INCORPORATING STATED PREFERENCE DATA IN TRAVEL DEMAND ANALYSIS

by

TAKAYUKI MORIKAWA

Master of Science in Transportation

Kogaku-syushi (M. Eng.), Transportation Engineering
Kyoto University, Kyoto, Japan (1983)

Kogaku-shi (B. Eng.), Transportation Engineering
Kyoto University, Kyoto, Japan (1981)

Submitted to the Department of Civil Engineering
in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy
at the
Massachusetts Institute of Technology
June 1989

© Takayuki Morikawa 1989. All rights reserved

The author hereby grants to MIT permission to reproduce and
to distribute copies of this thesis document in whole or in part.

Signature of Author ____________________________
Department of Civil Engineering
May 19, 1989

Certified by ____________________________ Moshe E. Ben-Akiva
Professor, Civil Engineering
Thesis Supervisor

Accepted by ____________________________ Ole S. Madsen
Chairman, Departmental Committee on Graduate Students

JUN 01 1989
INCORPORATING STATED PREFERENCE DATA IN TRAVEL DEMAND ANALYSIS

by

TAKAYUKI MORIKAWA

Submitted to the Department of Civil Engineering on May 19, 1989
in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

ABSTRACT

Traditionally estimation of travel demand models has relied on revealed preference (RP) data which are based on actual market choices made in observable situations. However, there has been an awakening interest in the use of stated preference (SP) data that are respondent's expressed preferences to hypothetical choice situations. For model estimation, SP surveys have the following advantages: i) choice set can be prespecified; ii) range of attributes can be extended; iii) multicollinearity among attributes can be avoided; iv) attributes that are not easily quantified, such as safety, reliability, and availability, can be incorporated; and v) attributes are measured without errors. Despite these advantages, SP data have not been widely used in model estimation due to uncertain reliability of the elicited preference information under hypothetical scenarios. The objective of this thesis is to develop methodologies for incorporating SP data in travel demand analysis with explicit considerations of the unknown reliability of SP data.

In order to exploit advantages of both RP and SP data, a methodology for combined estimation with RP and SP data is developed. In the proposed method the model that represents market behavior or RP data and the model that generates SP data are simultaneously estimated from RP and SP data. Accuracy of parameter estimates in the RP model can be gained by sharing some of its parameters with the SP model, while potential biases and errors specific to SP data are explicitly considered in the SP model. Two case studies demonstrated effectiveness and practicality of this methodology. For example, the parameter of travel time that was insignificant in the separate estimation of the RP model was successfully estimated by this combined estimation method.

Alternative estimators of discrete choice models for ranking data are also proposed. These explicitly consider the potential reliability problems in ranking data. An empirical analysis showed that decision protocols and reliability of preference information significantly vary according to the level of the rank. An strategy of pooling information from different depth of ranking is proposed.

A methodology for incorporating other psychometric data such as perceptual and attitudinal indicators is introduced. The proposed approach uses psychometric data as observable indicators of underlying latent variables and constructs a system of linear structural and measurement equations to identify these latent variables. The integrated system, composed of these linear equations and a discrete choice model, can be practically estimated by a two stage method. An empirical analysis supported the methodology.

Lastly, alternative strategies for explicit modeling of taste heterogeneity are compared using simulated ranking data by the Monte Carlo method.

Thesis Supervisor: Dr. Moshe E. Ben-Akiva
Title: Professor of Civil Engineering
Acknowledgments

This research was partially supported by the U.S. Department of Transportation, Urban Mass Transportation Administration (UMTA) grant to MIT "Ridership Attraction of Rail Compared with Bus." It was also partially supported by the UPS Doctoral Fellowship awarded through the Center for Transportation Studies (CTS) of MIT.

I would like to express my gratitude to all the people who have contributed towards the completion of this thesis.

Throughout my stay at MIT, I have been privileged to have Professor Moshe Ben-Akiva as my academic advisor and thesis supervisor. As the chairman of my doctoral committee he has provided me with valuable suggestions and insightful guidance. Professors Daniel McFadden and Nigel Wilson have also served as members of the committee and given me precious advice. Professor McFadden gave me very insightful suggestions from his extensive experiences in this research field. Professor Wilson's extensive comments on the draft copy significantly improved the clarity of this thesis.

Fumiaki Shiroishi and Denis Bolduc helped me with computation for Chapters 4 and 8, respectively.

I am particularly indebted to Professors Yoshimi Nagao and Katsuhiko Kuroda of Kyoto University for giving me the opportunity to study at MIT.

I learned many things, academic and non-academic, from my office mates, Bruno Bocca and Rohit Ramaswamy, who share the common research interest. Discussions with Rohit were extremely useful in conducting empirical analyses.

My sincere appreciation is expressed to the following people and organizations for providing me with survey data which are essential to all the case studies presented in the thesis: the Massachusetts Bay Transportation Authority (MBTA) and the Central Transportation Planning Staff (CTPS) for the Arborway Corridor Data (Chapter 4), Professor Katsutoshi Ohta and Mr. Satoshi Suzuki and their transportation group of the University of Tokyo for the Yokohama Data (Chapter 5), and Messrs. Albert Bouma, Mark Bradley and Erik Kroes for the Nijmegen - Randstad Corridor Data (Chapters 6 and 7).

Last, but never least, I would like to extend my sincere gratitude to my wife, Akiko. I attribute my physical and mental well being to her continual support. Without the understanding, encouragement, and enjoyment she has given me, this thesis would not exist.
# Table of Contents

Abstract .................................................................................................................. 2
Acknowledgments ................................................................................................... 3
Table of Contents ................................................................................................. 4
List of Figures ....................................................................................................... 8
List of Tables ....................................................................................................... 9

Chapter 1  Introduction ....................................................................................... 11

Chapter 2  Stated Preference Techniques: A Review ............................................ 19
  2.1 Definition of "Reliability" of Stated Preference Techniques .............................. 19
  2.2 Reliability of Conjoint Analysis ..................................................................... 21
    2.2.1 Definition of Conjoint Analysis ............................................................... 21
    2.2.2 Functional Form of Preference Models ................................................. 23
    2.2.3 Data Collection Methods ...................................................................... 24
    2.2.4 Construction of a Stimulus Set for the Full-Profile Method .................. 25
    2.2.5 Presentation of Stimuli ......................................................................... 27
    2.2.6 Measurement Scale for the Dependent Variable .................................. 28
    2.2.7 Estimation Methods ............................................................................. 29
    2.2.8 Other Studies Concerning Reliability of Conjoint Analysis .................. 29
  2.3 Stated Preference Techniques in Transportation Research ........................... 30
  2.4 Summary ..................................................................................................... 33

Chapter 3  Modeling with Stated Preference Data ................................................. 34
  3.1 Implications for Modeling with Stated Preference Data ............................... 34
    3.1.1 Validity of Stated Preferences ............................................................... 34
3.1.2 Relationship of Stated Preferences and Revealed Preferences .............. 37

3.2 Estimation Methods of Discrete Choice Models with Ranking Data .......... 40
  3.2.1 Expansion of Ranking Data into Choice Data ................................ 40
  3.2.2 Estimation of Multinomial Logit Models with Expanded Data .............. 41
  3.2.3 Alternative Estimators of Logit Models with Ranking Data ................. 42

3.3 Combined Estimation with Revealed Preference and Stated Preference Data ...... 45
  3.3.1 Combining Data from Multiple Data Sources .................................. 45
  3.3.2 Framework of Combined Estimation with RP and SP Data .................. 46
  3.3.3 Estimation Methods ................................................................. 48

3.4 Summary ......................................................................................... 50

Chapter 4 Estimation of Choice Models with Ranking Data ...................... 51
  4.1 Hypotheses Concerning the Reliability of Ranking Data ....................... 51
    4.1.1 General Hypotheses ................................................................ 51
    4.1.2 Description of the Survey Data ............................................. 52
    4.1.3 Tested Hypotheses .................................................................. 53

  4.2 Empirical Results ........................................................................... 54
    4.2.1 Test of H1 .............................................................................. 55
    4.2.2 Test for Equality of Coefficients across Data Sets ...................... 57
    4.2.3 Test of H2 .............................................................................. 60
    4.2.4 Test of H3 .............................................................................. 62
    4.2.5 Test of H4 .............................................................................. 64

  4.3 Discussion and Summary .................................................................. 64

Chapter 5 Combined Estimation with SP and RP Data - Case Study 1 ......... 67
  5.1 Framework for Combined Estimation with RP Data and Stated Intentions ...... 68
  5.2 Description of the Data Used for the Empirical Analysis ...................... 71
5.3 Model Estimation with the Before Data ......................................................... 72
  5.3.1 Estimation with RP Data ........................................................................ 72
  5.3.2 Estimation with SP Data ....................................................................... 76
  5.3.3 Combined Estimation with SP and RP Data .......................................... 77
5.4 Prediction Tests of the Estimated Models with the After Data................ 82
  5.4.1 Model Estimation with the After Data .................................................. 82
  5.4.2 Prediction Test of Before Models .......................................................... 84
5.5. Discussion and Summary .......................................................................... 86

Chapter 6 Combined Estimation with SP and RP Data - Case Study 2 ......... 88
  6.1 Description of the Survey Data ................................................................. 89
  6.2 Model Estimation with Separate Data Set ................................................. 91
    6.2.1 Estimation with RP Data .................................................................... 91
    6.2.2 MLE for Ordered Categorical Dependent Variable .......................... 93
    6.2.3 Estimation with SP Data .................................................................... 94
  6.3 Combined Estimation with RP and SP Data .............................................. 98
    6.3.1 Estimation with RP and SP1 Data ....................................................... 98
    6.3.2 Estimation with RP and SP2 Data ....................................................... 102
    6.3.3 Estimation with RP, SP1, and SP2 Data ......... ................................. 106
  6.4 Model Comparisons ................................................................................. 110
  6.5 Discussion and Summary ......................................................................... 116

Chapter 7 Incorporating Latent Variables in Choice Models ......................... 119
  7.1 Framework for Incorporating Perceptual and Attitudinal Data .................. 119
    7.1.1 General Framework for Choice Models .......................................... 119
    7.1.2 Example of Applying the General Framework to Mode Choice Models . 121
  7.2 Discrete Choice Model with Latent Explanatory Variables ...................... 125
Chapter 8  Simulation Study of Models with Taste Heterogeneity ............. 140
  8.1  Explicit Modeling of Taste Heterogeneity from Stated Preference Data 141
    8.1.1  Individual Specific Models and Market Segmentation .................. 141
    8.1.2  Taste Variation Models ........................................... 142
  8.2  Simulation Study .................................................................. 143
    8.2.1  Steps in the Monte Carlo Experiment ................................. 144
    8.2.2  Simulation Results ....................................................... 145
  8.3  Discussion and Summary ....................................................... 147

Chapter 9  Conclusions ................................................................. 150

References .................................................................................... 156
Appendix A ..................................................................................... 164
Appendix B ..................................................................................... 168
Appendix C ..................................................................................... 177
Appendix D ..................................................................................... 186
List of Figures

Figure 1.1  Consumer Decision Process................................................................. 12
Figure 1.2  Framework for Analysis of Consumer Behavior ................................. 13
Figure 3.1  The Relationship between Preferences and RP and SP Data .................. 39
Figure 6.1  Utility Function and Threshold Parameters........................................... 94
Figure 7.1  Framework for Incorporating Perceptual and Attitudinal Data .............. 120
Figure 7.2  Framework for Example Model............................................................ 123
Figure 7.3  Framework for the Estimated Model .................................................... 133
Figure 8.1  Change of Average Absolute Values of Bias......................................... 146
Figure 8.2  Change of Average MSE .................................................................. 146
# List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Characteristics of RP and SP Data</td>
<td>16</td>
</tr>
<tr>
<td>4.1</td>
<td>Estimation Results of Models 1 and 2</td>
<td>56</td>
</tr>
<tr>
<td>4.2</td>
<td>Base Models</td>
<td>57</td>
</tr>
<tr>
<td>4.3</td>
<td>Tests for Equality of Individual Coefficients</td>
<td>58</td>
</tr>
<tr>
<td>4.4</td>
<td>Estimation Results of Pooled Models</td>
<td>59</td>
</tr>
<tr>
<td>4.5</td>
<td>Tests for Equality of Coefficients</td>
<td>60</td>
</tr>
<tr>
<td>4.6</td>
<td>Estimation Results under H2</td>
<td>61</td>
</tr>
<tr>
<td>4.7</td>
<td>Tests for Equality of Coefficients under H2</td>
<td>61</td>
</tr>
<tr>
<td>4.8</td>
<td>Estimation Results under H3</td>
<td>63</td>
</tr>
<tr>
<td>4.9</td>
<td>Tests for Equality of Coefficients under H3</td>
<td>63</td>
</tr>
<tr>
<td>4.10</td>
<td>Estimation Results under H4</td>
<td>65</td>
</tr>
<tr>
<td>4.11</td>
<td>Test for Equality of Coefficients under H4</td>
<td>66</td>
</tr>
<tr>
<td>5.1</td>
<td>Comparison of Stated Intention and Actual Behavior</td>
<td>72</td>
</tr>
<tr>
<td>5.2</td>
<td>Estimation Results of RP Model</td>
<td>75</td>
</tr>
<tr>
<td>5.3</td>
<td>Estimation Results of SP Model</td>
<td>77</td>
</tr>
<tr>
<td>5.4</td>
<td>Combined Estimation with SP and RP Data</td>
<td>80</td>
</tr>
<tr>
<td>5.5</td>
<td>Comparison of the Estimation Results with the Before Data</td>
<td>82</td>
</tr>
<tr>
<td>5.6</td>
<td>Estimation Results with the After Data</td>
<td>84</td>
</tr>
<tr>
<td>5.7</td>
<td>Prediction Test of the Estimated Models</td>
<td>85</td>
</tr>
<tr>
<td>6.1</td>
<td>Estimation Results of RP Model</td>
<td>92</td>
</tr>
<tr>
<td>6.2</td>
<td>Estimation Results of SP1 Model</td>
<td>96</td>
</tr>
<tr>
<td>6.3</td>
<td>Estimation Results of SP2 Model</td>
<td>97</td>
</tr>
<tr>
<td>6.4</td>
<td>Joint Estimation Results with RP and SP1 Data</td>
<td>100</td>
</tr>
<tr>
<td>6.5</td>
<td>Two Step Sequential Estimation Results with RP and SP1 Data</td>
<td>101</td>
</tr>
<tr>
<td>6.6</td>
<td>Three Step Sequential Estimation Results with RP and SP1 Data</td>
<td>102</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

Background

This study has been motivated by recent reviews and reformulations of consumer behavior analysis by econometricians and market researchers. (See, for example, the reviews and frameworks in Cambridge Systematics, 1986; McFadden, 1986; and Ben-Akiva and Boccara, 1987.) The underlying idea is that consumer behavior should be analyzed in more detail using subjective attitudinal data if models with better predictive accuracy are to be obtained. This approach contrasts with the traditional treatment of consumer behavior, which regards the consumer as an "optimizing black box."

Adopting the view of most contemporary behavioral market researchers, the consumer decision process can be described as shown in Figure 1.1. In this diagram, ovals refer to unobservable or latent variables, while rectangular boxes represent observable variables. The relationship between the actual attributes of alternatives and observed behavior is represented by three groups of intervening factors: perceptions, attitudes and preferences. Perceptions are consumer's perceived values of attributes of alternatives which are usually influenced by his or her socioeconomic characteristics and market information, while attitudes are his or her perspectives of attributes. Preference is also a latent factor and represents desirability of alternative choices, which is usually expressed by a utility function. Traditionally, the latent factors enclosed by the dashed line have been treated as the black box.
Figure 1.1 Consumer Decision Process
Figure 1.2 Framework for Analysis of Consumer Behavior
Market researchers have attempted to analyze explicitly the latent psychological factors and have relied on various indicators of perceptions, attitudes and preferences as shown in Figure 1.2. Attitudinal and perceptual indicators usually represent the level of satisfaction or importance of attributes on a semantic scale. *Stated preference* (SP) data are collected by presenting hypothetical scenarios to the respondents and asking for their preferences. In contrast to this type of data, measurements based on actual market behavior are termed *revealed preference* (RP) data.

**Analysis of Stated Preference Data**

The analysis of SP data originated in mathematical psychology (Luce and Tukey, 1964), and this type of data and the associated analysis techniques are often called "conjoint measurement" and "conjoint analysis," respectively. Market researchers have been applying conjoint analysis since the early 1970's (e.g., Green and Rao, 1971 and Green and Srinivasan, 1978), with Cattin and Wittink (1982) estimating that about 1,000 commercial applications of conjoint analysis were carried out during the 1970's alone.

In the transportation research field methods of discrete choice analysis have been developed for the analysis of travel demand and applied to various aspects of travel behavior (see, for example, Ben-Akiva and Lerman, 1985). Although this approach has better theoretical properties than conventional aggregate demand models, the observed predictive accuracy is not always satisfactory, due to limitations of the available data. For example, individual travel behavior is usually described by using only travel time and travel cost as service attributes. However, travel behavior such as modal choice and route choice is almost certainly also affected by less quantifiable factors such as reliability, comfort, and safety. Measurement errors in variables and ambiguity in the definition of the individual choice sets are other sources of error in the description of behavior. Thus, a reasonable hypothesis is that data limitations hinder disaggregate behavioral models from exhibiting their theoretical superiority. In this
sense, attitudinal data provide an excellent complementary source of information for estimating travel behavior models because it adds information on intangible attributes and contains more precise descriptions of choice sets and explanatory variables.

One of the reasons that SP data are frequently used in market research is that the experimenter can control the choice scenarios. It implies the following advantages of SP data over RP data:

i) choice set can be prespecified;

ii) range of attributes can be extended;

iii) multicollinearity among attributes can be avoided;

iv) attributes that are not easily quantified, such as safety, reliability, and availability, can be incorporated; and

v) attributes are free from measurement errors.

These advantages enable one to elicit preferences with respect to totally new (non-existing) options. Another important advantage of SP data is that the experimenter can elicit several kinds of reasonable preference indicators such as rankings and ratings, while revealed preferences simply indicate the single preferred alternative. This implies that stated preferences convey a larger amount of information on individual preferences than revealed preferences. On the other hand, RP data have higher face validity than SP data because revealed preferences reflect actual market behavior. In other words, the unknown reliability is the principal and critical weakness of SP data. The advantages and disadvantages of both types of data are summarized in Table 1.1, which clearly shows that SP data and RP data are strongly complementary.
Table 1.1 Characteristics of RP and SP Data

<table>
<thead>
<tr>
<th>Revealed Preference Data</th>
<th>Stated Preference Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Based on actual market behavior</td>
<td>• Based on hypothetical scenarios</td>
</tr>
<tr>
<td>• Choice set is ambiguous</td>
<td>• Choice set is specified</td>
</tr>
<tr>
<td>• Attributes are subject to measurement errors</td>
<td>• Attributes are free from measurement errors but are subject to perception errors</td>
</tr>
<tr>
<td>• Range of attribute level is limited</td>
<td>• Range of attribute level can be extended</td>
</tr>
<tr>
<td>• Attributes may be highly correlated</td>
<td>• Correlation between attributes can be avoided or minimized</td>
</tr>
<tr>
<td>• Difficult to incorporate intangible attributes (e.g., reliability, comfort, safety, availability, security, image, information availability, and privacy)</td>
<td>• Can incorporate intangible attributes</td>
</tr>
<tr>
<td>• Cannot provide direct information on new (non-existing) alternatives</td>
<td>• Can elicit preferences for new (non-existing) alternatives</td>
</tr>
<tr>
<td>• Preference indicator is &quot;choice&quot; (most preferred alternative)</td>
<td>• Can elicit any reasonable preference indicators (e.g., rank, rate, and choice)</td>
</tr>
</tbody>
</table>

The Objective of This Research

In econometrics, because of the uncertain reliability of SP data, the traditional view has been that valid choice data result only from actual choices having been made and, therefore, SP data have rarely been used for quantitative analysis. At the other extreme, SP data have been used for estimation of choice models without any consideration of potential biases in the data.

16
Thus, the research question we face is how to assess the reliability of SP data and how to exploit the advantages of both SP and RP data in an integrated analysis. The objective of this research is to develop methodologies which explicitly consider the reliability of SP data for travel demand analysis. Two methodologies are presented. The first one estimates travel choice models from ranking data with explicit consideration of different reliability and decision protocols at each depth of the ranking. The second methodology combines SP and RP data to exploit the advantages of these complementary data sources. This is done with explicit treatment of the potential biases that may be present in SP data. Empirical studies are conducted using survey data to demonstrate the applicability of the proposed methodologies.

Outline of the Thesis

This thesis is composed of nine chapters. Chapter 2 reviews conjoint analysis and summarizes the results of prior research, focusing on how experimental settings affect the reliability of elicited preferences. Applications of SP techniques in transportation research are also reviewed in this chapter. Methodologies for analyzing SP data are presented in Chapter 3. First, the validity of stated preferences is discussed and then, two methods to incorporate SP data in model estimation are described. Chapter 4 presents a case study in which multinomial logit models are estimated from ranking data with explicit consideration of reliability problems. In Chapter 5, an empirical analysis of combined estimation with SP and RP data is presented, using before-and-after survey data of a new subway line. The before survey includes questions about current travel behavior (revealed preferences) and expressions of intention to switch one's commuting route by using the new subway line in the future (stated preferences). Another empirical application of the combined estimation method is presented in Chapter 6, in which intercity travel mode choice between car and rail is analyzed from revealed and stated preferences. The following two chapters address related research topics concerning utilization of SP data and other psychometric data. Chapter 7 proposes a methodology of incorporating
latent (unobservable) variables into demand modeling, in which the choice model includes latent explanatory variables such as convenience and comfort of travel. These latent variables are estimated by a set of equations which relate observable variables to the latent variables. This methodology can incorporate psychometric data from perceptual and attitudinal questions in a survey as observable indicators. An empirical application is also presented. In Chapter 8 the issue of heterogeneity of individual tastes is discussed. To capture taste heterogeneity among individuals, ranking or rating data can be used to estimate two types of models: individual-specific and parametric taste variation models. Simulation studies are carried out in order to compare the statistical properties of these two types of models. Finally, Chapter 9 summarizes the thesis and presents the conclusions.
Chapter 2

Stated Preference Techniques: A Review

The literature on SP techniques in market research and transportation applications is reviewed in this chapter with a focus on the reliability of SP techniques. In the first section, the concept of "reliability" is clarified while in the following section the reliability issues in conjoint analysis are described. More general issues on conjoint analysis are found in, for instance, Green and Rao (1971), Green and Srinivasan (1978), and Louviere (1988a). Commercial applications of conjoint analysis are reviewed by Cattin and Wittink (1982). The third section reviews applications of SP techniques in transportation research.

2.1 Definition of "Reliability" of Stated Preference Techniques

Green and Srinivasan (1978) were the first to discuss explicitly the reliability issue in conjoint analysis. They introduced three terms that convey distinct but related concepts: reliability, internal validity, and external validity. They define "reliability" as the stability of the experimentees responses to the task and they propose two approaches to its measurement:

1) The "test-retest reliability of the input preference judgments" directly examines the preference judgments (e.g., ranking and rating) of repeated experiments. After the main experiment, a subsample of the respondents are asked to do the same conjoint task against a subset of the original stimulus set.
ii) The "alternate forms method with spaced testing," compares the estimated model parameters. A subsample of the respondents are approached after a period of time and asked to perform the same conjoint task on a different set of stimuli. The second stimulus set ("alternate form") have the same number of alternatives but not the original one. Two estimates of the model parameters are obtained from the two data sets and their correlations are used to measure reliability.

They state that the second method is more rigorous because it takes into account four sources of error: inaccuracies in the preference judgments, variability in the set of constructed stimuli, errors in the estimation procedure, and lack of time stability, while the first method considers only the first source of error.

"Internal validity" can be reported in terms of the correlation between the observed versus estimated values of the preference judgment, or dependent variable. Data for the reliability tests can also be used for the internal validity test with the model parameters estimated from the first set of data used to predict the preference judgments for the second set.

"External validity" can be tested by comparing predictions against actual behavior. They mention that this type of validity is also referred to as "predictive validity," however, in later literature the predictive validity is usually calculated with respect to a "holdout" sample. Thus, most market researchers use the term "predictive validity" for cross-validation within the data rather than for validity against actual market behavior.

A recent review by Bateson et al (1988) identified over 30 studies which reported on the reliability of conjoint analysis. It concluded that it is difficult to make generalizations from this literature because of the plethora of procedures and approaches used as well as disagreement on the definition of reliability. It declares that one of the primary sources of confusion about the reliability of conjoint analysis is the lack of clarity over the definition of "reliability."

This thesis proposes a clearer definition of the "reliability" of stated preferences which includes two different concepts: "validity" and "stability." "Validity" concerns the consistency of stated preferences with actual market behavior and corresponds to "external validity" in
Green and Srinivasan's definition. "Stability" refers to the phenomenon that responses may fluctuate drastically according to the experimental settings and characteristics of the respondent and has a similar meaning to "reliability" in Green and Srinivasan's terminology. Validity and stability are somewhat similar concepts to "bias" and "random error," in statistics.

### 2.2 Reliability of Conjoint Analysis

#### 2.2.1 Definition of Conjoint Analysis

Since Luce and Tukey published the seminal paper in 1964, conjoint analysis has been subject to both theoretical development and practical application. Earlier theoretical development is well documented in Green and Rao (1971) and Green and Srinivasan (1978). Here, following the broad definition by Green and Srinivasan (1978), we use the term "conjoint analysis" to refer to any decompositional method that estimates the structure of consumer's preferences using his or her overall evaluation of a hypothetical alternative represented by a set of attributes.

The decompositional approach is to estimate a set of part-worths for the attributes from some subjective evaluation of the alternative, given a composition rule (e.g., an additive rule). On the other hand, the compositional approach such as expectancy-value models directly asks the respondent the values of part-worths. It is generally accepted that, for the purpose of prediction of consumer behavior, the decompositional approach is superior to the compositional approach because it is easier, and consequently more reliable, for the respondent to judge the preference of an alternative than to give an importance weight for each attribute.

Another important aspect of conjoint analysis is that the analysis is usually carried out at the individual level. In other words, individual-specific values of coefficient are estimated from
conjoint measurements. This is one of the key differences from ordinary econometric models estimated from revealed preferences.

Conjoint analysis is closely related to two other types of measurement developed in mathematical psychology: clinical judgments and functional measurement. Both these types of measurement involve subjective "rating" of a profile of attribute values and are analyzed by decompositional methods such as least squares and analysis of variance (ANOVA). Conjoint measurement sometimes refers strictly to "rank-order" or "paired-comparison" in order to distinguish it from these other types of measurement. However, in this thesis, conjoint measurement includes "rating" as well as the other preference indicators such as ranking and choice.

"Stated preference" is a broader concept than conjoint measurement since it includes any type of preference judgments in any hypothetical situation. Namely, any preference indications other than actual market behavior (revealed preference) can be called stated preferences. For example, consider the following questions: "A subway line connecting your town and the downtown area is planned. Would you use the subway for commuting downtown?" and "Would you buy this product if the price were 20% higher?" This type of data, called "stated behavioral intentions," is obviously stated preferences, but neither can be called conjoint measurements. Another type of SP data which does not fall into the conjoint measurement category is "matching data," or sometimes referred to as "transfer price data." To collect this type of data, the respondent is presented with two alternatives. For one alternative, one attribute, typically price or cost, is left blank and the respondent is asked to fill in the value of the attribute that would make him or her indifferent between the two alternatives. Accordingly, conjoint measurements can be defined as stated preferences under a fully hypothetical choice situation that reflects a well-established experimental design. Hence, conjoint analysis is included in the set of SP techniques.

The following sections summarize the conjoint analysis methodology, focusing on how each stage of the analysis affects reliability. Descriptions of the theoretical development are
largely based on Green and Srinivasan (1978). To avoid confusion, note that different terminology is used by econometricians and market researchers: so an "alternative" is called a "stimulus" or "profile" by market researchers; and a "coefficient" or "parameter" of an explanatory variable is referred to as a "part-worth" or "importance weight."

2.2.2 Functional Form of Preference Models

A general functional form of additive preference models is expressed as:

\[ P_{in} = \sum_{k=1}^{K} f_{kn}(x_{ik}) , \]  

(2.1)

where \( P_{in} \) is preference measurement of alternative \( i \) for individual \( n \), \( f_{kn} \) is the function denoting the part-worts of attribute \( k \) for individual \( n \), and \( x_{ik} \) is the value of attribute \( k \) of alternative \( i \). For instance, the ideal-point model presumes that the part-wort is negatively related to the squared distance from \( x_{ik} \) to the individual's ideal point \( z_{ik} \):

\[ P_{in} = \sum_{k=1}^{K} w_{kn}(x_{ik} - z_{ik})^2 , \]  

(2.2)

where

\[ f_{kn}(x_{ik}) = w_{kn}(x_{ik}^2 - 2x_{ik}z_{ik} + z_{ik}^2) , \]  

(2.3)

and \( w_{kn} \) are unknown weight parameters. The simplest, yet most frequently used, functional form is linear, as expressed below:

\[ P_{in} = \sum_{k=1}^{K} w_{kn}x_{ik} , \]  

(2.4)

where

\[ f_{kn}(x_{ik}) = w_{kn}x_{ik} . \]  

(2.5)

If interaction effects are expected a priori, these could be captured by adding pseudo-attributes of the form \( x_{ik}x_{il} \). Adding the disturbance term enables one to estimate the values of
\( w_{kn} \)'s by statistical methods such as linear regression, analysis of variance, and discrete choice models.

### 2.2.3 Data Collection Methods

Two basic data collection methods are used in conjoint analysis: the two-factor-at-a-time procedure and the full-profile approach. In the former procedure, also referred to as the "trade-off (matrix) analysis" (Johnson, 1974), combinations of the different levels of two factors, or attributes, are presented in a matrix form. The respondent is asked to rank the various combinations of pairs of factor levels from the most preferred to the least preferred.

The two-factor-at-a-time procedure is simple to apply and reduces information overload for the respondent. However, each stimulus is so simplified that the respondent may feel a lack of realism and may not be able to provide information on the pure trade-off between the two factors. For instance, given a trade-off matrix of travel time and seat comfort, the respondent may associate better seat comfort with higher fare and consider an implicit fare, consciously or unconsciously, in the response. Another drawback of this procedure is the large number of required judgments. For instance, with six factors each at four levels, the respondent could be asked to fill out 15 tables, each consisting of 16 cells.

The full-profile approach uses the complete set of factors and requires fewer judgments to be made by the respondent although each judgment is more complex. In this approach, the respondent is typically given a deck of cards. Each card defines values of all the attributes of an alternative. This approach can employ either rank orders or ratings, while the two-factor-at-a-time procedure provides only a set of rank orders. Here, "ranking" refers to the order of preferences of stimuli while "rating" requires some scale of preferences for each stimulus such as "very good" and "desirable." This topic will be discussed in more detail later. The main argument that seems to favor the full-profile approach is that it gives a more realistic description of alternatives than the two-factor-at-a-time procedure. But it is assumed that too many
attributes will decrease the respondent's ability to make comparisons and so in many cases the number of attributes is limited to between 3 and 6.

A number of empirical studies have compared the two approaches to data collection and reached conflicting conclusions. Montgomery et al (1977), in a study of job choice by MBA students, found that the two-factor-at-a-time procedure yielded higher predictive validity than the full-profile approach. Similarly, Alpert et al (1978) obtained better goodness-of-fit from the two-factor-at-a-time procedure in a study of commuters' choice of travel mode. Jain et al (1978) found that the two methods yielded approximately the same level of cross-validity for models of choosing checking accounts offered by various banks. Also, Oppedijk van Veen and Beazley (1977) concluded that the utilities determined by the two methods were roughly the same in the choice context of a durable good product class.

However, Segal (1982) compared both methods by the goodness-of-fit and preservation of a priori assumed monotonicity of coefficients and showed that the full-profile approach is superior overall.

2.2.4 Construction of a Stimulus Set for the Full-Profile Method

Since the coefficients are estimated at an individual level, the number of alternatives (profiles) presented to each respondent depends on the number of the coefficients to be estimated. Since too many alternatives will confuse the respondent and distract his or her interest, there is a matter of trade-off between the reliability of estimators and the validity of the raw data. It is usually difficult for the respondent to evaluate more than 30 alternatives. Malhotra (1982) investigated the effects of the number of profiles and the number of attributes. He concluded that using from 15 to 20 (or to 25) alternatives does not significantly affect the standard error of the parameters, especially if the number of attributes is small and that, therefore, conjoint analysis is a fairly robust procedure for assessing individual preferences.
In deciding the range of attribute values and interattribute correlation (e.g., between travel time and travel cost, or between horsepower and gas mileage of cars), two conflicting considerations arise. The use of stimulus descriptions similar to those that currently exist (similar in terms of ranges of attribute levels and interattribute correlations) will increase believability and consequently raise the validity of the preference judgments. If we make the ranges for attribute levels much larger than reality and/or decrease the magnitude of interattribute correlations to zero (as is implied by orthogonal designs), we may decrease believability and consequently lose the validity of the response. But orthogonal designs and/or larger ranges for attribute values have the advantage of improving the accuracy of the parameter estimates for a given level of validity of the preference judgments. Thus, the extreme strategy of using descriptions similar to those that exist has the disadvantage of the loss of accuracy of estimates, while the other extreme strategy of using an orthogonal design and/or ranges for attribute values much larger than reality has the disadvantage of decreasing the validity of the preference judgments.

Green and Srinivasan (1978) recommend that the ranges be made larger than reality, but not so large as to be unbelievable. Further, they recommend making interattribute correlations smaller (in absolute value) than the correlations that exist in real stimuli, but not to make them so small as to be unbelievable.

Since the full factorial design inflates the number of stimuli (e.g., with three attributes at three levels each and two attributes at two levels each, the total number of possible descriptions is $3^3 \times 2^2 = 108$), fractional factorial designs are used to reduce the number of combinations to a manageable size while at the same time maintaining orthogonality (Green, 1974). Various kinds of fractional factorial designs are discussed by Green et al (1978) and Louviere (1988a).

Reibstein et al (1988) evaluated comparative reliability scores for all the combinations of data collection methods (ranking, trade-off matrices, and paired comparison), stimulus sets, attribute sets, and types of products. They employed the test-retest procedure to obtain two sets of parameter estimates and calculated Chow's F-test statistics as the reliability measure.
They showed that the data collection procedure has a significant impact on the reliability score, independent of the stimulus and attribute sets.

2.2.5 Presentation of Stimuli

Three basic methods to present stimuli in the full-profile approach have been used: verbal description, paragraph description, and pictorial representation.

In a typical verbal description task, the respondent is given a deck of cards each of which describes a stimulus with the values of attributes and is asked either to rank order them or to rate them on a certain scale. The main advantages of this procedure are its simplicity and efficiency. The two-factor-at-a-time procedure has primarily used the verbal description.

The written paragraph description method (Hauser and Urban, 1977) has the advantage of providing a more realistic and complete description of the stimuli. The major disadvantage is that it limits the total number of descriptions to a small number, so that parameter estimates are likely to be inaccurate when estimated at the individual level.

Pictorial representation, which uses various kinds of visual aids or three dimensional models, provides several advantages:

i) Information overload is reduced;

ii) Higher homogeneity of perceptions is obtained across respondents;

iii) The task itself is more interesting and less tiring; and

iv) The alternatives are more realistic.

The primary disadvantages of this method are the increased cost and time in preparing the alternative descriptions and also the danger of conveying different information from that which the researcher intends.

Since factors such as learning, boredom, fatigue, and anchoring to earlier tasks are inherent in any experiment, the order of the tasks must be randomized to reduce the bias. Acito (1977) verified that the coefficient estimate of an attribute was affected by the order or position of the
attribute on the stimulus card. Eberts and Koeppel (1978), in their trade-off analysis for the choice of work schedule options, found that concise questions and randomized orders of alternatives are necessary due to substantial fatigue bias.

McFadden (1986) listed other factors influencing the accuracy of the responses, including presence of human observers or consultants, recall of past responses, and verbal versus key response (i.e., the respondent uses keys or buttons for response).

2.2.6 Measurement Scale for the Dependent Variable

Various alternatives for defining a measurement scale for the dependent variable can be roughly classified as nonmetric (e.g., "ranking," "choice," and "paired-comparison") and metric (e.g., "rating" and "matching").

In the "ranking" task, alternatives are ranked by the order of preference. As a special case of "ranking," the respondent simply chooses the most preferable alternative, which is called the "choice" task. The "rating" task requires the respondent to locate each alternative on a metric scale in accordance with preference. The scale usually corresponds to semantic desirability or choice probability (e.g., likelihood of purchase). In the "matching" task, or sometimes called "transfer price" task, the respondent is to state the amount of change in terms of an attribute (e.g., price) which makes the two alternatives equivalent. Conceptually, it measures the willingness-to-pay for the preferred alternative in terms of the reference attribute. Pessemier et al (1971) referred to this task as the "dollar metric" approach.

Although metric measurements contain more information than nonmetric ones, they are usually less reliable because of the respondent's limited ability to quantify preferences. Since it is extremely difficult for the respondent to quantify his or her preference in terms of choice probabilities, the elicited information from such rating data could be highly unreliable. It is also difficult to rank a large number of alternatives, so the reliability of information is likely to
increase at lower levels of the ranking. The reliability of estimators from ranking data is extensively discussed in the next chapter.

2.2.7 Estimation Methods

Desirable methods for estimating the coefficients depend on the type of dependent variable. For nonmetric dependent variables (e.g., ranks), estimation methods such as probabilistic discrete choice models (e.g., logit and probit models) (Ben-Akiva and Lerman, 1985) and non-statistical methods such as LINMAP (Pekelman and Sen, 1974) and MONANOVA (Kruskal, 1965) have been widely used. For metric dependent variables, linear regression analysis techniques such as OLS and GLS are most commonly used. Green and Srinivasan (1978) state that the estimation methods do not seem to differ very much in terms of their predictive validity.

2.2.8 Other Studies Concerning Reliability of Conjoint Analysis

Acito and Jain (1980) investigated how the number of profiles shown to a respondent affected the goodness-of-fit, the predictive validity for "holdout" stimuli, and the sign of parameters. They found that the reliability of conjoint measurement is significantly correlated with the educational level of the respondent.

Akaah and Korgaonkar (1983) compared the predictive validity of self-explicated, traditional conjoint, and hybrid conjoint models. The self-explicated model uses the information obtained by asking the respondent for the part-worth of each attribute. The hybrid model combines this information with conjoint measurement. The results indicate that the conjoint and hybrid models performed equally well and outperformed the self-explicated model. An expository review of hybrid models is given by Green (1984) and another comparison of methodological approaches was conducted by Leigh et al (1984). They
compared the test-retest reliability and the predictive validity of several alternative conjoint analysis techniques with that of self-explicated weights but failed to demonstrate greater reliability and validity for the conjoint analysis.

Srinivasan et al (1983) showed that predictive power can be improved by imposing constraints on parameters based on a priori knowledge of the ordering of part worths of attributes.

2.3 Stated Preference Techniques in Transportation Research

Transportation researchers have shown interest in using psychometric data for estimation of travel demand models. In a large number of studies, attitudinal and perceptual indicators were used as explanatory variables and/or relative importance weights (e.g., Recker and Golob, 1976; Johnson, 1978; Young and Morris, 1981; and Suzuki et al, 1986). Benjamin and Sen (1982) documented general discussions on incorporating psychometric data in transportation research. This section focuses on the utilization of SP data in travel demand analysis.

Although the SP techniques have been used in marketing and psychometric research for a long time, there were relatively few reports using those techniques in the transportation literature before 1980. One principal reason for this seems to be that "forecasting" has been emphasized in transportation demand analysis rather than consumer's preferences or trade-offs. More specifically, since the validity of data, or the consistency with actual behavior, is of great importance in predicting future behavior, transportation researchers have mostly relied on RP data. In the 1980's the increasing popularity of the disaggregate behavioral models provided an incentive to analyze travel behavior in more detail and there has been an awakening interest in using SP data for travel demand analysis. For example, the Journal of Transport Economic and Policy published a special issue titled "Stated Preference Methods in Transport Research" in January, 1988.
This special issue, including seven papers mostly from Europe, covers a wide range of topics in applying SP methods to transportation problems and reviews previous studies. Among others, Kroes and Sheldon (1988) and Louviere (1988b) give comprehensive overviews of SP methods and summaries of past works. Validity of SP data is treated by Fowkes and Wardman (1988), Wardman (1988), and Bradley (1988) as well as by the aforementioned two papers. All of them agree that the most important issue in using SP data is their external validity, and that such data showed "satisfactory" validity in the previous studies. However, they do not seem to be quite confident about those statements and call for more empirical works. Bates (1988) discussed econometric issues in estimating SP models and stated that special cares in the error terms of SP models are necessary because of potential heteroscedasticity and correlations. Hensher et al (1988) proposed to use SP data to explore the true functional form of the utility function. Experimental designs of SP surveys are discussed in detail in Fowkes and Wardman and Louviere.

Other important studies concerning SP techniques in the transportation field include the one by Lerman and Louviere (1978) who used SP rating data in order to identify the functional form of the utility expression for residential location choices and subsequently estimated the coefficients of the utility function using RP data. Their work is characterized by separately using SP and RP data for different purposes. The SP data are used only for the identification of the functional form because of their limited validity.

Bates (1983), in a general discussion on SP techniques, pointed out the importance of validity checks of SP data. He suggested that the most convincing way to assess the validity of SP data would be to compare two models of the same individual estimated from SP and RP data, respectively. In order to make this comparison, the researcher must conduct a survey which contains both SP and RP questions.

Louviere et al (1981) carried out such a comparison by conducting four surveys, at two cities at two different points in time, that included both SP and RP questions on travel mode choice. They estimated separate choice models from SP and RP data and compared those
models on the basis of predictive ability and consistency of the parameter estimates over time and space. They demonstrated that the SP models had a predictive ability equal to the RP models and that the parameter estimates of the SP models were also temporally and spatially stable and consistent with those of RP models. However, this result could be ascribed to the fact that most of the coefficient estimates of the RP models had large standard errors with presumably important attributes such as fare and operating cost turning out to be insignificant from the RP data, due to the multicollinearity and limited variation.

Intended and actual use of a new transit service was analyzed using before and after data by Couture and Dooley (1981) and Suzuki et al (1986). The before data include the intention of using the new transit service (stated preferences) and the current travel mode choice (revealed preferences), while the after data reveal the actual usage of the new transit service (revealed preferences). Both studies found that the prior intentions significantly overstated the actual use. Couture and Dooley found that respondents may have disregarded situational constraints (e.g., closeness to the transit stop) in stating their intended use. These two studies differ from the aforementioned work by Louviere et al in the following respect; the stated preferences analyzed by Couture and Dooley and Suzuki et al were the intended usage of a planned major change in the transportation system (the introduction of a new transit service), while Louviere et al set up purely hypothetical experiments using the conjoint technique.

Morisugi et al (1981) and Bovy and Bradley (1985) applied SP models to route choice behavior. The latter paper compares the estimation results from SP data of ranking, satisfaction rating, and choice probability rating. They showed that the model estimated from ranking data was most comparable to the RP model, and that choice probability rating was least reliable. In an application to intercity travel mode choice, Bates and Roberts (1983) used ranking and choice probability rating data and noted the different reliability of ranking information according to its depth.

Another response format for stated preferences, the "matching" task, in which the respondent states the amount of change in terms of an attribute (e.g., price), was used by Gunn
(1984) to investigate the value of time. Comparing with estimation results from RP data, he concluded that the respondents used the same decision process for stated and revealed preferences. Subjective value of time was also estimated by Horowitz (1981) from pairwise rating data.

Ben-Akiva and Boccara (1987) proposed an integrated framework for travel behavior analysis which incorporates various psychometric measurements including SP data and situational constraints. The framework for analysis of consumer behavior previously illustrated in Figure 1.2 follows their paper.

2.4 Summary

In market research the reliability of SP techniques, or conjoint analysis, has been extensively discussed. However, as Bateson et al (1988) suggested, one of the primary sources of confusion about the reliability of conjoint analysis is the lack of clarity over the definition of "reliability." We introduced two facets of reliability of stated preferences: "validity" and "stability." Most researchers realize that the crucial issue is the consistency of stated preferences with actual market behavior, that is, the validity of SP data. Nevertheless, in market research, most of the quantitative analyses on the "predictive validity" of conjoint analysis deal with the ability to reproduce conjoint measurements. Transportation researchers also have shown increasing interest in applying SP techniques to various choice contexts. Reflecting the fact that transportation practitioners are more "conservative" in the sense that they do not accept SP data at face value, they seem to be more concerned with the validity of SP data than market researchers. Although some studies which compared SP and RP models quantitatively have supported the reliability of SP data, persuasive conclusions have not yet been obtained.
Chapter 3

Modeling with Stated Preference Data

This chapter develops methodologies for modeling with SP data. First, the validity of
stated preferences is addressed by examining how the hypothetical setting affects the response
in comparison with preferences under actual market situations. Two methodologies which
explicitly consider the unknown validity of SP data are presented: estimation of choice models
from ranking data and combined estimation with SP and RP data.

3.1 Implications for Modeling with Stated Preference Data

3.1.1 Validity of Stated Preferences

Two concepts of validity and stability of stated preferences were introduced in the previous
chapter. It is obvious that the more serious concern in using SP data is the validity of the
elicited preference information. If stated preferences are significantly biased in an unknown
way, or are barely related to actual market behavior, they are of little value for predicting
demand, no matter how consistently demand models estimated from SP data reproduce the
stated preferences. Apparently, this is the main reason that many researchers and practitioners
hesitate to use SP data. As reviewed in the previous chapter, most research on the reliability of
demand models using SP data (e.g., conjoint analysis) has emphasized the "stability" or
"internal validity" of the analysis. This subsection discusses the "validity" of SP data, specifically, how stated preferences may differ from actual behavior.

(1) Decision protocols for stated and revealed preferences

A hard-core problem in collecting SP data is the indifference of the respondent to the experimental task. Since a hypothetical scenario does not generally affect the welfare of the respondent (unlike actual market behavior), the respondent may be so uninterested and careless that he or she might not make a rational decision. The "prominence hypothesis" is one explanation of such irrational responses. It assumes that the respondent may consider only one, or at most a few, attributes in evaluating the alternatives rather than considering the trade-offs among all the attributes given. Under this hypothesis, an alternative which is nearly dominated by other alternatives could be preferred. Consider, for instance, a product which incurs slightly less initial expense than other products but requires greater operating costs that eliminate the initial cost saving after a few uses. Suppose also that this product is substantially less attractive in terms of other attributes than other products. If the respondent takes into account only the initial expense, that product, which is obviously inferior to others for reasonable consumers, may be chosen.

Sometimes a hypothetical scenario does affect (or at least the respondent so believes) his or her welfare, but it affects him or her in a different way from the real market decision. The "policy-response bias" (Bonsall, 1985) arises when the respondent believes that he or she will benefit by responding in a certain way. For example, people may intentionally overstate their use of a planned new transportation system in order to promote its realization. Couture and Dooley (1981) and Suzuki et al (1986) verified this kind of overstatement through an analysis of before and after data.

On the other hand, the respondent may understate the use of a completely new type of product or service when some kind of "inertia" of the current actual choice influences the
response. Neslin (1976) defined "preference inertia" as a hesitation towards innovations. He found that respondents did not prefer an innovation to an existing product although the innovation was evaluated better in terms of its attributes. The "justification bias" is another potential problem. The respondent may want to justify past behavior and may respond in that way even to a hypothetical scenario. For instance, car commuters may deliberately increase the utility level of the car mode in their mind and give biased responses.

(2) Imperfect description of alternatives

In order to reduce the complexity of the task for respondents, the number of attributes describing each alternative is limited compared with the set of attributes which may be considered in real market situations. Then, when an SP model is used to forecast, some relevant variables might be omitted, resulting in inaccurate or biased forecasts.

When the name of an alternative is associated by the respondent with a certain image, there can be a discrepancy between the experimenter's intended alternative and the respondent's perception. Specifically, the respondent might take into consideration some attributes not described in the experiment but associated with the name of the alternative. The effects of those associated attributes will be compounded into the coefficient estimate of the alternative specific dummy variable. If those effects are correlated with any of the included attributes in some way, it causes other estimated parameters to be biased. Consider, for example, a mode choice between bus and rail. Some respondents who have a negative image of the bus mode might consider undesirable attributes of buses which are not presented in the description of the mode.
(3) Omission of situational constraints

In responding to hypothetical scenarios, the respondent may, consciously or unconsciously, ignore his or her situational constraints. In the study by Couture and Dooley (1981), none of the variables regarding situational constraints had significant coefficient estimates from SP data which had been collected before the introduction of a new transit, however these were the most important explanatory variables of the RP data collected after the introduction. The authors claimed this to be the primary cause of overprediction of usage of the new transit service.

On the other hand, SP methods are often used to elicit information on pure trade-offs among the presented attributes and the situational constraints should be ignored in this case. But it is, of course, necessary to introduce the situational constraints in forecasting.

3.1.2 Relationship of Stated Preferences and Revealed Preferences

As described thus far, SP data have the following weaknesses, compared with RP data, for inferring latent preferences.

i) Stated preferences are clearly related to the underlying preferences but do not necessarily replicate actual behavior.

ii) Stated preferences could be governed by a different decision making protocol from that of revealed preferences.

iii) Stated preferences could have bias and error structures different from RP data.

Thus, stated preferences may not have the same causal relationship to underlying latent preferences as revealed preferences do but clearly indicate some aspects of latent preferences. The linkages between stated preferences, revealed preferences, and the underlying latent preferences can be described by the two different models. The RP model represents actual market behavior. The SP model may not be capable of predicting market behavior but can be used to help identify the parameters of the RP model by providing additional information on
underlying preferences. As depicted in Figure 3.1, RP and SP data are generated from the underlying preferences by two different models. These linkages imply that SP data might not be valid for prediction but could help in the identification and estimation of the latent preferences that determine actual behavior. On the other hand, although RP data represent actual behavior it may not contain significant information to precisely identify the underlying preferences. Thus, RP and SP data, which have complementary characteristics, can be used to identify underlying latent preferences in this modeling framework.

The rest of this chapter presents two methodologies which incorporate SP data in demand analysis, explicitly considering the aforementioned weaknesses of SP data.
Figure 3.1 The Relationship between Preferences and RP and SP Data
3.2 Estimation Methods of Discrete Choice Models with Ranking Data

3.2.1 Expansion of Ranking Data into Choice Data

Discrete choice models such as logit and probit have been successfully applied to various choice contexts (Ben-Akiva and Lerman, 1985) including choices in hypothetical scenarios, namely, stated preferences. If the experimental task is to make a "choice," one can simply apply a discrete choice model to the SP data. However, estimating discrete choice models from ranking data, which is a common form of eliciting preference judgments in SP surveys, requires identifying the correct relation between ranking and choice probabilities. When choice behavior, by which each rank is determined, satisfies Luce's Choice Axiom (Luce, 1959), the probability of ranking is easily linked with choice probabilities and the multinomial logit (MNL) structure is the appropriate model, as described below. Empirical studies of this methodology are found in Beggs et al (1981) and Chapman and Staelin (1982). Alternative approaches, which do not require the MNL assumption but are computationally more burdensome, are presented in Falmagne (1978) and Barbara and Pattanaik (1986).

If the choice behavior underlying the ranking task follows Luce's Choice Axiom, the ranking of $J$ alternatives is equivalent to the following sequence of independent choice tasks: The alternative ranked first is chosen over all the other alternatives, the second ranked alternative is preferred to all others except the first ranked, and so on. The Luce and Suppes Ranking Choice Theorem (Luce and Suppes, 1965) states this decomposition of the ranking task in terms of choice probabilities:

$$P(1,2,\ldots,J) = P(1\{1,2,\ldots,J\}) \cdot P(2\{2,3,\ldots,J\}) \cdots P(J-1\{J-1,J\})$$

$$= \prod_{j=1}^{J-1} P(j\{j,\ldots,J\}) ,$$

(3.1)

where $P(1,2,\ldots,J)$ is the probability of observing the rank order of alternative 1 being preferred to alternative 2 being preferred to alternative 3, and so on, and $P(1\{1,2,\ldots,J\})$ is the
probability of alternative 1 being chosen from the set of alternatives \{1,...,J\}. This equation is equivalent to saying that the event of \( J \) ranked alternatives is composed of \( J-1 \) statistically independent choice events. Thus, the log-likelihood of a ranking observation is equal to the sum of \( J-1 \) log-likelihoods of choices as follows:

\[
\ln P(1,2,...,J) = \sum_{j=1}^{J-1} \ln P(j;\{J\}) .
\]

(3.2)

Thus, accepting the Rank Choice Theorem implies that an observation of \( J \) ranked alternatives can be "expanded" (or "exploded") into \( J-1 \) statistically independent choice observations. Note that the \( m \)th choice is made from \( J-m+1 \) alternatives.

### 3.2.2 Estimation of Multinomial Logit Models with Expanded Data

Luce's axiom also implies that the choice probabilities follow the structure of the logit model. This means that all the choice probabilities in (3.1) can be derived from the same logit model. However, in the following analysis we will maintain (3.1) but allow the choice probabilities to differ as a result of biases in SP data.

The choice probability, \( P(j;\{J\}) \), may be represented as a random utility model, as follows. Express the utility function of alternative \( j \) for individual \( n \) by:

\[
U_{jn} = \beta'x_{jn} + \epsilon_{jn} , \quad j=j,...,J ,
\]

(3.3)

where \( x_{jn} \) is a vector of attributes, \( \beta \) is a vector of unknown parameters, and \( \epsilon_{jn} \) is an independently and identically distributed Gumbel disturbance term. The probability that alternative \( j \) is preferred to alternatives \( j+1,...,J \) is given by the following multinomial logit model:

\[
P(j;\{J\}) = P(U_{jn} \geq U_{in} , \quad i=j,...,J)
= \frac{\exp(\mu(\beta'x_{jn}))}{\sum_{i=j}^{J} \exp(\mu(\beta'x_{in}))} ,
\]

(3.4)
where \( \mu \) is a positive scale parameter which is inversely proportional to the standard deviation of the disturbance terms and is usually normalized to be equal to one. Thus, if we assume that all the choice probabilities follow the same logit model, the probability of observing the rank order of alternative 1 being preferred to alternative 2, alternative 2 preferred to alternative 3, and so on, is given by the product of \( J-1 \) ordinary logit likelihood functions:

\[
P_n(1,2,...,J|\beta) = \prod_{j=1}^{J-1} \frac{\exp(\beta'x_{jn})}{\sum_{i=j}^{J} \exp(\beta'x_{in})}.
\]

For \( N \) observations of ranking data, the log-likelihood function for a logit model is:

\[
L(\beta) = \sum_{j=1}^{J-1} \sum_{n=1}^{N} \left[ \beta'x_{jn} - \ln \sum_{i=j}^{J} \exp(\beta'x_{in}) \right].
\]

This is how a multinomial logit (MNL) model is estimated with expanded choice data from ranking observations. Such an MNL model is often referred to as a "rank logit model" or an "exploded logit model" because its log-likelihood function is a summation of ordinary MNL log-likelihoods over all the decompositions of the expanded data.

### 3.2.3 Alternative Estimators of Logit Models with Ranking Data

The estimator described above assumes that the respondent accomplishes the ranking task by assigning a utility value to each alternative and then ranking the alternatives in a descending order of utility values. This may not be a realistic assumption in the context of an SP survey. For example, a respondent may find it more natural and therefore easier to choose the most preferable alternative than to assign ranks to inferior alternatives. Thus, a respondent may pay less attention to the task of ranking inferior alternatives. Consequently, the lower ranks may be less reliable than the higher ranks and it may be reasonable to assume that the reliability of ranking information decreases with decreasing rank. In other situations, the response may be
governed by different decision protocols according to the level of the rank. For instance, the top rank may be governed by the "policy-response bias" or by some inertia effect addressed in the previous section.

Chapman and Staelin (1982) referred to the problem of "increasing noise" and discussed how "deeply" the ranking data should be expanded. The "depth" is equal to the number of choice data sets created from the ranking data. They considered the trade-off between lower sampling variance versus greater bias with increasing depth and proposed an estimator based on the top $P$ ranks, obtained by maximizing the following log-likelihood function:

$$L_P(\beta) = \sum_{j=1}^{P} \sum_{n=1}^{N} \ln P_n(j_i \{j, \ldots, J\}), \quad 1 \leq P \leq J - 1.$$  

(3.7)

They suggested that one should plot a goodness-of-fit measure by gradually increasing the value of $P$ and stop expanding the data when the goodness-of-fit measure drops significantly. As a more formal test, they conducted likelihood ratio tests for equality of two coefficient vectors estimated for $P$ and $P+1$. This test is described in detail in Chapter 4.

The expanded logit estimator for $P > 1$ may not be valid because increasing noise or different decision protocols imply that the values of the coefficients are dependent on the depth of the rank. Since separate estimates of the coefficients, $\beta$, can be obtained by using any subset of choice data sets instead of pooling the expanded data, one can statistically test the rank dependence c.f. $\beta$.

Suppose that the utility function of an alternative depends on the depth of the rank. The utility function of alternative $i$ for the $m$th rank is expressed by:

$$U_{in}^m = \beta^m x_{in} + \epsilon_{in}^m, \quad m = 1, \ldots, J - 1 \text{ and } i = m, \ldots, J.$$  

(3.8)

This is a more general form of equation (3.3). $\beta^m$ can be estimated from the $m$th choice data set, $m = 1, \ldots, J - 1$. The log-likelihood of the $J - 1$ separate models is given by:

$$L(\beta^1, \ldots, \beta^{J-1}) = \sum_{j=1}^{J-1} \sum_{n=1}^{N} \left[ \beta_j^j x_{jn} - \ln \sum_{i=j}^{J} \exp(\beta_i^j x_{in}) \right].$$  

(3.9)
The null hypothesis, $\beta^1=\beta^2=\ldots=\beta^{l-1}$, can be tested by the classical likelihood ratio statistic based on the difference between (3.6) and (3.9).

The hypothesis that lower ranks are noisier than higher ranks is expressed by $\text{Var}(\varepsilon^m)<\text{Var}(\varepsilon^l)$ for $m<l$, or $\text{Var}(\varepsilon^m)=\mu_m^2\text{Var}(\varepsilon^l)$ where $0<\mu_m<1$. For $m<l$, the parameters of the $l$th choice data set, $\beta^l$, would be smaller in absolute values than those of the $m$th data set, $\beta^m$, because under this hypothesis:

$$\beta^l = \mu_m \beta^m. \hspace{1cm} (3.10)$$

In other words, normalizing the scale parameter of the $m$th data set to one requires a scale parameter for the $l$th data set, $\mu_m$, which is between zero and one.

Thus, under this hypothesis, the ranking data may be pooled by introducing rank dependent scale parameters as in equation (3.10) and jointly estimating the $\beta$'s and the $\mu$'s. The log-likelihood for this estimator is:

$$I_c(\beta, \mu) = \sum_{j=1}^{J-1} \sum_{n=1}^{N} \left[ \sum_{j=1}^{J} \mu_j \beta' x_{jn} - \ln \sum_{i=j}^{J} \exp(\mu_j \beta' x_{in}) \right], \hspace{1cm} (3.11)$$

where one of the $\mu$'s, for example, $\mu_1$, is set to one. Note that in this case the log-likelihood function is no longer a simple summation of ordinary MNL log-likelihoods. The null hypothesis,

$$\beta^1=\beta, \hspace{0.5cm} \beta^2=\mu_2 \beta, \hspace{0.5cm} \beta^3=\mu_3 \beta, \hspace{0.5cm} \ldots, \hspace{0.5cm} \beta^{l-1}=\mu_{l-1} \beta, \hspace{1cm} (3.12)$$

can be tested by a likelihood ratio test comparing the value of the restricted log-likelihood function given by (3.11) with the unrestricted log-likelihood in (3.9).

Hausman and Ruud (1987) estimated logit models with ranking data and included the scale parameters described above. Their results indicated the existence of an increasing noise pattern and hence they rejected an ordinary rank logit model (with an equal scale parameter for all ranks) in favor of a "heteroscedastic" rank logit model.

If different decision protocols are governing the ranking process at various dcspths, the estimates of $\beta$'s from different choice data sets will show significant fluctuations which are not
represented by a single scale parameter. Thus, it is essential to compare the coefficient estimates from different data sets and detect significant differences among depths of the ranking data in an attempt to infer a representative decision protocol for each depth of the ranking.

3.3 Combined Estimation with Revealed Preference and Stated Preference Data

3.3.1 Combining Data from Multiple Data Sources

Model estimation with multiple sources of data is known as "mixed estimation" in econometrics (Judge et al, 1982). Ben-Akiva (1987) presented a general framework for combining multiple data sources to correct sampling bias contained in a particular data source. In transportation research, data from different sources have been combined in the estimation of origin-destination (O-D) trip tables (e.g., Van Zuylen and Willumsen, 1980; Ben-Akiva, Macke, and Hsu, 1985; and Ben-Akiva and Morikawa, 1989).

These applications conform to the following framework. The population is stratified and unknown parameters of each stratum can be obtained from some source of data, which is called a direct measurement. The source of the direct measurement can be disaggregate data, the same measurement in the past, the fitted values of some behavioral model, etc. But the direct measurement is likely to be subject to large sampling error and bias such as nonresponse bias and sampling bias. An indirect measurement provides aggregate information on the unknown parameters of multiple population strata. In other words, an individual measurement is an observation of a transformation of multiple unknown parameters (e.g., a linear combination of the parameters of several strata). The sources of the indirect measurement include census data, traffic counts, ticket sales, etc., data with large sample sizes and little sampling bias. The idea is to alter the unknown parameters obtained from the direct measurement so that they fit the indirect measurement as well as possible. If the indirect measurement is assumed to be error-
free, it becomes a constrained estimation problem. If the indirect measurement as well as the
direct measurement is assumed to be stochastic, the estimation relies on a technique such as
joint maximum likelihood estimation. Thus, the underlying idea of using multiple data sources
is that different data sources have different levels of accuracy and different types of bias and
that combining them could exploit their advantages and overcome their disadvantages.

3.3.2 Framework of Combined Estimation with RP and SP Data

Needless to say, the main purpose of demand models is to predict future demand, especially in
conjunction with major changes in service levels (or product attributes) or introduction of an
entirely new service (or product). In transportation demand analysis, for example, these
contexts include fare increases, schedule changes, and introduction of a new transportation
mode. While one can easily set up hypothetical scenarios that are similar to the future
situations and ask for preference statements about them, the validity of SP data is questionable.
Specifically, there might be bias in stated preferences and the accuracy of SP data might
significantly differ from that of the RP data.

By combining RP and SP data the problem may be solved by exploiting the advantages of
both data sources and improving the accuracy of parameter estimates. This data combination
methodology has the same underlying idea as the application to the O-D table estimation,
although the model formulation is substantially different.

This research proposes a mixed estimator from SP and RP data. For the sake of
illustration, assume that the context is a choice of travel mode with two alternatives. Let \( \mathbf{x} \) and
\( \mathbf{w} \) denote vectors of explanatory variables (e.g., decision maker characteristics and alternative
attributes) in terms of the differences of the two alternatives. Suppose that the trade-offs
among certain attributes in \( \mathbf{x} \) cannot be estimated precisely with RP data alone but that SP
information on \( \mathbf{x} \) is available through a conjoint experiment or some other appropriate SP
questions. For instance, high correlation between travel cost and comfort level in RP data may
yield insignificant parameter estimates of either one of or both attributes. However, SP surveys with a design of low or no correlation between these attributes will provide clearer information on their trade-off.

Now we set up two types of model: RP and SP models. The RP model represents market behavior by some appropriate structure (e.g., random utility model with binary choices for this example), while SP response is modeled by the SP model. As discussed earlier, although SP response might not be valid for forecasting market behavior due to unknown bias and error, it often contains some useful information on trade-offs of attributes and the amount of information per individual is usually greater than in RP data. Thus, the role of SP data is illustrated by the following framework:

**The RP model**

- structural equation (utility function):

\[
U = \beta^{r}x^{rp} + \alpha^{r}w^{rp} + \varepsilon \tag{3.13}
\]

- measurement equation (choice indicator):

\[
d^{rp} = \begin{cases} 
1: & \text{if } U \geq 0 \\
0: & \text{otherwise} 
\end{cases} \tag{3.14}
\]

**The SP model**

- structural equation (utility function):

\[
\tilde{U} = \beta^{s}x^{sp} + \gamma^{s}z^{sp} + \nu \tag{3.15}
\]

- measurement equation (preference indicator, e.g., choice):

\[
d^{sp} = \begin{cases} 
1: & \text{if } \tilde{U} \geq 0 \\
0: & \text{otherwise} 
\end{cases} \tag{3.16}
\]

In the above equations, \(\alpha, \beta\) and \(\gamma\) are unknown parameters and superscripts "RP" and "SP" denote that corresponding data are obtainable from RP and SP data, respectively. Bias factors inherent to stated preferences are represented by \(z\) with the corresponding coefficients
\( \gamma \). For example, a variable representing the actual choice, if such information is available, included in \( z \) may capture the effect of justification bias. Sharing \( \beta \) in both models implies that trade-offs among attributes included in \( x \) are the same in both actual market behavior and the SP tasks. The random components of the utility functions, \( \epsilon \) and \( \nu \), are assumed to be independently distributed with zero means and the level of noise in the data sources is represented by the variance of the disturbance terms, \( \epsilon \) and \( \nu \). If RP and SP data have different noise levels, this can be expressed by:

\[
Var(\epsilon) = \mu^2 Var(\nu),
\]  
(3.17)

and if SP data contain more random noise than RP data, \( \mu \) will lie between 0 and 1.

\( \mu \) is also known to represent the "scale" of the model coefficients. That is, for instance, if the scale of the logit RP model is one, the RP and SP models are:

\[
P(d^{\text{RP}}=1) = \frac{1}{1+\exp(-\beta'x^{\text{RP}}-\alpha'w^{\text{RP}})},
\]
(3.18)

and,

\[
P(d^{\text{SP}}=1) = \frac{1}{1+\exp(\mu(-\beta'x^{\text{SP}}-\gamma'z^{\text{SP}}))}.
\]
(3.19)

3.3.3 Estimation Methods

The principle of combined estimation is that accuracy of parameter estimates can be gained by sharing some parameters (i.e., \( \beta \)) between the RP and SP models while the bias and random error specific to SP data are reflected by \( \gamma \) and \( \mu \), respectively. All of these parameters \( \alpha, \beta, \gamma \) and \( \mu \) are estimated either by i) joint MLE, using RP and SP data simultaneously, or ii) sequential MLE, using RP and SP data one after another.
(1) Joint estimation method

The joint estimators of $\alpha$, $\beta$, $\gamma$ and $\mu$ are obtained by maximizing the joint log-likelihood:

$$L(\alpha, \beta, \gamma, \mu) = \sum_{n=1}^{N_{RP}} d_{n}^{RP} \log \left[ F\left( \beta' x_{n}^{RP} + \alpha' w_{n}^{RP} \right) \right] + \sum_{n=1}^{N_{SP}} d_{n}^{SP} \log \left[ F\left( \mu(\beta' x_{n}^{SP} + \gamma' z_{n}^{SP}) \right) \right],$$

(3.20)

where $N_{RP}$ and $N_{SP}$ are the numbers of observations of RP and SP data sets, respectively, and $F$ is a cumulative density function of the random utility component such as:

$$F(\beta' x^{RP} + \alpha' w^{RP}) = \frac{1}{1 + \exp(\beta' x^{RP} - \alpha' w^{RP})},$$

(3.21)

for a binary logit model, and,

$$F(\beta' x_{n}^{RP} + \alpha' w_{n}^{RP}) = \Phi(\beta' x_{n}^{RP} + \alpha' w_{n}^{SP}),$$

(3.22)

for a binary probit model.

If one can assume statistical independence of $\varepsilon$ and $\nu$, or equivalently:

$$P(d^{RP}, d^{SP}) = P(d^{RP}) P(d^{SP}),$$

(3.21)

then joint estimation yields statistically consistent and efficient estimators. If this assumption does not hold, the joint estimators are still consistent but not fully efficient and the standard MLE variances are incorrect (see Amemiya, 1985).

The joint estimation requires using a general MLE program and programming the joint likelihood function for the model (as opposed to model specific MLE programs, e.g., MNL).

(2) Sequential estimation method

An alternative method to estimate all the parameters calibrates the RP and SP models sequentially. The advantage of this method is that one can use ordinary logit or probit estimation softwares. The sequential estimators are statistically consistent but not efficient.
**Step 1:**

Estimate the SP model (3.15) from SP data to obtain $\hat{\mu}$ and $\hat{\gamma}$. Form the fitted value $\hat{V} = \hat{\mu} \beta' x^{RP}$.

**Step 2:**

Estimate the following RP model with the fitted value $\hat{V}$ to obtain $\hat{\lambda}$ and $\hat{\alpha}$:

$$U = \lambda \hat{V} + \alpha' w^{RP} + \epsilon.$$  

(3.22)

Calculate $\hat{\mu} = 1/\hat{\lambda}$, $\hat{\beta} = \hat{\mu} / \hat{\mu}$, and $\hat{\gamma} = \hat{\mu} / \hat{\mu}$.

Accuracy of estimates $\hat{\alpha}$, $\hat{\beta}$ and $\hat{\gamma}$ can be improved by the following additional step.

**Step 3:**

Multiply $x^{SP}$ and $z^{SP}$ by $\hat{\mu}$ to obtain a modified SP data set. Pool the RP data and the modified SP data and then jointly estimate the two models jointly to obtain $\hat{\alpha}$, $\hat{\beta}$ and $\hat{\gamma}$.

**3.4 Summary**

The "validity" issues of SP data are discussed first. The principal factor of invalid stated preferences is the difference in decision protocols governing the SP task and actual market behavior. Considering the weaknesses of SP data in inferring actual behavior, we capture the relationship of stated preferences and the underlying latent preferences by a separate SP model. Two methodologies which incorporate SP data in demand analysis, explicitly considering the weaknesses of SP data, are proposed: estimation of discrete choice models from ranking data and combined estimation with SP and RP data.

An application of the rank logit model with actual data is demonstrated in the next chapter. Case studies of the combined estimation with SP and RP data are presented in Chapters 5 and 6.
Chapter 4

Estimation of Choice Models with Ranking Data

A theory for estimating discrete choice models from ranking data was described in Section 3.2. The objective of this chapter is to investigate the reliability of SP data through an empirical analysis of ranking data. Statistical tests of the following hypotheses concerning the reliability of SP data are performed:

i) The response to an SP survey is influenced by the actual market behavior;

ii) Each rank may be governed by a different decision protocol; and

iii) The magnitude of the random effects differ among the "depths" of ranking data.

The results of empirical analyses are presented in the second section. The last section contains concluding remarks.

4.1 Hypotheses Concerning the Reliability of Ranking Data

4.1.1 General Hypotheses

In previous studies discrete choice models have been estimated from ranking data using the methods described above. However, these studies did not explicitly consider the unknown reliability of the preference information elicited by this particular response format. This section postulates general hypotheses concerning the reliability of ranking data which should be taken
into account when using such data for model estimation. In the later section, these hypotheses are specified to the ranking data analyzed in this chapter.

The first hypothesis can be applied to any type of SP data:

H1: The response to an SP survey is influenced by actual market behavior.

This general hypothesis includes a variety of phenomena such as justification bias, inertia effects, and biased perceptions. The next two hypotheses are specific to ranking data.

H2: Each rank (e.g., first, second, etc.) may be governed by a different decision protocol.

For example, the first rank may be affected by a policy response bias or justification bias, while the rest of the ranks are based on trade-offs among attributes. Alternatively, the ranking, including or excluding the first rank, may be governed by a simple lexicographical decision rule — rank the alternatives on the basis of the most "important" attributes and resolve ties by using the second most "important" attribute, and so on.

H3: The magnitude of the random effects varies among "depths" of ranking data.

Typically, the first rank contains the least random noise and the noise level increases as the rank decreases because it is easier for the respondent to rank the most preferable alternative than to rank less preferable ones. In some situations, the least preferable alternative is also easier to identify and consequently the information on the last rank has less noise than middle ranks.

The natural extension of the above hypotheses yields the following hypothesis:

H4: Each ranking process may involve the compound effect of hypotheses H1, H2 and H3.

4.1.2 Description of the Survey Data

The SP data analyzed in this chapter were collected for a metropolitan transit agency as part of a study on the relative preferences for bus and light rail (or street car) services in Boston's
southwest corridor. The survey questions were designed to elicit passengers' preferences toward bus-rail transfers at specific points.

The questionnaire presented four transit service alternatives for the corridor, three of which included distinct transfer points between bus and rail, while the other alternative involved through service by rail without a transfer. Six sets of numerical values were specified for three attributes: wait time, travel time reliability, and in-vehicle ride time. Wait time by mode was specified on all the six types, while three included values of travel time reliability (i.e., for how many trips out of 20 you are more than 10 minutes late), and in-vehicle ride time was given in the other three types. Note that all the six versions of the questionnaires had the same design for the transfer points. A sample questionnaire is shown in Appendix A.

The questionnaires were distributed to travelers in buses and trolleys. Each participant was given one of the six versions of the questionnaire and was asked to rank the four alternatives. 840 usable questionnaires were returned.

4.1.3 Tested Hypotheses

The general hypotheses concerning the reliability of ranking data presented in 4.1.1 are now specified for the ranking data described above.

H1: The response to the SP questionnaire is influenced by the current actual choice. Namely, the perceived value of an attribute and/or its importance may depend on the current choice.

It is natural for the respondents to refer to the specific origin and destination of their most typical trip when they evaluate the alternatives that were presented to them. The data set includes information on the origin and the destination of the trip made on the day of the survey and that trip seems to be the typical daily trip for most respondents because work and school trips comprise 86.3% of the total trips. To test this hypothesis, the following dummy variables are specified:
\[ BUS_i = \begin{cases} 
1: & \text{if the current trip is made solely by bus (without transfer) with alternative } i \\ 
0: & \text{otherwise} 
\end{cases} \]

\[ RAIL_i = \begin{cases} 
1: & \text{if the current trip is made solely by rail (without transfer) with alternative } i \\ 
0: & \text{otherwise} 
\end{cases} \]

H1 implies that the dummy variables, \( BUS_i \) and \( RAIL_i \), have a significant effect on choice. This hypothesis also affects the perceived value of wait time. Since the wait time given is stratified by mode, the respondent may take into account only the part of the wait time which is relevant to the current trip. This value is labelled as the "perceived wait time," while the "total wait time" is the simple sum of wait times for bus and rail. It should be noted, however, that the coefficients of the variables that capture H1 may be biased if the utility function is not correctly specified. This may occur when the random utility components are correlated with the current choice.

H2: Each rank (i.e., first, second, and third) may be governed by a different decision protocol. More specifically, the first rank is principally governed by the most prominent attribute, i.e., the transfer point.

This hypothesis implies that choice models for different ranks have different coefficients for the attributes that reflect the transfer point.

H3: The magnitude of the random effects differs among the "depths" of ranking data. More specifically, the first rank is the most reliable and the third is the least reliable.

This hypothesis implies "depth" dependent scale parameters. Then, finally,

H4: Each ranking process may involve the compound effect of hypotheses H1, H2 and H3.

### 4.2 Empirical Results

The ranking data were expanded into three choice data sets: Data Set 1 (the first rank is chosen from all four alternatives), Data Set 2 (the second rank is chosen from the three alternatives left after eliminating the first ranked alternative), and Data Set 3 (the third rank is chosen from the
two alternatives left after eliminating the first and second ranked alternatives). The number of usable observations is 826 for Data Set 1, 815 for Data Set 2, and 805 for Data Set 3. Alternatives 1, 3 and 4 have a distinct transfer point from bus to rail (or vice versa) while alternative 2 consists of rail service without a transfer. Hence, three constants specific to alternatives 1, 3 and 4 represent dummy variables for specific transfer points. Wait time and ride time are given in minutes. As described in the previous section, the total wait time is the sum of wait time at both bus stop and rail station while the perceived wait time is the part of the total wait time relevant to the current trip. Travel time reliability is given in the number of trips out of 20 which are more than 10 minutes late.

4.2.1 Test of H1

Two models were estimated for every data set. Model 1 has the following seven explanatory variables:

1. transfer point 1 dummy (specific to alternative 1)
2. transfer point 3 dummy (specific to alternative 3)
3. transfer point 4 dummy (specific to alternative 4)
4. total wait time
5. perceived wait time
6. ride time
7. travel time reliability

Model 2 adds to Model 1 the bus and rail dummy variables which were defined in the previous section. The estimation results are shown in Table 4.1.

Model 2 shows that the bus and rail dummies have significantly positive coefficients for all the data sets, which strongly supports H1. In other words, the respondents are concerned with the need to transfer and the mode they would have to use for their current trip. The parameter estimates of the total and perceived wait time variables have small t-statistics due to
multicollinearity. Estimated models with either one of these wait time variables did not make a significant difference. This inconclusiveness may be ascribed to the insignificance of these wait time variables in all the models. Without a strict statistical justification, however, the perceived wait time will be used in the subsequent analyses. Thus, the models shown in Table 4.2 are used as the base models.

### Table 4.1 Estimation Results of Models 1 and 2

<table>
<thead>
<tr>
<th>Variable</th>
<th>Data set 1 Model 1</th>
<th>Data set 1 Model 2</th>
<th>Data set 2 Model 1</th>
<th>Data set 2 Model 2</th>
<th>Data set 3 Model 1</th>
<th>Data set 3 Model 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dummy 1</td>
<td>-0.952</td>
<td>-0.713</td>
<td>-0.392</td>
<td>0.0593</td>
<td>0.653</td>
<td>0.839</td>
</tr>
<tr>
<td></td>
<td>(-7.85)</td>
<td>(-4.41)</td>
<td>(-3.17)</td>
<td>(0.34)</td>
<td>(4.15)</td>
<td>(3.63)</td>
</tr>
<tr>
<td>Dummy 3</td>
<td>-0.425</td>
<td>-0.432</td>
<td>0.0945</td>
<td>0.532</td>
<td>0.215</td>
<td>0.249</td>
</tr>
<tr>
<td></td>
<td>(-4.55)</td>
<td>(-2.24)</td>
<td>(0.77)</td>
<td>(2.37)</td>
<td>(1.34)</td>
<td>(0.85)</td>
</tr>
<tr>
<td>Dummy 4</td>
<td>-1.52</td>
<td>-1.22</td>
<td>-0.751</td>
<td>-0.322</td>
<td>-0.135</td>
<td>0.0793</td>
</tr>
<tr>
<td></td>
<td>(-10.5)</td>
<td>(-6.67)</td>
<td>(-5.53)</td>
<td>(-1.87)</td>
<td>(-0.87)</td>
<td>(0.36)</td>
</tr>
<tr>
<td>Total wait time</td>
<td>-0.0140</td>
<td>-0.0540</td>
<td>0.0201</td>
<td>-0.0255</td>
<td>0.0327</td>
<td>-0.00142</td>
</tr>
<tr>
<td></td>
<td>(-0.32)</td>
<td>(-1.18)</td>
<td>(0.40)</td>
<td>(-0.50)</td>
<td>(0.53)</td>
<td>(-0.02)</td>
</tr>
<tr>
<td>Perceived wait time</td>
<td>-0.0753</td>
<td>-0.0272</td>
<td>-0.0823</td>
<td>-0.0324</td>
<td>0.0103</td>
<td>0.0543</td>
</tr>
<tr>
<td></td>
<td>(-2.26)</td>
<td>(-0.77)</td>
<td>(-2.36)</td>
<td>(-0.87)</td>
<td>(0.23)</td>
<td>(1.12)</td>
</tr>
<tr>
<td>Ride time</td>
<td>-0.0205</td>
<td>-0.0246</td>
<td>-0.0381</td>
<td>-0.0420</td>
<td>-0.0341</td>
<td>-0.0354</td>
</tr>
<tr>
<td></td>
<td>(-1.12)</td>
<td>(-1.34)</td>
<td>(-1.91)</td>
<td>(-2.07)</td>
<td>(-1.35)</td>
<td>(-1.38)</td>
</tr>
<tr>
<td>Travel time reliability</td>
<td>-0.166</td>
<td>-0.165</td>
<td>-0.0965</td>
<td>-0.100</td>
<td>0.0609</td>
<td>0.0629</td>
</tr>
<tr>
<td></td>
<td>(-3.82)</td>
<td>(-3.75)</td>
<td>(-2.01)</td>
<td>(-2.08)</td>
<td>(1.05)</td>
<td>(1.07)</td>
</tr>
<tr>
<td>Bus dummy</td>
<td>0.761</td>
<td></td>
<td>0.369</td>
<td></td>
<td>0.518</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(5.48)</td>
<td></td>
<td>(2.36)</td>
<td></td>
<td>(2.50)</td>
<td></td>
</tr>
<tr>
<td>Rail dummy</td>
<td>0.491</td>
<td></td>
<td>0.623</td>
<td></td>
<td>0.293</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.96)</td>
<td></td>
<td>(3.96)</td>
<td></td>
<td>(1.37)</td>
<td></td>
</tr>
<tr>
<td>( L(0) )</td>
<td>-1145.08</td>
<td>-1145.08</td>
<td>-895.37</td>
<td>-895.37</td>
<td>-558.39</td>
<td>-558.39</td>
</tr>
<tr>
<td>( L(\hat{\beta}) )</td>
<td>-1057.54</td>
<td>-1041.20</td>
<td>-861.21</td>
<td>-851.50</td>
<td>-531.52</td>
<td>-527.63</td>
</tr>
<tr>
<td>( \rho^2 )</td>
<td>0.0764</td>
<td>0.0907</td>
<td>0.0382</td>
<td>0.0490</td>
<td>0.0481</td>
<td>0.0551</td>
</tr>
<tr>
<td>( \overline{\rho}^2 )</td>
<td>0.0703</td>
<td>0.0829</td>
<td>0.0305</td>
<td>0.0391</td>
<td>0.0358</td>
<td>0.0393</td>
</tr>
<tr>
<td>( N )</td>
<td>826</td>
<td>826</td>
<td>815</td>
<td>315</td>
<td>805</td>
<td>805</td>
</tr>
</tbody>
</table>
Table 4.2 Base Models (t-statistics in parentheses)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Data set 1</th>
<th>Data set 2</th>
<th>Data set 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dummy 1</td>
<td>-0.767 (-4.94)</td>
<td>0.0398 (0.24)</td>
<td>0.838 (3.78)</td>
</tr>
<tr>
<td>Dummy 3</td>
<td>-0.486 (-2.58)</td>
<td>0.512 (2.31)</td>
<td>0.248 (0.88)</td>
</tr>
<tr>
<td>Dummy 4</td>
<td>-1.25 (-6.82)</td>
<td>-0.327 (-1.90)</td>
<td>0.0787 (0.36)</td>
</tr>
<tr>
<td>Perceived wait time</td>
<td>-0.0512 (-1.81)</td>
<td>-0.0432 (-1.44)</td>
<td>0.0537 (1.41)</td>
</tr>
<tr>
<td>Ride time</td>
<td>-0.0118 (-0.79)</td>
<td>-0.0361 (-2.24)</td>
<td>-0.0350 (-1.64)</td>
</tr>
<tr>
<td>Travel time reliability</td>
<td>-0.134 (-3.78)</td>
<td>-0.0840 (-2.38)</td>
<td>0.0638 (1.37)</td>
</tr>
<tr>
<td>Bus dummy</td>
<td>0.743 (5.40)</td>
<td>0.361 (2.33)</td>
<td>0.518 (2.50)</td>
</tr>
<tr>
<td>Rail dummy</td>
<td>0.452 (2.77)</td>
<td>0.609 (3.93)</td>
<td>0.292 (1.41)</td>
</tr>
</tbody>
</table>

L(0)  -1145.08 -895.37 -558.39
L(β)  -1041.89 -851.62 -527.63
ρ²    0.0764  0.0382  0.0481
σ²    0.0703  0.0305  0.0358
N     826    815    805

4.2.2 Test for Equality of Coefficients across Data Sets

Table 4.2 shows that the parameter estimates vary substantially across the data sets, suggesting that the data sets should not be pooled without more detailed consideration of the differences in decision making protocols among the data sets. The equality of individual parameters from two data sets was examined by using the following asymptotically normal test statistic:

\[
\frac{\hat{\beta}_k^1 - \hat{\beta}_k^2}{\sqrt{\text{Var}(\hat{\beta}_k^1) + \text{Var}(\hat{\beta}_k^2)}}
\]

where \( \beta_k \) is the coefficient of the \( k \)th attribute, \( \hat{\beta}_k^1 \) is the estimate of \( \beta_k \) from Data Set 1 and \( \hat{\beta}_k^2 \) is the estimate of \( \beta_k \) from Data Set 2. The test results are presented in Table 4.3, which shows
that for the estimates from Data Set 1 versus Data Set 2 only the transfer point dummies are significantly different between the two data sets. The parameters of Dummy 1, perceived wait time and ride time are significantly different between Data Sets 2 and 3.

Table 4.3 Tests for Equality of Individual Coefficients
(* indicates significant difference at the 5% level)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Data sets 1 vs. 2</th>
<th>Data sets 2 vs. 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dummy 1</td>
<td>-3.51*</td>
<td>-2.86*</td>
</tr>
<tr>
<td>Dummy 3</td>
<td>-3.43*</td>
<td>0.74</td>
</tr>
<tr>
<td>Dummy 4</td>
<td>-3.67*</td>
<td>-1.46</td>
</tr>
<tr>
<td>Perceived wait time</td>
<td>-0.20</td>
<td>-2.00*</td>
</tr>
<tr>
<td>Ride time</td>
<td>1.11</td>
<td>-0.04</td>
</tr>
<tr>
<td>Reliability</td>
<td>-1.00</td>
<td>-2.52*</td>
</tr>
<tr>
<td>Bus dummy</td>
<td>1.84</td>
<td>-0.61</td>
</tr>
<tr>
<td>Rail dummy</td>
<td>-0.70</td>
<td>1.22</td>
</tr>
</tbody>
</table>

The test for equality of the coefficient vectors across two or more data sets can be performed by comparing likelihood value of the pooled model with that of the separate models. The likelihood ratio test statistic is given by,

\[-2(L_R - L_U),\]

where

\( L_R \): log-likelihood for the restricted model; and

\( L_U \): log-likelihood for the unrestricted model.

This test statistic is \( \chi^2 \) distributed with \( K_U - K_R \) degrees of freedom where \( K_U \) and \( K_R \) are the numbers of estimated parameters in the unrestricted and restricted models, respectively.
The log-likelihood for the unrestricted model is given by the sum of log-likelihood of the models estimated separately from Data Set 1 and Data Set 2 (Case 1), Data Set 2 and Data Set 3 (Case 2), and all the data sets (Case 3). The restricted models estimated from pooled data are shown in Table 4.4.

Table 4.4 Estimation Results of Pooled Models  
(t-statistics in parentheses)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Data sets 1 &amp; 2</th>
<th>Data sets 2 &amp; 3</th>
<th>Data sets 1,2 &amp; 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dummy 1</td>
<td>-0.451 (-4.01)</td>
<td>0.338 (2.56)</td>
<td>-0.180 (-1.82)</td>
</tr>
<tr>
<td>Dummy 3</td>
<td>-0.102 (-0.72)</td>
<td>0.416 (2.42)</td>
<td>-0.0476 (-0.38)</td>
</tr>
<tr>
<td>Dummy 4</td>
<td>-0.839 (-6.93)</td>
<td>-0.186 (-1.42)</td>
<td>-0.646 (-6.30)</td>
</tr>
<tr>
<td>Perceived wait time</td>
<td>-0.0479 (-2.34)</td>
<td>-0.00855 (-0.37)</td>
<td>-0.0286 (-1.60)</td>
</tr>
<tr>
<td>Ride time</td>
<td>-0.0223 (-2.05)</td>
<td>-0.0372 (-2.92)</td>
<td>-0.0266 (-2.76)</td>
</tr>
<tr>
<td>Travel time reliability</td>
<td>-0.110 (-4.43)</td>
<td>-0.0270 (-0.97)</td>
<td>-0.0690 (-3.18)</td>
</tr>
<tr>
<td>Bus dummy</td>
<td>0.554 (5.47)</td>
<td>0.431 (3.53)</td>
<td>0.537 (5.96)</td>
</tr>
<tr>
<td>Rail dummy</td>
<td>0.495 (4.49)</td>
<td>0.492 (4.04)</td>
<td>0.441 (4.61)</td>
</tr>
<tr>
<td>L(0)</td>
<td>-2040.45</td>
<td>-1453.76</td>
<td>-2598.84</td>
</tr>
<tr>
<td>L(\hat{\beta})</td>
<td>-1911.18</td>
<td>-1400.48</td>
<td>-2478.91</td>
</tr>
<tr>
<td>p²</td>
<td>0.0634</td>
<td>0.0366</td>
<td>0.0461</td>
</tr>
<tr>
<td>p²</td>
<td>0.0594</td>
<td>0.0311</td>
<td>0.0431</td>
</tr>
<tr>
<td>N</td>
<td>1641</td>
<td>1620</td>
<td>2446</td>
</tr>
</tbody>
</table>

The likelihood ratio tests, shown in Table 4.5, reject the null hypothesis that the coefficients are equal between any combination of data sets. This means that the three sets of parameters estimated from the three data sets are significantly different and therefore none of these three data sets can be pooled to estimate the base model using the simple rank logit
estimator. The following analyses concern strategies of combining the data sets by explicitly testing hypotheses H2, H3 and H4.

Table 4.5 Tests for Equality of Coefficients

<table>
<thead>
<tr>
<th>Case</th>
<th>Test Statistics</th>
<th>Degrees of Freedom</th>
<th>P-Value</th>
<th>$\chi^2_{dm}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Data 1 vs. 2</td>
<td>35.34</td>
<td>8</td>
<td>0.00002</td>
<td>15.51</td>
</tr>
<tr>
<td>2. Data 2 vs. 3</td>
<td>42.46</td>
<td>8</td>
<td>0.00000</td>
<td>15.51</td>
</tr>
<tr>
<td>3. Data 1 vs. 2 vs. 3</td>
<td>115.54</td>
<td>16</td>
<td>0.00000</td>
<td>26.30</td>
</tr>
</tbody>
</table>

4.2.3 Test of H2

The statistical tests presented in the previous section suggested that principal differences between parameter estimates from separate data sets lie in the alternative specific constants, or Dummies 1, 3 and 4, as postulated by H2. This hypothesis is tested by estimating restricted models from pooled data which allow for separate alternative specific constants for each data set. Table 4.6 shows the estimation results. The test results are presented in Table 4.7. As shown, the null hypothesis of equality of coefficients except alternative specific constants cannot be rejected for Cases 1 and 2, which supports H2, while it is rejected for Case 3. These results indicate that Data Sets 1 and 2 or Data Sets 2 and 3 can be combined if we estimate the transfer point dummies specific to each data set. However, even with this strategy, all three data sets cannot be combined.
Table 4.6 Estimation Results under H2
(t-statistics in parentheses)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Data sets 1 &amp; 2</th>
<th>Data sets 2 &amp; 3</th>
<th>Data sets 1, 2 &amp; 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dummy 1 (Data 1)</td>
<td>-0.648 (-4.93)</td>
<td>-0.615 (-4.94)</td>
<td></td>
</tr>
<tr>
<td>Dummy 3 (Data 1)</td>
<td>-0.287 (-1.90)</td>
<td>-0.298 (-2.14)</td>
<td></td>
</tr>
<tr>
<td>Dummy 4 (Data 1)</td>
<td>-1.18 (-7.68)</td>
<td>-1.16 (-7.90)</td>
<td></td>
</tr>
<tr>
<td>Dummy 1 (Data 2)</td>
<td>-0.106 (-0.75)</td>
<td>0.0243 (0.16)</td>
<td>-0.0833 (-0.62)</td>
</tr>
<tr>
<td>Dummy 3 (Data 2)</td>
<td>0.258 (1.52)</td>
<td>0.402 (2.15)</td>
<td>0.250 (1.57)</td>
</tr>
<tr>
<td>Dummy 4 (Data 2)</td>
<td>-0.413 (-2.84)</td>
<td>-0.330 (-2.15)</td>
<td>-0.396 (-2.85)</td>
</tr>
<tr>
<td>Dummy 1 (Data 3)</td>
<td></td>
<td>0.831 (4.81)</td>
<td>0.727 (4.57)</td>
</tr>
<tr>
<td>Dummy 3 (Data 3)</td>
<td></td>
<td>0.384 (1.84)</td>
<td>0.234 (1.27)</td>
</tr>
<tr>
<td>Dummy 4 (Data 3)</td>
<td></td>
<td>0.0669 (0.41)</td>
<td>0.00257 (0.02)</td>
</tr>
<tr>
<td>Perceived wait time</td>
<td>-0.0474 (-2.31)</td>
<td>-0.00585 (-0.25)</td>
<td>-0.0254 (-1.42)</td>
</tr>
<tr>
<td>Ride time</td>
<td>-0.0228 (-2.08)</td>
<td>-0.0358 (-2.79)</td>
<td>-0.0259 (-2.66)</td>
</tr>
<tr>
<td>Travel time reliability</td>
<td>-0.110 (-4.39)</td>
<td>-0.0273 (-0.98)</td>
<td>-0.0702 (-3.21)</td>
</tr>
<tr>
<td>Bus dummy</td>
<td>0.574 (5.64)</td>
<td>0.422 (3.42)</td>
<td>0.558 (6.12)</td>
</tr>
<tr>
<td>Rail dummy</td>
<td>0.514 (4.62)</td>
<td>0.482 (3.91)</td>
<td>0.457 (4.69)</td>
</tr>
<tr>
<td>L(0)</td>
<td>-2040.45</td>
<td>-1453.76</td>
<td>-2598.84</td>
</tr>
<tr>
<td>L(β)</td>
<td>-1897.32</td>
<td>-1384.67</td>
<td>-2433.93</td>
</tr>
<tr>
<td>ρ²</td>
<td>0.0701</td>
<td>0.0475</td>
<td>0.0635</td>
</tr>
<tr>
<td>ρ²</td>
<td>0.0648</td>
<td>0.0400</td>
<td>0.0581</td>
</tr>
<tr>
<td>N</td>
<td>1641</td>
<td>1620</td>
<td>2446</td>
</tr>
</tbody>
</table>

Table 4.7 Tests for Equality of Coefficients under H2

<table>
<thead>
<tr>
<th>Case</th>
<th>Test Statistics</th>
<th>Degrees of Freedom</th>
<th>P-Value</th>
<th>χ²₀.₀₅</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Data 1 vs. 2</td>
<td>7.62</td>
<td>5</td>
<td>0.17846</td>
<td>11.07</td>
</tr>
<tr>
<td>2. Data 2 vs. 3</td>
<td>10.84</td>
<td>5</td>
<td>0.05465</td>
<td>11.07</td>
</tr>
<tr>
<td>3. Data 1 vs. 2 vs. 3</td>
<td>25.58</td>
<td>10</td>
<td>0.00435</td>
<td>18.31</td>
</tr>
</tbody>
</table>

61
4.2.4 Test of H3

Hypothesis H3 is expressed by:
\[ \beta^{m+1} = \mu_m \beta^m, \quad m=1, 2, \] (4.1)
where
- \( \beta^m \): vector of parameters of Data Set \( m \); and
- \( \mu_m \): scale parameter.

A value of \( \mu_m \) between zero and one implies that Data Set \( m+1 \) contains more random noise than Data Set \( m \). In this case, the parameters of Data Set \( m \) are greater in absolute magnitude than the parameters of Data Set \( m+1 \). The value of \( \mu_m \) is estimated jointly with \( \beta^m \) by combining the data from Data Sets \( m \) and \( m+1 \) and imposing the restriction in (4.1) above. This joint estimation yields statistically efficient estimators although it requires the programming of the joint likelihood function with a general MLE program. Values of \( \mu_m \) can also be estimated graphically using standard MNL estimators. Given a trial value of \( \mu_m \), Data Set \( m+1 \) is transformed by multiplying all the variables by \( \mu_m \) and then using the standard rank logit estimator. The likelihood values with respect to different values of \( \mu_m \) are plotted and the value which gives the maximum likelihood is selected. The output from this estimation procedure gives the correct coefficients and log-likelihood values. However, the calculated variances of the estimated \( \beta^m \) parameters are understated. This analysis employs the joint estimation method.

As shown in Table 4.8, all the \( \mu \)'s are accurately estimated to be between zero and one and \( \mu_1 \mu_2 \) is significantly less than \( \mu_1 \) in the estimation with all data sets, all of which support H3.

Table 4.9 reports the results of a likelihood ratio test for equality of coefficients between two data sets. The unrestricted models are given in Table 4.2 and the restricted models are in Table 4.8. The test results reject the equality of the coefficients in both cases, indicating that the data sets cannot be combined by introducing different scale parameters.
### Table 4.8 Estimation Results under H3
(t-statistics in parentheses)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Data sets 1 &amp; 2</th>
<th>Data sets 2 &amp; 3</th>
<th>Data sets 1, 2 &amp; 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dummy 1</td>
<td>-0.624 (-4.52)</td>
<td>0.315 (2.14)</td>
<td>-0.518 (-3.81)</td>
</tr>
<tr>
<td>Dummy 3</td>
<td>-0.222 (-1.32)</td>
<td>0.472 (2.46)</td>
<td>-0.206 (-1.24)</td>
</tr>
<tr>
<td>Dummy 4</td>
<td>-1.13 (-7.09)</td>
<td>-0.251 (-1.69)</td>
<td>-1.15 (-7.25)</td>
</tr>
<tr>
<td>Perceived wait time</td>
<td>-0.0578 (-2.32)</td>
<td>-0.0175 (-0.69)</td>
<td>-0.0523 (-2.13)</td>
</tr>
<tr>
<td>Ride time</td>
<td>-0.0226 (-1.72)</td>
<td>-0.0408 (-2.86)</td>
<td>-0.0262 (-2.02)</td>
</tr>
<tr>
<td>Travel time reliability</td>
<td>-0.138 (-4.40)</td>
<td>-0.0413 (-1.35)</td>
<td>-0.129 (-4.14)</td>
</tr>
<tr>
<td>Bus dummy</td>
<td>0.714 (5.74)</td>
<td>0.451 (3.32)</td>
<td>0.738 (5.99)</td>
</tr>
<tr>
<td>Rail dummy</td>
<td>0.589 (4.19)</td>
<td>0.563 (4.09)</td>
<td>0.603 (4.35)</td>
</tr>
<tr>
<td>μ1</td>
<td>0.572 (7.00)</td>
<td>0.570 (7.00)</td>
<td>0.728 (4.50)</td>
</tr>
<tr>
<td>μ2</td>
<td></td>
<td></td>
<td>0.728 (4.50)</td>
</tr>
<tr>
<td>μ1μ2</td>
<td>0.245 (3.09)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L(0)</td>
<td>-2040.45</td>
<td>-1453.76</td>
<td>-2598.84</td>
</tr>
<tr>
<td>L(β)</td>
<td>-1902.84</td>
<td>-1399.68</td>
<td>-2456.99</td>
</tr>
<tr>
<td>ρ²</td>
<td>0.0674</td>
<td>0.0372</td>
<td>0.0546</td>
</tr>
<tr>
<td>ρ²</td>
<td>0.0630</td>
<td>0.0310</td>
<td>0.0507</td>
</tr>
<tr>
<td>N</td>
<td>1641</td>
<td>1620</td>
<td>2446</td>
</tr>
</tbody>
</table>

### Table 4.9 Tests for Equality of Coefficients under H3

<table>
<thead>
<tr>
<th>Case</th>
<th>Test statistics</th>
<th>Degrees of freedom</th>
<th>P-Value</th>
<th>χ² obs</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Data 1 vs. 2</td>
<td>18.66</td>
<td>7</td>
<td>0.00932</td>
<td>14.07</td>
</tr>
<tr>
<td>2. Data 2 vs. 3</td>
<td>40.86</td>
<td>7</td>
<td>0.00000</td>
<td>14.07</td>
</tr>
<tr>
<td>3. Data 1 vs. 2 vs. 3</td>
<td>71.70</td>
<td>14</td>
<td>0.00000</td>
<td>23.69</td>
</tr>
</tbody>
</table>
4.2.5 Test of H4

The final test, concerned with the compound effect of different decision protocols and heteroscedasticity, was carried out by combining data sets using different scale parameters and transfer point dummies. The results by the joint estimation are given in Table 4.10. As shown in Table 4.11, the equality of model parameters across the data sets cannot be rejected for Case 1, but is rejected for Cases 2 and 3. According to these test results, Data Sets 1 and 2 can be combined with this strategy, but Data Set 3 may not be used due to too much instability or some unknown biases.

4.3 Discussion and Summary

The empirical analysis of ranking data reported in this chapter indicates significant differences among choice models for different ranks, clearly demonstrating the potential for bias in simple pooling of ranking data. Alternative pooling strategies with explicit hypotheses on the biases in ranking data were also tested.

The hypothesis that the respondents refer to the current usage pattern in ranking hypothetical alternatives was strongly supported. The bus and rail dummy variables which reflect this hypothesis show significant explanatory power in all the models and all the data sets. Exclusion of these variables would cause the problem of "omitted variables" and hence the rest of parameter estimates would be biased. This finding suggests that SP data are likely to be affected by actual choices.

It was shown that choice data for ranks 1 and 2 can be combined if we make allowances for different sets of alternative specific constants, which correspond to transfer point dummies in these data. For example, the coefficients of transfer point dummies estimated from rank 1 data are greater in magnitude than those from rank 2. This may imply that the transfer point plays a
Table 4.10 Estimation Results under H4  
(t-statistics in parentheses)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Data sets 1 &amp; 2</th>
<th>Data sets 2 &amp; 3</th>
<th>Data sets 1,2 &amp; 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dummy 1 (Data 1)</td>
<td>-0.659 (-4.74)</td>
<td></td>
<td>-0.634 (-4.62)</td>
</tr>
<tr>
<td>Dummy 3 (Data 1)</td>
<td>-0.318 (-1.94)</td>
<td></td>
<td>-0.317 (-1.97)</td>
</tr>
<tr>
<td>Dummy 4 (Data 1)</td>
<td>-1.18 (-7.25)</td>
<td></td>
<td>-1.16 (-7.23)</td>
</tr>
<tr>
<td>Dummy 1 (Data 2)</td>
<td>-0.186 (-1.00)</td>
<td>0.0543 (0.34)</td>
<td>-0.154 (-0.85)</td>
</tr>
<tr>
<td>Dummy 3 (Data 2)</td>
<td>0.318 (1.40)</td>
<td>0.449 (2.21)</td>
<td>0.318 (1.43)</td>
</tr>
<tr>
<td>Dummy 4 (Data 2)</td>
<td>-0.586 (-2.52)</td>
<td>-0.311 (-1.93)</td>
<td>-0.555 (-2.45)</td>
</tr>
<tr>
<td>Dummy 1 (Data 3)</td>
<td></td>
<td>1.13 (2.63)</td>
<td>1.83 (1.78)</td>
</tr>
<tr>
<td>Dummy 3 (Data 3)</td>
<td></td>
<td>0.516 (1.71)</td>
<td>0.580 (1.07)</td>
</tr>
<tr>
<td>Dummy 4 (Data 3)</td>
<td></td>
<td>0.0328 (0.15)</td>
<td>-0.198 (-0.48)</td>
</tr>
<tr>
<td>Perceived wait time</td>
<td>-0.0532 (-2.25)</td>
<td>-0.0160 (-0.61)</td>
<td>-0.0448 (-1.96)</td>
</tr>
<tr>
<td>Ride time</td>
<td>-0.0232 (-1.87)</td>
<td>-0.0398 (-2.69)</td>
<td>-0.0267 (-2.19)</td>
</tr>
<tr>
<td>Travel time reliability</td>
<td>-0.127 (-4.12)</td>
<td>-0.0438 (-1.40)</td>
<td>-0.113 (-3.78)</td>
</tr>
<tr>
<td>Bus dummy</td>
<td>0.677 (5.33)</td>
<td>0.448 (3.17)</td>
<td>0.703 (5.53)</td>
</tr>
<tr>
<td>Rail dummy</td>
<td>0.574 (4.03)</td>
<td>0.563 (3.88)</td>
<td>0.583 (4.15)</td>
</tr>
<tr>
<td>$\mu_1$</td>
<td>0.734 (4.29)</td>
<td></td>
<td>0.749 (4.29)</td>
</tr>
<tr>
<td>$\mu_2$</td>
<td></td>
<td>0.683 (2.52)</td>
<td></td>
</tr>
<tr>
<td>$\mu_1 \mu_2$</td>
<td></td>
<td></td>
<td>0.346 (1.81)</td>
</tr>
<tr>
<td>$L(0)$</td>
<td>-2040.45</td>
<td>-1453.76</td>
<td>-2598.84</td>
</tr>
<tr>
<td>$L(\hat{\beta})$</td>
<td>-1896.49</td>
<td>-1384.27</td>
<td>-2430.66</td>
</tr>
<tr>
<td>$p^2$</td>
<td>0.0706</td>
<td>0.0478</td>
<td>0.0647</td>
</tr>
<tr>
<td>$p^2$</td>
<td>0.0647</td>
<td>0.0395</td>
<td>0.0586</td>
</tr>
<tr>
<td>N</td>
<td>1641</td>
<td>1620</td>
<td>2446</td>
</tr>
</tbody>
</table>

more important role in choosing the most preferred alternative than does ranking the second most preferred alternative.
Table 4.11 Test for Equality of Coefficients under H4

<table>
<thead>
<tr>
<th>Case</th>
<th>Test statistics</th>
<th>Degrees of freedom</th>
<th>P-Value</th>
<th>$\chi^2_{0.05}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Data 1 vs. 2</td>
<td>5.96</td>
<td>4</td>
<td>0.20216</td>
<td>9.49</td>
</tr>
<tr>
<td>2. Data 2 vs. 3</td>
<td>10.04</td>
<td>4</td>
<td>0.03976</td>
<td>9.49</td>
</tr>
<tr>
<td>3. Data 1 vs. 2 vs. 3</td>
<td>19.04</td>
<td>8</td>
<td>0.01465</td>
<td>15.51</td>
</tr>
</tbody>
</table>

The hypothesis that the stability of ranking information decreases with decreasing rank was supported by the results that the scale parameter estimates were between zero and one and their magnitude decreases as the rank lowers. However, the inclusion of the scale parameter does not by itself justify combining data from different ranks.

The final strategy tested was to combine the data sets with different scale parameters and different sets of alternative specific constants. However, not even this strategy permitted combining the data set from the third rank. It was concluded, therefore, that choice data from the third rank do not provide additional accuracy to the estimates from ranks 1 and 2 with any of the strategies that were tested.

Although the specific empirical results presented in this chapter may not be applicable to other ranking data, the hypotheses tested are directly relevant to any ranking data. This empirical case study provides a demonstration of the kinds of analyses which are needed for reliable model estimation from ranking data. Neither blindly using ranking data for model estimation nor throwing them away because of unknown reliability is an appropriate strategy. With explicit concern for their reliability, ranking data can be used to estimate choice models.
Chapter 5

Combined Estimation with SP and RP Data

- Case Study 1 -

Chapter 1 states that RP data and SP data have complementary characteristics and in Section 3.3, a methodology of combining these complementary data sources for model estimation was developed. This chapter presents a case study of the combined estimation method to show its applicability. The survey used in this case study was conducted to forecast ridership of a new subway line and it included questions about the current commuting route (RP) and the intention of switching to the new subway line (SP). Another survey was conducted after the opening of the subway in order to collect actual ridership data. The following hypotheses on the characteristics of the stated intention data are postulated and tested:

i) The stated intention data have more random noise than the RP data; and

ii) The stated intentions contain a bias toward the use of the new subway line.

Commuter's route choice models are estimated from the before data using both RP and stated intention data. The after data are used to evaluate the forecasting accuracy of each model estimated from the before data.
5.1 Framework for Combined Estimation with RP Data and Stated Intentions

Most of discrete choice models, which have been widely applied to model travel choices using RP data, predict stationary choice probabilities. In a few applications the choice model represents the decision to switch to a new alternative or to retain one's existing choice (see, for example, Ben-Akiva and de Palma, 1986). Switching behavior is of interest in conjunction with drastic transportation service changes such as the introduction of a new transit system (e.g., Kawakami and Hirobata, 1984). Information about switching behavior can be obtained from a panel survey or from a one-time survey with retrospective questions. Particularly useful are panel surveys conducted before and after major changes in the transportation system.

An alternative source of data on switching behavior is a survey of stated intentions. The respondents are presented with a planned or a potential future change in the transportation system and are asked if and how they intend to modify their current choices in response to the change. For example, a survey could be used to project a change in ridership patterns by asking respondents to state whether or not they would use a newly planned subway line.

This section develops a method for combining stated intention data and RP data to estimate mode switching models. This method follows the framework for combined estimation with RP and SP data presented in Chapter 3. The system contains two models: the RP and stated intention models.

The RP Model:

Consider a choice model with the following stationary random utility function:

\[ U_i = \alpha' \mathbf{w}_i + \beta' \mathbf{x}_i + \epsilon_i = V_i + \epsilon_i, \ i = 1, \ldots, I, \]  

(5.1)

and the revealed choice given by:

\[ d_i = \begin{cases} 1, & \text{if } U_i = \max_{j=1, \ldots, I} \{U_j\} \\ 0, & \text{otherwise} \end{cases} \]

(5.2)
for $i = 1, \ldots, I$, where $w_i$ and $x_i$ are vectors of attributes of alternative $i$, $\alpha$ and $\beta$ are vectors of unknown parameters, and $e_i$, $i = 1, \ldots, I$, are IID random disturbances. Equations (5.1) and (5.2) constitute the model of actual travel behavior, or equivalently, the process generating the RP data.

**The Stated Intention Model:**

The respondent is presented with a new alternative which is not currently available and is asked whether he or she intends to switch from his or her present choice to the new option. Assume that the utility function governing the response to this question is different from the one that governs the RP data and is expressed as follows:

$$
\bar{U}_i = \beta'x_i + \gamma'z_i + v_i = \bar{V}_i + v_i,
$$

(5.3)

where $z_i$ is a vector of bias factors for alternative $i$ and $\gamma$ is the corresponding vector of coefficients. Note that the vector of attributes $w_i$ which appears in the utility function in equation (5.1) is omitted from equation (5.3). This vector represents attributes which affect the actual choice but are not taken into consideration in expressing the intention to switch or not to switch to the new alternative. The response indicator is given by,

$$
\tilde{d} = \begin{cases} 
1, & \text{if } \bar{U}_s - \max_{j=1,\ldots,I} \{\bar{U}_j\} \geq \delta \\
0, & \text{otherwise}
\end{cases}
$$

(5.4)

where $s$ is the index of the new alternative and $\delta$ is a threshold parameter. If the current choice is known and is denoted by $r$ the response indicator is,

$$
\tilde{d} = \begin{cases} 
1, & \text{if } \bar{U}_s - \bar{U}_r \geq \delta \\
0, & \text{otherwise.}
\end{cases}
$$

(5.5)

In the switching model it is assumed that the new alternative is chosen if its utility is greater than the utility of the current mode by more than some threshold value $\delta$. For actual behavior
we expect $\delta$ to be positive reflecting an inertia effect. However, for stated intention data $\delta$ could be negative because of a tendency to overstate the choice of a new alternative.

The random components of the utility functions, $\varepsilon$ and $\nu$, are assumed to be independently distributed with zero means. The random errors in the stated intention data are represented by the variance of $\nu$ being greater than the variance of $\varepsilon$. It is expressed by,

$$Var(\varepsilon) = \mu^2 Var(\nu),$$ \hspace{1cm} (5.6)

where $\mu$ is expected to be between 0 and 1. $\mu$ represents the "scale" of the stated intention model relative to the RP model.

Since the RP model and stated intention model share an unknown parameter vector $\beta$, joint estimation of both models can identify: the SP bias factor coefficients $\gamma$, the threshold parameter $\delta$, and the scale parameter $\mu$, as well as the utility coefficients $\alpha$ and $\beta$. The joint log-likelihood for a sample of $N$ observations is given by,

$$L(\alpha, \beta, \gamma, \mu, \delta) = \sum_{n=1}^{N} d_{rn} \log[F_r(V_{1n}, ..., V_{rn})] + \sum_{n=1}^{N} \tilde{d}_n \log[G(\mu(V_{sn} - V_{rn} - \delta))],$$ \hspace{1cm} (5.7)

where $F_r$ is a function expressing the choice probability of alternative $r$ such as multinomial logit (MNL):

$$F_r(V_{1n}, ..., V_{ln}) = \frac{\exp(V_{rn})}{\sum_{j=1}^{l} \exp(V_{jn})},$$ \hspace{1cm} (5.8)

and $G$ is a binary choice probability such as binary logit:

$$G(\mu(V_{sn} - V_{rn} - \delta)) = \frac{1}{1 + \exp(-\mu(V_{sn} - V_{rn} - \delta))}.$$ \hspace{1cm} (5.9)
5.2 Description of the Data Used for the Empirical Analysis

A transportation research group at The University of Tokyo, Japan, conducted before and after surveys to investigate the ridership of a new subway line, which opened in March, 1985 in Yokohama, Japan. The before survey was conducted three months prior to the opening of the subway line and the after survey was carried out six months after the opening.

Self-administered questionnaires were distributed to randomly selected households in the vicinity of a selected station of the new subway line. The questionnaires were collected at home by survey personnel. The numbers of respondents to the before and the after surveys are 564 and 1,201 with response rates of 70.0% and 80.7%, respectively.

The before survey included the following questions about the regular trip to work or school:

- the regularly used transit route from home to the place of work or school;
- an alternate transit route for the same trip; and
- intentions of using the new subway line.

The after survey included the following questions about the trip to work or school:

- the route from home to work or school using the new subway line;
- the route for the same trip without using the new subway line; and
- satisfaction levels with the service attributes of both routes.

Questionnaires of the before and after surveys are shown in Appendices B and C, respectively (translated into English).

Exploratory data analysis revealed a discrepancy between the stated intention to use the subway in the before data and actual use in the after data. 42.9% of the respondents in the before data expressed an intention to use the new subway line, while only 34.5% of the respondents in the after data were actually using it (Suzuki et al, 1986). Although the before and the after surveys were conducted independently, 100 observations reported the same personal attributes and destinations in the two surveys. The cross tabulation of these 100
individuals with respect to reported intention to use the subway and actual behavior is shown in Table 5.1. There are 17 respondents who expressed the intention to switch to the subway but did not switch and 9 respondents who stated no intention of switching but actually did. This analysis suggests that the stated intention data are biased by an overstatement of the intention to switch to the new subway line.

Table 5.1 Comparison of Stated Intention and Actual Behavior

<table>
<thead>
<tr>
<th>Reported Intention of Subway Use</th>
<th>Actual Behavior</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>no</td>
<td>37</td>
<td>9</td>
</tr>
<tr>
<td>yes</td>
<td>17</td>
<td>37</td>
</tr>
<tr>
<td>total</td>
<td>54</td>
<td>46</td>
</tr>
</tbody>
</table>

5.3 Model Estimation with the Before Data

5.3.1 Estimation with RP Data

A binary choice data set was created from the RP information by regarding the regular route and the alternate route as chosen and unchosen options, respectively. The data set consists of 107 respondents whose principal commuting mode is commuter rail for both regular and
alternative routes. For most respondents, the regular route and the alternate route differ only in the boarding stations of the commuter rail and/or the access modes to the boarding stations. Calculated values of the service levels (e.g., travel time and wait time) of each route were used in the data set.

In this data set, information is only available about two alternatives, namely, the regular and the alternate routes. In this survey, the regular and alternate routes can be considered as the best and second best alternatives, respectively. However, if the true choice set contains other routes besides the reported two routes, then the estimator of a binary choice model will be biased.

Suppose that the true choice set of individual $n$ consists of $J_n$ ($J_n \geq 2$) alternatives and denotes the regular and the alternate routes as alternatives 1 and 2, respectively. For an MNL model with scale parameter $\mu$, the choice probability of alternative 1 from a full choice set is given by:

$$P_n(1) = \frac{\exp(\mu V_{1n})}{\sum_{j=1}^{J_n} \exp(\mu V_{jn})}$$
$$= \frac{\exp(\mu V_{1n})}{\exp(\mu V_{1n}) + \sum_{j=2}^{J_n} \exp(\mu V_{jn})}$$
$$= \frac{\exp(\mu V_{1n})}{\exp(\mu V_{1n}) + \exp(\mu V_{2n}) \sum_{j=2}^{J_n} \exp(\mu (V_{jn} - V_{2n}))}$$
$$= \frac{\exp(\mu V_{1n})}{\exp(\mu V_{1n}) + \exp(\mu V_{2n} + \ln \sum_{j=2}^{J_n} \exp(\mu (V_{jn} - V_{2n})))}$$
$$= \frac{\exp(\mu V_{1n})}{\exp(\mu V_{1n}) + \exp \left[ \mu \left( V_{2n} + \frac{1}{\mu} \ln(1 + e^{\mu (V_{3n} - V_{2n})} + e^{\mu (V_{4n} - V_{2n})} + \ldots + e^{\mu (V_{J_n} - V_{2n})}) \right) \right]} . \quad (5.10)$$

Assume that $U_{2n}, U_{3n}, \ldots, U_{J_n}$ are randomly distributed with the same mean, which implies that $V_{2n} = V_{3n} = \ldots = V_{J_n}$. Then, equation (5.10) is reduced to,
\[
P_n(1) = \frac{\exp(\mu V_{1n})}{\exp(\mu V_{1n}) + \exp\left[\mu\left(V_{2n} + \frac{1}{\mu}\ln(J_n-1)\right)\right]}
= \frac{1}{1 + \exp\left[\mu(V_{2n} - V_{1n} + \frac{1}{\mu}\ln(J_n-1))\right]}.
\]

(5.11)

The correction term (the last term in the exponential) in the above equation has the following intuitive explanation. Because of the assumption of identical systematic utilities among alternatives \(2, \ldots, J_n\), these alternatives are distinguished by the random component of the utility function. Hence, the expected value of the maximum utility among those alternatives must have some positive additional term which reflects the number of alternatives available. The correction term implies that the random utility of the alternate route for an individual with large number of available alternatives has a large value.

In this case study, the number of alternatives in the full choice set, \(J_n\), can be approximated by the number of available access modes, as described by the question: "What vehicles that your household owns are usually available for your commuting trip?" If alternative 2 is the second best alternative, this correction method will somewhat overestimate the utility values of the rest of alternatives due to the assumption of \(V_{2n} = V_{3n} = \ldots = V_{J_n n}\). In this survey, however, the reported alternate route is not necessarily the second best alternative because the questionnaire does not explicitly ask for the second best choice but simply asks for another route the respondent sometimes uses.

The RP model has the following explanatory variables and the estimation results are shown in Table 5.2.

- **Bus dummy (busdum)** = \(1\): if the principal access mode is bus
  \(0\): otherwise

- **Bike dummy (bikedum)** = \(1\): if the principal access mode is bicycle or moped
  \(0\): otherwise

- **Car dummy (cardum)** = \(1\): if the principal access mode is car, motorcycle or taxi
  \(0\): otherwise

- **Walk time (walkt)** = walking time (minutes)
Access in-vehicle time \((\text{accivt})\) = in-vehicle (bus, bike, moped, car, and motorcycle) time of access trip (minutes)

Number of transfers \((\text{xfem})\) = the number of transfers

Table 5.2 Estimation Results of RP Model

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient estimate</th>
<th>Standard error</th>
<th>t-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bus dummy</td>
<td>-2.99</td>
<td>1.00</td>
<td>-2.99</td>
</tr>
<tr>
<td>Bike dummy</td>
<td>-5.77</td>
<td>1.32</td>
<td>-4.37</td>
</tr>
<tr>
<td>Car dummy</td>
<td>-8.48</td>
<td>1.55</td>
<td>-5.84</td>
</tr>
<tr>
<td>Walk time</td>
<td>-0.335</td>
<td>0.0933</td>
<td>-3.59</td>
</tr>
<tr>
<td>Access in-vehicle time</td>
<td>0.00769</td>
<td>0.0481</td>
<td>0.16</td>
</tr>
<tr>
<td>Number of transfers</td>
<td>-1.12</td>
<td>0.410</td>
<td>-2.73</td>
</tr>
</tbody>
</table>

\(L(0) = -113.51\) \hspace{1cm} \(L(\hat{\beta}) = -74.14\) \hspace{1cm} \(\rho^2 = 0.347\) \hspace{1cm} \(\bar{\rho}^2 = 0.294\) \hspace{1cm} \(N = 107\)

The configuration of the three access mode dummies implies that the "base" access mode is "walk." Since the regular and alternate routes use the same principal mode (commuter rail), neither the total travel time nor total travel cost had a significant coefficient. Another reason may be that the commuting cost is usually provided by the employer in Japan. Thus, the model includes only the mode dummies and travel times of the access trip and the number of transfers.

The estimated parameters show that walk access is most preferred followed by bus, bicycle or moped over the other access modes, i.e., car, motorcycle, and taxi.
5.3.2 Estimation with SP Data

Stated preferences, namely, the intention of using the new subway line, were employed to create a binary choice data in the following way. If the respondent expressed the intention of switching to the route with subway, then he or she is considered to choose the subway route over the currently used route. Otherwise, the respondent is considered to choose the current route over the subway route. Thus, a binary choice data set can be created. In the sample 60 out of 107 respondents expressed an intention to switch to the subway route. Note that the subway would be used mostly as an access mode to the existing commuter rail.

The subway route specific constant is introduced in the model so that it may capture i) additional attributes of the subway route that are not included in the model; ii) the response bias toward the new subway route; and iii) the threshold value $\delta$. The access in-vehicle time includes subway and bus ride time and captures the effect of the savings of in-vehicle access time by switching from bus to subway. The dummy variable for car, motorcycle or taxi access was set to be minus infinity because no respondent reported a choice of any of these modes in the SP data.

This model can also be corrected for the implicit choice set and the corrected utility function of alternative 1 (current route) is given by,

$$\tilde{U}_1 = \tilde{V}_1 + \frac{1}{\mu} \ln(J_n) + \nu_{1n}.$$  \hspace{1cm} (5.12)

Estimation results of the model are shown in Table 5.3. The subway constant has a significant positive coefficient, which may reflect the overstated use of the subway route, implying a negative threshold value $\delta$.

Comparing the two models reveals that the parameter estimates from the stated intention data have smaller magnitudes than those from the RP data except for the parameter of the number of transfers.
Table 5.3 Estimation Results of SP Model

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient estimate</th>
<th>Standard error</th>
<th>t-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bus dummy</td>
<td>-1.57</td>
<td>0.882</td>
<td>-1.78</td>
</tr>
<tr>
<td>Bike dummy</td>
<td>-2.89</td>
<td>0.900</td>
<td>-3.21</td>
</tr>
<tr>
<td>Car dummy</td>
<td>(\infty)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Walk time</td>
<td>-0.238</td>
<td>0.0719</td>
<td>-3.31</td>
</tr>
<tr>
<td>Access in-vehicle time</td>
<td>-0.0856</td>
<td>0.0386</td>
<td>-2.22</td>
</tr>
<tr>
<td>Number of transfers</td>
<td>-1.32</td>
<td>0.498</td>
<td>-2.65</td>
</tr>
<tr>
<td>Subway route constant</td>
<td>0.974</td>
<td>0.583</td>
<td>1.67</td>
</tr>
</tbody>
</table>

\(L(0) = -97.28\) \(L(\hat{\beta}) = -59.76\) \(\rho^2 = 0.386\) \(\bar{\rho}^2 = 0.324\) \(N = 107\)

5.3.3 Combined Estimation with SP and RP Data

Following the framework developed in Section 5.1, the RP and stated intention models are expressed as follows.

The RP Model:

\[
U_{rn} = \beta^*x_{rn} + \alpha^*w_{rn} + \varepsilon_{rn} \tag{5.13}
\]

\[
U_{an} = \beta^*x_{an} + \alpha^*w_{an} + \ln(J_{n-1}) + \varepsilon_{an} \tag{5.14}
\]

\[
U_{rn} \geq U_{an} \text{, for all } n \tag{5.15}
\]

The Stated Intention Model:

\[
\tilde{U}_{rn} = \beta^*x_{rn} + \gamma z_{rn} + \delta + \ln(J_{n}) + \nu_{rn} \tag{5.16}
\]
\[ \tilde{U}_{sn} = \beta^r x_{sn} + \gamma^s z_{sn} + \nu_{sn} \]  

(5.17)

\[ \tilde{d}_n = \begin{cases} 
1: \text{if } \tilde{U}_{sn} - \tilde{U}_{rn} \geq 0 \\
0: \text{otherwise} 
\end{cases} \]  

(5.18)

In the above equations, \( x, w, \) and \( z \) are vectors of independent variables and subscripts \( r, a, \) and \( s \) represent the regular route, alternate route, and subway route, respectively. \( \tilde{d} \) is the choice indicator of the stated intentions. Inequality (5.15) holds because alternative \( r \) is always the chosen one. Both models are described by a binary discrete choice model such as logit or probit.

Note that the stated intention model has a subway constant which also captures the effect of the threshold value \( \delta \).

Specifications of the systematic utility functions for the two models are shown below.

**Specification of the RP Model:**

\[ U_{rn} = \beta_1 \text{busdum}_{rn} + \beta_2 \text{bikedum}_{rn} + \beta_3 \text{walkt}_{rn} + \beta_4 \text{accivt}_{rn} + \beta_5 x_{fern} + \alpha \text{cardum}_{rn} + \epsilon_{rn} \]

\[ \equiv V_{rn} + \epsilon_{rn} \]  

(5.19)

\[ U_{an} = \beta_1 \text{busdum}_{an} + \beta_2 \text{bikedum}_{an} + \beta_3 \text{walkt}_{an} + \beta_4 \text{accivt}_{an} + \beta_5 x_{fern}_{an} + \alpha \text{cardum}_{an} + \ln(J_n - 1) + \epsilon_{an} \]

\[ \equiv V_{an} + \ln(J_n - 1) + \epsilon_{an} \]  

(5.20)

**Specification of the Stated Intention Model:**

\[ \tilde{U}_{rn} = \beta_1 \text{busdum}_{rn} + \beta_2 \text{bikedum}_{rn} + \beta_3 \text{walkt}_{rn} + \beta_4 \text{accivt}_{rn} + \beta_5 x_{fern}_{rn} + \delta + \ln(J_n) + \nu_{1n} \]

\[ \equiv \tilde{V}_{rn} + \ln(J_n) + \nu_{rn} \]  

(5.21)
\[ \begin{align*}
    \tilde{U}_{sn} &= \beta_1 \text{busdum}_{sn} + \beta_2 \text{walkdum}_{sn} + \beta_3 \text{walkt}_{sn} + \beta_4 \text{accivt}_{sn} \\
    &\quad + \beta_5 \text{xfer}_{sn} + \gamma_1 \text{cardum}_{sn} + \gamma_2 + \nu_{sn} \\
    &\equiv \tilde{V}_{sn} + \nu_{sn}
\end{align*} \] (5.22)

Note that \( \beta \)'s are common to both models while \( \alpha \) and \( \gamma \) are specific to the RP and stated intention models, respectively. \( \gamma_1 \) is set to be minus infinity for the reason mentioned earlier. The subway specific constant \( \gamma_2 \) and the threshold parameter \( \delta \) cannot be distinguished in this specification. Consequently \( \delta \) is set to zero and the subway constant reflects the term \( \gamma_2 - \delta \).

The following proposition on the random terms is introduced:

\[ \text{Var}(\varepsilon) = \mu^2 \text{Var}(\nu), \] (5.23)

where

\[ 0 \leq \mu \leq 1. \]

Estimation of discrete choice models requires an arbitrary normalization of the scale parameter. Setting the scale parameter of the RP model to one yields the following choice model:

\[ P_n(r) = \frac{1}{1 + \exp(V_{an} + \ln(J_n - 1) - V_{rn})}. \] (5.24)

Since the variance of a Gumbel variable is inversely proportional to the square of the Gumbel scale parameter, the choice probability for the stated intention model is given by:

\[ Q_n(r) = \frac{1}{1 + \exp(\mu(\tilde{V}_{sn} - \tilde{V}_{rn} - \ln(J_n)))}. \] (5.25)

Intuitively, \( \mu \) scales down the explanatory variables in the stated intention model because the reported intentions are more random than actual choices. Estimation solely from the stated intention data can not separately identify \( \mu \) and \( \beta \)'s but yields estimates of \( \mu \beta \)'s, which are the scaled-down values of \( \beta \)'s. This can be observed in Tables 5.2 and 5.3, where the estimates from the stated intention data are generally smaller in magnitude than the estimates from the RP data.
The likelihood function for the estimation of this model is given by:

\[
L(\alpha, \beta, \gamma, \mu) = \prod_{n=1}^{N} P_{n}(r)Q_{n}(s)^{\tilde{\alpha}_n} Q_{n}(r)^{1-\tilde{\alpha}_n} \\
= \prod_{n=1}^{N} \left( \frac{1}{1+\exp(V_{an}+\ln(J_{n}-1)-V_{rn})} \right) \left( \frac{1}{1+\exp(\mu(V_{rn}+\ln(J_{n}))-\tilde{V}_{sn})} \right) \left( \frac{1}{1+\exp(\mu(V_{sn}-\tilde{V}_{r}-\ln(J_{n})))} \right)^{1-\tilde{\alpha}_n},
\]

(5.26)

where subscript \( n \) represents the \( n \)th observation and the sample size is \( N \).

The estimation results are given in Table 5.4. It shows that \( \mu \) was successfully estimated between 0 and 1 with a small standard error (strongly supporting the proposition on the disturbance terms of the two equations). All the model parameters except for the access in-vehicle time are accurately estimated. The positive coefficient estimate of the subway constant indicates the bias toward the subway route.

Table 5.4 Combined Estimation with SP and RP Data

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient estimate</th>
<th>Standard error</th>
<th>t-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bus dummy</td>
<td>-2.54</td>
<td>0.904</td>
<td>-2.81</td>
</tr>
<tr>
<td>Bike dummy</td>
<td>-5.73</td>
<td>1.26</td>
<td>-4.55</td>
</tr>
<tr>
<td>Car dummy</td>
<td>-8.81</td>
<td>1.42</td>
<td>-6.21</td>
</tr>
<tr>
<td>Walk time</td>
<td>-0.365</td>
<td>0.0855</td>
<td>-4.27</td>
</tr>
<tr>
<td>Access in-vehicle time</td>
<td>-0.0630</td>
<td>0.0387</td>
<td>-1.63</td>
</tr>
<tr>
<td>Number of transfers</td>
<td>-1.69</td>
<td>0.392</td>
<td>-4.31</td>
</tr>
<tr>
<td>Subway route constant</td>
<td>1.22</td>
<td>0.542</td>
<td>2.25</td>
</tr>
<tr>
<td>( \mu )</td>
<td>0.559</td>
<td>0.174</td>
<td>3.21</td>
</tr>
</tbody>
</table>

\( L(0) = -210.80 \) \( L(\beta) = -135.77 \) \( R^2 = 0.356 \) \( \bar{R}^2 = 0.318 \) \( N = 214 \)
As shown in Table 5.5, the coefficient estimates of this combined model are very similar to those of the RP model which indicates that the joint estimation successfully replicates the RP model. The equality of those parameter estimates can be tested by comparing likelihood value of the pooled model with that of the separate models. The likelihood ratio test statistic is given by,

\[-2(L_R - L_U),\]

where

\[L_R : \text{log-likelihood for the restricted model; and}\]
\[L_U : \text{log-likelihood for the unrestricted model}.\]

This test statistic is \(\chi^2\) distributed with \(K_U - K_R\) degrees of freedom where \(K_U\) and \(K_R\) are the numbers of estimated parameters in the unrestricted and restricted models, respectively. The log-likelihood for the unrestricted model is given by the sum of log-likelihoods of the models estimated separately from RP data and stated intention data, while the combined estimation yields the restricted model. Namely, the test statistic is,

\[-2\{ -135.51 - (-74.14 - 59.76) \} = 3.22 , \tag{5.27}\]

with degrees of freedom 6+6-8=4, which gives the P-value 0.5217. Therefore, the null hypothesis can not be rejected.
Table 5.5 Comparison of the Estimation Results with the Before Data
(t-statistics in parentheses)

<table>
<thead>
<tr>
<th></th>
<th>RP model</th>
<th>SP model</th>
<th>RP+SP model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bus dummy</td>
<td>-2.99</td>
<td>-1.57</td>
<td>-2.54</td>
</tr>
<tr>
<td></td>
<td>(-2.99)</td>
<td>(-1.78)</td>
<td>(-2.81)</td>
</tr>
<tr>
<td>Bike dummy</td>
<td>-5.77</td>
<td>-2.89</td>
<td>-5.73</td>
</tr>
<tr>
<td></td>
<td>(-4.37)</td>
<td>(-3.21)</td>
<td>(-4.55)</td>
</tr>
<tr>
<td>Car dummy</td>
<td>-8.48</td>
<td>-∞</td>
<td>-8.81</td>
</tr>
<tr>
<td></td>
<td>(-5.84)</td>
<td></td>
<td>(-6.21)</td>
</tr>
<tr>
<td>Walk time</td>
<td>-0.335</td>
<td>-0.238</td>
<td>-0.365</td>
</tr>
<tr>
<td></td>
<td>(-3.59)</td>
<td>(-3.31)</td>
<td>(-4.27)</td>
</tr>
<tr>
<td>Access in-vehicle time</td>
<td>0.00769</td>
<td>-0.0856</td>
<td>-0.0630</td>
</tr>
<tr>
<td></td>
<td>(0.16)</td>
<td>(-2.22)</td>
<td>(-1.63)</td>
</tr>
<tr>
<td>Number of transfers</td>
<td>-1.12</td>
<td>-1.32</td>
<td>-1.69</td>
</tr>
<tr>
<td></td>
<td>(-2.73)</td>
<td>(-2.65)</td>
<td>(-4.31)</td>
</tr>
<tr>
<td>Subway route constant</td>
<td></td>
<td>0.974</td>
<td>1.22</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.67)</td>
<td>(2.25)</td>
</tr>
</tbody>
</table>

\[ \mu \]

\[ L(0) \]

\[ L(\beta) \]

\[ \rho^2 \]

\[ \rho^2 \]

\[ N \]

0.559

-113.51

-97.28

-210.80

-74.14

-59.76

-135.77

0.347

0.386

0.356

0.294

0.324

0.318

107

107

214

5.4 Prediction Tests of the Estimated Models with the After Data

5.4.1 Model Estimation with the After Data

The after data were collected six months subsequent to the opening of the subway line. The survey questionnaire asked for the frequencies of using a commuting route with the new subway line and another route without the subway. Because of the ambiguity of the question,
there seemed to have been a confusion among the respondents in reporting these frequencies. Some respondents gave the frequency of using a route assuming that they had to choose that route, while other respondents reported actual frequencies of use. Hence, identifying actual choices in the after situation had to rely on other data items such as answers to questions about the reasons for not using the subway. Thus, we expect this data to have significant errors. Selecting individuals whose principal mode was rail produced 428 observations, of which 254 respondents (59.35%) were identified to be users of a route with the subway.

A choice model with the same specification as the before models was estimated with the results shown in Table 5.6. The subway route constant is not significant, which may indicate that there is no bias toward the subway route in the after data. Because of the potential large noise in the dependent variables, the fit of the model turned out to be poor compared with before models.

In order to test whether the after data were generated from the same model as the before data, the two were combined allowing for a different scale parameter for the after model. The scaled after model is joined with the combined before model. Consequently, this model has two scale parameters, \( \mu_1 \) for the before stated intention data and \( \mu_2 \) for the after data. Since the subway route dummy and the access in-vehicle time, including subway ride time, may capture biases in the before data, the parameters of these variables are separately estimated from the before and after data in the combined estimation. Tables 5.5 and 5.6 show that the parameter estimates of the bike dummy from the before data and the after data seem to be significantly different even after the proper correction of the model scale. This may be ascribed to the difference of the sampling area used in the before and after surveys. Therefore, the parameters of this variable are also separately estimated from the before and after data. As shown in the second column of Table 5.6, \( \mu_2 \) is smaller than \( \mu_1 \), indicating that the after data contain large random errors. With this strategy of uniting before and after data, equality of coefficients between the combined before model and the after model was accepted according to a likelihood ratio test (the test statistic is 4.46 with degrees of freedom 3, giving P-value 0.2159).
5.4.2 Prediction Test of Before Models

A prediction test of the models estimated with the before data was carried out in the following way. Predicted choice probabilities are calculated with the values of the explanatory variables from the after data and the coefficient estimates from the before data to yield a predicted probability \( \hat{P}_n(s) \). Then, the predicted market share of the subway route is,

\[
\sum_{n=1}^{N} \hat{P}_n(s) .
\]

Table 5.6 Estimation Results with the After Data
(t-statistics in parentheses)

<table>
<thead>
<tr>
<th></th>
<th>After model</th>
<th>Before+After model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bus dummy</td>
<td>-0.399</td>
<td>-2.00 (-2.53)</td>
</tr>
<tr>
<td>Bike dummy (Before)</td>
<td></td>
<td>-5.42 (-4.80)</td>
</tr>
<tr>
<td>Bike dummy (After)</td>
<td>0.228</td>
<td>0.0733 (0.04)</td>
</tr>
<tr>
<td>Car dummy</td>
<td>-1.79</td>
<td>-8.37 (-6.16)</td>
</tr>
<tr>
<td>Walk time</td>
<td>-0.0768</td>
<td>-0.327 (-4.40)</td>
</tr>
<tr>
<td>Access in-vehicle time (Before)</td>
<td></td>
<td>-0.0657 (-1.78)</td>
</tr>
<tr>
<td>Access in-vehicle time (After)</td>
<td>-0.0754</td>
<td>-0.275 (-3.47)</td>
</tr>
<tr>
<td>Number of transfers</td>
<td>-0.942</td>
<td>-2.03 (-5.41)</td>
</tr>
<tr>
<td>Subway constant (Before)</td>
<td></td>
<td>1.45 (2.94)</td>
</tr>
<tr>
<td>Subway constant (After)</td>
<td>-0.112</td>
<td>-0.917 (-1.41)</td>
</tr>
<tr>
<td>( \mu_1 )</td>
<td></td>
<td>0.561 (3.39)</td>
</tr>
<tr>
<td>( \mu_2 )</td>
<td></td>
<td>0.313 (3.88)</td>
</tr>
<tr>
<td>( L(\theta) )</td>
<td>-296.67</td>
<td>-507.47</td>
</tr>
<tr>
<td>( L(\hat{\theta}) )</td>
<td>-257.92</td>
<td>-395.66</td>
</tr>
<tr>
<td>( \hat{\rho}^2 )</td>
<td>0.131</td>
<td>0.220</td>
</tr>
<tr>
<td>( \hat{\rho}^2 )</td>
<td>0.107</td>
<td>0.197</td>
</tr>
<tr>
<td>( N )</td>
<td>428</td>
<td>642</td>
</tr>
</tbody>
</table>
Log-likelihood values are also calculated using the observed choice in the after data as a goodness-of-fit measure.

These predictions are performed for several models as shown in Table 5.7. The "SP bias adjusted" model is obtained by omitting the subway route constant from the model. If this constant reflects an overstatement of the subway route usage, removing it from the model will yield a better fit to the after data and will better predict the actual usage. As shown in the table, all the before models without this bias adjustment overestimated the share of the subway route by as much as 20%. The correction of the bias substantially reduces this overestimation.

The overestimation of the subway route share may have resulted from the following causes. As discussed earlier, the format of the after survey questionnaire is so ambiguous that the dependent variable in the after data might not precisely reflect the actual choice. Another potential explanation is that a relatively large number of respondents are captive to the previous (non-subway) route. The after data revealed that captive travelers do not use the subway for reasons which include disliking subway, unfamiliarity with the subway route, and habitual usage of the previous route. The overprediction might have resulted from the inability of the

Table 5.7 Prediction Test of the Estimated Models

<table>
<thead>
<tr>
<th>Model</th>
<th>Log-likelihood</th>
<th>Predicted share of subway route (observed share = 59.35%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>After model</td>
<td>-260.59</td>
<td>59.35%</td>
</tr>
<tr>
<td>RP model</td>
<td>-367.64</td>
<td>69.70%</td>
</tr>
<tr>
<td>SI model</td>
<td>-365.90</td>
<td>81.97%</td>
</tr>
<tr>
<td>SI model*</td>
<td>-298.63</td>
<td>68.22%</td>
</tr>
<tr>
<td>RP+SI combined model</td>
<td>-443.85</td>
<td>82.89%</td>
</tr>
<tr>
<td>RP+SI combined model*</td>
<td>-355.71</td>
<td>68.68%</td>
</tr>
</tbody>
</table>

Note: * SP bias adjusted
before models to capture this effect.

The bias adjusted stated intention model had the closest predicted share and achieved the best goodness-of-fit value (log-likelihood). This suggests that if a potential bias (such as overstatement of new options) can be corrected, stated intention data can have good predictive validity.

5.5. Discussion and Summary

An empirical analysis of the combined estimation with RP and SP data was conducted in this chapter. Switching models for commuting routes were estimated from stationary RP data and stated intentions of switching. Since RP and stated intention data have complementary characteristics, model estimation can benefit from the strategy of combining both types of data with explicit consideration of their advantages and disadvantages. The case study presented in this chapter demonstrates the effectiveness and practicality of this strategy. Namely, we found that:

i) combining the stated intention data with the RP data increased the accuracy of parameter estimates of the model;

ii) a statistical test showed that the model for the stated intention data, if properly scaled, had the same coefficients as the RP model;

iii) the stated intention data contained more random noise; and

iv) the utility threshold value for switching routes estimated from the stated intentions was negative, implying that the respondents overstated their switching to the new alternative.

On the first point, the RP model did not produce a successful estimate of the coefficient of access in-vehicle time. However, the combined estimators did estimate a coefficient with the correct sign. A bias toward the route with the new subway line was detected by incorporating
SP data with RP data. Namely, the models estimated with these strategies had a significantly positive coefficient for the subway route specific constant, while that coefficient estimated from the after data was not significant. Predictive tests using the after data revealed that removing that constant term from the models substantially improved the prediction of the subway usage.

The combined estimation with RP and SP data supports the hypothesis that the SP data contain greater noise than the RP data and this was verified by estimating a scale parameter between zero and one with notable accuracy.

In the following chapter, another case study for combined estimation with RP and SP data is presented.
Chapter 6

Combined Estimation with SP and RP Data

- Case Study 2 -

This chapter presents a case study of combined estimation with SP and RP data on mode choice behavior of intercity travelers. The survey includes two SP experiments in addition to RP information. This case study demonstrates the following characteristics of SP data:

i) an SP survey can elicit pure information on trade-offs among attributes and consequently may contain less random noise than RP data; and

ii) knowledge of the actual choice can be used to estimate the combined effect of a justification bias and a preference inertia.

Mode choice models are estimated using different combinations of RP and SP data. The combined models are shown to be superior to those estimated on RP or SP data alone. They produced more accurate estimates and more reasonable values of time and elasticities. This case study is a clear demonstration of the effectiveness and practicality of combined estimation with RP and SP data.
6.1 Description of the Survey Data

The survey was conducted during 1987 by the Hague Consulting Group for the Netherlands Railways to assess factors which influence the choice between rail and car for intercity travel. The following description of the survey data is based on a report by Bradley et al (1988).

The City of Nijmegen, in the eastern part of the Netherlands near the border with West Germany, was selected as the data collection site. This city has typical rail connections with the major cities in the western metropolitan area called the Randstad which contains Amsterdam, Rotterdam and The Hague. Trips from Nijmegen to the Randstad takes approximately two hours by both rail and car. The sample consisted of residents of Nijmegen who:

- made a trip in the previous three months to Amsterdam, Rotterdam or The Hague;
- did not use a yearly rail pass, or other type of pass which would eliminate the marginal cost of a rail trip;
- had the possibility of using a car, namely, possessed a driver's license and had a car available in the household; and
- had the possibility of using rail, namely, did not have very heavy baggage, were not handicapped, and did not need to visit multiple destinations.

Qualifying residents of Nijmegen were identified in a random telephone survey and requested to participate in a home interview. 235 interviews were conducted out of the 365 people who were reached by telephone and satisfied the above criteria.

The home interview survey consisted of three parts:

i) the characteristics of an intercity trip to the Randstad made within the previous three months (RP data);

ii) SP experiment of a choice between rail and car (SP1 data); and

iii) SP experiment of a choice between two different rail services (SP2 data).
The entire home interview was administered using lap-top micro-computers, so the respondent replied to the questions appearing on the computer screen. The main advantage of using a portable computer is that it can create a desirable SP experimental design on site based on the service levels of the actual trip.

The RP data have 235 observations each including the following items:

- mode used (rail or car)
- trip purpose
- cost (for both chosen and unchosen modes)
- in-vehicle travel time (for both chosen and unchosen modes)
- access and egress time (for both chosen and unchosen modes)
- number of transfers for rail mode
- subjective ratings of latent travel characteristics (e.g., relaxation, reliability) (for both chosen and unchosen modes)
- socioeconomic characteristics of the respondents (e.g., age, sex)

The SP experiments were framed in the context of the actual trip observed in the RP data and used the full-profile pairwise comparison method. The respondent was shown two hypothetical alternatives (rail versus car in the first experiment, and two different rail services in the second experiment) at a time, each of which was described by the following four attributes: travel cost, travel time, number of transfers (only for rail), and comfort level (only for rail). Then, the respondent was asked which mode would be chosen for the particular intercity trip reported in the RP question in terms of a five point rating scale: 1) definitely choose alternative 1, 2) probably choose alternative 1, 3) not sure, 4) probably choose alternative 2, and 5) definitely choose alternative 2. Each respondent was presented with several pairs at each experiment. Since the main concern of the survey was the mode switching behavior from car to rail or vice-versa, the attribute levels were set so that rail became more attractive and/or car became less attractive to the car users (those who actually used a car for the intercity trip) and set in the opposite direction for the rail users. The order of presenting the
alternatives to the respondent was randomized so as to minimize the potential response bias.
SP1 data (rail vs. car) include 1,628 comparisons (an average of 7 comparisons per
respondent), while SP2 data (rail vs. rail) contain 2,965 comparisons (an average of 12-13
comparisons per respondent).

The questions for this home interview are presented in Appendix D.

6.2 Model Estimation with Separate Data Set

6.2.1 Estimation with RP Data

The home interview first asked about a recent intercity trip from Nijmegen to the Randstad.
The respondents were requested to report the characteristics of that trip and those of a trip to the
same destination but with the other mode. So, the attribute values of both modes were
provided by the respondents rather than calculated from network data.

Binary probit models were estimated from this RP data. The latent preferences are
expressed by a utility function:

\[ U = \beta'x + \varepsilon, \quad (6.1) \]

where

\[ \beta = \text{vector of unknown parameters}; \]
\[ x = \text{vector of explanatory variables}; \text{and} \]
\[ \varepsilon \sim N(0,1). \]

The revealed preferences are expressed by,

\[ y = \begin{cases} 1, & \text{if } U \geq 0 \\ 0, & \text{otherwise} \end{cases} \quad (6.2) \]
The dependent variable takes the value of one if rail was chosen and zero if car was chosen. Accordingly, all the explanatory variables are in terms of differences between rail and car, more specifically, the values for rail minus the values for car. The model includes the following variables:

- a constant;
- travel cost per person by rail less travel cost by car (Guilders);
- in-vehicle travel time by rail less in-vehicle travel time by car (minutes);
- access and egress time by rail less parking time by car (minutes);
- number of transfers by rail; and
- non-work trip dummy = 1 if the trip is not business related, 0 otherwise.

The estimation results are shown in Table 6.1. The in-vehicle travel time has the wrong sign although it is not statistically significant. The rail specific constant has a large positive value, which might have been affected by the sampling strategy. Specifically, if the sample for the home interview was not drawn randomly but intentionally included more rail users, the parameter of the rail specific constant will be overestimated.

Table 6.1 Estimation Results of RP Model

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient estimate</th>
<th>Standard error</th>
<th>t-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rail constant</td>
<td>1.74</td>
<td>0.318</td>
<td>5.47</td>
</tr>
<tr>
<td>Cost per person</td>
<td>-0.0257</td>
<td>0.00585</td>
<td>-4.39</td>
</tr>
<tr>
<td>In-vehicle time</td>
<td>0.000548</td>
<td>0.00347</td>
<td>0.16</td>
</tr>
<tr>
<td>Out-of-vehicle time</td>
<td>-0.0263</td>
<td>0.00595</td>
<td>-4.42</td>
</tr>
<tr>
<td>Number of transfers</td>
<td>-0.344</td>
<td>0.133</td>
<td>-2.59</td>
</tr>
<tr>
<td>Non-work trip dummy</td>
<td>-0.900</td>
<td>0.227</td>
<td>-3.97</td>
</tr>
</tbody>
</table>

\[ L(0) = -162.84 \quad L(\hat{\theta}) = -114.84 \quad \rho^2 = 0.295 \quad \bar{\rho}^2 = 0.258 \quad N = 235 \]
6.2.2 MLE for Ordered Categorical Dependent Variable

As mentioned in Section 6.1, two types of SP data were collected from each respondent: SP1 data are for choice between rail and car, while SP2 data are for choice between two rail alternatives. Since the choice indicator, or dependent variable, is a five point preference rating, an ordinary binary discrete choice model is inappropriate. The probit model with the ordered categorical dependent variable (McKelvey and Zavoina, 1975), or the ordered probit model may be used for this type of ordered categorical dependent variable, as described below.

The utility function for this model is given in equation (6.1). The stated preference is a categorical dependent variable, denoted by \( y \), and represented by:

\[
\begin{align*}
  y = 1; & \text{ if the response is "definitely choose alternative 1";} \\
  y = 2; & \text{ if the response is "probably choose alternative 1";} \\
  y = 3; & \text{ if the response is "not sure";} \\
  y = 4; & \text{ if the response is "probably choose alternative 2";} \text{ and} \\
  y = 5; & \text{ if the response is "definitely choose alternative 2".}
\end{align*}
\]

Since there are five categories in the dependent variables, we can identify four threshold values, \( \theta_1, \theta_2, \theta_3, \text{ and } \theta_4 \), in the utility scale such that:

\[
\begin{align*}
  P(y=1) &= P(\theta_4 \leq U) , \\
  P(y=2) &= P(\theta_3 \leq U < \theta_4) , \\
  P(y=3) &= P(\theta_2 \leq U < \theta_3) , \\
  P(y=4) &= P(\theta_1 \leq U < \theta_2) , \\
  P(y=5) &= P(U < \theta_1) .
\end{align*}
\]

Note that if the utility function has an intercept, then one of the four threshold parameters is not identifiable, so arbitrarily one of them is usually fixed to be zero.

Since the variable \( y \) in this case study is symmetrical with respect to category 3, it is reasonable to allow for only two threshold parameters and construct symmetrical threshold values around zero as depicted in Figure 6.1.
Accordingly, the probability that the choice indicator falls into each category is given by:

\[ P(y=1) = P(\theta_2 \leq U) = 1 - \Phi(\theta_2 - \beta'x) , \]  \hspace{1cm} (6.6) \\
\[ P(y=2) = P(\theta_1 \leq U < \theta_2) = \Phi(\theta_2 - \beta'x) - \Phi(\theta_1 - \beta'x) , \]  \hspace{1cm} (6.9) \\
\[ P(y=3) = P(\theta_1 \leq U < \theta_1) = \Phi(\theta_1 - \beta'x) - \Phi(-\theta_1 - \beta'x) , \]  \hspace{1cm} (6.10) \\
\[ P(y=4) = P(-\theta_2 \leq U < -\theta_1) = \Phi(-\theta_1 - \beta'x) - \Phi(-\theta_2 - \beta'x) , \]  \hspace{1cm} (6.11) \\
\[ P(y=5) = P(U < -\theta_2) = \Phi(-\theta_2 - \beta'x) . \]  \hspace{1cm} (6.12) \\

An MLE is used to jointly estimate the \( \beta \) and the \( \theta \)'s.

### 6.2.3 Estimation with SP Data

(1) Estimation with SP1 data

The first SP experiment was designed to collect information on mode switching behavior (from rail to car, or vice-versa) by presenting to the respondents hypothetical rail and car modes which are described by in-vehicle travel time and travel cost. However, since the respondent was instructed to refer to the trip reported in the RP questions, he or she may have considered
additional attributes such as out-of-vehicle time and the number of transfers that would have been required for the trip in evaluating the hypothetical alternatives. These additional attributes have the same values as reported in the RP questions. Thus, the model estimated from SP1 data includes these additional trip attribute variables which are not specified in the SP experiment.

There may also be a bias in the stated preferences toward the mode actually used, reflecting the inertia effect, justification of past behavior, or omitted attributes that are not captured by the included variables. This bias can be estimated by including a dummy variable which indicates the actual choice.

Accordingly, the estimated model has the following explanatory variables:

- a constant;
- travel cost by rail less travel cost per person by car (Guilders), as specified in the SP experiment;
- in-vehicle travel time by rail less in-vehicle travel time by car (minutes), as specified in the SP experiment;
- access and egress time by rail less parking time by car (minutes);
- number of transfers by rail;
- non-work trip dummy = 1 if the trip is not business related, 0 otherwise; and
- inertia dummy = 1 if rail was actually used, 0 if car was actually used.

The estimated coefficients of these variables and threshold parameters, $\theta_{11}$ and $\theta_{12}$, are shown in Table 6.2. The coefficients of all the variables are estimated with the expected signs. The rail specific constant is now negative but the coefficient of the inertia dummy is large enough to cancel it out. This implies that the intercept is negative for car users and positive for rail users, which may indicate the inertia effect of the actual choice in the SP experiment. The parameter of the non-work trip dummy is not significant, which implies that in the SP context the respondent did not consider the factors associated with the purpose of the trip. This can be interpreted as the evidence of the proposition that SP methods can extract pure information on
the trade-offs among presented attributes because the respondents do not take into consideration the various situational constraints that affect actual choices. The monetary value of in-vehicle travel time is 0.45 Guilders/min. or $13/hour. The threshold parameters are accurately estimated.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient estimate</th>
<th>Standard error</th>
<th>t-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rail constant</td>
<td>-0.966</td>
<td>0.122</td>
<td>-7.91</td>
</tr>
<tr>
<td>Cost per person</td>
<td>-0.00719</td>
<td>0.00136</td>
<td>-5.29</td>
</tr>
<tr>
<td>In-vehicle time</td>
<td>-0.00324</td>
<td>0.00119</td>
<td>-2.72</td>
</tr>
<tr>
<td>Out-of-vehicle time</td>
<td>-0.00661</td>
<td>0.00206</td>
<td>-3.22</td>
</tr>
<tr>
<td>Number of transfers</td>
<td>-0.0528</td>
<td>0.0322</td>
<td>-1.64</td>
</tr>
<tr>
<td>Non-work trip dummy</td>
<td>-0.0250</td>
<td>0.0775</td>
<td>-0.32</td>
</tr>
<tr>
<td>Inertia dummy</td>
<td>1.49</td>
<td>0.0222</td>
<td>21.5</td>
</tr>
<tr>
<td>$\theta_{11}$</td>
<td>0.0805</td>
<td>0.00959</td>
<td>8.40</td>
</tr>
<tr>
<td>$\theta_{12}$</td>
<td>0.478</td>
<td>0.0222</td>
<td>21.5</td>
</tr>
</tbody>
</table>

$L(0) = -2620.16 \quad L(\hat{\beta}) = -1645.91 \quad \hat{\rho}^2 = 0.372 \quad \overline{\rho}^2 = 0.369 \quad N = 1628$

(2) Estimation with SP2 data

The second SP experiment presented to the respondent two different rail services and asked for preferences in terms of a five point rating scale. Each rail service was described by in-vehicle travel time, travel cost, number of transfers, and ride comfort. Ride comfort is actually a
package of different aspects such as seating room and availability, quietness, smoothness of ride, heating/ventilation, and minibar service, but it is presented at only two levels and therefore coded as 0 for better level and 1 for worse level. Additional attributes from the RP data such as out-of-vehicle time are not included in the model because the alternatives are both rail trips and these attributes are equal for the two alternatives. For the same reason, a mode specific constant is also not included in this model. Consequently, this model only includes the four explanatory variables specified in the SP2 experiment as listed below:

- fare (Guilders);
- in-vehicle travel time (minutes);
- number of transfers; and
- ride comfort level (the less, the better).

The coefficients of these variables are jointly estimated with the threshold parameters, \( \theta_{21} \) and \( \theta_{22} \), using the MLE of the ordered probit model. As shown in Table 6.3, all the coefficients and the threshold parameters are estimated with very small standard errors. The calculated value of in-vehicle travel time is 0.2 Guilders/min. or \$6/hour, which seems reasonable.

### Table 6.3 Estimation Results of SP2 Model

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient estimate</th>
<th>Standard error</th>
<th>t-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost per person</td>
<td>-0.0807</td>
<td>0.00318</td>
<td>-25.3</td>
</tr>
<tr>
<td>In-vehicle time</td>
<td>-0.0159</td>
<td>0.00136</td>
<td>-11.7</td>
</tr>
<tr>
<td>Number of transfers</td>
<td>-0.156</td>
<td>0.0319</td>
<td>-4.91</td>
</tr>
<tr>
<td>Comfort</td>
<td>-0.500</td>
<td>0.0337</td>
<td>-14.8</td>
</tr>
<tr>
<td>( \theta_{21} )</td>
<td>0.0180</td>
<td>0.00298</td>
<td>6.04</td>
</tr>
<tr>
<td>( \theta_{22} )</td>
<td>0.277</td>
<td>0.0106</td>
<td>26.1</td>
</tr>
</tbody>
</table>

\( L(0) = -4774.58 \quad L(\hat{\beta}) = -3266.96 \quad \rho^2 = 0.315 \quad \hat{\rho}^2 = 0.314 \quad N = 2965 \)
6.3 Combined Estimation with RP and SP Data

6.3.1 Estimation with RP and SP1 Data

Following the framework presented in Chapter 3, the RP and SP models are specified as follows:

\[ U = \beta'x + \epsilon, \]  \hspace{1cm} \text{(6.13)}

**Specification of the RP model:**

\[ \tilde{U}_i = \beta'x + \gamma_1z + \nu, \]  \hspace{1cm} \text{(6.14)}

**Specification of the SP1 model:**

\[ \text{Relationship of the random terms:} \]
\[ \text{Var}(\epsilon) = \mu_1^2 \text{Var}(\nu), \]  \hspace{1cm} \text{(6.15)}

where \( x \) is the vector of the variables included in the RF model presented earlier, \( \tilde{x} \) contains a subset of \( x \) which is included in the SP model, \( z \) is a vector of SP specific variables, \( \beta \) and \( \gamma_1 \) are vectors of coefficients, \( \epsilon \) and \( \nu \) are normally distributed random terms, and \( \mu_1 \) is a parameter that represents the relative scale of the two models. The value \( \tilde{x} \) is obtained from \( x \) by omitting the rail constant and the non-work dummy. The vector \( z \) includes the following variables:

- a rail specific constant;
- a non-work trip dummy; and
- an inertia dummy.

Note that \( \beta \)'s are estimated from both RP and SP1 data, while the rail specific constants and the parameter of the non-work dummy are separately estimated from each data, and the coefficient of the inertia variable is estimated from SP1 data. The likelihood for the RP model is expressed by an ordinary binary probit model, while that for the SP model is expressed by the
ordered probit model with threshold parameters. The unknown parameter vectors $\beta$, $\gamma_1$, $\theta$ and the scalar $\mu_1$ can be estimated either by joint or sequential methods as described in Chapter 3.

Table 6.4 shows the results of the joint estimation. All the coefficients have the expected sign. $\mu_1$ is estimated between 0 and 1, which indicates a greater variance of the random utilities in the SP data. The calculated value of in-vehicle time is 0.15 Guilders/min. or $4.20$/hour.

The results of a sequential estimation are shown in Tables 6.5 and 6.6. The two step sequential estimators are obtained from Steps 1 and 2 (see Section 3.3.3) and the additional step (Step 3) produces more accurate estimators, i.e., the three step sequential estimators. It is shown that the three step method yields closer estimates to the values of the joint estimators than the two step method.
<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient estimate</th>
<th>Standard error</th>
<th>t-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rail constant (RP)</td>
<td>1.61</td>
<td>0.292</td>
<td>5.50</td>
</tr>
<tr>
<td>Rail constant (SP)</td>
<td>-2.96</td>
<td>0.860</td>
<td>-3.44</td>
</tr>
<tr>
<td>Cost per person</td>
<td>-0.0240</td>
<td>0.00486</td>
<td>-4.94</td>
</tr>
<tr>
<td>In-vehicle time</td>
<td>-0.00463</td>
<td>0.00241</td>
<td>-1.92</td>
</tr>
<tr>
<td>Out-of-vehicle time</td>
<td>-0.0250</td>
<td>0.00534</td>
<td>-4.68</td>
</tr>
<tr>
<td>Number of transfers</td>
<td>-0.245</td>
<td>0.0895</td>
<td>-2.73</td>
</tr>
<tr>
<td>Non-work trip dummy (RP)</td>
<td>-0.900</td>
<td>0.216</td>
<td>-4.18</td>
</tr>
<tr>
<td>Non-work trip dummy (SP)</td>
<td>-0.129</td>
<td>0.263</td>
<td>-0.49</td>
</tr>
<tr>
<td>Inertia dummy (SP)</td>
<td>4.84</td>
<td>1.16</td>
<td>4.16</td>
</tr>
<tr>
<td>$\theta_{11}$</td>
<td>0.0803</td>
<td>0.00953</td>
<td>8.42</td>
</tr>
<tr>
<td>$\theta_{12}$</td>
<td>0.476</td>
<td>0.0219</td>
<td>21.7</td>
</tr>
<tr>
<td>$\mu_1$</td>
<td>0.294</td>
<td>0.0689</td>
<td>4.27</td>
</tr>
</tbody>
</table>

$L(0) = -2783.00$  $L(\hat{\beta}) = -1763.56$  $\rho^2 = 0.366$  $\overline{\rho^2} = 0.362$  $N = 1863$
Table 6.5 Two Step Sequential Estimation Results with RP and SP1 Data

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient estimate</th>
<th>Standard error</th>
<th>t-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rail constant (RP)</td>
<td>1.31</td>
<td>0.242</td>
<td>5.40</td>
</tr>
<tr>
<td>Rail constant (SP)</td>
<td>-2.69</td>
<td>0.340</td>
<td>-7.91</td>
</tr>
<tr>
<td>Cost per person</td>
<td>-0.0200</td>
<td>0.00379</td>
<td>-5.29</td>
</tr>
<tr>
<td>In-vehicle time</td>
<td>-0.00903</td>
<td>0.00332</td>
<td>-2.72</td>
</tr>
<tr>
<td>Out-of-vehicle time</td>
<td>-0.0184</td>
<td>0.00574</td>
<td>-3.22</td>
</tr>
<tr>
<td>Number of transfers</td>
<td>-0.147</td>
<td>0.0898</td>
<td>-1.64</td>
</tr>
<tr>
<td>Non-work trip dummy (RP)</td>
<td>-0.919</td>
<td>0.204</td>
<td>-4.51</td>
</tr>
<tr>
<td>Non-work trip dummy (SP)</td>
<td>-0.0697</td>
<td>0.216</td>
<td>-0.32</td>
</tr>
<tr>
<td>Inertia dummy</td>
<td>4.15</td>
<td>0.0619</td>
<td>21.5</td>
</tr>
<tr>
<td>$\theta_{11}$</td>
<td>0.0805</td>
<td>0.00959</td>
<td>8.40</td>
</tr>
<tr>
<td>$\theta_{12}$</td>
<td>0.478</td>
<td>0.0222</td>
<td>21.5</td>
</tr>
<tr>
<td>$\mu_1$</td>
<td>0.359</td>
<td>0.0568</td>
<td>5.40</td>
</tr>
</tbody>
</table>

$L(0) = -2783.00 \quad L(\hat{\beta}) = -1766.36 \quad \rho^2 = 0.365 \quad \bar{\rho}^2 = 0.361 \quad N = 1863$
### Table 6.6 Three Step Sequential Estimation Results with RP and SP1 Data

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient estimate</th>
<th>Standard error</th>
<th>t-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rail constant (RP)</td>
<td>1.47</td>
<td>0.238</td>
<td>6.17</td>
</tr>
<tr>
<td>Rail constant (SP)</td>
<td>-2.38</td>
<td>0.296</td>
<td>-8.03</td>
</tr>
<tr>
<td>Cost per person</td>
<td>-0.0212</td>
<td>0.00308</td>
<td>-6.91</td>
</tr>
<tr>
<td>In-vehicle time</td>
<td>-0.00479</td>
<td>0.00223</td>
<td>-2.15</td>
</tr>
<tr>
<td>Out-of-vehicle time</td>
<td>-0.0223</td>
<td>0.00410</td>
<td>-5.45</td>
</tr>
<tr>
<td>Number of transfers</td>
<td>-0.212</td>
<td>0.0736</td>
<td>-2.87</td>
</tr>
<tr>
<td>Non-work trip dummy (RP)</td>
<td>-0.905</td>
<td>0.215</td>
<td>-4.21</td>
</tr>
<tr>
<td>Non-work trip dummy (SP)</td>
<td>-0.0899</td>
<td>0.212</td>
<td>-0.42</td>
</tr>
<tr>
<td>Inertia dummy</td>
<td>3.98</td>
<td>0.219</td>
<td>18.2</td>
</tr>
<tr>
<td>$\theta_{11}$</td>
<td>0.0804</td>
<td>0.00955</td>
<td>8.42</td>
</tr>
<tr>
<td>$\theta_{12}$</td>
<td>0.477</td>
<td>0.0220</td>
<td>21.7</td>
</tr>
<tr>
<td>$\mu_1$</td>
<td>0.359</td>
<td>0.0568</td>
<td>5.40</td>
</tr>
</tbody>
</table>

$L(0) = -2783.00$  
$L(\hat{\beta}) = -1763.89$  
$\rho^2 = 0.366$  
$\bar{\rho}^2 = 0.362$  
$N = 1863$

### 6.3.2 Estimation with RP and SP2 Data

The RP and SP models are specified as follows:

**Specification of the RP model:**

$$U = \beta'x + \epsilon,$$

(6.16)

**Specification of the SP2 model:**

$$\tilde{U}_2 = \beta'\tilde{x} + \gamma_2'\tilde{x} + \zeta,$$

(6.17)
Relationship of the random terms:

\[
\text{Var}(\varepsilon) = \mu_2^2 \text{Var}(\zeta),
\]

(6.18)

where \( \mathbf{x} \) is the vector of the variables included in the RP model presented earlier, \( \bar{x} \) contains a subset of \( \mathbf{x} \) which is included in the SP2 model, \( \bar{z} \) is a vector of SP specific variables, \( \beta \) and \( \gamma_2 \) are vectors of coefficients, \( \varepsilon \) and \( \zeta \) are normally distributed random terms, and \( \mu_2 \) is a parameter that represents the relative scale of the two models. \( \beta \) is estimated from both RP and SP2 data, while \( \gamma_2 \) is estimated from SP2 data.

Table 6.7 shows the estimation results. All the estimates have the expected signs and have very small standard errors. \( \mu_2 \) is estimated to be greater than 1, which indicates that SP2 data have less random noise than RP data. The calculated value of in-vehicle time is 0.2 Guilder/min. or $5.60/hour, which seems more reasonable than the one calculated from the RP and SP1 data.

The two step and three step sequential estimation results are presented in Tables 6.8 and 6.9, respectively. Since the scale parameter estimates from the sequential method almost coincided with the joint estimate, all the other estimates and the goodness-of-fit of the joint and sequential methods are nearly identical. However, comparing calculated standard errors from both estimation methods, we notice that those of the sequential estimation are underestimated by the factor of 10 to 500%.
Table 6.7 Joint Estimation Results with RP and SP2 Data

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient estimate</th>
<th>Standard error</th>
<th>t-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rail constant (RP)</td>
<td>1.44</td>
<td>0.244</td>
<td>5.93</td>
</tr>
<tr>
<td>Cost per person</td>
<td>-0.0242</td>
<td>0.00500</td>
<td>-4.85</td>
</tr>
<tr>
<td>In-vehicle time</td>
<td>-0.00474</td>
<td>0.00102</td>
<td>-4.64</td>
</tr>
<tr>
<td>Out-of-vehicle time</td>
<td>-0.0258</td>
<td>0.00561</td>
<td>-4.60</td>
</tr>
<tr>
<td>Number of transfers</td>
<td>-0.0487</td>
<td>0.0135</td>
<td>-3.61</td>
</tr>
<tr>
<td>Comfort (SP)</td>
<td>-0.150</td>
<td>0.0322</td>
<td>-4.67</td>
</tr>
<tr>
<td>Non-work trip dummy (RP)</td>
<td>-0.867</td>
<td>0.213</td>
<td>-4.08</td>
</tr>
<tr>
<td>$\theta_{21}$</td>
<td>0.0180</td>
<td>0.00300</td>
<td>6.04</td>
</tr>
<tr>
<td>$\theta_{22}$</td>
<td>0.277</td>
<td>0.0106</td>
<td>26.1</td>
</tr>
<tr>
<td>$\mu_2$</td>
<td>3.33</td>
<td>0.699</td>
<td>4.77</td>
</tr>
</tbody>
</table>

$L(0) = -4934.82$  $L(\hat{\beta}) = -3385.63$  $\rho^2 = 0.314$  $\widetilde{\rho^2} = 0.312$  $N = 3200$
Table 6.8 Two Step Sequential Estimation Results with RP and SP2 Data

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient estimate</th>
<th>Standard error</th>
<th>t-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rail constant (RP)</td>
<td>1.44</td>
<td>0.243</td>
<td>5.93</td>
</tr>
<tr>
<td>Cost per person</td>
<td>-0.0242</td>
<td>0.000952</td>
<td>-25.3</td>
</tr>
<tr>
<td>In-vehicle time</td>
<td>-0.00476</td>
<td>0.000407</td>
<td>-11.7</td>
</tr>
<tr>
<td>Out-of-vehicle time</td>
<td>-0.0258</td>
<td>0.00560</td>
<td>-4.61</td>
</tr>
<tr>
<td>Number of transfers</td>
<td>-0.0467</td>
<td>0.00955</td>
<td>-4.91</td>
</tr>
<tr>
<td>Comfort (SP)</td>
<td>-0.150</td>
<td>0.0101</td>
<td>-14.8</td>
</tr>
<tr>
<td>Non-work trip dummy (SP)</td>
<td>-0.919</td>
<td>0.204</td>
<td>-4.51</td>
</tr>
<tr>
<td>$\theta_{21}$</td>
<td>0.0180</td>
<td>0.00298</td>
<td>6.04</td>
</tr>
<tr>
<td>$\theta_{22}$</td>
<td>0.277</td>
<td>0.0106</td>
<td>26.1</td>
</tr>
<tr>
<td>$\mu_2$</td>
<td>3.34</td>
<td>0.688</td>
<td>4.85</td>
</tr>
</tbody>
</table>

$L(0) = -4934.82 \quad L(\hat{\beta}) = -3385.66 \quad \rho^2 = 0.314 \quad \bar{\rho}^2 = 0.312 \quad N = 3200$
Table 6.9 Three Step Sequential Estimation Results with RP and SP2 Data

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient estimate</th>
<th>Standard error</th>
<th>t-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rail constant (RP)</td>
<td>1.44</td>
<td>0.226</td>
<td>6.37</td>
</tr>
<tr>
<td>Cost per person</td>
<td>-0.0242</td>
<td>0.000937</td>
<td>-25.8</td>
</tr>
<tr>
<td>In-vehicle time</td>
<td>-0.00473</td>
<td>0.000402</td>
<td>-11.8</td>
</tr>
<tr>
<td>Out-of-vehicle time</td>
<td>-0.0258</td>
<td>0.00549</td>
<td>-4.70</td>
</tr>
<tr>
<td>Number of transfers</td>
<td>-0.0486</td>
<td>0.00947</td>
<td>-5.13</td>
</tr>
<tr>
<td>Comfort (SP)</td>
<td>-0.150</td>
<td>0.0101</td>
<td>-14.9</td>
</tr>
<tr>
<td>Non-work trip dummy (SP)</td>
<td>-0.870</td>
<td>0.209</td>
<td>-4.16</td>
</tr>
<tr>
<td>( \theta_{21} )</td>
<td>0.0180</td>
<td>0.00298</td>
<td>6.04</td>
</tr>
<tr>
<td>( \theta_{22} )</td>
<td>0.277</td>
<td>0.0106</td>
<td>26.1</td>
</tr>
<tr>
<td>( \mu_2 )</td>
<td>3.34</td>
<td>0.688</td>
<td>4.85</td>
</tr>
</tbody>
</table>

\[ L(0) = -4934.82 \quad L(\hat{\beta}) = -3385.63 \quad \rho^2 = 0.314 \quad \bar{\rho}^2 = 0.312 \quad N = 3200 \]

6.3.3 Estimation with RP, SP1, and SP2 Data

Since two independent SP data sets are used, the system is composed of the principal model and two auxiliary models:

**Specification of the RP model:**
\[ U = \beta'x + \varepsilon, \quad (6.19) \]

**Specification of the SP1 model:**
\[ \tilde{U}_1 = \beta'x + \gamma_1'z_1 + \nu, \quad (6.20) \]

**Specification of the SP2 model:**
\[ \tilde{U}_2 = \beta'x + \gamma_2'z_2 + \zeta, \quad (6.21) \]
Relationship of the random terms:

\[ \text{Var}(\varepsilon) = \mu_1^2 \text{Var}(v) = \mu_2^2 \text{Var}(\zeta) \]  \hspace{1cm} (6.22)

As shown in Table 6.10, all the parameters are accurately estimated with the expected signs. The value of in-vehicle time is calculated to be 0.2 Guilders/min. or $5.60/hour. The sequential estimation was performed in the following way: first, estimate \( \bar{U}_2 \) from the SP2 data, then estimate \( \bar{U}_1 \) with the SP1 data and the fitted value of \( \bar{V}_2 \), lastly estimate \( U \) with the RP data and the fitted value of \( \bar{V}_1 \). The three step method requires an additional step of pooling three data sets using \( \hat{\mu}_1 \) and \( \hat{\mu}_2 \) obtained in the former steps and jointly estimating all the utility functions. Tables 6.11 and 6.12 show the estimation results by the two step and three step sequential methods, respectively. It is shown that parameter estimates by the three estimation methods are very similar but standard errors of the sequential estimation are substantially underestimated.
Table 6.10 Joint Estimation Results with RP, SP1 and SP2 Data

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient estimate</th>
<th>Standard error</th>
<th>t-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rail constant (RP)</td>
<td>1.43</td>
<td>0.240</td>
<td>5.95</td>
</tr>
<tr>
<td>Rail constant (SP)</td>
<td>-3.18</td>
<td>0.835</td>
<td>-3.81</td>
</tr>
<tr>
<td>Cost per person</td>
<td>-0.0250</td>
<td>0.00462</td>
<td>-5.41</td>
</tr>
<tr>
<td>In-vehicle time</td>
<td>-0.00497</td>
<td>0.000973</td>
<td>-5.11</td>
</tr>
<tr>
<td>Out-of-vehicle time</td>
<td>-0.0247</td>
<td>0.00507</td>
<td>-4.86</td>
</tr>
<tr>
<td>Number of transfers</td>
<td>-0.0520</td>
<td>0.0138</td>
<td>-3.89</td>
</tr>
<tr>
<td>Comfort</td>
<td>-0.156</td>
<td>0.0301</td>
<td>-5.16</td>
</tr>
<tr>
<td>Non-work trip dummy (RP)</td>
<td>-0.864</td>
<td>0.211</td>
<td>-4.09</td>
</tr>
<tr>
<td>Non-work trip dummy (SP)</td>
<td>-0.0745</td>
<td>0.259</td>
<td>-0.29</td>
</tr>
<tr>
<td>Inertia dummy</td>
<td>4.90</td>
<td>1.12</td>
<td>4.36</td>
</tr>
<tr>
<td>$\theta_{11}$</td>
<td>0.0801</td>
<td>0.00951</td>
<td>8.43</td>
</tr>
<tr>
<td>$\theta_{12}$</td>
<td>0.476</td>
<td>0.0218</td>
<td>21.8</td>
</tr>
<tr>
<td>$\theta_{21}$</td>
<td>0.0180</td>
<td>0.00300</td>
<td>6.04</td>
</tr>
<tr>
<td>$\theta_{22}$</td>
<td>0.277</td>
<td>0.0106</td>
<td>26.1</td>
</tr>
<tr>
<td>$\mu_1$</td>
<td>0.298</td>
<td>0.0663</td>
<td>4.50</td>
</tr>
<tr>
<td>$\mu_2$</td>
<td>3.22</td>
<td>0.609</td>
<td>5.30</td>
</tr>
</tbody>
</table>

$L(0) = -7554.98 \quad L(\hat{\beta}) = -5034.21 \quad R^2 = 0.334 \quad \bar{R}^2 = 0.332 \quad N = 4828$
Table 6.11 Two Step Sequential Estimation Results with RP, SP1 and SP2 Data

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient estimate</th>
<th>Standard error</th>
<th>t-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rail constant (RP)</td>
<td>1.40</td>
<td>0.231</td>
<td>6.07</td>
</tr>
<tr>
<td>Rail constant (SP)</td>
<td>-3.24</td>
<td>0.336</td>
<td>-9.67</td>
</tr>
<tr>
<td>Cost per person</td>
<td>-0.0259</td>
<td>0.00102</td>
<td>-25.3</td>
</tr>
<tr>
<td>In-vehicle time</td>
<td>-0.00510</td>
<td>0.000436</td>
<td>-11.7</td>
</tr>
<tr>
<td>Out-of-vehicle time</td>
<td>-0.0232</td>
<td>0.00682</td>
<td>-3.40</td>
</tr>
<tr>
<td>Number of transfers</td>
<td>-0.0501</td>
<td>0.0102</td>
<td>-4.91</td>
</tr>
<tr>
<td>Comfort</td>
<td>-0.160</td>
<td>0.0108</td>
<td>-14.8</td>
</tr>
<tr>
<td>Non-work trip dummy (RP)</td>
<td>-0.857</td>
<td>0.209</td>
<td>-4.10</td>
</tr>
<tr>
<td>Non-work trip dummy (SP)</td>
<td>-0.0611</td>
<td>0.241</td>
<td>-0.24</td>
</tr>
<tr>
<td>Inertia dummy</td>
<td>4.94</td>
<td>0.241</td>
<td>20.5</td>
</tr>
<tr>
<td>$\theta_{11}$</td>
<td>0.0801</td>
<td>0.00951</td>
<td>8.43</td>
</tr>
<tr>
<td>$\theta_{12}$</td>
<td>0.476</td>
<td>0.0218</td>
<td>21.8</td>
</tr>
<tr>
<td>$\theta_{21}$</td>
<td>0.0180</td>
<td>0.00298</td>
<td>6.04</td>
</tr>
<tr>
<td>$\theta_{22}$</td>
<td>0.277</td>
<td>0.0106</td>
<td>26.1</td>
</tr>
<tr>
<td>$\mu_1$</td>
<td>0.297</td>
<td>0.0501</td>
<td>5.94</td>
</tr>
<tr>
<td>$\mu_2$</td>
<td>3.12</td>
<td>0.510</td>
<td>6.11</td>
</tr>
</tbody>
</table>

$L(0) = -7554.98 \quad L(\beta) = -5034.37 \quad \rho^2 = 0.334 \quad \bar{\rho}^2 = 0.332 \quad N = 4828$
Table 6.12 Three Step Sequential Estimation Results with RP, SP1 and SP2 Data

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient estimate</th>
<th>Standard error</th>
<th>t-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rail constant (RP)</td>
<td>1.44</td>
<td>0.207</td>
<td>6.96</td>
</tr>
<tr>
<td>Rail constant (SP)</td>
<td>-3.19</td>
<td>0.295</td>
<td>-10.8</td>
</tr>
<tr>
<td>Cost per person</td>
<td>-0.0258</td>
<td>0.000976</td>
<td>-26.4</td>
</tr>
<tr>
<td>In-vehicle time</td>
<td>-0.00513</td>
<td>0.000425</td>
<td>-12.1</td>
</tr>
<tr>
<td>Out-of-vehicle time</td>
<td>-0.0247</td>
<td>0.00432</td>
<td>-5.73</td>
</tr>
<tr>
<td>Number of transfers</td>
<td>-0.0537</td>
<td>0.0101</td>
<td>-5.33</td>
</tr>
<tr>
<td>Comfort</td>
<td>-0.161</td>
<td>0.0107</td>
<td>-15.0</td>
</tr>
<tr>
<td>Non-work trip dummy (RP)</td>
<td>-0.861</td>
<td>0.207</td>
<td>-4.15</td>
</tr>
<tr>
<td>Non-work trip dummy (SP)</td>
<td>-0.0656</td>
<td>0.249</td>
<td>-0.26</td>
</tr>
<tr>
<td>Inertia dummy</td>
<td>4.92</td>
<td>0.237</td>
<td>21.7</td>
</tr>
<tr>
<td>$\theta_{11}$</td>
<td>0.0802</td>
<td>0.00951</td>
<td>8.43</td>
</tr>
<tr>
<td>$\theta_{12}$</td>
<td>0.476</td>
<td>0.0218</td>
<td>21.9</td>
</tr>
<tr>
<td>$\theta_{21}$</td>
<td>0.0180</td>
<td>0.00298</td>
<td>6.04</td>
</tr>
<tr>
<td>$\theta_{22}$</td>
<td>0.277</td>
<td>0.0106</td>
<td>26.1</td>
</tr>
<tr>
<td>$\mu_1$</td>
<td>0.297</td>
<td>0.0501</td>
<td>5.94</td>
</tr>
<tr>
<td>$\mu_2$</td>
<td>3.12</td>
<td>0.510</td>
<td>6.11</td>
</tr>
</tbody>
</table>

$L(0) = -7554.98 \quad L(\tilde{\beta}) = -5034.23 \quad \rho^2 = 0.334 \quad \tilde{\rho}^2 = 0.332 \quad N = 4828$

6.4 Model Comparisons

The three combined models estimated by the joint and sequential methods are compared with the three separate models in Tables 6.13 and 6.14.
### Table 6.13 Comparison of Parameter Estimates by Joint Estimation
(t-statistics in parentheses)

<table>
<thead>
<tr>
<th></th>
<th>RP</th>
<th>SP1</th>
<th>SP2</th>
<th>RP+SP1</th>
<th>RP+SP2</th>
<th>RP+SP1+SP2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rail constant (RP)</td>
<td>1.74</td>
<td>1.61</td>
<td>1.44</td>
<td>1.43</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(5.47)</td>
<td>(5.50)</td>
<td>(5.93)</td>
<td>(5.95)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rail constant (SP)</td>
<td>-0.966</td>
<td>-2.96</td>
<td>-3.18</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(-7.91)</td>
<td>(-3.44)</td>
<td>(-3.81)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cost per person</td>
<td>-0.0257</td>
<td>-0.00719</td>
<td>-0.0240</td>
<td>-0.0242</td>
<td>-0.0250</td>
<td></td>
</tr>
<tr>
<td>(-4.39)</td>
<td>(-5.29)</td>
<td>(-25.3)</td>
<td>(-4.94)</td>
<td>(-4.85)</td>
<td>(-5.41)</td>
<td></td>
</tr>
<tr>
<td>In-vehicle time</td>
<td>0.000548</td>
<td>-0.00324</td>
<td>-0.00463</td>
<td>-0.00474</td>
<td>-0.00497</td>
<td></td>
</tr>
<tr>
<td>(0.16)</td>
<td>(-2.72)</td>
<td>(-11.7)</td>
<td>(-1.92)</td>
<td>(-4.64)</td>
<td>(-5.11)</td>
<td></td>
</tr>
<tr>
<td>Out-of-vehicle time</td>
<td>-0.0263</td>
<td>-0.00661</td>
<td>-0.0250</td>
<td>-0.0258</td>
<td>-0.0247</td>
<td></td>
</tr>
<tr>
<td>(-4.42)</td>
<td>(-3.22)</td>
<td>(-4.68)</td>
<td>(-4.60)</td>
<td>(-4.86)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of transfers</td>
<td>-0.344</td>
<td>-0.0528</td>
<td>-0.245</td>
<td>-0.0487</td>
<td>-0.0520</td>
<td></td>
</tr>
<tr>
<td>(-2.59)</td>
<td>(-1.64)</td>
<td>(-4.91)</td>
<td>(-2.75)</td>
<td>(-3.61)</td>
<td>(-3.89)</td>
<td></td>
</tr>
<tr>
<td>Comfort</td>
<td>-0.500</td>
<td>-0.150</td>
<td>-0.156</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(-14.8)</td>
<td>(-4.67)</td>
<td>(-5.16)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-work trip dummy (RP)</td>
<td>-0.900</td>
<td>-0.900</td>
<td>-0.870</td>
<td>-0.864</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(-3.97)</td>
<td>(-4.18)</td>
<td>(-4.08)</td>
<td>(-4.09)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-work trip dummy (SP)</td>
<td>-0.0250</td>
<td>-0.129</td>
<td>-0.0745</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(-0.32)</td>
<td>(-0.49)</td>
<td></td>
<td>(-0.29)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inertia dummy</td>
<td>1.49</td>
<td>4.84</td>
<td>4.90</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(17.5)</td>
<td>(4.16)</td>
<td></td>
<td>(4.36)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>θ₁₁</td>
<td>0.0805</td>
<td>0.0803</td>
<td>0.0801</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(8.40)</td>
<td>(8.42)</td>
<td></td>
<td>(8.43)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>θ₁₂</td>
<td>0.478</td>
<td>0.476</td>
<td>0.476</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(21.5)</td>
<td>(21.7)</td>
<td></td>
<td>(21.8)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>θ₂₁</td>
<td>0.0180</td>
<td>0.0180</td>
<td>0.0180</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(6.04)</td>
<td>(6.04)</td>
<td></td>
<td>(6.04)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>θ₂₂</td>
<td>0.277</td>
<td>0.277</td>
<td>0.277</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(26.1)</td>
<td>(26.1)</td>
<td></td>
<td>(26.1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>µ₁</td>
<td>0.294</td>
<td></td>
<td>0.298</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(4.27)</td>
<td></td>
<td></td>
<td>(4.50)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>µ₂</td>
<td></td>
<td>3.33</td>
<td>3.22</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(4.77)</td>
<td></td>
<td>(5.30)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ᴦ²</td>
<td>0.258</td>
<td>0.369</td>
<td>0.314</td>
<td>0.362</td>
<td>0.312</td>
<td>0.332</td>
</tr>
</tbody>
</table>
Table 6.14  Comparison of Parameter Estimates by Three Step Sequential Estimation
(t-statistics in parentheses)

<table>
<thead>
<tr>
<th></th>
<th>RP</th>
<th>SP1</th>
<th>SP2</th>
<th>RP+SP1</th>
<th>RP+SP2</th>
<th>RP+SP1+SP2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rail constant (RP)</td>
<td>1.74</td>
<td>1.47</td>
<td>1.44</td>
<td>1.44</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(5.47)</td>
<td>(6.17)</td>
<td>(6.37)</td>
<td>(6.96)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rail constant (SP)</td>
<td>-0.966</td>
<td>-2.38</td>
<td>-3.19</td>
<td>-0.0258</td>
<td>-0.0258</td>
<td></td>
</tr>
<tr>
<td>(-7.91)</td>
<td>(-8.03)</td>
<td>(-10.8)</td>
<td>(-26.4)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cost per person</td>
<td>-0.0257</td>
<td>-0.0807</td>
<td>-0.0212</td>
<td>-0.0242</td>
<td>-0.0258</td>
<td></td>
</tr>
<tr>
<td>(-4.39)</td>
<td>(-25.3)</td>
<td>(-6.91)</td>
<td>(-25.8)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>In-vehicle time</td>
<td>0.000548</td>
<td>-0.0159</td>
<td>-0.00479</td>
<td>-0.00473</td>
<td>-0.00513</td>
<td></td>
</tr>
<tr>
<td>(0.16)</td>
<td>(-11.7)</td>
<td>(-2.15)</td>
<td>(-11.8)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Out-of-vehicle time</td>
<td>-0.0263</td>
<td>-0.0223</td>
<td>-0.0258</td>
<td>-0.0247</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(-4.42)</td>
<td>(-5.45)</td>
<td>(-4.70)</td>
<td>(-5.73)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of transfers</td>
<td>-0.344</td>
<td>-0.212</td>
<td>-0.0486</td>
<td>-0.0537</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(-2.59)</td>
<td>(-2.87)</td>
<td>(-5.13)</td>
<td>(-5.33)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comfort</td>
<td>-0.500</td>
<td>-0.150</td>
<td>-0.161</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-14.8)</td>
<td>(-14.9)</td>
<td>(-15.0)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-work trip dummy (RP)</td>
<td>-0.900</td>
<td>-0.905</td>
<td>-0.870</td>
<td>-0.861</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(-3.97)</td>
<td>(-4.21)</td>
<td>(-4.16)</td>
<td>(-4.15)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-work trip dummy (SP)</td>
<td>-0.0250</td>
<td>-0.899</td>
<td></td>
<td></td>
<td>-0.0656</td>
<td></td>
</tr>
<tr>
<td>(-0.32)</td>
<td>(-0.42)</td>
<td></td>
<td></td>
<td></td>
<td>(-0.26)</td>
<td></td>
</tr>
<tr>
<td>Inertia dummy</td>
<td>1.49</td>
<td>3.98</td>
<td>4.92</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(17.5)</td>
<td>(18.2)</td>
<td>(20.7)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>θ_{11}</td>
<td>0.0805</td>
<td>0.0804</td>
<td>0.0802</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(8.40)</td>
<td>(8.42)</td>
<td>(8.43)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>θ_{12}</td>
<td>0.478</td>
<td>0.477</td>
<td>0.476</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(21.5)</td>
<td>(21.7)</td>
<td>(21.9)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>θ_{21}</td>
<td>0.0180</td>
<td>0.0180</td>
<td>0.0180</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(6.04)</td>
<td>(6.04)</td>
<td>(6.04)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>θ_{22}</td>
<td>0.277</td>
<td>0.277</td>
<td>0.277</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(26.1)</td>
<td>(26.1)</td>
<td>(26.1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>μ₁</td>
<td>0.359</td>
<td></td>
<td>0.297</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(5.40)</td>
<td></td>
<td>(5.94)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>μ₂</td>
<td></td>
<td>3.34</td>
<td>3.12</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(4.85)</td>
<td>(6.11)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>\bar{R}^2</td>
<td>0.258</td>
<td>0.369</td>
<td>0.314</td>
<td>0.362</td>
<td>0.312</td>
<td>0.332</td>
</tr>
</tbody>
</table>
The first step in evaluating the usefulness of the combined estimator is to inspect the estimated coefficients of the separate RP, SP1 and SP2 models. A comparison of equivalent coefficients among these three models reveals large differences in the scales of the estimated utilities; the scale of the SP2 model is about three times greater than the scale of the RP model and the scale of the RP model is about three times greater than the scale of the SP1 model. This observation is verified by the results of the combined estimators. The ratio of the scale parameters of the SP1 and the RP models is given by $\mu_1$ with an estimated value of about 0.3. The ratio of the scales of the SP2 and the RP models is given by $\mu_2$ with an estimated value of about 3.2. These results indicate that the respondents were able to sharply discriminate between alternative rail services in the SP2 experiments. On the other hand, the stated choices between rail and car alternatives in the SP1 experiment were subject to significantly greater unexplained variance. Thus, a simple SP experiment such as SP2, may yield reliable information about trade-offs among attributes. However, using this information to predict choices between rail and car requires additional data and an estimator that allows for scale differences among different data sources.

The most convincing demonstration of the important role that SP data can play in model estimation is provided by the estimated coefficient of the in-vehicle travel time variable. In the RP model this coefficient has an incorrect sign and is not significantly different from zero. (This is not an unusual occurrence in the estimation of mode choice models from RP data and is due to the limited variability of the difference between car and train in-vehicle time.) In the SP models the coefficients of in-vehicle time are negative and significantly different from zero. Thus, a combined estimator that controls for the difference in scales yields a usable negative coefficient of approximately -0.0005 which can now be used to predict the effects of changes in in-vehicle travel times.

The preference bias in the SP1 data toward the mode actually chosen was detected by the *inertia* variable. In the RP+SP1+SP2 model, for example, the rail specific constant estimated for the SP data is -3.77 for car users and 2.34 (= -3.77 + 6.11) for rail users. Thus, rail users
have an SP rail constant of 2.34, which is greater than the RP value of 1.36, while for car users the SP rail constant is -3.77 and this is significantly smaller than the RP value. This indicates that car users have a greater preference bias toward their current mode than rail users. In other words, car users have a greater inertia or exhibit a greater justification bias than rail users.

In order to test if SP data produce the same models as RP data, likelihood ratio tests for the combined RP and SP models were performed. The likelihood value of an unrestricted model is given by the sum of likelihoods of separately estimated models, while a combined model gives the likelihood of the corresponding restricted model. Note that this is not a test of strict equality between the RP and SP parameters. It is a test of the restrictions which are embedded in the combined model. According to the test results shown in Table 6.15, the null hypothesis of the combined RP and SP models cannot be rejected for any of the three cases at the 2% level of significance. This indicates that the information from the stated preferences as used in the combined model convey the same choice information as the revealed preferences. In other words, the introduction of parameters that capture justification bias, inertia effects and differences in random errors permitted the pooling of SP and RP data.

<table>
<thead>
<tr>
<th>Case</th>
<th>Test Statistics</th>
<th>Degrees of Freedom</th>
<th>P-Value</th>
<th>$\chi^2_{0.05}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. RP vs. SP1</td>
<td>5.62</td>
<td>3</td>
<td>0.13163</td>
<td>11.35</td>
</tr>
<tr>
<td>2. RP vs. SP2</td>
<td>7.66</td>
<td>2</td>
<td>0.02171</td>
<td>9.21</td>
</tr>
<tr>
<td>3. RP vs. SP1 vs. SP2</td>
<td>13.00</td>
<td>5</td>
<td>0.02338</td>
<td>15.09</td>
</tr>
</tbody>
</table>
Arc elasticities of rail share with respect to 10% increase in rail fare, rail in-vehicle time, car cost, or car in-vehicle time were calculated by using the sample enumeration method in the following way. First, the initial, or before-the-change, rail share is calculated by,

$$S_1 = \frac{1}{N} \sum_{n=1}^{N} \frac{1}{1+\exp(-\hat{\beta}'x_{1n})},$$  \hspace{1cm} (6.23)$$

where

- $S_1 =$ initial market share of rail;
- $\hat{\beta} =$ vector of parameter estimates; and
- $x_{1n} =$ vector of attributes from the RP data.

Likewise, the new, or after-the-change, rail share is,

$$S_2 = \frac{1}{N} \sum_{n=1}^{N} \frac{1}{1+\exp(-\hat{\beta}'x_{2n})},$$  \hspace{1cm} (6.24)$$

where

- $S_2 =$ new market share of rail; and
- $x_{2n} =$ modified vector of attributes (i.e., increasing one attribute in $x_{1n}$ by 10%).

Then, the arc elasticity for a 10% change in one attribute is calculated by,

$$\frac{10(S_2-S_1)}{S_1}.$$  \hspace{1cm} (6.25)$$

Table 6.16 shows the elasticities calculated from the three models estimated separately from each data set, the three jointly estimated combined models, and the three combined models estimated by the three step sequential method. The RP model has the wrong sign for the in-vehicle travel time elasticities because the parameter estimate of that variable had the wrong sign. All of the three combined models show reasonable and stable values.
Table 6.16 Elasticities of Rail Share

<table>
<thead>
<tr>
<th>Model</th>
<th>rail cost</th>
<th>rail time</th>
<th>car cost</th>
<th>car time</th>
</tr>
</thead>
<tbody>
<tr>
<td>RP</td>
<td>-0.643</td>
<td>0.0402</td>
<td>0.305</td>
<td>-0.0397</td>
</tr>
<tr>
<td>SP1</td>
<td>-0.415</td>
<td>-0.530</td>
<td>0.234</td>
<td>0.571</td>
</tr>
<tr>
<td>SP2</td>
<td>-1.50</td>
<td>-0.939</td>
<td>1.18</td>
<td>1.12</td>
</tr>
<tr>
<td>RP+SP1 (joint est.)</td>
<td>-0.605</td>
<td>-0.337</td>
<td>0.288</td>
<td>0.342</td>
</tr>
<tr>
<td>RP+SP2 (joint est.)</td>
<td>-0.628</td>
<td>-0.356</td>
<td>0.297</td>
<td>0.360</td>
</tr>
<tr>
<td>RP+SP1+SP2 (joint est.)</td>
<td>-0.648</td>
<td>0.373</td>
<td>0.288</td>
<td>0.342</td>
</tr>
<tr>
<td>RP+SP1 (seq. est.)</td>
<td>-0.557</td>
<td>-0.361</td>
<td>0.265</td>
<td>0.366</td>
</tr>
<tr>
<td>RP+SP2 (seq.est.)</td>
<td>-0.627</td>
<td>-0.355</td>
<td>0.296</td>
<td>0.360</td>
</tr>
<tr>
<td>RP+SP1+SP2 (seq.est.)</td>
<td>-0.664</td>
<td>-0.382</td>
<td>0.313</td>
<td>0.388</td>
</tr>
</tbody>
</table>

6.5 Discussion and Summary

An empirical analysis of combined estimation with RP and SP data was presented. The data for this analysis was obtained from a survey of intercity travelers in The Netherlands and concerned the choice between rail and car. The survey included the usual RP data and two SP experiments: rail vs. car (SP1) and rail vs. rail (SP2).

Three combined models were estimated: RP data combined with SP1 (rail vs. car) data, RP data combined with SP2 (rail vs. rail) data, and RP data combined with both SP1 and SP2 data. These combined models were compared against the three models that were separately estimated from the three data sets. The RP model could not successfully identify an important parameter (the coefficient of in-vehicle travel time), which is a typical problem encountered in
estimating models from RP data. This is usually caused by lack of variation in the data and/or misspecification of the model. However, obtaining an acceptable model specification is often very difficult because the actual behavior is influenced by numerous related attributes while the available data are limited. Furthermore, even if the correct model specification was known, estimation of model parameters could fail because of limitations in the data. SP experiments present simplified hypothetical choice contexts and, therefore, may provide useful information on trade-offs among attributes.

The case study provided a clear demonstration of the usefulness of the combined models. Specifically, the coefficient of the in-vehicle travel time variable was successfully estimated by combining RP and SP data. The SP2 experiment for rail vs. rail choice provided information on the trade-offs among attributes with the least random noise. On the other hand, SP data are often not reliable because of the oversimplified hypothetical circumstances. This problem was mitigated by using additional variables from the RP data in using the SP data.

A potential bias in the SP data was captured by the introduction of the inertia variable in the models. This variable captured the preference bias toward the mode actually chosen. As discussed above, it was found that car users had a greater inertia or habitual effect in choosing a travel mode.

Likelihood ratio test results gave a statistical justification of combining RP and SP data. The combined models also produced the most reasonable and consistent values for intuitive measurements such as value of time and elasticities.

Three alternative estimation methods for combined models were presented: joint estimation, two step sequential estimation, and three step sequential estimation. Although yielding statistically efficient estimators, joint estimation requires an ad hoc MLE program, while sequential estimation can be performed by ordinary econometric packages. This case study shows that sequential methods, especially the three step method, produced very similar estimates to those of the joint estimation, whereas standard errors of parameters reported by sequential methods were substantially underestimated.
This case study has clearly demonstrated the effectiveness and practicality of the methodology to combine RP and SP data. It was shown that this estimation strategy exploits the advantages of both revealed and stated preferences.
Chapter 7

Incorporating Latent Variables in Choice Models

An integrated framework for analysis of consumer behavior was presented in Chapter 1. The preceding part of this thesis has focused on utilizing stated preference data which is one type of psychometric data included in the framework. This chapter presents a methodology for utilizing other types of psychometric data such as perceptual and attitudinal indicators in discrete choice analysis. An empirical analysis is also performed using the intercity travel mode choice data described in Chapter 6.

7.1 Framework for Incorporating Perceptual and Attitudinal Data

7.1.1 General Framework for Choice Models

As reviewed in Chapter 2, most empirical choice studies in transportation demand analysis have made little use of perceptual and attitudinal data although such data could be used to improve the definition of attributes and taste heterogeneity. In practical applications, it would be useful to develop a model that integrates perceptions, attitudes and choice behavior if it is shown that such a model is empirically practical and provides useful additional information.
Figure 7.1 Framework for Incorporating Perceptual and Attitudinal Data
Figure 7.1 shows a general framework for incorporating perceptual and attitudinal data in a choice model, which is a part of the integrated framework for analyzing consumer behavior presented in Figure 1.2. In this framework the psychometric data are linked to latent perceptions and attitudes by measurement relations. Since latent perceptions, attitudes, and utility cause actual behavior, they are linked by structural relations. The next section demonstrates how this general framework can be formulated by presenting an application example to a travel mode choice model.

7.1.2 Example of Applying the General Framework to Mode Choice Models

This section adapts an example of incorporating psychometric data in a mode choice model from McFadden (1986) and Ben-Akiva and Boccara (1987) to show a prototype of the methodology proposed in the subsequent sections.

Consider a travel mode choice model with two alternatives. Assume that the utility of each alternative depends on generalized travel cost (which is a function of measurable attributes such as travel time and travel cost) and the non-measured attribute of travel time reliability. Cost consciousness — a latent attitudinal variable — is hypothesized to be an important determinant of preferences. In this model, assume that perceptions and attitudes can be scaled metrically. The following notation is used:

\[ c = \text{travel cost}; \]
\[ z = \text{a vector of other service attributes (such as in-vehicle travel time, walking distance, scheduled headway, etc.)}; \]
\[ q^* = \text{perceived travel time reliability (a latent variable)}; \]
\[ q = \text{a vector of indicators of perceived travel time reliability (responses to one or more questions about perceptions of service reliability)}; \]
\[ w = \text{a vector of socioeconomic characteristics}; \]
\[ a^* = \text{cost consciousness (a latent variable)}; \]
a = a vector of indicators of cost consciousness (responses to one or more attitudinal questions about the importance of the cost of travel);

\( u^* = \) relative utility of the two modes of travel; and

\( d = \) choice of travel mode (revealed preference).

This model is schematically represented in Figure 7.2 as a special case of the general framework shown in Figure 7.1.

The linkages between the variables can, for example, be formulated by the following set of linear structural equations:

\[ q^* = B_1 z + \zeta_1 \] \hspace{1cm} (7.1)

\[ a^* = B_2 w + \zeta_2 \] \hspace{1cm} (7.2)

\[ u^* = (\gamma_1 + \gamma_2 a^*) c + \gamma_3 q^* + \Gamma z + \nu \] \hspace{1cm} (7.3)

and the following three measurement equations which express the relationships with the indicators:

\[ q = \Lambda_1 q^* + \varepsilon_1 \] \hspace{1cm} (7.4)

\[ a = \Lambda_2 a^* + \varepsilon_2 \] \hspace{1cm} (7.5)

\[ d = \begin{cases} 
1, & \text{if } u^* \geq 0 \\
-1, & \text{if } u^* < 0 
\end{cases} \] \hspace{1cm} (7.6)

where \( B_1, B_2, \Gamma, \Lambda_1, \) and \( \Lambda_2 \) denote vectors of unknown parameters, \( \gamma_1, \gamma_2, \) and \( \gamma_3 \) are scalar unknown parameters, \( \zeta_1, \zeta_2, \) and \( \nu \) are disturbances and \( \varepsilon_1 \) and \( \varepsilon_2 \) are vectors of disturbances.

Equations (7.1) and (7.2) relate the latent attitude and perception variables to the measured causes and the measurement equations (7.4) and (7.5) express the measured indicators as functions of the latent variables. These four equations constitute a special case of a structural equation model with latent variables (e.g., Joreskog, 1973; Everitt, 1984). Equation (7.3) describes the utility, which is, up to an additive disturbance, a linear function of the attributes when the coefficient of travel cost is a linear function of cost consciousness \( a^* \).
Figure 7.2 Framework for Example Model
The model outlined in equations (7.1) - (7.6) applies to the choice behavior of an individual. Assume that the unknown parameters are common to all individuals in the market segment and that the unspecified variations in choice behavior across individuals are captured by the random disturbances in these equations. Thus, this model is homogeneous and can be estimated from pooled data across individuals.

Consider the reduced form utility model obtained by substituting in (7.3) expressions (7.1) and (7.2) to obtain:

\[ u^* = (\gamma_1 + \gamma_2 B_2 w) c + (\gamma_3 B_1 + \Gamma) z + v_1, \quad (7.7) \]

where

\[ v_1 = \gamma_2 c \zeta_2 + \gamma_3 \zeta_1 + \nu. \quad (7.8) \]

The utility is now expressed as a function of observable variables but note that the random utility component \( v_1 \) combines the original utility disturbance with the disturbances from equations (7.1) and (7.2). Thus, if \( v_1 \) has a parametric distribution, then equations (7.6) and (7.7) define a discrete choice model with the choice probability expressed as a function of measured socioeconomic variables and service attributes. The parameters in this equation combine the effects of perceptions and attitudes but they do not permit the identification of the separate contributions of these factors. This reduced form model does not exploit the information that may be obtained in the indicators of perceptions and attitudes. Note, however, that if the disturbance \( \zeta_2 \) is significant then the random utility element is not independent of the included variables.

Given data on perceptual and attitudinal indicators, it is possible to first estimate (7.1), (7.2), (7.4), and (7.5) using a technique for a structural equation model with latent variables such as LISREL, as described in the next section. Fitted values of \( a^* \) and \( q^* \), obtained in this fashion can be substituted into the utility function (7.3) and a discrete choice model can now be applied to equations (7.3) and (7.6) to obtain estimates of the utility parameters. This procedure contains a degree of statistical inconsistency due to the use of fitted values of \( a^* \) and \( q^* \) in the choice model. It is possible, however, to adopt a consistent estimator for the choice
model that adjusts for these measurement errors in the values of the variables (Train, McFadden and Goett, 1986 and McFadden, 1986). A fully efficient estimator can be obtained by jointly estimating (7.1) - (7.6). However, this approach could not rely on existing estimation procedures and would be computationally more demanding.

7.2 Discrete Choice Model with Latent Explanatory Variables

7.2.1 Review of Linear Structural Equations with Latent Variables

A linear structural equation model is a set of linear equations used to specify phenomena in terms of presumed cause-and-effect variables (see, for example, Duncan, 1975, and Bielby and Hauser, 1977). In its most general form, the model allows for variables that cannot be measured directly, or latent variables. Computer software for estimating structural equation models has been developed by Joreskog and Sorbom (1984) and is known as the LISREL (Linear Structural Relationships) system.

A model with latent variables is composed of two types of linear equations: structural equations and measurement equations. The structural equations specify relationships between cause-and-effect variables. Since cause-and-effect variables are sometimes not directly observable (e.g., comfort), or latent, identifying these latent variables requires observable indicators. The measurement equations relate latent variables and their indicators. The framework of the model is given by:

**Structural Equation:**

\[ \eta = B\eta + \zeta \]  

(7.9)

**Measurement Equation:**

\[ Y = \Lambda\eta + \varepsilon \]  

(7.10)
where
\[ \eta = m \times 1 \text{ vector of latent variables}; \]
\[ Y = n \times 1 \text{ vector of observed variables}; \]
\[ B = m \times m \text{ matrix of structural coefficients}; \]
\[ \Lambda = n \times m \text{ matrix of measurement coefficients}; \]
\[ \zeta = m \times 1 \text{ vector of disturbances; and} \]
\[ \epsilon = n \times 1 \text{ vector of measurement errors}. \]

The random variables, \( \eta, \zeta, \epsilon, \) and \( Y \), are assumed normally distributed with mean zero and covariance matrices:
\[
\text{Cov}(\eta) = \Phi \quad \text{Cov}(\zeta) = \Psi \quad \text{Cov}(\epsilon) = \Theta \quad \text{Cov}(Y) = \Sigma. \tag{7.11}
\]

Then,
\[
\Phi = E[\eta \eta']
\quad = E[(I-B)^{-1}\zeta'(I-B)^{-1}]
\quad = (I-B)^{-1}\Psi(I-B)^{-1}', \tag{7.12}
\]

and
\[
\Sigma = E[YY']
\quad = E[(\Lambda \eta + \epsilon)(\Lambda \eta + \epsilon)']
\quad = \Lambda E[\eta \eta'] \Lambda' + E[\epsilon \epsilon']
\quad = \Lambda \Phi \Lambda' + \Theta. \tag{7.13}
\]

Substituting (7.12) into (7.13) gives:
\[
\Sigma = \Lambda \Phi \Lambda' + \Theta
\quad = \Lambda (I-B)^{-1}\Psi(I-B)^{-1}' \Lambda' + \Theta. \tag{7.14}
\]

Given a sample of observations of \( Y \), we calculate the sample covariance \( S \) and search for a fitted value of \( \Sigma \), denoted by \( \hat{\Sigma} \), which is expressed in terms of estimated values of the unknown parameters, \( B, \Lambda, \Psi, \) and \( \Theta \) as in equation (7.14). An iterative procedure is used to find the values of the parameters that minimize the discrepancy between \( \hat{\Sigma} \) and \( S \).
7.2.2 The Model

The framework of the linear structural equation with latent variables model is now combined with a discrete choice model with latent explanatory variables. A binary choice model is used in the following illustration to simplify the exposition, where all the variables in the utility are expressed in terms of differences between the two alternatives. The framework is given by:

**Structural Equations:**

\[ u^* = a + b'x + c'x^* + \nu \]
\[ x^* = Bz + \zeta \]

where

- \( u^* \) = latent utility;
- \( x \) = vector of observable explanatory variables;
- \( x^* \) = vector of latent explanatory variables;
- \( z \) = vector of observable variables that influence \( x^* \);
- \( a, b, c, B \) = arrays of unknown parameters;
- \( \nu \) = random component of utility distributed \( N(0,1) \);
- \( \zeta \) = vector of disturbances distributed \( MVN(0,\Psi) \);

**Measurement Equations:**

\[ d = \begin{cases} 
1, & \text{if } u^* \geq 0 \\
-1, & \text{if } u^* < 0 
\end{cases} \]
\[ Y = \Lambda x^* + \epsilon \]

where

- \( Y \) = vector of observed indicators of \( x^* \);
- \( \Lambda \) = matrix of unknown parameters; and
\[ \varepsilon = \text{vector of disturbances distributed } \text{MVN}(0, \Theta). \]

The framework proposed above contains two models; one is a binary discrete choice model that consists of a structural equation (7.15) and a measurement equation (7.17), and the other is a linear system with latent variables that consists of a structural equation (7.16) and a measurement equation (7.18).

### 7.2.3 Covariance Structure

From equations (7.15) - (7.18), the joint distribution of \( Y, x^* \) and \( u^* \) given \( z \) and \( x \) is:

\[
\begin{bmatrix}
Y \\
x^* \\
u^*
\end{bmatrix} = \text{MVN}(M_1, \Omega_1),
\]

(7.19)

where

\[
M_1 = \begin{bmatrix}
\Lambda Bz \\
Bz \\
a + b'x + c'Bz
\end{bmatrix}
\]

(7.20)

and

\[
\Omega_1 = \begin{bmatrix}
\Lambda \Psi \Lambda' + \Theta & \Lambda \Psi & \Lambda \Psi c \\
\Psi \Lambda' & \Psi & \Psi c \\
c'\Psi \Lambda' & c'\Psi & 1 + c'\Psi c
\end{bmatrix}.
\]

(7.21)

Then, the conditional distribution of \( x^* \) and \( u^* \) given \( Y, z \) and \( x \) is (see, for example, Johnson and Wichern, 1988, for the proof):

\[
\begin{bmatrix}
x^* \\
u^*
\end{bmatrix} = \text{MVN}(M_2, \Omega_2),
\]

(7.22)

where

\[
M_2 = \begin{bmatrix}
Bz + \Psi \Lambda'[\Lambda \Psi \Lambda' + \Theta]^{-1}(Y - \Lambda Bz) \\
a + b'x + c'\left(Bz + \Psi \Lambda'[\Lambda \Psi \Lambda' + \Theta]^{-1}(Y - \Lambda Bz)\right)
\end{bmatrix}
\]

(7.23)

and

\[
\Omega_2 = \begin{bmatrix}
\Psi - \Psi \Lambda'[\Lambda \Psi \Lambda' + \Theta]^{-1} \Lambda \Psi & \Psi c - \Psi \Lambda'[\Lambda \Psi \Lambda' + \Theta]^{-1} \Lambda \Psi c \\
c'\Psi - c'\Psi \Lambda'[\Lambda \Psi \Lambda' + \Theta]^{-1} \Lambda \Psi & 1 + c'\Psi c - c'\Psi \Lambda'[\Lambda \Psi \Lambda' + \Theta]^{-1} \Lambda \Psi c
\end{bmatrix}.
\]

(7.24)
Hence, the choice probability of the discrete choice model given \( Y, z \) and \( x \) is:

\[
P(d|x, Y, z) = \Phi \left( d \frac{a+b'x+c' \left( Bz + \Psi \Lambda' (\Lambda \Psi \Lambda' + \Theta)^{-1} (Y - \Lambda Bz) \right)}{\sqrt{1 + c' \Psi c - c' \Psi \Lambda' (\Lambda \Psi \Lambda' + \Theta)^{-1} \Lambda \Psi c}} \right). \tag{7.25}
\]

### 7.2.4 Estimation Method

Since the measurement equation of the choice model, (7.17), is non-linear, the whole system of equations (7.15) - (7.18) cannot be estimated simultaneously with an existing program such as LISREL. Instead, the practical two step estimation method described below will yield consistent but not fully efficient estimators.

1. Use a LISREL type estimator to estimate (7.16) and (7.18) and calculate the fitted values:

\[
\begin{align*}
\hat{x}^* &= \hat{B}z + \hat{\Psi} \Lambda' (\Lambda \hat{\Psi} \Lambda' + \hat{\Theta})^{-1} (Y - \Lambda \hat{B}z), \tag{7.26} \\
\hat{\omega} &= \hat{\Psi} - \hat{\Psi} \Lambda' (\Lambda \hat{\Psi} \Lambda' + \hat{\Theta})^{-1} \Lambda \hat{\Psi}. \tag{7.27}
\end{align*}
\]

2. Use a probit MLE to estimate the model of (7.25) using \( \hat{x}^* \) and \( \hat{\omega} \), namely, estimate \( a \), \( b \) and \( c \) using the following choice probability:

\[
P(d|x, Y, z) = \Phi \left( d \frac{a+b'x+c'\hat{x}^*}{\sqrt{1 + c'\hat{\omega} c}} \right). \tag{7.28}
\]

The second step requires a correction of the standard maximum likelihood covariance estimates for the use of fitted values of \( x^* \) and \( \omega \). The corrected covariance matrix of the parameter estimates is calculated from the following equation (see McFadden, 1989):

\[
\Sigma_{\text{cor}} = \Sigma_{\text{std}} + \Sigma_{\text{std}} C \Sigma_{\text{step1}} C' \Sigma_{\text{std}}, \tag{7.29}
\]

where

\[
\begin{align*}
\Sigma_{\text{cor}} &= \text{corrected covariance matrix of the step 2 estimates;} \\
\Sigma_{\text{std}} &= \text{uncorrected covariance matrix of the step 2 estimates;} \\
\Sigma_{\text{step1}} &= \text{covariance matrix of the step1 estimates; and}
\end{align*}
\]
C is a matrix estimated by the sample mean of:

\[ v_1v_2 \frac{\phi(v)^2}{\Phi(v)\Phi(-v)}, \]

(7.30)

where

\[ v = \frac{a + b'x + c'x^*}{\sqrt{1 + c'}}, \]

(7.31)

\[ v_1 = \text{vector of derivatives of } v \text{ with respect to } a, b \text{ and } c; \] and

\[ v_2 = \text{vector of derivatives of } v \text{ with respect to the parameter estimated in step 1.} \]

7.3 Empirical Analysis

7.3.1 Specification of Model

The RP data of the Nijmegen Randstad intercity travel mode choice survey are used for an empirical study of the model. Questions for several perceptual indicators as well as ordinary travel attributes were included in a survey of the choice of rail versus car for an intercity trip. The survey consisted of a sample of 235 individuals. (For more detailed descriptions of the survey, see Chapter 6 and Appendix D). The five perceptual indicators, which are i) - v) listed below, are described by five point ratings such as 1) very poor, 2) poor, 3) neutral, 4) good, and 5) very good, and the overall evaluation of the mode is rated by a 10 point scale:

i) relaxation during the trip (Relax);

ii) reliability of the arrival time (Relia);

iii) flexibility of choosing departure time (Flex);

iv) ease of traveling with children and/or heavy baggage (Ease);

v) safety during the trip (Safety); and

vi) overall rating of the mode (Overall).
In addition to the observable trip attributes used in the RP model in Chapter 6, namely, travel cost per person, in-vehicle travel time, out-of-vehicle travel time, the number of transfers by rail, non-work trip dummy, and rail constant, two latent explanatory variables are included in the model: *convenience* and *ride comfort*. These variables are specified as follows:

- **convenience**

\[
x_1^* = \beta_1 ovtime + \beta_2 ovtime^2 + \beta_3 acp + \beta_4 egrpt + \beta_5 xfern + \beta_6 nonwork + \beta_7 female + \beta_8 age + \zeta_1,
\]

(7.32)

- **ride comfort**

\[
x_2^* = \beta_9 ivtime + \beta_{10} ivtime^2 + \beta_{11} costivt + \beta_{12} first + \beta_{13} nonwork + \beta_{14} female + \beta_{15} age + \zeta_2,
\]

(7.33)

where

- `ovtime` = out-of-vehicle travel time (minutes);
- `acpt` = 1 if access to rail is by public transit, 0 otherwise;
- `egrpt` = 1 if egress from rail is by public transit, 0 otherwise;
- `xfern` = number of transfers;
- `nonwork` =1 if non-work trip, 0 otherwise;
- `female` = 1 if female, 0 if male;
- `age` = 1 if 40 years old or older, 0 otherwise;
- `ivtime` = in-vehicle travel time (minutes);
- `costivt` = `costpp` `ivtime`;
- `costpp` = travel cost per person (Guilder); and
- `first` = 1 if rail trip is made by first class, 0 otherwise.

Note that (7.32) and (7.33) construct the structural equations of the LISREL model.

The aforementioned six perceptual indicators are used as observable variables in the measurement equation of the LISREL, and are specified as:
\[
\begin{bmatrix}
Y_1 (Relax) \\
Y_2 (Relia) \\
Y_3 (Flex) \\
Y_4 (Ease) \\
Y_5 (Safety) \\
Y_6 (Overall)
\end{bmatrix} = \begin{bmatrix}
0 & 1 \\
1 & 0 \\
\lambda_{31} & 0 \\
\lambda_{41} & 0 \\
0 & \lambda_{52} \\
\lambda_{61} & \lambda_{62}
\end{bmatrix} \begin{bmatrix}
x_1^* \\
x_2^*
\end{bmatrix} + \begin{bmatrix}
\epsilon_1 \\
\epsilon_2 \\
\epsilon_3 \\
\epsilon_4 \\
\epsilon_5 \\
\epsilon_6
\end{bmatrix},
\]

(7.35)

where each column of \( \Lambda \) has an element normalized to be one to fix the scale of each latent variable.

The framework of the model described above is schematically represented in Figure 7.3.

### 7.3.2 Estimation Results

First, the LISREL part was estimated by MLE. As shown in Table 7.1, most parameters have expected signs and fairly large t-statistics.
Figure 7.3 Framework for the Estimated Model
Table 7.1 Parameter Estimates of the LISREL Model
(t-statistics in parentheses)

\[ \hat{\Lambda} = \begin{bmatrix}
(\text{convenience}) & (\text{comfort}) \\
0 & 1 & (\text{relax}) \\
1 & 0 & (\text{relia}) \\
1.49 & 0 & (\text{flex}) \\
1.06 & 0 & (\text{ease}) \\
0 & 0.690 & (\text{safe}) \\
2.50 & 2.26 & (\text{overall}) \\
\end{bmatrix} \]

\[ \hat{\Psi} = \begin{bmatrix}
0.251 & 0 \\
(3.40) & (0) \\
0 & 0.298 \\
(2.47) & (0) \\
\end{bmatrix} \]

\[ \hat{\Theta} = \text{diag} \begin{bmatrix}
1.33 \\
(8.46) \\
0.913 \\
(9.28) \\
0.786 \\
(6.70) \\
1.40 \\
(9.72) \\
0.673 \\
(8.63) \\
1.58 \\
(2.91) \\
\end{bmatrix} \]
The fitted values of $x_1^*$ and $x_2^*$ are calculated by (7.26), and then the binary probit model with these fitted variables and calculated $\Omega$ by (7.27) is estimated by MLE using (7.28). The estimation results of the models with and without these latent variables are shown in Tables 7.2 and 7.3, respectively. They show that the latent variables have significant positive parameters (since the standard errors are not corrected, t-statistics are likely to be slightly overestimated), and consequently the goodness-of-fit measure of the model with the latent variables is significantly better than that of the model without them. This indicates that these latent variables are important factors in choosing a travel mode. In addition, the magnitude of parameter estimates for observable variables in the model with the latent variables is overall greater, which implies that the random utility component has a smaller variance.

Judging from estimation results of both the LISREL type model and discrete choice model, it can be concluded that this empirical study is a clear demonstration of effectiveness and practicality of the method proposed above to incorporate latent variables in a choice model.

Fitted values of $x_1^*$ and $x_2^*$ can also be calculated separately from structural and measurement equations. Especially, using only the structural equation has a meaning because the psychometric indicators are usually not available when the model is used to forecast. This is equivalent to infinitely noisy psychometric indicators, or $\Theta$ is positive infinite, and hence the variance correction factor in the choice model, $\omega$, is given by $\Psi$ (see equation (7.27)). Estimation results with the latent variables fitted separately from structural and measurement equations are shown in Tables 7.4 and 7.5, respectively. The latent variables fitted from the structural equations (Table 7.4) have parameter estimates with the correct sign but are insignificant and the fit of the model is substantially poorer than the model shown in Table 7.2.

On the other hand, the model with the latent variables fitted from the measurement equations show a slightly better fit than the model shown in Table 7.2. These two results indicate that the structural equations of the model are so noisy that the fitted values from the structural equations, $\hat{B}z$, do not play a role in equation (7.26). In addition, since variables used in the structural equations of the model are also included in the utility function of the
choice model, the fitted values from the structural equation do not provide much additional information about the utility function.

The high explanatory power of the latent variables fitted from the measurement equations may have resulted from the following hypothesis: those psychometric indicators are influenced by actual choice. In other words, this hypothesis states that the respondent overstates the values of the psychometric indicators of the chosen mode to justify his or her behavior, and, as a result, the perceptual indicators may contain information on the actual choice. This reversed relation of cause and effect is known as cognitive dissonance in psychology. Consequently, the latent variables which are linear combinations of the perceptual indicators have large explanatory power on the actual choice.

Table 7.2 Estimation Results of Probit Model with Latent Variables

(Latent Variables Extracted from Both the Structural and Measurement Equations)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient estimate</th>
<th>Standard error</th>
<th>t-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rail constant</td>
<td>1.87</td>
<td>0.434</td>
<td>4.30</td>
</tr>
<tr>
<td>Cost per person</td>
<td>-0.0345</td>
<td>0.0112</td>
<td>-3.08</td>
</tr>
<tr>
<td>In-vehicle time</td>
<td>0.00866</td>
<td>0.00576</td>
<td>1.50</td>
</tr>
<tr>
<td>Out-of-vehicle time</td>
<td>-0.0173</td>
<td>0.00842</td>
<td>-2.06</td>
</tr>
<tr>
<td>Number of transfers</td>
<td>-0.466</td>
<td>0.185</td>
<td>-2.52</td>
</tr>
<tr>
<td>Non-work trip dummy</td>
<td>-1.33</td>
<td>0.325</td>
<td>-4.08</td>
</tr>
<tr>
<td>Convenience*</td>
<td>1.78</td>
<td>0.481</td>
<td>3.69</td>
</tr>
<tr>
<td>Ride comfort*</td>
<td>1.26</td>
<td>0.433</td>
<td>2.91</td>
</tr>
</tbody>
</table>

$L(0) = -162.84$  $L(\hat{\beta}) = -87.07$  $\rho^2 = 0.465$  $\overline{\rho^2} = 0.416$  $N = 235$
Table 7.3 Estimation Results of Model without Latent Variables (same as Table 6.1)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient estimate</th>
<th>Standard error</th>
<th>t-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rail constant</td>
<td>1.74</td>
<td>0.318</td>
<td>5.47</td>
</tr>
<tr>
<td>Cost per person</td>
<td>-0.0257</td>
<td>0.00585</td>
<td>-4.39</td>
</tr>
<tr>
<td>In-vehicle time</td>
<td>0.000548</td>
<td>0.00347</td>
<td>0.16</td>
</tr>
<tr>
<td>Out-of-vehicle time</td>
<td>-0.0263</td>
<td>0.00595</td>
<td>-4.42</td>
</tr>
<tr>
<td>Number of transfers</td>
<td>-0.344</td>
<td>0.133</td>
<td>-2.59</td>
</tr>
<tr>
<td>Non-work trip dummy</td>
<td>-0.900</td>
<td>0.227</td>
<td>-3.97</td>
</tr>
</tbody>
</table>

$L(0) = -162.84 \quad L(\hat{\beta}) = -114.84 \quad \rho^2 = 0.295 \quad \bar{\rho}^2 = 0.258 \quad N = 235$

Table 7.4 Estimation Results of Probit Model with Latent Variables

(Latent Variables Extracted from the Structural Equations)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient estimate</th>
<th>Standard error</th>
<th>t-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rail constant</td>
<td>1.86</td>
<td>1.14</td>
<td>1.63</td>
</tr>
<tr>
<td>Cost per person</td>
<td>-0.0282</td>
<td>0.0169</td>
<td>-1.67</td>
</tr>
<tr>
<td>In-vehicle time</td>
<td>0.00515</td>
<td>0.00986</td>
<td>0.52</td>
</tr>
<tr>
<td>Out-of-vehicle time</td>
<td>-0.0229</td>
<td>0.0168</td>
<td>-1.37</td>
</tr>
<tr>
<td>Number of transfers</td>
<td>-0.429</td>
<td>0.354</td>
<td>-1.21</td>
</tr>
<tr>
<td>Non-work trip dummy</td>
<td>-1.29</td>
<td>0.958</td>
<td>-1.35</td>
</tr>
<tr>
<td>Convenience*</td>
<td>1.38</td>
<td>1.67</td>
<td>0.82</td>
</tr>
<tr>
<td>Ride comfort*</td>
<td>1.49</td>
<td>2.59</td>
<td>0.58</td>
</tr>
</tbody>
</table>

$L(0) = -162.84 \quad L(\hat{\beta}) = -112.85 \quad \rho^2 = 0.307 \quad \bar{\rho}^2 = 0.258 \quad N = 235$
Table 7.5 Estimation Results of Probit Model with Latent Variables

(Latent Variables Extracted from the Measurement Equations)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient estimate</th>
<th>Standard error</th>
<th>t-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rail constant</td>
<td>2.09</td>
<td>0.453</td>
<td>4.61</td>
</tr>
<tr>
<td>Cost per person</td>
<td>-0.0376</td>
<td>0.0116</td>
<td>-3.24</td>
</tr>
<tr>
<td>In-vehicle time</td>
<td>0.00807</td>
<td>0.00561</td>
<td>1.44</td>
</tr>
<tr>
<td>Out-of-vehicle time</td>
<td>-0.0222</td>
<td>0.00826</td>
<td>-2.69</td>
</tr>
<tr>
<td>Number of transfers</td>
<td>-0.499</td>
<td>0.186</td>
<td>-2.68</td>
</tr>
<tr>
<td>Non-work trip dummy</td>
<td>-1.35</td>
<td>0.312</td>
<td>-4.30</td>
</tr>
<tr>
<td>Convenience*</td>
<td>1.88</td>
<td>0.484</td>
<td>3.89</td>
</tr>
<tr>
<td>Ride comfort*</td>
<td>1.26</td>
<td>0.442</td>
<td>2.84</td>
</tr>
</tbody>
</table>

\[ L(0) = -162.84 \quad L(\hat{\beta}) = -86.87 \quad R^2 = 0.467 \quad \bar{R}^2 = 0.418 \quad N = 235 \]

7.4 Discussion and Summary

A methodology for incorporating psychometric data in discrete choice models and its empirical application to travel mode choice were presented in this chapter. The proposed approach is a combined system of linear equations (a LISREL type model), which identify latent variables using observable indicators, and non-linear equations, which express the discrete choice model with the latent variables. Psychometric data such as perceptual and attitudinal indicators are used as observable indicators of the latent variables in the LISREL type model.

The proposed methodology is characterized by its consistency with the framework of consumer's behavior shown in Figure 1.2. It is a totally different approach from the one in
which psychometric data are simply added to the utility function, namely, the approach that some transportation researchers have attempted (see the review in Chapter 2).

The empirical analysis demonstrated that the latent variables were successfully estimated from the LISREL type model and had a large explanatory power in the choice model. Although, this is a clear demonstration of effectiveness and practicality of the proposed method, more empirical studies in wider varieties of applications are needed to support the methodology conclusively.
Chapter 8

Simulation Study of Models

with Taste Heterogeneity

*Taste heterogeneity* refers to the variation of part-worhgs or importance of attributes across individuals. Since SP data such as ranking and rating data contain greater amounts of information on preferences per individual than RP data, models explicitly considering taste heterogeneity can be estimated from such data. In market research, model parameters are usually estimated for each individual using conjoint data. Although parameter estimates by such estimation scheme are not accurate due to the limited degrees of freedom, the rationale is that by aggregating the prediction of individual specific models these errors will cancel out. In this chapter, statistical properties of two alternative models -- individual specific and taste variation model -- are compared using simulated data sets in order to test the validity of this premise.
8.1 Explicit Modeling of Taste Heterogeneity from Stated Preference Data

8.1.1 Individual Specific Models and Market Segmentation

The utility function of random utility models is expressed by:

\[ U_{in} = \beta'x_{in} + \varepsilon_{in}, \]  

(8.1)

where

- \( U_{in} \) = utility of alternative \( i \) for individual \( n \);
- \( x_{in} \) = vector of attributes of alternative \( i \) faced by individual \( n \);
- \( \beta \) = vector of unknown parameters; and
- \( \varepsilon_{in} \) = disturbance term.

This type of model assumes that variation of underlying preferences among individuals is captured by socioeconomic variables included in the model and that tastes, which are represented by \( \beta \), are homogeneous among them. When tastes, or the values of \( \beta \), are suspected to vary significantly among individual, the method of market segmentation has been employed. Specifically, the population is stratified into several segments such that tastes are homogeneous within each segment, and the model is separately estimated for each segment. The extreme case of the market segmentation is an individual specific model, where each segment consists of a single individual.

Estimating individual specific models with accuracy requires a large amount of preference information per individual. "Ranking" or "rating" data potentially contain rich information on individual preferences; for example, a full ranking data set for 16 alternatives can be expanded into 15 sequential choice data sets if the Ranking Choice Theorem (Luce and Suppes, 1965) holds, as described in Chapter 3. Hence, this data set can identify up to 15 individual specific parameters. In marketing research, this type of "individual specific" model has been estimated from conjoint data (Green, 1984). Obviously, the accuracy of the estimates is limited because of the limited number of replications possible per individual. The rationale postulated is that
accuracy will be gained by aggregating the individual parameter estimates through the population.

In order to improve the accuracy of the estimates without sacrificing the representation of taste heterogeneity, market researchers have attempted to develop general methods of market segmentation (Green and DeSarbo, 1979; Moore, 1980; Currim, 1981; Hagerty, 1985; and Ogawa, 1987). Specifically, they try to find general rules for pooling individuals with similar tastes so that a model of homogeneous taste can be estimated reliably at each market segment. The most common technique employed to group similar individuals is cluster analysis (e.g., Currim, 1981 and Hagerty, 1985). As Montgomery and Wittink (1980) reported, however, the predictive ability of market share is not dramatically improved compared with a homogeneous taste model for the whole market, while individual specific models are superior in terms of internal validity.

8.1.2 Taste Variation Models

The alternative approach to represent taste heterogeneity assumes that the taste parameters can be expressed by some socioeconomic variables of the decision maker and disturbance terms. Namely, this method parameterizes variation of β, which substantially reduces the total number of parameters to be estimated compared with individual specific models, as described below.

Consider a conjoint experiment with an $R \times K$ design matrix $X$, where $K$ is the number of attributes and $R$ is the number of alternatives presented to a respondent $n$, $n=1,...,N$. The individual specific model estimates the following utility function directly:

$$ U_n = X_n \beta_n + \epsilon_n, \quad (8.2) $$

where

- $U_n = R \times 1$ vector of utilities of $R$ alternatives for individual $n$;
- $\beta_n = K \times 1$ vector of the unknown taste parameters for individual $n$; and
- $\epsilon_n = R \times 1$ vector of IID disturbances.
Thus, $K N$ unknown parameters are to be estimated from $R N$ observations. Obviously, the necessary condition for identification is $R \geq K$.

The *taste variation model* parameterizes the coefficient vector, $\beta_n$, as follows:

$$\beta_n = Z_n'b + \nu_n,$$  \hfill (8.3)

where

- $b = M \times K$ matrix of unknown parameters;
- $Z_n = M \times 1$ vector of characteristic of individual $n$; and
- $\nu_n = K \times 1$ vector of IID disturbances.

$b$ is a "deep" parameter matrix which is assumed to be common in the population. Hence, the number of unknown parameters to be estimated is $K M$, which is much smaller than $K N$ because $M$ is significantly smaller than $N$. The mean value of $\beta_n$ can be calculated by (8.3) using the estimate of $b$. Since (8.3) includes a random disturbance term, this type of model is often called the *random coefficient model*.

The random coefficient model requires a complicated estimation method because it does not have an IID random utility component. That is, substituting (8.3) into (8.2) gives:

$$U_n = X_n(Z_n'b + \nu_n) + \varepsilon_n$$
$$= X_nZ_n'b + X_n\nu_n + \varepsilon_n$$
$$= X_nZ_n'b + \varepsilon_n^*,$$  \hfill (8.4)

then, $\varepsilon_n^*$ can no longer be a vector of IID disturbances. For the sake of simplicity, $\nu_n$ is set to be zero in the following simulation study so that the logit assumption holds.

### 8.2 Simulation Study

In this section statistical properties of individual specific and taste variation models are compared by estimating both types of models from simulated data. For a given deep parameter matrix $b$ and generated individual characteristics $Z_n$, $n=1,\ldots,N$, individual specific taste
parameters $\beta_n$'s are calculated from (8.3). Then, for a given design matrix $X$, ranking data are generated with an assumed distribution of $\varepsilon_n$'s. $\beta_n$, $n=1,...,N$, are estimated by both individual specific and taste variation models and, then, statistical properties of the two estimators such as biases and mean square errors (MSE) are computed. This experiment is repeated with different sample sizes to investigate the effect of sample size on the statistical properties. The following sub-section describes each step of this Monte Carlo experiment in more detail.

### 8.2.1 Steps in the Monte Carlo Experiment

#### I. Data generation

1. Generate individual characteristic vectors $Z_n$, $n=1,...,N$, from an assumed distribution.

2. Using an assumed "true" deep parameter matrix $b$, calculate "true" individual-specific parameters $\beta_n$, $n=1,...,N$, using $\beta_n = bZ_n$.

3. Calculate the values of $U_n$, $n=1,...,N$, using $\beta_n$, the design matrix $X$, and randomly generated vectors $\varepsilon_n$ from a known distribution, using $U_n = X\beta_n + \varepsilon_n$.

4. Using an assumed preference model, map the values of $U_n$ into an $R \times 1$ vector of expression of preferences, $Y_n$, $n=1,...,N$ (e.g., ranking of alternatives).

#### II. Estimation of individual specific model

1. Given $X$ and $Y_n$, obtain estimates $\hat{\beta}_n$, $n=1,...,N$ using an appropriate estimation method (e.g., rank logit).

2. Compute the following properties:

$$\text{Bias}(\hat{\beta}_n) = \frac{1}{N} \sum_{n=1}^{N} (\hat{\beta}_n - \beta_n),$$

$$\text{MSE}(\hat{\beta}_n) = \frac{1}{N} \sum_{n=1}^{N} (\hat{\beta}_n - \beta_n)^2.$$
III. Estimation of taste variation model

1. Estimate the deep parameter $\mathbf{b}$ using (8.4) by an appropriate estimation method.

2. Calculate:
   \[ \hat{\beta}_n = \hat{\mathbf{b}} \mathbf{Z}_n. \]  
   (8.7)

3. Compute Bias and MSE using (8.5) and (8.6).

Repeat the above steps for different $N$. It is also desired to repeat the experiment for the same $N$ to obtain the average values of bias and MSE for a given $N$.

8.2.2 Simulation Results

Table 8.1 summarizes the settings of the Monte Carlo experiments performed. Upon completion of the experiments, the means of the absolute values of Bias and MSE over the ten parameters and five repetitions were calculated for individual specific and taste variation models. In Figures 8.1 and 8.2 these values are plotted against the sample size.

### Table 8.1 Settings of Monte Carlo Experiment

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R$ (number of alternatives presented to each individual)</td>
<td>16</td>
</tr>
<tr>
<td>$K$ (number of utility coefficients)</td>
<td>10</td>
</tr>
<tr>
<td>$M$ (number of individual characteristics)</td>
<td>2</td>
</tr>
<tr>
<td>$N$ (sample sizes)</td>
<td>20, 40, 60, 80, 100, 120, 140, 160, 180, 200</td>
</tr>
<tr>
<td>Number of repetitions for each sample size</td>
<td>5</td>
</tr>
<tr>
<td>Type of preference expression</td>
<td>ranking</td>
</tr>
<tr>
<td>Estimation method</td>
<td>rank logit with fully expanded data</td>
</tr>
</tbody>
</table>
Figure 8.1 Change of Average Absolute Values of Bias

Figure 8.2 Change of Average MSE
The results show that the bias of the individual specific model is about ten times greater than that of the taste variation model, while the ratio for MSE is more than one hundred. Another important observation is that the bias and MSE of the individual specific model stay at the same magnitude over the whole range of sample size because as \( N \) increases so does the number of unknown parameters. On the other hand, the bias and MSE of the taste variation model decrease as the sample size increases, especially for sample sizes in the range of 20 to 40. This drop of "badness-of-fit" measures is not very significant beyond the sample size of 40.

8.3 Discussion and Summary

Two types of models, individual specific and taste variation, which explicitly account for taste heterogeneity, were estimated from the simulated data sets in order to compare statistical properties of estimators from both models. In market research, conjoint data are generally used to estimate individual specific models, while in transportation research individual characteristic variables are added and/or the population is stratified into market segments to reflect preference variation by individuals. The latter method often cannot adequately represent taste heterogeneity because of fixed taste parameters. Furthermore, the market segmentation method requires multiple models to be estimated and, therefore, it loses degrees of freedom. The individual specific model can be seen as an ultimate market segmentation because a single decision making unit can be defined as a market segment. Hence, this approach requires estimation of as many models as there are individuals in the sample, which implies enormous loss of degrees of freedom.

An alternative approach that explicitly considers taste heterogeneity allows taste parameters to vary according to some assumed distribution or parametric function, which is referred to as the taste variation model. Although this method minimizes loss of degrees of freedom for a
given sample, it requires a reasonable specification of the parametric function for the taste parameters. Moreover, if random fluctuation is assumed in the taste parameters, it is necessary to employ a very complicated estimation method such as multinomial probit to estimate the model.

This chapter investigated statistical properties of individual specific and taste variation models by means of a Monte Carlo experiment. For known taste parameters, ranking data were generated, then the parameters were estimated by both models to compare bias and MSE.

It was found that magnitude of bias and MSE of the individual specific model is substantially greater than those of the taste variation model in the range of sample sizes tested, i.e., from 20 to 200. This implies that the rationale of gaining accuracy by aggregating individual specific parameter estimates is rejected. Also, those measures did not change by sample size for the individual specific model, while they are significantly improved by increasing sample size for the taste variation model. Furthermore, the marginal improvement of bias and MSE of the taste variation model sharply decreased beyond a sample size of 40. This may indicate that, for this particular model and generated data, 40 is a good choice of sample size to obtain adequately accurate estimates by a taste variation model.

These findings suggest that individual specific models yield unstable estimates due to limited replications per individual and that sufficiently reliable estimates can be obtained by taste variation models with a limited sample size (e.g., 40) provided that an appropriate specification for the taste parameters is known.

The statistical properties of individual specific models that this simulation study verified is well known among econometricians. However, market researchers often ignore these properties and estimate individual specific models using conjoint data.

Simulation study is an efficient method to compare statistical properties and predictive validity of different estimators. The study presented in this chapter is still limited in the sense that the experiment was carried out with only one experimental setting including the design matrix and probability distribution of random variables. Rating data instead of ranking data
could be generated and then a linear regression instead of the rank logit could be used to estimate models. More simulation studies are needed to find more concrete conclusions on statistical properties of both types of models.
Chapter 9

Conclusions

For the past decade the development of methodologies of demand modeling in transportation research as well as in market research has focused on discrete choice models such as logit and probit models which are based on individual behavior. It is widely recognized that the theoretical background of this type of model is superior to that of more traditional transportation models (such as the gravity model) in the sense that behavioral disaggregate models are consistent with a theory of consumer's behavior and estimated from individual level data. In empirical applications, however, such models do not always yield satisfactory results in terms of predictive ability and model transferability. In other words, the theoretical superiority of this type of models has not shown its full expected advantages in the practical level.

This problem seems to be principally caused by the limitation of input data. For instance, a model of travel mode choice is usually estimated from the information on travel characteristics such as travel time, cost, and access and egress conditions, and on some socioeconomic characteristics of the decision maker. However, the actual choice decision may be greatly affected by the factors such as travel time reliability, information availability, and ride comfort. Especially, when the estimated model is used to predict the demand of a new service, the traditional type of input data often cannot yield a satisfactory prediction. Here, the "traditional" type of data are based on the past actual behavior and the data items are limited to quantifiable attributes. The "new" or "non-traditional" type of data can be based on hypothetical scenarios
and include attributes which are not easily quantifiable. These data types are referred to as *psychometric data*. The traditional data can be represented by *revealed preference* (RP) data, while *stated preference* (SP) data belong to the new type. This thesis has focused on the utilization of the new type of data. The major findings obtained are summarized below.

1. Combined estimation with RP and SP data

The main objective of this thesis was to develop a methodology for incorporating SP data in demand analysis with explicit consideration of the advantages and disadvantages of each. As discussed in Chapter 1, RP and SP data have complementary characteristics. RP data have high face-validity, or reliability, because they are based on actual behavior, while SP data are more operational because the hypothetical scenarios can be arbitrarily created. Practitioners, especially in the transportation field, have made little use of SP data simply because reliability of stated preferences is unknown. The proposed methodology presented in Chapter 3 is based on the recognition that stated preferences are clearly related to underlying latent preferences but do not necessarily reflect actual market behavior. Therefore, stated preferences are treated as indicators of underlying latent preferences rather than being used directly to express causal relations of market behavior. Hence, the proposed method uses SP data to build measurement equations and RP data to construct structural equations. Furthermore, the measurement equations explicitly consider potential biases and random errors specific to SP data. Joint estimation of structural and measurement equations with RP and SP data would yield more reliable and accurate estimates of the structural parameters than the estimates from only RP data.

Two empirical case studies using this methodology were presented in Chapters 5 and 6. In the first case study, stated intention of switching the commuting route in conjunction with an introduction of a subway line is combined with RP data to estimate route choice models. This stated preference is not a choice from a purely hypothetical choice set but a choice from an
implicit choice set which consists of an existing option plus a new option. The combined estimation revealed that the SP information contained more random noise than the RP information and an overstatement (or bias) for using the new subway line.

The second case study which analyzes intercity travel mode choice has a clearer configuration of RP and SP data. Besides the information of an actual mode choice (RP data), the survey included two conjoint experiments which yielded two sets of SP data. Three models were estimated separately from the three data sets (one RP and two SP) and three combined models were also estimated from three different combinations of the data sets. It was found that the conjoint data on car versus rail choice contained more random errors than the RP data, while the other conjoint data which compared two rail services had less noise than the RP data. The inertia effect in the conjoint task, by which the respondent tends to choose the mode actually chosen, was also found by including in the measurement equations a dummy variable which indicates the actual choice. Furthermore, although the parameter of in-vehicle travel time was not accurately estimated from only the RP data, combining them with SP data enabled it to be successfully estimated. Thus, this case study was a clear demonstration of effectiveness and practicality of the methodology of combined estimation with RP and SP data.

(2) Special estimators for SP data

One important feature of stated preferences is that they can be expressed by various response formats such as ranking and rating, and, therefore, the amount of information on preference per respondent may be richer than that contained in RP data. Appropriate estimators, some of which were developed in this thesis, for discrete choice models can take advantage of that rich information. Alternative estimators from ranking data were proposed in Chapter 3 with a case study of commuting route choice presented in Chapter 4. These estimators explicitly consider different reliability of choice information at different depth of the ranking. A methodology for seeking the optimal strategy of pooling information from different depth of the ranking was
presented with an empirical application. The data analyzed in the case study showed that reliability of preference information decreased as the ranking went down.

In estimating from the rating conjoint data in Chapter 6, an estimator for an ordered categorical dependent variable was employed. Threshold values on the utility scale for ordered categories of the dependent variable are jointly estimated with the model coefficients.

Another special estimator that takes advantages of SP data was proposed in Chapter 8, which explicitly accounts for taste heterogeneity by allowing for variation of parameter values by individual. Two approaches were compared in terms of their statistical properties by the Monte Carlo method. One approach estimates the model separately with the data from a single individual, which is called the individual specific model. The other approach assumes a functional form of taste parameters and estimates "deep" parameters in that function which are common in the population, which is called the taste variation model. The simulation study verified that parameter estimates from the individual specific models are unstable due to the limited degrees of freedom and biases and mean square errors do not decrease even if the sample size increases. The taste variation models produced much more stable parameter estimates and biases, and the mean square errors decreases significantly with a large sample size.

(3) Utilization of other psychometric data

In Chapter 7 a methodology for incorporating a wider variety of psychometric data was introduced, followed by an empirical analysis. This approach uses psychometric data as observable indicators of underlying latent variables and constructs a system of linear structural and measurement equations to identify these latent variables. The whole system is composed of this system of linear equations and a system of non-linear equations representing a discrete choice model that includes the latent variables as explanatory variables. This whole system formulates the path diagram of consumer's behavior presented in Figure 1.2. Since it includes
non-linear equations, the practical estimation is performed sequentially; first, the system of linear equations is estimated to obtain estimates of the latent variables, then, the discrete choice model is estimated using the fitted values of the latent variables. The empirical analysis used the RP data of intercity travel mode choice analyzed in Chapter 6 which include respondent's personal ratings on some aspects of the trip. Two latent travel characteristic variables, convenience and ride comfort, were specified by these perceptual indicators and some observable attributes. The two stage estimation method successfully identified these latent variables from the system of linear equations and the fitted values of the latent variables had significant explanatory power in the discrete choice model. This empirical result is encouraging for further investigation of this promising field of research.

Thus, this thesis explored effective methodologies to utilize non-traditional data sources focusing on SP data in order to exploit advantages of different types of data. The principal contribution is the development of the combined estimator for discrete choice models from RP and SP data. Although the effectiveness and practicality of the methodology were demonstrated through two case studies, more empirical analyses are required to justify the methodology for a wider variety of applications. Since SP data could have different types of bias according to the survey environments, specification of the bias parameter requires ad hoc consideration of the survey context and, therefore, accumulation of empirical results are especially required. In the theoretical aspect, the assumption that the random utility components of the RP and SP models are statistically independent should be further investigated. The proposed combined estimator assumes the independence of these random components, and if this assumption is violated, then the estimator is not fully efficient and the standard MLE variances of the parameter estimates will be understated.

The extension of the methodology of combined estimation with RP and SP data is found in the incorporation of latent variables with psychometric data described in Chapter 7. A challenging but worthwhile future task is to develop a joint estimation method of the LISREL
type model and the discrete choice model. That joint estimation method will be able to combine SP data, and then the whole system would faithfully follow the path diagram of consumer's behavior that many researchers have attempted to model.

The final word of this thesis would be that non-traditional data could provide valuable information for developing and estimating demand models with explicit considerations of their statistical properties such as biases and random errors.
REFERENCES


APPENDIX A

Survey Questionnaire for the Arborway Data

Please answer the following questions for your entire trip in this direction, from the place it started (home, work, school, etc.) to your final destination.

1. Where did your trip begin?
   City/Town/Neighborhood___________________________________________
   Location________________________________________________________
   (Nearest intersection or major landmark)

2. Where is the final destination of your trip?
   City/Town/Neighborhood___________________________________________
   Location________________________________________________________

3. Why did you make this trip?
   1) To or from work
   2) To or from school
   3) To of from shopping
   4) To or from lunch
   5) To or from medical appointment
   6) Other reasons

4. If you are making this trip for work or school, what time do you usually leave home to go to work?______________ leave work?__________
5. Imagine you could decide what improvements should be made to MBTA services in this area. Tell us which single thing on this list is most important to you by writing a '1' next to it. Show your second most important consideration with a '2,' and so on until you put '7' next to the least important.

___ Fare is low
___ No need to transfer to another vehicle
___ Short travel time
___ Vehicles usually arrive on schedule
___ Service runs often
___ Vehicles are not too crowded
___ Vehicles are clean

6. Example of four possible inbound Arborway services are shown below. Please tell us which of the four you would like most by writing '1' in the small box beneath it. Then show us your second, third, and last choices by writing '2,' '3,' and '4' in the other small boxes. Your preference for these services will tell us what you think is most important about transit service.
THE ROUTE IS
ARBORWAY  BRIGHAM  PARK STREET

BUS  [TRANSFER]  GREEN LINE

YOU WAIT
5 MINUTES FOR THE BUS
4 MINUTES FOR THE GREEN LINE
FROM ARBORWAY TO PARK STREET
YOU RIDE 27 MINUTES IN THE VEHICLES

RANK

THE ROUTE IS
ARBORWAY  PARK STREET

GREEN LINE

YOU WAIT
2 MINUTES FOR THE GREEN LINE
FROM ARBORWAY TO PARK STREET
YOU RIDE 33 MINUTES IN THE VEHICLES

RANK

THE ROUTE IS
ARBORWAY  COLEY  PARK STREET

BUS  [TRANSFER]  GREEN LINE

YOU WAIT
5 MINUTES FOR THE BUS
2 MINUTES FOR THE GREEN LINE
FROM ARBORWAY TO PARK STREET
YOU RIDE 30 MINUTES IN THE VEHICLES

RANK

THE ROUTE IS
ARBORWAY  HEATH  PARK STREET

BUS  [TRANSFER]  GREEN LINE

YOU WAIT
3 MINUTES FOR THE BUS
2 MINUTES FOR THE GREEN LINE
FROM ARBORWAY TO PARK STREET
YOU RIDE 27 MINUTES IN THE VEHICLES

RANK
7. How often do you use the MBTA Arborway Line?
   1) Four or more times a week
   2) Between one and three times a week
   3) Less than once a week
   4) Never

8. Are you...
   1) under 16 years old?
   2) between 16 and 64 years old?
   3) 65 years old or older?

9. Do you have any comments that would help us improve transit service in the Arborway corridor?

   ___________________________________________________________
APPENDIX B

Survey Questionnaire for the Yokohama Before Data

Questionnaire for Household

1. How many of your household are
   in total?________
   less than 6 years old?________
   going to elementary or junior high?_______
   commuting to the work place or school (high school and college)?_______
   not commuting (excluding children)?_______

2. How many vehicles does your household own?
   bicycle________
   moped________
   motorcycle__ ____
   car________

3. Which bus stop is closest to your house?
   __________________

4. What is your approximate household income?
   1) less than ¥ 3,000,000
   2) ¥ 3,000,000 - 3,990,000
   3) ¥ 4,000,000 - 4,990,000
   4) ¥ 5,000,000 - 5,990,000
Questionnaire for Commuting Individual

1. What is your age, sex, and purpose of trip?

<Age>
   1) less than 20 years
   2) twenties
   3) thirties
   4) forties
   5) fifties
   6) sixties or older

<Sex>
   1) male
   2) female

<Purpose of trip>
   1) school trip
   2) work trip

If the purpose is work, who pays your commuting cost?
   1) Employer pays all
   2) Employer pays ______%
3) I pay all
4) Others

2. What vehicles does your household own that are usually available for your commuting?

(Check all applicable)

1) none
2) bicycle
3) moped
4) motorcycle
5) car
6) truck
7) others

3. Where is your destination?

Prefecture/City/Town/Street

4. What time do you usually leave home?

1) before 7 am
2) 7 am - 8 am
3) 8 am - 9 am
4) 9 am - 4 pm
5) 4 pm - 7 pm
6) after 7 pm
7) depends
5. Please give the commuting route and mode you regularly use (see the example).

<Example>
home -- (bicycle) -- [Rokkakubashi-kitamachi] -- (bus) -- [Yokohama Station] -- (Japan Railway) -- [Tokyo Station] -- (Japan Railway) -- [Ochanomizu Station] -- (walk) -- destination

<Your Route>
home -- ( ) -- [ ] -- ( ) -- [ ] -- ( ) -- [ ] -- ( ) -- destination

6. Please give the total travel time, delay in the total travel time, and congestion level for the route you gave in question 5.

Total travel time from home to the destination is ____ hours ____ minutes on the average.
The trip sometimes takes about ____ minutes more than the average travel time.

If you use rail or bus, give the most congested point and congestion level.

The most congested point is from [ ] to [ ].

Congestion level

1) Almost everyone can take a seat.
2) Almost everyone can hold a handrail.
3) Very crowded but can open a newspaper or magazine.
4) Absolutely no space to move.
7. Please give the satisfaction level of your regular commuting route with respect to the following attributes.

<Example>

<table>
<thead>
<tr>
<th>satisfaction level*</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>overall satisfaction</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>total travel time</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<Your Response>

<table>
<thead>
<tr>
<th>satisfaction level*</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>overall satisfaction</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>total travel time</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>delay in travel time</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>in-vehicle travel time**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>walking time**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>transfer and wait time**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fare**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>congestion**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>comfort in vehicle**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Satisfaction level corresponds:

1 = absolutely satisfied
2 = satisfied
3 = more or less satisfied
4 = not sure
5 = more or less dissatisfied
6 = dissatisfied
7 = absolutely dissatisfied

** Answer only if you use rail or bus.

8. Please give another commuting route and mode you sometimes use.

<Example>

a) I sometimes use Toyoko-line (commuter rail) instead of bus to go to Yokohama Station.
home -- (bicycle) --> [Sigaraki Station] -- (commuter rail) --> [Yokohama Station] -- (Japan Railway) --> [Tokyo Station] -- (Japan Railway) --> [Ochanomizu Station] -- (walk) --> destination

b) I sometimes walk to the station.
home -- (walk) --> [Sigaraki Station] -- (commuter rail) --> [Yokohama Station] -- (Japan Railway) --> [Tokyo Station] -- (Japan Railway) --> [Ochanomizu Station] -- (walk) --> destination

c) I sometimes drive to my destination.
home -- (car) --> [parking lot] -- (walk) --> destination

<Your Route>

home -- ( ) --> [ ] -- ( ) --> [ ] -- ( ) --> [ ] -- ( ) --> destination

9. What will you do when the new subway line between Yokohama and Shin-yokohama opens next March with the following service level:
- running time and fare from each station to Yokohama

<table>
<thead>
<tr>
<th>Station</th>
<th>Travel Time</th>
<th>Fare</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mitsuzawa Shita Machi</td>
<td>2 min.</td>
<td>¥ 120</td>
</tr>
<tr>
<td>Mitsuzawa Kami Machi</td>
<td>4 min.</td>
<td>¥ 120</td>
</tr>
<tr>
<td>Katakura Cho</td>
<td>6 min.</td>
<td>¥ 140</td>
</tr>
<tr>
<td>Kishinekoen</td>
<td>8 min.</td>
<td>¥ 140</td>
</tr>
<tr>
<td>Shinyokohama</td>
<td>11 min.</td>
<td>¥ 160</td>
</tr>
</tbody>
</table>

- see the attached map for the new bus routes

• Which station will be your boarding station, if you use the subway?

• How will you go to that station?
  1) walk
  2) bicycle
  3) moped
  4) motorcycle
  5) car
  6) bus
  7) others (       )

• If you use the subway, how do you think travel time and fare to Yokohama Station will change?

- Travel time
  1) will be longer
  2) will be unchanged
  3) will be shorter
- fare
  1) will be higher
  2) will be unchanged
  3) will be lower

10. Considering the service levels given below, do you think you will use the subway line for commuting?
  1) yes, I will use it. --> please go to question 11
  2) no, I will not use it. --> please go to question 12

11. If you answered "no" to question 10, please give the reasons for not using it. (up to 3 reasons)
  1) The nearest subway station is too far.
  2) Total travel time will become longer.
  3) Wait time will become longer.
  4) Walking time will become longer.
  5) Total cost will become higher.
  6) There will be too many transfers.
  7) Another mode will be more convenient.
  8) Subway is inconvenient when I carry luggage.
  9) others (   )

12. If you answered "yes" to question 10, please give the commuting route and mode with the subway line and give the reasons of using it.
• commuting route and mode
<Example>
home -- (walk) -- [Kishine Koen Station] -- (subway) -- [Yokohama Station] -- (Japan Railway) -- [Tokyo Station] -- (Japan Railway) -- [Ochanomizu Station] -- (walk) -- destination

<Your Route>

home -- ( ) -- [ ] -- ( ) -- [ ] -- ( ) -- [ ] -- ( ) -- destination

• reasons for using subway (up to 3 reasons)

1) The subway station is close.

2) Total travel time will become shorter.

3) Total cost will become lower.

4) Wait time will become shorter.

5) Arrival time will less fluctuate.

6) There will be less risk of accidents.

7) Subway will be less crowded.

8) Subway will be more comfortable.

9) others ( )
APPENDIX C

Survey Questionnaire for the Yokohama After Data

Household Questionnaire

1. How many are in your household:
   total? ________
   commuting to work? ________
   commuting to school (high school and college)? ________

2. How many vehicles does your household own?
   bicycle ________
   moped ________
   motorcycle ________
   car ________

3. Which bus stop and subway station is closest to your house, respectively? How long does it take to get there?
   - bus stop
     ________________ ________________ min. by walking
   - subway station
     ________________ ________________ min. by ________________

4. What is your approximate household income?
   1) less than ¥ 3,000,000
   2) ¥ 3,000,000 - 3,990,000
3) ¥ 4,000,000 - 4,990,000
4) ¥ 5,000,000 - 5,990,000
5) ¥ 6,000,000 - 6,990,000
6) ¥ 7,000,000 - 7,990,000
7) ¥ 8,000,000 - 8,990,000
8) ¥ 9,000,000 - 9,990,000
9) ¥ 10,000,000 or more

Questionnaire for Commuting Individual

1. What is your age, sex, and purpose of trip?

<Age>

1) less than 20 years
2) twenties
3) thirties
4) forties
5) fifties
6) sixties or older

<Sex>

1) male
2) female

<Purpose of trip>

1) school trip
2) work trip

If the purpose is work, who pays your commuting cost?

1) Employer pays all
2) Employer pays _____%

3) I pay all

2. What is the availability of the following vehicles for your commuting?
   - bicycle
     1) not available
     2) sometimes available
     3) always available
   - moped
     1) not available
     2) sometimes available
     3) always available
   - motorcycle
     1) not available
     2) sometimes available
     3) always available
   - car
     1) not available
     2) sometimes available
     3) always available

3. Where is your destination? Is that the same as before the subway opened?
   - address of destination
     Prefecture/City/Town/Street__________________________________________
   - change in destination
     1) the same as before the opening
     2) not the same as before the opening
4. Do you leave home at fixed time?
   - now
     1) non fixed time
     2) fixed time -- What time? ____________
   - before the opening of the subway
     1) non fixed time
     2) fixed time -- What time? ____________

5. Does your work (school) begin at fixed time? If so, how many minutes advance do you try to arrive at your working place (school) before the work (school) begins?
   1) non fixed time
   2) fixed time -- What time? _______ -- How many minutes advance? _______

6. If you used the new subway line for commuting, what would be the route, travel time, frequency of usage, and overall evaluation of the route?

   - route
     home -- ( ) -- [ ] -- ( ) -- [ ] -- ( ) -- [ ] -- ( ) -- [ ] -- ( ) -- > destination

   - travel time
     1) total time ________min.
     2) in-vehicle time ________min.
     3) wait time ________min.
     4) transfer time ________min.
     5) walk time ________min.
     Total travel time is sometimes ________minute longer than the average.
- level of congestion in the subway
  1) There are vacant seats.
  2) Everyone can hold a handrail.
  3) Everyone cannot hold a handrail but can move.
  4) Absolutely no space to move.

- frequency of usage
  1) almost everyday
  2) 3 - 4 times a week
  3) 1 - 2 times a week
  4) 2 - 3 times a month
  5) rarely use

- overall satisfaction
  1) absolutely satisfied
  2) more or less satisfied
  3) not sure
  4) more or less dissatisfied
  5) absolutely dissatisfied

7. If you did not use the new subway line for commuting, what would be the route, travel time, frequency of usage, and overall evaluation of the route?

- route
  
  home -- ( ) -- [ ] -- ( ) -- [ ] -- ( ) -- [ ] -- ( ) -- destination
- travel time
  1) total time _______ min.
  2) in-vehicle time _______ min.
  3) wait time _______ min.
  4) transfer time _______ min.
  5) walk time _______ min.
  Total travel time is sometimes _______ minute longer than the average.

- level of congestion in the bus
  1) There are vacant seats.
  2) Everyone can hold a handrail.
  3) Everyone cannot hold a handrail but can move.
  4) Absolutely no space to move.

- frequency of usage
  1) almost everyday
  2) 3 - 4 times a week
  3) 1 - 2 times a week
  4) 2 - 3 times a month
  5) rarely use

- overall satisfaction of the route
  1) absolutely satisfied
  2) more or less satisfied
  3) not sure
  4) more or less dissatisfied
  5) absolutely dissatisfied
7. Please give your satisfaction level of the route with subway and the route without subway with respect to the following attributes.

<table>
<thead>
<tr>
<th></th>
<th>route with subway</th>
<th>route without subway</th>
</tr>
</thead>
<tbody>
<tr>
<td>cost</td>
<td>1 2 3 4 5</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>total travel time</td>
<td>1 2 3 4 5</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>waiting time</td>
<td>1 2 3 4 5</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>walking time</td>
<td>1 2 3 4 5</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>fluctuation of arrival time</td>
<td>1 2 3 4 5</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>ease of transfers</td>
<td>1 2 3 4 5</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>number of transfers</td>
<td>1 2 3 4 5</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>ease of using stairways</td>
<td>1 2 3 4 5</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>ease of boarding and alighting</td>
<td>1 2 3 4 5</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>environments in vehicle (e.g., air-conditioning)</td>
<td>1 2 3 4 5</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>privacy</td>
<td>1 2 3 4 5</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>accessibility in bad water</td>
<td>1 2 3 4 5</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>risk of accidents and crimes</td>
<td>1 2 3 4 5</td>
<td>1 2 3 4 5</td>
</tr>
</tbody>
</table>

* Satisfaction level corresponds:

1 = absolutely satisfied
2 = more or less satisfied
3 = not sure
4 = more or less dissatisfied
5 = absolutely dissatisfied
9. Before the subway line opened, had you been using the same route as the one you provided in question 7?

1) yes, the same route --> please go to question 10.

2) no, not the same route --> please describe the route below.

- route
  home --> ( ) --> [ ] --> ( ) --> [ ] --> ( ) --> [ ] --> ( ) --> destination

- travel time
  1) total time ________ min.
  2) in-vehicle time ________ min.
  3) wait time ________ min.
  4) transfer time ________ min.
  5) walk time ________ min.

Total travel time is sometimes ________ minute longer than the average.

10. If you rarely use the subway line for commuting, please give the reasons for not using it.
(Please check all applicable.)

1) I do not like subway.

2) The subway is not running when I go to or come back from the work place (school).

3) The subway line is away from my direction.

4) I do not know well about the subway.

5) I must use some other mode.

6) I am used to using some other mode.

7) others ( )
11. In choosing your commuting route and mode, which service attributes do you take into account most? (up to 3 items)

1) cost
2) travel time
3) waiting time
4) walking distance
5) fluctuation of arrival time
6) ease of transferring
7) number of transfers
8) ease of using stairways
9) ease of boarding and alighting
10) environments in vehicle (e.g., noise and air-conditioning)
11) privacy
12) accessibility in bad weather
13) risk of accidents and crimes
APPENDIX D

Survey Questionnaire for the Nijmegen-Randstad Data

NS NIJMEGEN-RANDSTAD STUDY

This research is being undertaken to find out how the travellers from Nijmegen to the Randstad choose between car and train.

In order to do this we need to determine what the important reasons are to choose the car or the train, and under what circumstances the travellers might change their choice of travel mode.

SECTION 1

Questions about your journey (1)

FIRST, WE WOULD LIKE TO ASK YOU A FEW QUESTIONS ABOUT THE JOURNEY WHICH WE SPOKE WITH YOU ABOUT ON THE TELEPHONE.

1. WANN
How long ago did you make this journey?
   1) less than 1 week ago
   2) 1 to 4 weeks ago
   3) 1 month ago or longer

2. BEST
Where did you travel to then?
1) Amsterdam
2) Rotterdam
3) Den Haag

3. VERV
What mode of travel did you use?
1) car driver
2) car passenger
3) train

4. VERT
Where did you begin this journey?
1) your home
2) your work
3) your study
4) other (specify)

5. MOTI
What was the main purpose for the journey?
1) to regular work place
2) employer's business
3) school/university
4) shopping
5) visit
6) recreation
7) other (specify)
SECTION 2

Questions about your journey (2)

NOTE: Here REIS1 set to chosen mode, REIS2 to unchosen

6. GROEP
How many people made this journey with you (yourself not included)?
   1) none (alone)
   2) with 1 person
   3) with 2 people
   4) with 3 people
   5) with 4 people
   6) with 5 or more people

7. BEPAALD
For this journey, did you have to be at your destination in #BEST# at a certain time or didn't that matter very much?
   1) a certain time
   2) didn't matter very much

IF #BEPAALD# IS NOT EQUAL TO 1, GO TO VTIJD.

8. BTLJ
By what time did you have to be there? (e.g. enter half three as 1530)

9. VTIJD
At about what time did you leave from #VERT#? (e.g. enter half three as 1530)
SECTION 3

Questions about your #REIS1# journey (1)

10. FREQ1
How often do you travel to #BEST# by #REIS1#?

1) 4 times/week or more
2) 1 to 3 times per week
3) 1 to 3 times per month
4) less than once per month
5) less than once per year

IF #REIS1# IS EQUAL TO 1, GO TO AKOST1.

NOTE: 11-18 if train was used

11. VOOR1
How did you travel from #VERT# to the train station in Nijmegen?

1) by foot
2) bike/moped
3) bus/tram/metro
4) auto, as driver
5) auto, as passenger
6) other (specify)

12. TATIJD1
How long did this take you? (e.g. enter 1 hour 5 min. as 65)
13. KLAS1
Did you buy a 1st class or 2nd class ticket?
   1) 1st class
   2) 2nd class

14. TKOST1
How much did a #KLAS1# RETURN for this journey to #BEST# cost? (e.g. enter 19.99 as 1999)

15. OVER1
Did you have to change trains during your journey to #BEST#?
   1) yes, change once
   2) yes, change twice
   3) yes, change 3 times
   4) no, no interchange
   5) other (specify)

16. TTIJD1
And how long did the train journey take from the station in Nijmegen to the station in #BEST#, including any interchanges? (e.g. enter 1 hour 5 min. as 65)

17. NATR1
How did you travel further from the station in #BEST# to your destination?
   1) by foot
   2) bike/moped
   3) bus/tram/metro
   4) auto, as driver
5) auto, as passenger
6) other (specify)

18. TETIJD1
And how long did this take?

IF #REIS1# IS EQUAL TO 2, GO TO RED1.

NOTE: 19-22 if car was used

19. AKOST1
How much did your car journey cost you to #BEST# AND BACK AGAIN? (e.g. ente. 9.99 as 1999)

20. ATIJD1
And how long did the car journey take from #VERT# in Nijmegen to your parking place in #BEST#? (e.g. enter 1 hour 5 min. as 65)

21. PARK1
How did you park at your destination in #BEST#?

1) along the street, free
2) at a parking meter
3) in a parking garage
4) on private property
5) other (specify)

22. AETIJD1
And how long did you have to walk from your parking place to your destination?
23. RED1

What was your most important reason for making this journey by #REIS1#?

24. ARED1

Was there another important reason?

1) yes (specify)
2) no

SECTION 4

Questions about your journey by #REIS1# (2)

NOW WE WOULD LIKE YOU TO EVALUATE A NUMBER OF DIFFERENT ASPECTS
OF YOUR JOURNEY BY #REIS1# TO #BEST#.

25. ONTS1

How would you rate...

- the possibility to RELAX during the journey in the #REIS1#?

1) very bad
2) bad
3) neutral
4) good
5) very good

26. ZEKER1

- the RELIABILITY to be on time in the #REIS1# at #BEST#?

1) very bad
2) bad
3) neutral
4) good
5) very good

27. VRIJ1
- the FREEDOM to choose the time you begin your journey?
  1) very bad
  2) bad
  3) neutral
  4) good
  5) very good

28. GEMAK1
- the EASE of travelling with children and/or heavy baggage?
  1) very bad
  2) bad
  3) neutral
  4) good
  5) very good

29. VEILIG1
- the SAFETY during the journey?
  1) very bad
  2) bad
  3) neutral
  4) good
5) very good

30. WELKE1
Which of the aspects of the #REIS1# journey mentioned above was the most important for you?
1) relaxation
2) reliability
3) freedom
4) ease
5) safety

31. CIJFER1
If you were to score the #REIS1# as a MEANS of making this journey to #BEST#, what score would you give it? (1=worst, 10=best)

SECTION 5

Questions about the journey by #REIS2# (1)

WOULD WOULD NOW LIKE TO FIND OUT HOW YOU THINK THIS JOURNEY WOULD HAVE GONE IF YOU HAD TRAVELLED BY #REIS2#.

32. FREQ2
How often do you travel to #BEST# by #REIS2#?
1) 4 times/week or more
2) 1 to 3 times per week
3) 1 to 3 times per month
4) less than once per month
5) less than once per year
6) never

33. VTIJD2
You left #VERT# to begin this journey at #VTIJD#. What time would you have left if you had travelled by #REIS2# instead. (e.g. enter half three as 1530)

IF #REIS2# IS EQUAL TO 1, GO TO AKOST2.
NOTE: 34-41 if car was used

34. VOOR2
How would you have gone from #VERT# to the train station in Nijmegen?
    1) by foot
    2) bike/moped
    3) bus/tram/metro
    4) auto, as driver
    5) auto, as passenger
    6) other (specify)

35. TATIJD2
About how long would this have taken you? (e.g. enter 1 hour 5 min. as 65)

36. KLAS2
Would you have bought a 1st class or 2nd class ticket?
    1) 1st class
    2) 2nd class
37. TKOST2
About how much would a #KLAS2# RETURN to #BEST# have cost you? (e.g. enter 19.99 as 1999)

38. OVER2
Do you know if you would have had to change trains during the journey to #BEST#?

1) yes, change once
2) yes, change twice
3) yes, change 3 times
4) no, no interchange
5) don't know
6) other (specify)

39. TTIJD2
And about how long would this train journey from the station in Nijmegen to the station in #BEST# have taken, including interchanges? (e.g. enter 1 hour 5 min. as 65)

40. NATR2
How would you have travelled further from the station in #BEST# to your destination?

1) by foot
2) bike/moped
3) bus/tram/metro
4) auto, as driver
5) auto, as passenger
6) other (specify)
41. TETIJD2

And about how long would this have taken you?

IF #REIS2# IS EQUAL TO 2, GO TO RED2.

NOTE: 42-45 if train was used

42. AKOST2

About how much would this car journey cost you to #BEST# AND BACK AGAIN? (e.g. enter 19.99 as 1999)

43. ATIJD2

And about how long would the car journey take from #VERT# in Nijmegen to your parking place in #BEST#? (e.g. enter 1 hour 5 min. as 65)

44. PARK2

How would you have parked at your destination in #BEST#?

1) along the street, free
2) at a parking meter
3) in a parking garage
4) on private property
5) don't know
6) other (specify)

45. AETIJD2

And about how long would you have to walk from your parking place to your destination?
46. RED2
What was your most important reason for NOT making this journey by #REIS2#?

47. ARED2
Was there another important reason?
1) yes (specify)
2) no

SECTION 6

Questions about your journey by #REIS2# (2)

NOW WE WOULD LIKE YOU TO EVALUATE A NUMBER OF DIFFERENT ASPECTS OF YOUR JOURNEY TO #BEST# IF IT HAD BEEN BY #REIS2#.

48. ONTS2
How would you rate...
- the possibility to RELAX during the journey in the #REIS2#?
  1) very bad
  2) bad
  3) neutral
  4) good
  5) very good

49. ZEKER2
- the RELIABILITY to be on time in the #REIS2# at #BEST#?
  1) very bad
2) bad
3) neutral
4) good
5) very good

50. VRIJ2
- the FREEDOM to choose the time you begin your journey?
  1) very bad
  2) bad
  3) neutral
  4) good
  5) very good

51. GEMAK2
- the EASE of travelling with children and/or heavy baggage?
  1) very bad
  2) bad
  3) neutral
  4) good
  5) very good

52. VEILIG2
- the SAFETY during the journey?
  1) very bad
  2) bad
  3) neutral
  4) good
5) very good

53. WELKE2
Which of the aspects of the #REIS2# journey mentioned above is the most important for you?

1) relaxation
2) reliability
3) freedom
4) ease
5) safety

54. CIJFER2
If you were to score the #REIS2# as a MEANS of making this journey to #BEST#, what score would you give it? (1=worst, 10=best)

SECTION 7

Give your preference (1)

ON THE FOLLOWING SCREENS WE WILL GIVE YOU A NUMBER OF POSSIBLE CHANGES FOR THE TRAIN SERVICE FOR YOUR JOURNEY TO #BEST#.

ON EACH SCREEN, WE GIVE TWO POSSIBLE TRAIN JOURNEYS. PLEASE LOOK AT EACH OF THEM CAREFULLY AND COMPARE THEM.

THEN, PLEASE STATE WHICH TRAIN OPTION YOU WOULD HAVE PREFERRED FOR THE JOURNEY TO #BEST# WHICH YOU HAVE DESCRIBED?
NOTE: The train vs. train preferences given here

SECTION 8

Give your preferences (2)

WE KNOW THAT YOU DID NOT USE THE #REIS2# FOR THE JOURNEY WE HAVE BEEN DISCUSSING WE WOULD NOW LIKE TO GET AN IDEA OF THE CIRCUMSTANCES UNDER WHICH YOU PROBABLY WOULD HAVE TAKEN THE #REIS2#.

ON THE FOLLOWING SCREENS WE AGAIN GIVE YOU A NUMBER OF POSSIBLE CHANGES FOR YOUR JOURNEY. ON EACH SCREEN WE SHOW YOU A POSSIBLE TRIP BY TRAIN AND A POSSIBLE TRIP BY CAR. PLEASE COMPARE THESE CAREFULLY AND THEN TELL US WHICH MEANS OF TRANSPORT YOU WOULD HAVE PREFERRED.

PLEASE KEEP THIS JOURNEY TO #BEST# IN MIND, PARTICULARLY THE REASON YOU MADE THIS JOURNEY AND YOUR REASONS FOR TAKING THE TRAIN OR TAKING THE CAR.

NOTE: The train vs. car choices made here

SECTION 9

Questions about yourself
2) scholar/student
3) without work
4) disabled
5) pensioned
6) other (specify)

59. POSIHH

What is your role in your household?
1) main money earner
2) other family member
3) other resident
4) other (specify)

60. AANT

How many people reside in this household (INCLUDING YOURSELF)?

61. OPMERK

Do you have any further remarks about your travel or about this interview?
1) yes (please note)
2) no
FINALLY, WE WOULD LIKE TO ASK A FEW QUESTIONS ABOUT YOURSELF AND YOUR HOUSEHOLD.

55. GESL
Sex .......
   1) male
   2) female

56. LEEF
What is your age?
   1) 25 or younger
   2) 26 to 40
   3) 41 to 59
   4) 60 or older

57. WERKZ
Are you employed for 20 hours per week or more?
   1) yes, private employer
   2) yes, public employer
   3) yes, self-employed
   4) no

IF #WERKZ# IS LESS THAN 4, GO TO POSIHH.

58. BEROEP
Are you .......
   1) housewife/husband