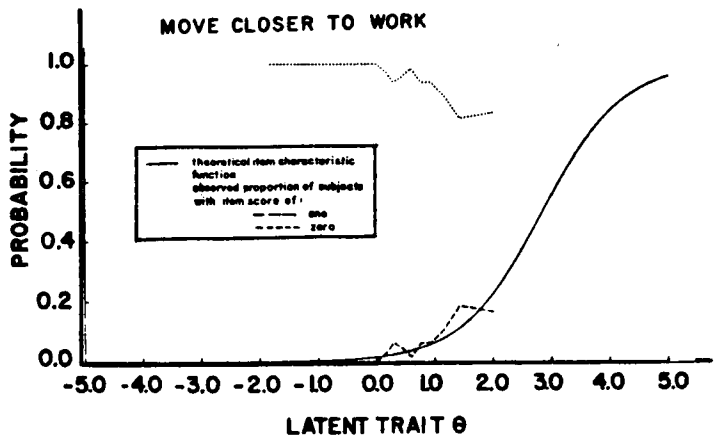

DRIVE SLOWER
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ

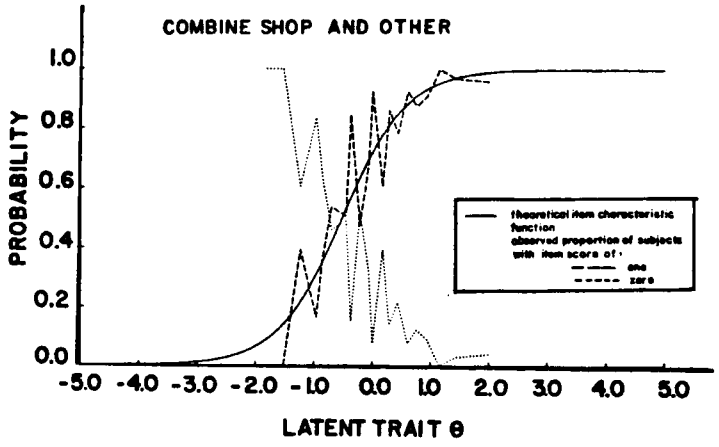

CARPOOL TO WORK
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ

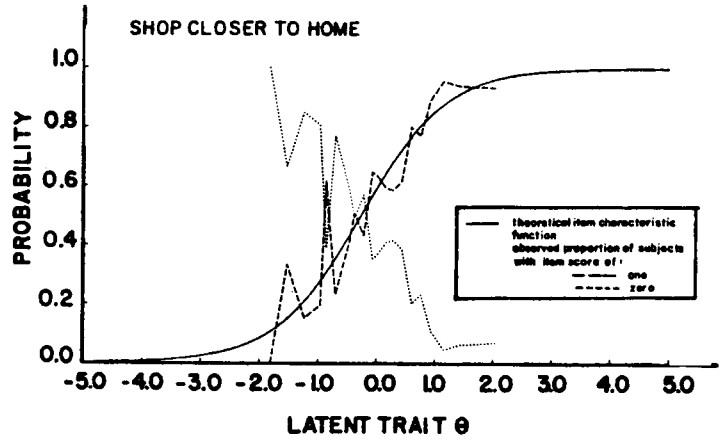

BUS OR SUBWAY TO WORK
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ

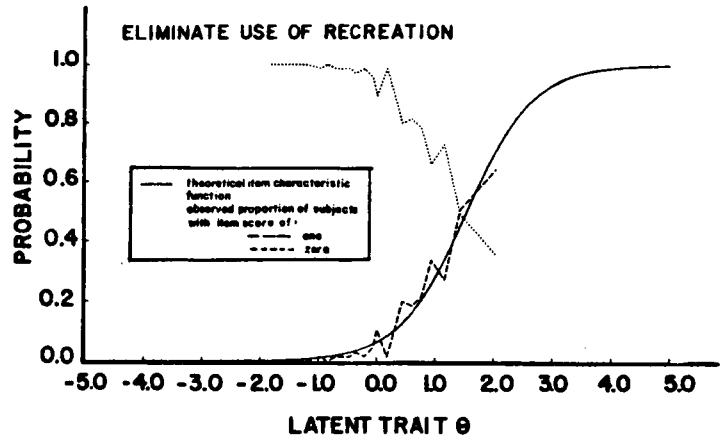

WALK OR BIKE TO WORK
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ

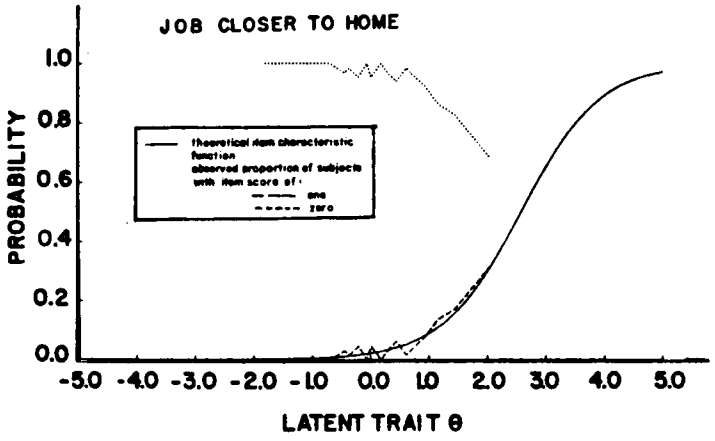

TUNE CAR
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ

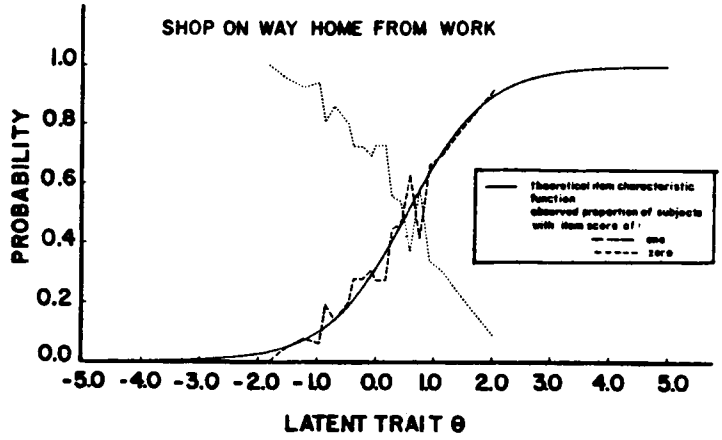

MOVE CLOSER TO WORK
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ

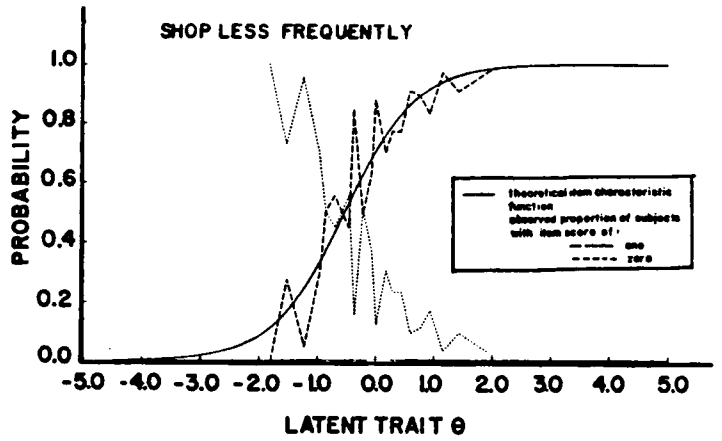

COMBINE SHOP AND OTHER
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ
theoretical item characteristic function
observed proportion of subjects with item score of
one
zero

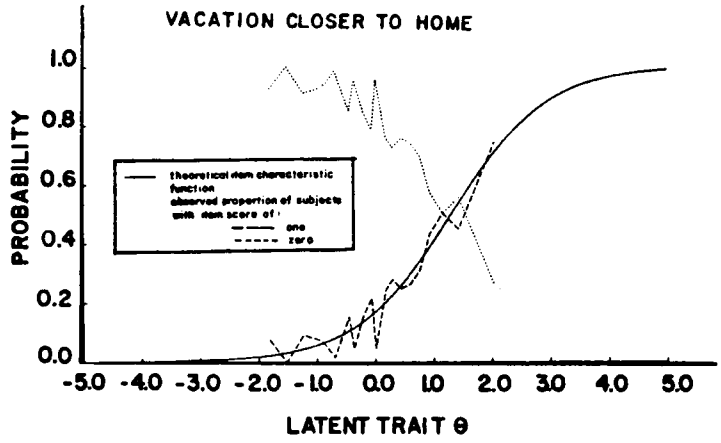

SHOP CLOSER TO HOME
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ
theoretical item characteristic function
observed proportion of subjects with item score of
one
zero

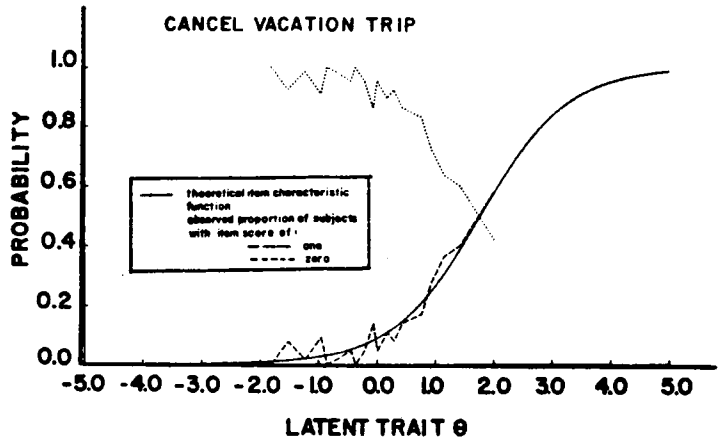

ELIMINATE USE OF RECREATION
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ
theoretical item characteristic function
observed proportion of subjects with item score of
one
zero

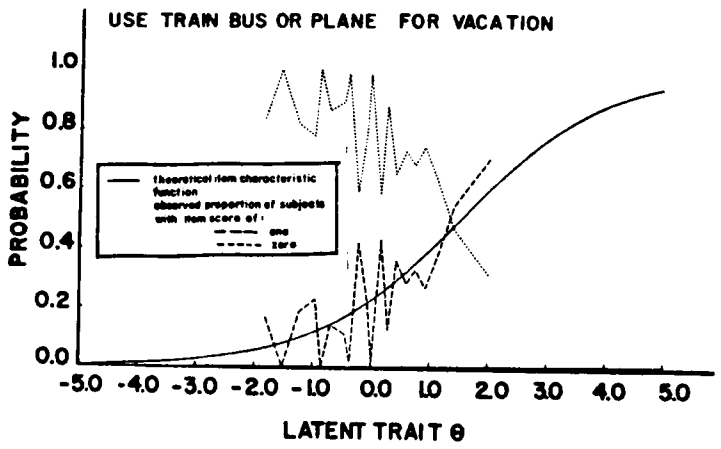

JOB CLOSER TO HOME
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ

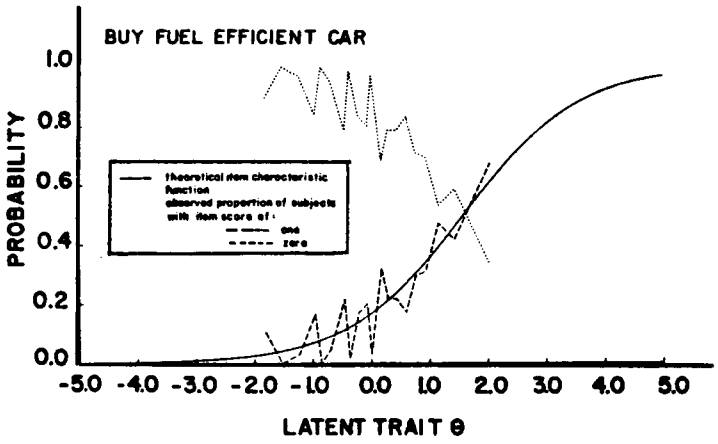

SHOP ON WAY HOME FROM WORK
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ

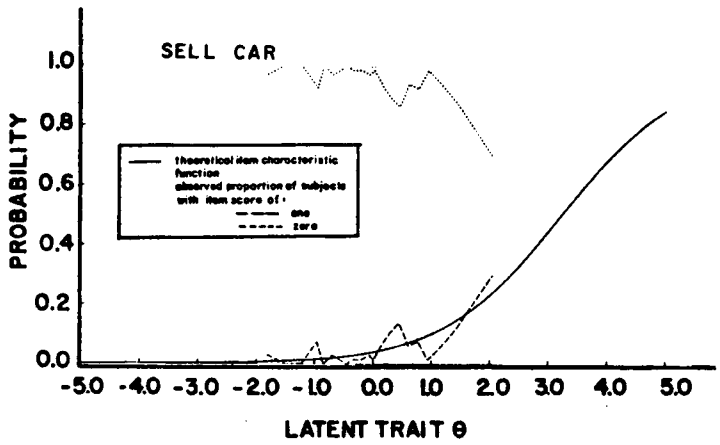

SHOP LESS FREQUENTLY
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ

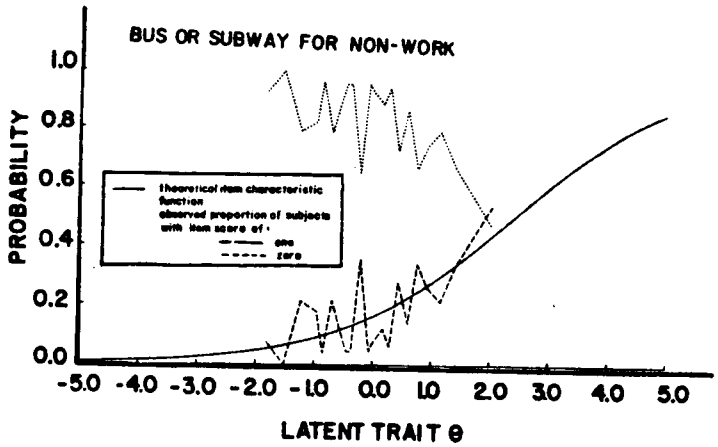

VACATION CLOSER TO HOME
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ


CANCEL VACATION TRIP
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ


USE TRAIN BUS OR PLANE FOR VACATION
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ


BUY FUEL EFFICIENT CAR
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ


SELL CAR
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ


BUS OR SUBWAY FOR NON-WORK
PROBABILITY
1.0
0.8
0.6
0.4
0.2
0.0
-5.0 -4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0 5.0
LATENT TRAIT θ

VITA

Jeffrey William Trombly was born in Chazy, New York, on November 1, 1954. He attended elementary schools in that town and was graduated from Chazy Central Rural School in June 1972. The following September he entered the State University of New York College of Arts and Sciences at Plattsburgh, and in June 1976 he received a Bachelor of Arts Degree with a major in Geography. In the fall of 1976 he entered The University of Tennessee, Knoxville, and began study toward the Master of Science in Planning Degree. This degree was awarded in June 1978.

After working for three years for the Capital District Transportation Committee in Albany, New York, and The New York State Department of Transportation, Mr. Trombly re-entered the Graduate School of The University of Tennessee, Knoxville, in September 1981. He received the Doctor of Philosophy Degree with a major in Civil Engineering in March 1985.

The author is a member of Phi Kappa Phi, the American Planning Association, the Transportation Research Board, and the Institute of Transportation Engineers. Mr. Trombly will be employed by the New York State Department of Transportation after graduation.