Railroad Marketing Support System
Based On the Freight Choice Model

by
Jaikue Park

M.S., Massachusetts Institute of Technology, Cambridge (1992)
M.S., Purdue University, West Lafayette (1989)
B.A., Yonsei University, Seoul (1985)

Submitted to the Department of Civil and Environmental Engineering
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY
in
Transportation and Logistics System Analysis
at the
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 1995

© 1995 Massachusetts Institute of Technology
All rights reserved

Signature of Author

Department of Civil and Environmental Engineering
July 7, 1995

Certified by

Moshe E. Ben-Akiva
Professor of Civil and Environmental Engineering
Thesis Supervisor

Certified by

Denis Bolduc
Associate Professor of Economics, Laval University
Thesis Co-supervisor

Accepted by

Chairman, Departmental Committee on Graduate Students

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

OCT 25 1995

LIBRARIES
Railroad Marketing Support System based on the Freight Choice Model

by
Jai-Kue Park

Submitted to the Department of Civil and Environmental Engineering
on July 7, 1995 in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Transportation and Logistics System Analysis

ABSTRACT

We propose an information system to assist railroad managers in making effective marketing decisions. The system is designed to perform three tasks: collect market information, analyze market response, and generate marketing recommendations. Of the three tasks, the ability to analyze market response is the most critical but least developed. Thus, this thesis focuses on developing a model that analyzes shippers' modal selection.

Based on a literature review, we conclude that a freight choice model should focus on shippers' three objectives: total logistics costs, customer service quality and strategic comparative advantage. We develop freight choice models that reflect such behavioral theory. First, a total logistics cost model is developed both with and without heterogeneity of discount rate. During this development process, we propose an MDI estimator that minimizes the discrepancy of information between actual and predicted shares. Secondly, a model that estimates the impacts of service quality is developed. Because service perceptions are unobservable, we extract perceptions from perceptual indicators and estimate their importance by jointly minimizing information discrepancy in shares and maximizing the likelihood of the observed indicators. Thirdly, we apply the total logistics cost model and the service quality model to the combined data set of revealed preference and stated preference. The use of simulation enabled us to overcome a numerical evaluation problem inherent to complex model specifications.

Lastly, we discuss how estimates of demand models may be used for actual decisions of service design, pricing and industry analysis. We propose a measure of shippers' willingness to pay for improved service based on own-elasticities and propose a market strength map for analyzing industry structure based on cross-elasticities.

Overall, this thesis shows the feasibility of building railroad marketing support system based on the freight choice model. Our research advances the freight choice model by directly reflecting a behavioral theory, utilizing multiple sources of data, and applying advanced statistical technologies.

Thesis Supervisors: Moshe E. Ben-Akiva¹ and Denis Bolduc²
Title¹: Professor of Civil and Environmental Engineering, M.I.T.
Title²: Associate Professor of Economics, Laval University, Quebec, Canada
ACKNOWLEDGMENTS

I have a vision that someday marketing managers will execute and update marketing plans real-time while monitoring market information through multiple computer screens, just as financial vice-presidents order trading, or railroad dispatchers control train movements. I hope that my study can make a step toward this vision of marketing engineering. I have been fortunate to have the opportunity to study under the pioneers in this area.

I wish to thank Professor Ben-Akiva who has been a wonderful advisor, both academically and personally. I have always admired his dedication to research. My thesis is just a small reflection of his continuing research on discrete choice modeling. I thank Professor Bolduc who went over all the materials in the thesis and helped me improve my programming substantially. If there were any mistake, that's because I was not a good listener. Professor Little is a true pioneer of marketing revolution utilizing a marketing information system. He opened my eye about scientific model-based marketing. I also thank Professor Bernstein who taught me the concept of network equilibrium and encouraged me always.

I deeply thank Carl Martland who always looks for positive aspects of a person and encourage him to develop the strengths. I thank the American Association of Railroads Affiliated Laboratory for funding my research, Professor Sussman for teaching me how to approach transportation issues, Professor Barnhart for providing great interactions during my T.A.-ship and her lecture of the large-scale optimization, and Professors Gakenheimer, Salvucci, and Humphrey for evaluating my performance in the urban planning class highly.

I thank everybody who helped make my study at M.I.T. enjoyable: Dinesh, Ashok, John, Qi, Joan, Amalia, Adriana, Geoffrey and Andras in the Demand Modeling Group; Oh-Kyung, Dae-Ki, Prodyut, Sige, Oki, Yoshi, Shenoi, Kazi, Owen and Jiang in the Asian Transportation Interest Group; Marty, Yan, and Bill in the Freight Transportation Seminar; Chris, Michelle, Marcus in the Logistics Interest Group and John, Wen-Cheng, Gordon, Dale, Jeff, Nicola and all other friends in the MST program. I also would like to thank Peter Metz, Tim and Nancy at the Center and Peter and Julie at Professor Ben-Akiva's office for their kindness and consideration.

I thank Professors Se-Chul Oh, Jong-Rak Jung, Anant Balakrishnan, Suresh Chand, R. Chatterjee, Caroline Woo and Chairman Jae-Yeun Lee who wrote me recommendations with strong trust. I also thank Professors Gu-Hyun Jung, Wann Yu, Wujin Chu, Se-Young Kim, Mr. Donovan, Dr. Yeun-Chan Ha, Dr. McCullough and Mr. Byung-Woong Yoo for continuous encouragement. I thank Professors Tae-Hyun Kim, Tetsuo Yai, Young-Bae Kim and Mr. Kwang-Ho Lee who provided valuable advices during their stay at M.I.T.

Finally, I thank my wife, parents, parents-in-law and all other family members who patiently supported my study. My study could never have been completed without strong encouragement from Dr. Daeshik Kim, my father-in-law. Especially, I dedicate this thesis to my wife, Susan, who is positive, cheerful, affectionate, considerate, and sincere. She has been my love-a-lot care-bear and will always be (although I have been a grumpy).
To My Dearest Wife
# TABLE OF CONTENTS

i. Title Page ................................................................. 1

ii. Abstract ........................................................................ 2

iii. Acknowledgments .......................................................... 3

iv. Dedication ...................................................................... 4

v. Table of Contents .......................................................... 5

vi. List of Appendices ......................................................... 8

vii. List of Tables ............................................................... 8

viii. List of Figures ............................................................ 10

I. Introduction ................................................................... 11

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

2. Research Motivation ...................................................... 11

3. Railroad Marketing Support System ............................... 15

4. Freight Choice Model .................................................... 20

5. Thesis Outline ............................................................. 23
II. Literature Review

1. Introduction .................................................. 26
2. Behavioral Theories of Modal Selection ................. 26
3. Factors that Influence Modal Selection ................. 35

III. The Conceptual Framework

1. Introduction .................................................. 49
2. Three Stage Development of The Logistics System .... 49
3. Case Study ..................................................... 66
4. Freight Choice Model ........................................ 69
5. Three Generations of Freight Choice Models .......... 72

IV. The Total Logistics Cost Model ......................... 74

1. Introduction .................................................. 74
2. Modal Share Model .......................................... 75
3. Total Logistics Cost Model ................................ 85
4. The Fixed Rate Model ....................................... 99
5. The Random Rate Model .................................... 106
6. The Agent Effect Model .................................... 119
7. Weaknesses of Total Logistics Cost Model .......... 122

V. The Service Quality Perception Model ................. 123

1. Introduction .................................................. 123
2. Behavioral Framework ...................................... 123
3. Modeling Perception Generation ....................... 128
4. Modeling Alternative-specific Perception ............. 139
5. Combination of Perception Generation and Modal Selection 147
6. Issues related to Perceptual Modeling ................. 155
VI. Stated Preference Data .................................................................................. 163
  1. Introduction .................................................................................................. 163
  2. SP Model with Fixed Rate ............................................................................ 164
  3. SP Model with Random Rates ...................................................................... 169
  4. SP Model with Agent Effects ........................................................................ 172
  5. SP Model with Both Effects .......................................................................... 174
  6. The Combined Model of RP and SP ............................................................... 180

VII. Applications of a Freight Choice Model ..................................................... 185
  1. Introduction .................................................................................................. 185
  2. Own Elasticity to Service Attributes ............................................................. 185
  3. Market Segmentation ................................................................................... 192
  4. Willingness to pay for Service Quality ........................................................ 199
  5. Cross-elasticity and Market Strength Map .................................................. 206

VIII. Conclusion .................................................................................................. 214
  1. Significance of Research ............................................................................. 214
  2. Research Contributions ............................................................................... 216
  3. Future Research .......................................................................................... 220
LIST OF APPENDICES

1. Case Study: Gillette ................................................. 228
2. Statistical Property of MDI Estimators ......................... 247
3. Calculation of Safety Stock Holding Costs ................. 251
4. Gauss-Hermite Quadrature ......................................... 253
5. Delta Method .......................................................... 254
6. Simulation Method for Parameter Estimation .............. 255
7. Stated Preference Data ............................................. 272

LIST OF TABLES

2.1. The 1994 Intermodal Index ................................. 41
4.1. Total Logistics Cost table ........................................ 87
4.2. The RP Model with i.i.d. Probit Probability ........... 101
4.3. The RP Model with Randomly-distributed Discount Rate 116
4.4. The RP Model with Agent Effects ...................... 121
5.1. The Sequential Model Conditional on lef ............ 132
5.2. The Full MLE Conditioning on rsp ................... 134
5.3. The Full MLE with Three Perceptual Indicators ........ 137
5.4. The Simulated MLE of Alternative-specific Perceptions 146
5.5. The Sequential Model for Service Quality Perception ... 149
5.6. The Joint Model for Service Quality Perception ........ 153
5.7. Average Statistics of Total Logistics Costs and Service Quality 154
5.8. Comparison of the Effects of Total Logistics Costs and Service Quality 155
6.1. The Fixed Discount Rate SP Model 168
6.2. The SP Model with Random Rates 171
6.3. The SP Model with Agent Effects 174
6.4. The SP Model with Combined Effects (Error Component Model) 176
6.5. The SP Model with Combined Effects (Rate Decomposition Model) 179
6.6. The Combined Model of RP and SP data 184
7.1. Own Elasticity of Choice Probability to Freight Rate 188
7.2. Own Elasticity of Choice Probability to On-time Delivery 189
7.3. Own Elasticity of Choice Probability to Transit Time 190
7.4. Own Elasticity of Choice Probability to Loss and Damage 191
7.5. Own Elasticity of Choice Probability to Equipment Usability 191
7.6. Shippers’ Willingness to Pay for an Improvement in On-time Delivery 201
7.7. Shippers’ Willingness to Pay for an Improvement in Transit Time 202
7.8. Shippers’ Willingness to Pay for an Improvement in Loss and Damage 203
7.9. Shippers’ Willingness to Pay for an Improvement in Equipment Usability 203
7.10. Cross-elasticities of Freight rate 209
7.11. Cross-elasticities of On-time Delivery 210
7.12. Cross-elasticities of Transit Time 211
7.13. Cross-elasticities of Loss and Damage 211
LIST OF FIGURES

1.1. Growth in Intercity Freight Revenues (Truck vs. Rail) 13
1.2. Growth in Intermodal Freight 14
1.3. Rail Share of Automotive Traffic 14
1.4. Conceptual Process of Marketing Planning 17
1.5. System Configuration of Railroad Marketing Support System 19
2.1. Process-based Industrial Buying Model 33
4.1. The Revealed Preference Model 75
4.2. The Distribution of Perceived Transit Time 91
4.3. Inconsistent Estimates when Intercepts are Heterogeneous 109
4.4. Inconsistent Estimates when Slope is Heterogeneous 109
4.5. Heterogeneity of Discount Rate 114
5.1. The Modal Selection Model with Perception Generation 147
6.1. The Stated Preference Model 164
6.2. Example of Stated Preference Experiment 165
6.3. Distribution of Stated Preference 166
6.4. The Combined Model of RP and SP 180
7.1. The Market Strength Map 213
8.1. A Framework for Freight Choice Models 217
8.2. A General Framework for Static Freight Choice Models 222
8.3. A General Framework for Dynamic Freight Choice Models 224
8.4. Freight Choice Model for Policy Makers 226
Chapter 1. Introduction

1.1. Introduction

This thesis focuses on designing and developing a railroad marketing support system (RMSS). We believe that a decision support system is necessary for railroad marketing to utilize new growth opportunities. As railroads become more customer-oriented, the need for demand analysis grows stronger. An information system can assist railroad marketing managers by providing a structured data base, interpreting the data with advanced statistical techniques, and integrating the knowledge of marketing, customer service and sales managers.

The RMSS proposed in this thesis is a model-oriented information system that utilizes a demand response function estimated from shippers’ survey. We begin in section 1.2 with research motivation concerning the necessity for RMSS. In section 1.3, we introduce the concept, desired properties and system configuration of RMSS. In section 1.4, we introduce the freight choice model as a key component of RMSS. Section 1.5 contains an outline of the thesis.

1.2. Research Motivation

Railroads underwent a period of downsizing in the 1980s despite the continued rapid growth of inter-city freight volume (Figure 1.1). After passage of the Staggers Act in 1980, railroad companies were able to pursue a survival strategy based upon elimination or sale of unprofitable local lines, terminal consolidation, reduction of employees, heavier axle loads and computer control systems. Reviewing the performance of the U.S. railroads since deregulation, Martland (1989) concluded that while downsizing improved financial
performance by making railroads focus on large shippers and bulk traffic, it failed to improve either the industry's cost competitiveness or productivity. He suggested that an alternative strategy of pursuing revenue growth through the development of innovative services and aggressive marketing programs could improve railroad competitiveness.

In fact, innovative service and aggressive marketing seem to have finally reversed the trend of declining rail share, as shown in the growth of intermodal and automotive freight (Figure 1.2). The intermodal operation of unit trains that move double-stacked containers from the west to the east coast is growing in double digits annually and appears profitable thanks to a low cost structure (Welty 1994a). The intermodal operation of domestic freight using piggyback is also growing due to a shortage in the number of long distance drivers. Strategic partnerships between railroads and trucking companies spur the growth of the intermodal market (Welty 1994d). The rail share of automotive freight has grown due to reliable delivery schedules and the development of equipment that can protect finished paint and reduce vibration during transportation (Figure 1.3, Welty 1994e). Such examples show that the growth of railroads may occur through attention to customer needs, technology trends and industry structure.

Many railroad presidents now emphasize that railroads should switch from a survival strategy to a growth strategy (e.g. Welty 1994b). Accordingly, we may expect that mergers among railroads will soon emphasize an opportunity of growing revenue (e.g. trans-continental line, seamless service and single source of contact) rather than an opportunity of decreasing operating costs (e.g. rationalizing parallel tracks and terminals). The 1990s present a good opportunity to a railroad company that can utilize market data efficiently, since shippers are increasingly willing to work with carriers to redesign their logistics processes and networks (e.g. strategic partnership or outsourcing) in order to improve cost competitiveness or service quality (Welty 1994c, also refer to Chapter 3).
Source: Transportation in America

Figure 1.1. Freight Revenues (Truck Intercity vs. Rail)
Figure 1.1. Growth in Intermodal Freight

Figure 1.3. Rail Share of Automotive Traffic
1.3. Railroad Marketing Support System (RMSS)

The Concept of RMSS

Experts predict that companies able to utilize market information will find new business opportunities, while those unable to do so will not survive (Blattberg, Glazer and Little 1994). They recommend companies to recognize information technology as an enabler which will transform the firm into a truly customer-driven organization, and to invest aggressively in building marketing information systems, judging the investment on revenue growth potential rather than on cost projections. In fact, advanced railroads have already started to change their organizations towards such direction. Below is an example at Union Pacific:

At Union Pacific, all new employees who aspire to a management position must first become a "data integrity analyst." Union Pacific carries 13,000 shipments a day on 700 trains running on 19,000 miles of track. Since 1986, Union Pacific has been getting rid of much of the paperwork to handle shipping orders, bills of lading, and invoices from its customers and their shipping agents. Instead, a computerized electronic data interchange (EDI) system is used to update data. Now, some two-thirds of all the railroad's client communications - up from just 3% eight years ago - are managed via EDI from a single customer center in St. Louis rather than through the 40 offices that formerly handled the unwieldy paper flow.

Empowered by EDI, the data integrity analysts keep tabs on all of the customers' contacts with the railroad. They create a detailed electronic profile for each shipper that permits customer service representatives to facilitate order taking or to resolve questions. They also provide the information that dispatchers in Omaha use to track shipments and that clerks in accounting rely on for accurate billing information.

Just as valuable as the huge improvement in efficiency that EDI has wrought [...], are the "fabulously rich" strategic uses that Union Pacific can make of the amassed data. The railroad's goal is to mine that treasure-trove to be able to offer customers higher value-added services tailored to their needs. [...] A manager uses the EDI customer profiles to build new databases that might, say, help uncover evolving patterns in a shipper's usage of the railroad's services and sell the shipper a new service.¹

¹ Richman (1994), p. 59
The example shows the direction that top managers of railroad companies pursue. Railroad managers now must focus on analyzing shipment data in order to understand customers' transportation needs and discover a new revenue stream, rather than focus on handling paperwork for shipments. A system that supports railroad managers with summary reports on the trends of major variables is desired. If the system can analyze freight data, explain shippers' modal shares and predict how shippers will respond to railroad marketing decisions, it will greatly help railroad managers. This research purports to design such a decision support system, which henceforth we will call a railroad marketing support system (RMSS).

**Desired Properties of RMSS**

Although there may be an extensive list of desirable properties for an RMSS, we believe that at minimum, an RMSS should be consistent, decision-analytic and proactive.

*Consistent*  
RMSS should be consistent with the overall marketing strategy and should support the whole marketing planning process which may be classified into two stages (Figure 1.4). At the first stage, top managers analyze customer needs, company resources and competitive structure, and determine marketing strategy. At the second stage, local managers establish specific marketing plans. In order to support this process, a railroad needs to construct three databases: customer, company and competitor DB. It is also desired that RMSS provides automated marketing reports on tactical areas such as service design, sales force management and pricing.
**Decision-Analytic**

Instead of providing simple average statistics, RMSS should be problem-specific and decision-analytic. By utilizing advanced marketing models, RMSS should allow managers to run different scenarios and to quantify expected revenue before making marketing decisions. Moreover, by internalizing railroad marketing managers' know-how's, RMSS provide specific guidelines to marketing decisions. A model-based system is recommended as superior to a data-oriented system since it promises to analyze market response with advanced statistical analysis and document expert knowledge of marketing managers in terms of quantitative models (Little 1979). In particular, we design RMSS around the disaggregate choice model. Disaggregate models analyze individual shipper behaviors and can help railroads effectively develop customer-tailored services. A model-based system also becomes inexpensive to build, since the advancement of computer and communication technology enables companies to acquire, process and analyze market information at reasonable costs.

**Proactive**

Not only should RMSS provide model-based analytic reports, but also its recommendations should be proactive to market changes. Recent
developments in game theory suggest that if a company makes a commitment to a strategy, follower companies will set their strategies after accepting the committed strategy as fixed. If a pioneer can determine followers' reactions, it can command higher profits than followers (Tirole 1989). For example, if a railroad could predict how truckers would price intermodal services after they were given the railroad's intermodal prices, it can earn extra profit by devising an intermodal price carefully. Since such models utilize shippers' response to intermodal price, the estimation of demand function is necessary for the development of a proactive decision support system.

System Configuration

In order to build a model-based proactive RMSS that is consistent with marketing activities, a railroad needs to monitor market response constantly, to forecast demand response accurately, and to determine the best marketing plans. RMSS can support such activities by utilizing three systems: marketing information system, marketing analysis system, and marketing report system, as shown in Figure 1.5.²

The marketing information system constantly monitors market situations, collects all the data necessary to conduct market analysis, and structures the data within marketing databases. Input data range widely from aggregate statistics on the general economy and on modal competition to disaggregate market survey results and railroad marketing managers' judgments. The input data are structured in three separate databases (DB), i.e., customer DB, competitor DB, and performance DB, so that they can be used for strategic analysis. Customer DB must be categorized by the industry in which customers belong to. Expected freight volumes are then adjusted by the growth rate of the industry. Competitor DB is composed of information about service performance of other modes (e.g., truck, barge, etc.) and of competing railroads. Performance DB includes diverse records of operational performance and customer service. The three DB's allow analysis results of RMSS to be consistent with overall marketing activities.

² This system modifies the system recommended by Little (1979).
The marketing analysis system operates two modules: the marketing statistics bank and the marketing model bank. The marketing statistics bank analyzes market information by using various statistical methods such as regression analysis, factor analysis, cluster analysis, discriminant analysis, and discrete choice analysis. Its primary roles are to estimate demand response functions by market segments and to predict how shippers will respond to railroad marketing decisions. The marketing model bank, on the other hand, stores various mathematical models that process market statistics and demand response functions, and in turn, provide recommendations for various marketing decisions such as service design, pricing and sales effort allocation. Recent developments in marketing science models allow marketing managers to check their judgments with model-based recommendations.

The marketing report system receives the results of marketing models and generates marketing reports that marketing managers can use for day-to-day operations.
By automating the process of generating routine marketing reports, this system greatly reduces the workload of marketing managers and allows them to focus on understanding customers' problems and on devising customized total transportation solutions.

Among the three systems, the ability to analyze market response is the most critical in order for RMSS to be decision-analytic and proactive, but has been the least developed. Therefore, the primary focus of this thesis is given to the market analysis system. In particular, we have focused on the marketing statistics banks which utilize the statistical analysis of discrete choice modeling. We hope that the methodology developed in this thesis can provide a tool to analyze shippers' behaviors. When companies actually build this system, efforts should be made to elaborate the marketing information system and the marketing report system. Railroads have to determine how to collect market data in quantitative terms, how to store and structure the database, what report forms are desired by managers, and how often marketing reports should be updated.

1.4. Freight Choice Modeling

We propose that railroads use the discrete choice modeling approach as the core of railroad marketing support system (RMSS). If we can identify true relationships among variables, we can forecast a shipper's modal selection by monitoring his network characteristics. We will also be able to predict the direction and magnitude of changes in modal share corresponding to changes in service performance. Discrete choice models allow such individual-level predictions.

The basic framework of the discrete choice model is well-summarized in Ben-Akiva and Lerman (1985). The model has been widely used by public agencies involved in passenger, and to a lesser extent, freight transportation. The applications of discrete choice modeling to freight data were made by Roberts, Brigham and Miller (1977), Chiang, Roberts and Ben-Akiva (1980), McFadden, Winston and Boersch-Supan (1985), Temple, Barker \& Sloane / Strategic Planning Associates (1991), and Vieira (1992). For
summary, refer to Winston (1983). The freight choice modeling approach has many advantages which support its use.

First, we can represent a behavioral theory by adopting this approach. For instance, the choice modeling approach explicitly models how shippers value logistics. The model shows how modal choices change depending on total logistics costs and service quality perceptions. Using this approach, rail companies can judge railroad performance in terms of shippers' criteria rather than in terms of rail industry standards.

Secondly, the choice modeling approach allows prediction of market response at the disaggregate shipper level. Since the approach looks at shipment-specific situations where an individual shipper selects a particular mode for a given shipment, it does not involve bias from aggregating data and allows a researcher to analyze modal choices at shipper and shipment level. Since railroad customers are composed of a relatively small number of large shippers, RMSS will help railroads design customer-tailored services by predicting how each of its customers will respond to new marketing programs.

Thirdly, the choice modeling approach reflects competition among different modes. It estimates choice probabilities as a function of relative utilities of available modes. Using this model, rail companies can understand differential effects of freight rate and service quality due to the competitive structure. Thus, the model will be useful in general merchandise transportation markets where competition with trucking is fierce.

Fourthly, the choice modeling approach can also be used to predict demands for new services which may not exist today but may be under consideration by railroad managers. Techniques such as conjoint experiments are available to analyze demand for new services. Combination of such data with actual share data will greatly enhance the area to which the choice modeling approach can be applied.
Despite all the advantages, few railroads have used the choice modeling approach for analyzing freight demand. There may be many reasons for little usage among practitioners.

First, as an advanced method, the freight choice model is difficult to understand. Railroad managers will be reluctant to rely on model recommendations if they cannot understand the model. In order to overcome mistrust, model users should be educated about the behavioral concept, the estimation algorithm and application examples of discrete choice models.

Second, the freight choice model requires a large database (i.e. shipper-specific data on all available modes). In order to build it, railroad companies should be committed to collecting data at the shipper level and on both their own and competitors' performance. The collection, administration and analysis of data is neither easy nor inexpensive. Yet, this should be viewed as a strength rather than a weakness. By incorporating a greater amount of information, the model not only improves the reliability of predicted modal shares but also builds comparative advantage against other carriers.

Third, few marketing decision models utilize freight choice models. Railroad managers would not want to invest in estimating a model, if they do not know how to use estimation results. Although we will show some direction about how freight choice models can be utilized in making marketing decisions, further research is necessary. Mathematical models that support various marketing decisions should be developed simultaneously with structuring the data collection process and developing a freight choice model.

We believe that the benefits of proper demand analysis will outweigh the costs of education, data collection and model development. Thus, we have pursued building a

---

3 Only a few reports are available on railroads' applications of discrete choice model to freight data. The ARES project by Burlington Northern Railroad, which we will discuss in Chapter 7, is the best known application. Except the ARES project, the researches cited in the previous page are almost all of its applications.
marketing support system based on demand analysis. In particular, our major efforts are placed on developing a freight choice model that predicts shippers' modal selection given network characteristics. This thesis presents such efforts in the following order:

1.5 Outline of the Thesis

Chapter 2 reviews previous research on behavioral theories and factors that influence modal or carrier selection. Behavioral theories are categorized into task-oriented, perception-oriented and process-oriented models. Factors are also reviewed in three categories (inbound, outbound, and intra-company shipment). The review suggests that desired service differ depending on job characteristics (transportation, purchasing, production, or marketing) and shipment characteristics.

Most of the previous research we reviewed in Chapter 2 lacks a strategy-oriented view which becomes, we believe, the most important perspective in modal selection. We thus perform detailed case studies and propose the three-stage development model of the logistics system in Chapter 3. The model argues that the role of logistics systems develops from a simple order processing system through a value-adding system to a strategic differentiation system. Accordingly, shippers' objective function of selecting carriers changes from cost minimization through customer service maximization to comparative advantage sustenance. This behavioral theory serves the basic framework of this thesis.

Chapter 4 introduces a freight choice model which assumes that shippers minimize total logistics costs. Using the minimum discrimination information (MDI) estimation, we estimate the importance of cost components and predict modal shares using shipment environment data. We also note that total logistics costs are influenced greatly by the discount rate that shippers apply in calculating capital costs of holding inventory, and that the discount rate varies among shippers. We discuss the estimation of a freight choice model with such heterogeneity of discount rate.
Chapter 5 extends the cost-based model into incorporating service quality perception. The discussion includes how to quantify service quality and how to estimate its importance in modal choice. Although service perceptions are unobservable, we extract perceptions from indicators and estimate their importance in modal selection by maximizing the likelihood of observing modal shares and perceptual indicators jointly.

Chapter 6 points to the fact that we can collect more data by asking shippers to make hypothetical choices. This approach provides stated preference data. By combining revealed and stated preference data, we can improve external validity and increase statistical efficiency. We discuss how to combine both RP and SP data sets with and without heterogeneity of discount rate.

Chapter 7 discusses applications of a freight choice model based mainly on the analysis of elasticities. We discuss how to design customer-tailored service by utilizing disaggregate elasticities, how to prioritize projects by utilizing the estimates of shippers' willingness to pay for service improvements and how to select areas for strategic differentiation by utilizing market strength maps.

In chapter 8, we conclude by discussing research contributions and by proposing areas for future research.

All the discussions in this thesis will be based on a survey that a major U.S. railroad company administered in 1988. The data set is one of the largest shipper surveys available and contains both RP and SP data. The survey included 166 transportation managers who ship one of five commodities: paper, aluminum, pet food, plastics and tires. Each shipper rated the performance of available transportation modes for their two largest corridors. While they are mostly truck users, their volumes are large; all shippers ship at least a $1 million worth of their commodity annually. Thus, they can be potential customers of railroads. For detailed descriptions of the survey, refer to Vieira (1992). Before running the model, we screened the data and dropped 21 out of 166 respondents.
who did not report the transit time or freight rate of the mode they use. The following table shows the average modal share (%) by product type:

<table>
<thead>
<tr>
<th>PRODUCT</th>
<th>Truck</th>
<th>Rail</th>
<th>Intermodal</th>
<th>Total Obs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paper</td>
<td>67.79</td>
<td>28.92</td>
<td>3.29</td>
<td>97</td>
</tr>
<tr>
<td>Aluminum</td>
<td>89.23</td>
<td>8.96</td>
<td>1.81</td>
<td>69</td>
</tr>
<tr>
<td>Pet food</td>
<td>87.17</td>
<td>6.97</td>
<td>5.85</td>
<td>40</td>
</tr>
<tr>
<td>Plastics</td>
<td>58.79</td>
<td>38.25</td>
<td>2.96</td>
<td>48</td>
</tr>
<tr>
<td>Tires</td>
<td>94.42</td>
<td>2.56</td>
<td>3.06</td>
<td>36</td>
</tr>
<tr>
<td>Overall</td>
<td>77.38</td>
<td>19.41</td>
<td>3.21</td>
<td>290</td>
</tr>
</tbody>
</table>
Chapter 2.
Literature Review

2.1. Introduction

In this Chapter, we review previous research on how logistics managers select freight modes. Section 2.2 reviews behavioral theories of modal selection and categorizes them into task-based, perception-based and process-based models. Many industry-sponsored researches simply ignore behavioral theories, and instead make a list of service attributes with no underlying theory, and ask transportation managers to rate their importance. Section 2.3 reviews the results of such stated importance approach. While transportation-specific, their results are subject to response bias and lack of behavioral theory.

2.2. Behavioral Theory of Industrial Purchase

In order to develop efficient logistics systems, shippers typically evaluate potential carriers, compare competing offers, determine the optimal modal mix, and select appropriate modes and carriers. During the process, shippers are influenced by both rational and perceptual factors. Moreover, logistics managers, as role players within an organization, may also consider evaluations by other departments when they select carriers. We can categorize shippers' modal choices into task-based, perception-based and process-based models, as shown in the table below. No single model can explain the carrier selection process fully. But the discussion of the models will point out variables that can be used to advantage by marketing managers of carriers, and will indicate the direction that can be used to design research analysis and measurement techniques.
1. Task-oriented models
   o Single Agent Models
     - Transportation cost model
     - Total logistics cost model
     - Discrete choice model
     - Neoclassical microeconomic model
   o Game-theoretic Models
   o Management-based Models
2. Perception-oriented models
3. Process-oriented model

2.2.1. Task-oriented Models

Task-oriented models argue that shippers make rational purchase decisions given the choice situations. By rationality, these models assume that shippers can write down their objective functions clearly, and that they can control variables that determine the objective functions. Most of these models are microeconomic models and can be classified into single agent models and game-theoretic models depending on whether interactions between shippers and carriers are explicitly considered.

1) Single Agent Models

A. Transportation Cost Model

The simplest model of modal selection is to assume that logistics managers minimize only transportation costs (this is still the basis of many optimization models, McGinnis 1989). Classical economic theory states that price will stabilize at the long-term average cost, which is a function of production technology. Average costs are broken down into fixed and variable costs. Variable costs in transportation usually depend on shipment distance or shipment size. Since truck and rail differ in terms of initial capital investments and operating structure, economists argue that there are certain thresholds that make a certain mode efficient or inefficient. For instance, the model may suggest that truck is more efficient than rail when shipment distance is less than a certain threshold.
(e.g. 600 miles) or when shipments weigh less than a certain threshold (e.g. 100,000 pounds). Refer to the below figure:

![Graph showing truck and rail costs vs. distance.](image)

Nevertheless, Roth (1977) found, using a 1972 Census data set, that rail is more competitive than truck for shipments weighing more than 60,000 pounds even for short distance. Morton (1972) also found by analyzing the Census of Transportation data for 1967 that rails and trucks compete for all distances, and for shipments between 10,000 and 90,000 pounds. Therefore, the usefulness of this deterministic model is limited. The limitation may come from its assumptions such as (1) carriers price at long-term average cost, (2) all carriers of the same mode have the same cost structure, and (3) shippers consider only cost advantages, ignoring non-price attributes such as service quality. Obviously, transportation markets are differentiated by service/price packages, and shippers consider more than just transportation costs.

The transportation cost model tends to emphasize freight rate and availability of service. For railroad or intermodal services that operate on fixed rail networks, the availability of terminals or sidings may be important constraints that influence shippers' choice sets. But transportation costs alone cannot explain shippers' modal selections completely. Shippers may save transportation costs by using the cheapest transportation. But they have to make large investments in inventory and packaging and may lose business that cannot be served well by these low-cost means. Also, even when shippers
use premium transportation, they may be able to reduce warehousing costs by having fewer distribution centers.

B. Total Logistics Cost Model

A more advanced model assumes that shippers minimize total logistics costs that include transportation costs as well as inventory holding costs (Meyer, Peck, Stenason and Zwick 1959 and Ann Friedlaender 1969). This model recognizes that several individual items interact in complex ways to determine total logistics costs and that an attempt to minimize any single cost element can actually increase total costs. For instance, if transit time is long, shippers incur high inventory holding costs during the long in-transit time. If transit time is unpredictable, shippers have to hold high safety stocks which increase safety stock holding costs. Since the total logistics cost model emphasizes inventory costs relative to transportation costs, the model is also called an inventory-theoretic model. For more discussion on total logistics costs, refer to Chapter 4.

The best known application of this model is the ShipSmart software that was actively promoted by Burlington Northern Railroad to its customers and which is probably the most frequently used model for modal selection. The software adopts the framework of total logistics costs developed by Roberts (1976) and makes normative recommendations on modal selection as well as on shipment size. Its recommendations are based on the calculation of the most cost-efficient transportation mode, assuming that total logistics costs can be defined exactly based on shipment characteristics (e.g. product price, demand projections, shipment distance, and pre-determined discount rate). This deterministic approach is also limited since it assumes that (1) all carriers of the same mode charge the same price and provide the same performance and that (2) shippers consider only cost advantages, ignoring non-price attributes such as perceptions on service quality.
C. Discrete Choice Model

In many occasions, normative recommendations provided by cost models do not coincide with actual choices. A major problem has been that it is usually difficult to define and measure total logistics costs exactly. A statistical approach that estimates discount rate and the importance of cost variables is desired. As a result, the concept of minimizing total logistics costs is combined with a new econometric modeling methodology which explains choices among discrete alternatives as decision makers’ efforts to optimize random utilities.

As can be seen in the citations, this approach has been researched mostly at M.I.T. Roberts conducted pioneering studies (1976, 1977), applying the concept of total logistics costs to the estimation of rail shares in response to different public policies. Chiang (1979) extended the model by estimating the joint decision of modal selection and shipment size. McFadden, Winston and Boersch-Supan (1985) estimated the joint decision using a non-random sample. Vieira (1992) estimated the value of service and market segmentation in modal selection. He also compared the results of a linear model and a total logistics cost model and claimed that the latter is superior in terms of data fit. Since this approach is considered the state-of-art model of freight choice (Winston 1983), this thesis will adopt and further elaborate this approach.

D. Neo-classical Microeconomic Model

This model tries to explain modal splits by assuming that a firm minimizes its transportation-related costs. Suppose that a firm's cost ($C$) is a function of total output ($Y$), freight rate ($P$), shipment characteristics ($Q$), and factor prices other than freight rate ($w$), then

$$C = C(Y, P, Q, w).$$

The demand for transportation mode ($D$) is derived by using Shephard's lemma (Winston 1983), i.e.
\[ D_{in} = \frac{\partial C_{tn}}{\partial P_{tn}} \]

where \( i \) denotes mode, \( t \) denotes time, and \( n \) denotes shipper. If we further assume that the firm's technology can be characterized by a translog cost function that represents a local second-order approximation of an arbitrary cost function, the modal split model is greatly simplified while still allowing the cross-elasticities of freight rates to be estimated).

\[
S_{in} = \frac{P_{in} D_{in}}{C_{tn}} = \frac{\partial \ln C_{in}}{\partial \ln P_{tn}} = \alpha_0 + \sum_i \beta_i \ln P_{in} + \sum_k \gamma_{ik} \ln Q_{kn} + \sum_h \delta_{ih} \ln w_{hn} + \phi_i \ln Y_{in}
\]

where \( S_{in} \) is the share of shipments shipped through mode \( i \) among total shipments made by shipper \( n \) at time \( t \). Since this model depends on the assumption that production costs have a form of a Cobb-Douglas function, McCullough (1994) extends it into the generalized trans-log function and compares price elasticities of different modes. By including interaction terms among variables, the generalized form is more robust to the form of cost functions. But, this approach still contains the drawback of not reflecting inventory holding costs, effects of shipment size and frequency, or perceptions on service quality.

2) Game-theoretic Models

Game-theoretic models have been under active research recently. Game-theoretic models are suited to modeling organizational relationship between carriers and shippers. Game theory models predict decisions of carriers and shippers in the equilibrium state where no party can profit by deviation, given the game rules. The characteristics of equilibrium depend on the level of competition, industry capacity, number of players and game rules. Yet, most models focus on gaining insight into long-term equilibrium and we have not seen practical models yet.
One application is a lateral relationship where one firm moves first and the other firm follows (Tirole 1989). For instance, if there are plenty of carriers who are willing to undercut, shippers will make freight decisions first and carriers will follow. On the other hand, if rail capacity and truck drivers for long-distance are in shortage, carriers may determine freight rate first and shippers accept the rate. Another example is the intermodal pricing. If a railroad can exercise monopoly power on certain routes and trucking companies must take the railroad’s intermodal price, then the railroad can anticipate trucking companies’ responses to its price levels and calculate the optimum price level that maximizes its profits given trucking companies’ responses. Its results depend heavily on the assumptions that the first mover knows the exact profit function of the follower and that the follower believes the first mover’s commitment on the original decision. Another application is the reciprocity relationship. For instance, if shippers and carriers have close business relationships, they are likely to repeat business deals, i.e. "If you purchase rail tracks from my company, I will use your transportation service for shipping my steel." When the game between a shipper and a carrier occurs in many periods, the game becomes a dynamic interaction game.

3) Management-based Models

Management-based models argue that savings in logistics costs from keeping economic order quantity are small relative to the loss that production departments would incur when they have to stop production lines or to the loss that marketing departments would incur when they have to turn away customers due to inventory shortage. The tasks of logistics managers are to ensure smooth flow of inbound and outbound goods so that they can best support production and marketing efforts. With this perspective, shipment size and frequency are pre-determined by production and marketing plans such as Material Requirement Planning, Just-in-time Production (Miller 1994), and Distribution Requirement Planning (Martin 1990). Freight mode and carrier are selected so that they can support pre-specified inventory plans with least costs and without interruptions.
2.2.2. Perception-oriented Models

Perceptions of service quality can influence carrier choice. For instance, logistics managers may select a mode that is convenient (e.g. providing pickup and delivery, designating single source of contact and real-time shipment tracing) or familiar (e.g. knowledge about companies' mission, strategic directions or business process). Recent popularity of core carrier programs may represent efforts to institutionalize and exploit such perceptions. However, little literature is available for empirical results about the effects of perceptions on carrier selection. More research is desired in this area.

2.2.3. Process-oriented Models

The third approach to carrier selection is process-based models. There are several process-based models (e.g. Sheth model, Webster and Wind model, Anderson and Chambers model, and Choffray and Lilien model) which can be applied to the explicit modeling of the core-carrier selection process. Since they are all similar, we will combine them into one framework (mostly by modifying the Sheth model) and discuss the modified model here. For greater details, refer to Figure 2.1 and Webster (1991).

Figure 2.1. Process-based Industrial Buying Model
The key concepts are information sources, individual, product and company characteristics, joint decisions, and conflict resolution. Information sources include salesmen, exhibitions, trade shows, direct mail, press releases, journal advertising, professional and technical conferences, trade news, worth-of-mouth and others. Individuals differ in terms of specialized education, role orientation, and life style, and thus, have different expectations and perceptions. In addition, task-specific factors (e.g. time pressure, perceived risk, and type of purchase) and company-specific factors (e.g. organization orientation, organization size, and degree of concentration) also influence the individual buying process. Since transportation service would influence the performance of multiple individuals such as purchasing agents, engineers and users, carrier selection involves joint decisions which usually require conflict resolution. Finally, a carrier is selected and its performance is evaluated by each individual and provides (dis)satisfaction which influences carrier selection in the next period.

Sheth distinguishes between autonomous decisions that are delegated to a single individual, and joint decisions that are made collectively by the participants in the decision process. In the context of carrier selection, periodic evaluation and selection of core carriers would be an example of joint decisions. Given core carriers selected for each mode, the selection of a carrier for a particular shipment is an example of autonomous decision. Railroad marketing departments should therefore make two separate efforts, one to improve the prospect of being chosen as a core carrier of major customers, and the other to improve the frequency of being utilized for day-to-day shipments.

For autonomous decisions, Sheth pointed out the importance of perceptual distortion, which is the extent to which each participant modifies information to make it consistent with his existing beliefs and previous experience. For joint decisions, he pointed out perceived risk as another important variable. There are two type of risk. One is performance risk which is associated with the extent to which the product meets the buyer's expectations with respect to actual performance. The other is perceptual risk which is associated with how other relevant persons would react to the decision and how
the buyer himself feels about the outcome. The greater the uncertainty and the more significant the consequences, the higher the degree of perceived risk. Buyers can adopt several tactics for reducing perceived risk, including gathering information, avoiding a decision, passing responsibility on to other persons, minimizing the investment of time and money in the decision, or simply reducing goals. If conflicts occur during joint decision process, conflicts may be resolved through problem solving, persuasion, bargaining, or politicking.

This process model implies that when a traffic manager selects a mode, he will consider not only task-specific characteristics but also evaluations by other departments. It also suggests that the goal and risk tolerance of individual managers and those of a company may be different. The inclusion of process-oriented variables into a choice model is recommended.

2.3. Factors that Influence Shippers' Modal Selection

If the development of a behavioral theory has received major attention in economic and industrial marketing literature, the measurement of the importance of service attributes has received major attention in industry-sponsored logistics researches.

Self-stated importance approach has been the most popular way to measure the importance of service attributes. By providing a survey that contains a list of service attributes that may influence modal selection, railroad managers or third-party consultants directly ask shippers to rate the importance of the given attributes. Based on the average value, we can order service attributes in the order of importance. Its popularity comes from several convenience factors, e.g. it is easy to understand for both railroad managers and shippers. Since shippers can easily understand the survey context, it is easy to collect data. Also, since the analysis needs to calculate only the averages, it is easy to analyze the collected data.
We categorize these types of researches in terms of inbound, outbound and intra-company shipments. Most of them focused on the outbound movements. Few studied the differences of transportation needs, expectations and evaluations between inbound and outbound shipments, or between shippers and receivers. Few also studied the differences in the decision criteria of modal selection and carrier selection.

2.3.1. Outbound Movements

For a survey of the criteria for carrier selection, we may look at McGinnis's report (McGinnis 1978, 1989, 1990). He reviews 14 empirical studies in total and studies what variables were found to be important to making a transportation choice. The seven important attributes he found are as follow:

1) Freight rate (costs, charges, rates)
2) Reliability (reliability of delivery time)
3) Transit time (time-in-transit, speed, delivery time)
4) Over, short, and damaged (loss, damage, claims processing, tracing)
5) Shipper market considerations (customer service, user satisfaction, market competitiveness, market influences)
6) Carrier considerations (availability, capability, reputation, special equipment)
7) Product characteristics (perishability, packaging requirements, new products)

<table>
<thead>
<tr>
<th>Variable (from the above list)</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>Before 1980</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Saleh 1970</td>
<td>v</td>
<td>v</td>
<td>v</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bardi 1971</td>
<td>v</td>
<td>v</td>
<td>v</td>
<td>v</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Evans &amp; Southard 1974</td>
<td>v</td>
<td></td>
<td>v</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jones 1975</td>
<td>v</td>
<td>v</td>
<td>v</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gilmour 1976</td>
<td>v</td>
<td>v</td>
<td>v</td>
<td>v</td>
<td>v</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stock &amp; LaLonde 1977</td>
<td>v</td>
<td>v</td>
<td>v</td>
<td>v</td>
<td>v</td>
<td>v</td>
<td>v</td>
</tr>
<tr>
<td>McGinnis 1978</td>
<td>v</td>
<td>v</td>
<td>v</td>
<td>v</td>
<td>v</td>
<td></td>
<td></td>
</tr>
<tr>
<td>After 1980</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chow &amp; Poist 1984</td>
<td>v</td>
<td>v</td>
<td>v</td>
<td>v</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Burning &amp; Lynagh 1984</td>
<td>v</td>
<td>v</td>
<td>v</td>
<td>v</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Burdg &amp; Daley 1985</td>
<td>v</td>
<td>v</td>
<td>v</td>
<td>v</td>
<td>v</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brand &amp; Grabner 1987</td>
<td>v</td>
<td>v</td>
<td>v</td>
<td>v</td>
<td>v</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hayuth 1985</td>
<td>v</td>
<td>v</td>
<td>v</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quinn 1987</td>
<td>v</td>
<td>v</td>
<td>v</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bardi et al. 1989</td>
<td>v</td>
<td>v</td>
<td>v</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Variables marked with "v" are those that were reported to be significantly important. Overall, freight rate, reliability and transit time were consistently cited as being important. Reliability, in particular, was rated as being the most important variable in all studies. Transit time was rated frequently as being more important than freight rates. On the other hand, shipper market considerations, carrier considerations and product characteristics received mixed ratings. In particular, the low importance given to customer satisfaction indicates either that logistics managers still think in terms of operation efficiency or that survey questionnaires did not properly address the dimension of customer satisfaction.

As an example of a study for understanding modal selection, we can cite McGinnis’ (1978) survey. He surveyed attitudinal indicators from 351 U.S. traffic executives and reports factor analysis results on the data. He extracts the following seven attitudinal factors (The percentage inside each parenthesis indicates the portion of total variation among indicators explained by the factor) (McGinnis 1978).

<table>
<thead>
<tr>
<th>Factor</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speed and reliability</td>
<td>14.8%</td>
</tr>
<tr>
<td>Loss and damage</td>
<td>9.4%</td>
</tr>
<tr>
<td>Inventories</td>
<td>8.4%</td>
</tr>
<tr>
<td>Freight rates</td>
<td>7.5%</td>
</tr>
<tr>
<td>Market competitiveness</td>
<td>5.6%</td>
</tr>
<tr>
<td>Company policy and customer influence</td>
<td>5.2%</td>
</tr>
<tr>
<td>External market influences</td>
<td>4.7%</td>
</tr>
</tbody>
</table>

As an example of carrier selection determinants among motor carriers, Bardi et al. (1989) extracts four factors from 18 determinants by analyzing a survey of 296 U.S. companies. Their results are as follow:

<table>
<thead>
<tr>
<th>Factor 1: rate-related</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Door-to-door transportation rates or costs</td>
<td>(1.41)</td>
</tr>
<tr>
<td>Willingness of carrier to negotiate rate changes</td>
<td>(2.16)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Factor 2: customer-service</th>
<th></th>
</tr>
</thead>
</table>

37
<table>
<thead>
<tr>
<th>Factor 3: claim handling and follow-up</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Claim processing</td>
<td>(2.34)</td>
</tr>
<tr>
<td>Freight loss and damage</td>
<td>(2.04)</td>
</tr>
<tr>
<td>Shipment tracing</td>
<td>(2.12)</td>
</tr>
<tr>
<td>Pick-up and delivery service</td>
<td>(1.90)</td>
</tr>
<tr>
<td>Shipment expediting</td>
<td>(2.05)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Factor 4: Special equipment availability and service flexibility</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Equipment availability</td>
<td>(1.88)</td>
</tr>
<tr>
<td>Special equipment</td>
<td>(3.48)</td>
</tr>
<tr>
<td>Quality of operating personnel</td>
<td>(2.64)</td>
</tr>
<tr>
<td>Line-haul service</td>
<td>(2.39)</td>
</tr>
<tr>
<td>Scheduling flexibility</td>
<td>(2.27)</td>
</tr>
</tbody>
</table>

The number inside the left parenthesis is the mean importance rating on a scale where 1 represents the highest importance and 5 represents the lowest importance. The number inside the right parenthesis indicates the change in emphasis due to deregulation on a rating scale of very high change (1) to very low change (5). The results seem to indicate that rate and customer service-related attributes are important and that they become more important after deregulation.

### 2.3.2. Inbound Movements

Inbound materials are very interesting for railroad marketing departments, since many inbound movements utilize railroads. Transportation expenditures for inbound materials are substantial. A survey of 678 U.S. companies covering 29 industry groups finds that the average annual total expenditure on inbound transportation was $13.4 million in 1990, ranging from a minimum of $60,000 to a maximum of $501 million. The majority of them (74%) respond that transportation expenditures are between 1 and 10 percent of the total dollars spent by the purchasing department (Gentry 1991).

An important characteristic of inbound movements is that multiple parties, e.g. distribution, purchasing and manufacturing managers, are involved in selecting carriers.
Their involvement includes choosing the mode or carrier, determining the transportation price, and rating carriers' performance levels. The survey reports that the responsibility for inbound transportation decisions lie

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>between purchasing and traffic/distribution</td>
<td>(34 %)</td>
</tr>
<tr>
<td>in the buyer in purchasing department</td>
<td>(27 %)</td>
</tr>
<tr>
<td>outside of purchasing department</td>
<td>(15 %)</td>
</tr>
<tr>
<td>in someone (other than buyer) in purchasing</td>
<td>(13 %)</td>
</tr>
<tr>
<td>between purchasing and sale/marketing</td>
<td>( 7 %)</td>
</tr>
<tr>
<td>in a transportation committee that include purchasing</td>
<td>( 4 %)</td>
</tr>
</tbody>
</table>

High involvement of the purchasing department appears notable for inbound transportation. Such high involvement may occur for various reasons. Companies may find it beneficial to integrate transportation and purchasing expertise. Increasing trends toward downsizing companies and outsourcing transportation services may be another reason. On the other hand, the complexity of freight rate/service options due to the deregulation of transportation services may penalize high involvement by the purchasing department if the purchasing department is not well educated about the available options. For example, purchasing managers tend to emphasize service factors highly when they evaluate modal and carrier selections. Purchasing managers reported that

1. The most influential factor in determining the choice of mode is

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>the required delivery date</td>
<td>(37 %)</td>
</tr>
<tr>
<td>the cost of transportation</td>
<td>(17 %)</td>
</tr>
<tr>
<td>reliability and service quality</td>
<td>(14 %)</td>
</tr>
<tr>
<td>shipment size</td>
<td>(11 %)</td>
</tr>
<tr>
<td>transit time</td>
<td>(10 %)</td>
</tr>
<tr>
<td>type of item being shipped</td>
<td>( 7 %)</td>
</tr>
<tr>
<td>possibility of damage</td>
<td>( 2 %)</td>
</tr>
<tr>
<td>availability of service</td>
<td>( 2 %)</td>
</tr>
</tbody>
</table>

2. The most influential factor in determining the choice of carrier is

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>on-time deliveries</td>
<td>(25 %)</td>
</tr>
<tr>
<td>rates</td>
<td>(18 %)</td>
</tr>
<tr>
<td>geographical coverage</td>
<td>(10 %)</td>
</tr>
<tr>
<td>intransit time</td>
<td>(10 %)</td>
</tr>
<tr>
<td>------------------------</td>
<td>--------</td>
</tr>
<tr>
<td>shipment tracing</td>
<td>( 6 %)</td>
</tr>
<tr>
<td>care in handling</td>
<td>( 6 %)</td>
</tr>
<tr>
<td>financial condition of carriers</td>
<td>( 5 %)</td>
</tr>
<tr>
<td>door-to-door deliveries</td>
<td>( 4 %)</td>
</tr>
<tr>
<td>through vehicle routing</td>
<td>( 3 %)</td>
</tr>
<tr>
<td>type of equipment</td>
<td>( 3 %)</td>
</tr>
<tr>
<td>convenient schedules</td>
<td>( 3 %)</td>
</tr>
<tr>
<td>claim handling</td>
<td>( 2 %)</td>
</tr>
<tr>
<td>insurance coverage</td>
<td>( 1 %)</td>
</tr>
<tr>
<td>EDI capabilities</td>
<td>( 1 %)</td>
</tr>
<tr>
<td>shipment consolidation</td>
<td>( 1 %)</td>
</tr>
<tr>
<td>consolidation/breaking capability</td>
<td>( 1 %)</td>
</tr>
</tbody>
</table>

In order to satisfy a purchasing department which has great influence in the choice of a mode and a carrier for inbound transportation, railroad companies should emphasize quality service over low rates. Railroads should also note that the involvement of the purchasing department does not cease after a carrier is chosen. The department also evaluates carriers' performances regularly and influences day-to-day dispatching as well.

The spread of just-in-time (JIT) practices in manufacturing increases the importance of service quality. A survey of 77 trucking companies reports that service quality becomes more important in selecting a carrier following the implementation of JIT manufacturing programs. More specifically, carriers respond that they believe that shippers emphasize the following attributes more after shippers implemented JIT than before they did (Millen and Lieb 1990):

<table>
<thead>
<tr>
<th>On-time performance</th>
<th>(93 %)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Responsiveness</td>
<td>(86 %)</td>
</tr>
<tr>
<td>Shipment tracing capability</td>
<td>(83 %)</td>
</tr>
<tr>
<td>Route network</td>
<td>(62 %)</td>
</tr>
<tr>
<td>Price</td>
<td>(50 %)</td>
</tr>
<tr>
<td>Terminal proximity</td>
<td>(43 %)</td>
</tr>
<tr>
<td>Special equipment</td>
<td>(40 %)</td>
</tr>
<tr>
<td>Intermodal offering</td>
<td>(12 %)</td>
</tr>
</tbody>
</table>
2.3.3. Intra-company Movements

Companies typically move finished goods from plants to distribution. These intra-company movements comprise a large portion of total transportation movements. Since their shipments are usually large and over a long distance, these shipments tend to utilize full truck-load, intermodal or rail. Unfortunately, very little is known about the characteristics of these movements. Mercer conducted a survey which compared the performance of intermodal and trucking services (Mercer 1994). The survey is a part of their 5-year project of constructing the Intermodal Index. Since they selected shippers who ship 100 or more full trailer-loads or container-loads per year and at least 10% of their loads 500 miles or more, majority of the surveyed movements will be intra-company shipments. In order to understand how differently traffic managers and plant managers (or distribution center) evaluate service performance of each freight mode, they surveyed traffic managers and plant managers separately on different forms.

Their results are in Table 2.1. Traffic managers tend to emphasize freight rate more than delivery time reliability, while purchasing managers emphasize delivery time reliability and shipment tracing more than freight rate for inbound shipments. Managers who receive shipments (e.g. outside managers in purchasing, production and marketing departments) tend to emphasize delivery time reliability and shipment tracing. Top managers may also have different view on what to emphasize: cost reduction, customer satisfaction or time-based competition. Also, traffic managers seem to evaluate carrier performance more positively than other managers. Although there was no examination into why this is the case, we presume that traffic managers interact directly with carriers, understand the constraints and situations that carriers face for day-to-day operations, and thus, become more sympathetic to carriers than non-traffic managers do.
<table>
<thead>
<tr>
<th>Traffic/Transport Manager Survey Criteria</th>
<th>SI</th>
<th>TP</th>
<th>IP</th>
<th>Plant / D.C. Manager Survey Criteria</th>
<th>TP</th>
<th>IP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quality of delivery</td>
<td>4.6</td>
<td>4.2</td>
<td>3.7</td>
<td>On-time delivery</td>
<td>3.7</td>
<td>3.4</td>
</tr>
<tr>
<td>Service reliability</td>
<td>4.6</td>
<td>4.1</td>
<td>3.7</td>
<td>Equip.availability</td>
<td>3.8</td>
<td>3.5</td>
</tr>
<tr>
<td>Equip. availability</td>
<td>4.5</td>
<td>4.0</td>
<td>3.6</td>
<td>Equip. availability</td>
<td>3.8</td>
<td>3.5</td>
</tr>
<tr>
<td>Low risk of service failure</td>
<td>4.5</td>
<td>4.1</td>
<td>3.6</td>
<td>Equip. size</td>
<td>4.0</td>
<td>3.8</td>
</tr>
<tr>
<td>Door-to-door transit time</td>
<td>4.4</td>
<td>4.2</td>
<td>3.6</td>
<td>Accuracy of paperwork</td>
<td>3.9</td>
<td>3.6</td>
</tr>
<tr>
<td>Low likelihood of damage</td>
<td>4.4</td>
<td>4.1</td>
<td>3.7</td>
<td>Overall ease of doing business</td>
<td>3.9</td>
<td>3.5</td>
</tr>
<tr>
<td>Quality of pickup</td>
<td>4.4</td>
<td>4.2</td>
<td>3.7</td>
<td>Overall quality of customer service</td>
<td>3.7</td>
<td>3.5</td>
</tr>
<tr>
<td>Overall ease of doing business</td>
<td>4.3</td>
<td>4.1</td>
<td>3.7</td>
<td>Freight bill accuracy</td>
<td>3.8</td>
<td>3.7</td>
</tr>
<tr>
<td>Price</td>
<td>4.3</td>
<td>3.9</td>
<td>3.9</td>
<td>Freight bill accuracy</td>
<td>3.8</td>
<td>3.7</td>
</tr>
<tr>
<td>Freight bill accuracy</td>
<td>4.3</td>
<td>4.2</td>
<td>4.0</td>
<td>Equip. quality</td>
<td>3.8</td>
<td>3.6</td>
</tr>
<tr>
<td>Equipment quality</td>
<td>4.2</td>
<td>4.1</td>
<td>3.8</td>
<td>Equip. appearance</td>
<td>3.6</td>
<td>3.4</td>
</tr>
<tr>
<td>Intransit shipment information</td>
<td>4.1</td>
<td>3.8</td>
<td>3.7</td>
<td>Interior cleanliness</td>
<td>3.6</td>
<td>3.5</td>
</tr>
<tr>
<td>Equipment dimension/size</td>
<td>3.7</td>
<td>4.1</td>
<td>3.7</td>
<td>Door opening</td>
<td>4.1</td>
<td>3.8</td>
</tr>
<tr>
<td>Quality of salespeople</td>
<td>3.7</td>
<td>3.8</td>
<td>3.7</td>
<td>Quality of salespeople</td>
<td>3.8</td>
<td>3.7</td>
</tr>
<tr>
<td>EDI capabilities</td>
<td>3.2</td>
<td>3.3</td>
<td>3.5</td>
<td>Quality of salespeople</td>
<td>3.8</td>
<td>3.7</td>
</tr>
<tr>
<td>Overall performance</td>
<td>4.1</td>
<td>3.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1. SI = Stated importance (1=not important, 5=very important)
2. TP= Truck performance (1=poor, 5=excellent)
3. IP = Intermodal performance (1=poor, 5=excellent)
2.3.4. Major Findings

Overall, the stated importance research finds that freight rate, transit time, and delivery time reliability are the most important attributes for all of outbound, inbound and intra-company shipments, whereas service flexibility and customer satisfaction are less important. In addition to the papers cited above, there has been extensive research done by Kearney, Ohio State University, Pennsylvania State University and Michigan State University. Using average statistics of the importance that shippers report about service attributes, they report similar conclusions as those which we summarized in the previous section.

In addition, research indicates that service quality became more important after deregulation. Commonly we believed that freight rate would influence modal selection in a greater degree after regulation, since carriers who competed on service differentiation during regulation can compete on price when deregulation allowed them to write secret contracts with their customers. Many economists (e.g. Meyer, Peck, Stenason and Zwick 1959 and Friedlaender 1969) argue that service differentiation might be inefficient and that shippers should be given choices among different price levels. Accordingly, McGinnis (1989) reports that 12% of the respondents exhibited an orientation toward price minimization, and Quinn (1987) finds that 26% of his respondents placed more emphasis on price than service. Many other empirical results, however, suggest that service attributes continue to be important. For instance, the continued growth of trucking after deregulation suggests that shippers are willing to pay a premium rate for a high-quality service. There can be many explanations for this progress.

The first view is that deregulation has not changed carriers' practices and/or shippers' priorities much. Those who hold this view believe that price flexibility before deregulation was not as rigid as we think it now to be and that shippers could get special rates depending on origin-destination, commodity, packaging, volume, and vehicle routing. This view, however, does not explain why price fell so much after deregulation.
Holders of this view also believe that shippers require a minimum level of service and are unwilling to trade off critical service issues with costs. Since shippers were more interested in service quality than in freight rate always, they believe that the alleged economic loss and traffic misallocation due to lack of price competition before deregulation was small. This view, while plausible, does not explain why service sensitivity has increased after deregulation.

The second view is that once price fell dramatically (in real terms) after deregulation, service became relatively more important than before. For instance, inventory holding costs become relatively important as transportation costs decrease. Moreover, the trend toward just-in-time delivery requires a high service quality from transportation carriers. Shippers are increasingly sending shipments in smaller sizes and in greater frequency than before. They, at the same time, require shipments to be delivered within shorter time windows, i.e. not just every Tuesday but between 9:30 to 10:00 a.m. of every Tuesday. Morton (1972) reported that shippers are willing to pay a premium rate of about 15 to 20 % more for the service advantages of shipping by truck. He did not, however, specify whether such premium is for convenience of pick-up and delivery, for faster transit time, for reliable delivery schedule, or for a pure truck-intrinsic preference. The service premium may have increased after deregulation.

The third view is that deregulation increased service differentiation. By having flexibility in negotiating price/service packages, both shippers and carriers no longer need to be limited by tariffs and rates. If carriers can find niche markets that need services differentiated with shipper- or product-specific considerations, carriers can command a local monopoly within the niche markets. The relative importance of freight rate or time may become of less concern. Carriers can also serve market segments that best fit its company resources. As many differentiated services tailored to specific market segments became available, service sensitivity may have increased after deregulation.
Finally, we can categorize attributes in terms of the importance and the railroad’s relative advantage to truck. For example, the Intermodal Index indicates the table below:

<table>
<thead>
<tr>
<th>Adv. of rail Importance</th>
<th>Low</th>
<th>High</th>
</tr>
</thead>
</table>
| High                    | Quality of delivery / pick-up  
                         Service reliability  
                         Equipment availability  
                         Low risk of transit time  
                         Door-to-door transit time  
                         Low likelihood of damage  
                         Overall ease of doing business | Quality of customer service  
                         Price  
                         Freight bill accuracy |
| Low                     | Payment of damage claims  
                         Equipment dimension / size | Equipment quality  
                         Real-time shipment tracing  
                         Quality of salespeople  
                         EDI capability |

The table suggests that although railroads have been investing heavily in establishing EDI links, real-time shipment tracing, single source of contact, or equipment modernization, shippers seem to consider them either unimportant or difficult to understand. Railroads may be making wrong investments, by thinking from an internal or engineering perspective rather than from a customer’s perspective. Railroads' recent investments on AC/DC locomotives that provide higher horse power may be criticized in this aspect. If what shippers really need are mini-trains that move point-to-point frequently with a small number of consolidation, the purchase of high-power locomotives at high prices may not be justified except in terms of cost reduction of bulk shipments.

2.3.5. Weakness of The Stated Importance Approach

Relative to the consensus about which service attributes are important, there is a big disagreement about the trade-offs among service attributes. Such disagreement may be the direct result of noise contained in self-stated importance data. The self-stated importance approach, in general, can be classified into three different types depending on
how to solicit importance: importance rating, constant sum, and importance ranking. All three approaches have problems as discussed below:

1) The importance rating approach directly asks respondents to rate the importance of an attribute on an interval scale. This approach assumes that importance can be ratio-scaled in a Euclidean metric. The use of this approach has been popular in marketing research, since it is easy to understand and easy to administer. Much research has been conducted to understand how many intervals respondents can differentiate, whether each interval should have a clear and precise wording or not, whether an even-numbered or an odd-numbered scale is desirable, and whether the scale should be balanced or not.

This approach is the least reliable and is useful only for obtaining insights at the exploratory stage. One problem is that there is no way of knowing whether responses are equally-spaced. For example, is a rating of 7 more important than a rating of 6 as much as 6 is than 5? Secondly, there is a lot of heterogeneity in the responses. The author found, from many survey results, that a respondent who marks high (or low) importance for the first attribute tend to mark high (or low) importance for all attributes. Third, some people who have no opinion often make the error of marking 4 (i.e. neutral or so-so importance).

2) The constant-sum approach asks respondents to allocate a fixed sum (e.g. 100) among attributes. By forcing trade-offs among attributes' importance weights, this approach tries to improve reliability of importance estimates. But, respondents may become easily distracted when the number of attributes is large, e.g. greater than 7 or so.

3) The importance ranking approach asks respondents to rank attributes by order of importance. Researchers try to discover the importance of attributes by counting the number of respondents who gave the highest rank to the attribute or by calculating a weighted sum of the ranks given to the attributes. The problem is that importance ratings cannot be inferred from rankings. That is, the second-ranked attribute may be as important
as the first-ranked attribute or can be unimportant relative to the first-ranked attribute by a wide margin.

In general, the self-stated importance approach suffers from a lack of external validity both in terms of data collection and data analysis as shown below:

1) Noise in Data Collection

In constructing the 1994 Intermodal Index, Mercer (1994) reported that shippers rated services to be important criteria in selecting modes, and yet that most shippers answered that they did not have formal measurement systems for carrier evaluation. Lack of a measurement system implies that shippers responses will be based on perceptions which typically have a lot of measurement errors. In particular, if shippers rate the importance of attributes without a realistic scenario, it is likely that the respondents may ignore important choice contexts or may not seriously compute tradeoffs between price and service.

2) Justification bias:

Respondents may try to give rational reasons for their choices. For example, traffic managers might use a trucking carrier because they are accustomed to its services, because the carrier understands their needs, or because they receive special treatment by having a close contact with the carrier. Yet, they may say that they chose the trucking carrier because its reliability is very important to them. This tendency for respondents to over-state generic or rational dimensions and to under-state specific or perceptual dimensions is called justification bias.

3) Policy bias:

Respondents may try to influence carriers’ policy by stating that service quality attributes are very important, even though they are not willing to pay much in actual choices for improvements in service quality. This tendency can be large when shippers are asked to rate the importance of service attributes without making any trade-offs.
4) Individual heterogeneity: A respondent who gives a high (or low) rating on the first attribute tends to give high (or low) ratings to all remaining attributes. Moreover, all respondents tend to rate high importances for most of the attributes. For instance, the 1994 Intermodal Index reports that shippers rated 14 out of 17 attributes as having similar average importance (e.g. between 4.6 and 4.1 in a scale of 1 to 5). Looking at the averages of stated importance without considering their standard errors can be misleading.

5) Phrasing effects: Because of the unreliability of importance measures, some marketing researchers argued that a desirability scale rather than an importance scale should be used. Their argument is based on a belief that people will not overstate the desirability of generic and rational attributes (Clancy and Shulman 1991). Some researchers favored a liking scale. In particular, Kano (1984) proposed to ask respondents twice: how much they would like if an attribute is provided, and how much they would dislike if the attribute is not provided. Based on the two ratings in two scales, he classifies customer requirements into one of 5 categories: must-be, attractive, one-dimensional, indifferent, and reverse attributes. We believe that neither approach is useful. The “importance” of a service attribute is not a well-defined concept. It is more meaningful to compare attributes in terms of demand elasticity than in terms of stated importance.

6) Comparability The studies we reviewed consist of different survey methods, including structured and open-ended questionnaires as well as mail and personal interviews, and also with different survey objectives such as understanding intermodal usage patterns, deregulation effects, or benefits of electronic data interchange. The studies also consisted of different survey respondents in terms of industry and geography. The average importance will differ depending on the characteristics of shipper segments. Small shippers with high-valued products, or those with little experience using rail service would be very different from large shippers.
Chapter 3.
The Conceptual Framework

3.1. Introduction

Most behavioral theories we reviewed in Chapter 2 do not provide reliable specification of the objective function of shippers when they manage their logistics systems and select carriers. In particular, they disregard a major perspective, i.e. the role of a logistics system in implementing management strategy. Without incorporating the strategic perspective that logistics managers emphasize highly, a freight choice model can omit many important variables in designing logistics systems. Therefore, in section 3.2, we review managerial literature and develop a theory of the three-stage development of logistics systems. The model suggests that shippers' major objectives in making modal selections are to minimize total logistics costs, maximize customer service and sustain comparative advantage. Section 3.3 tests the external validity of the three-stage development model by conducting case studies of actual modal selection.

Once shippers' objectives of managing their logistics systems are defined, we can specify shippers' modal selection process in quantitative terms and estimate the mathematical relationships by employing the random utility model. Section 3.4 introduces the general concept of the random utility model and apply it to modal selection. In section 3.5, we categorize freight choice models into three generations based on the three-stage development model.

3.2. Three Stage Development of The Logistics System

We will first review major trends in shippers' and carriers' business environment. We then discuss how these trends affect the shippers' logistics systems and their transportation implications. The following chart summarizes the structure of this section:
In the below analysis, our primary focus is on the general merchandise and intermodal markets that include automobile, appliances, tires, etc. These markets have been the main forces behind the growth of intercity rail freight movements. For example, intermodal services such as trailer-on-flat-car, container-on-flat-car, and unit trains for double-stacked containers have shown double-digit growth, whereas movements of bulk commodities such as coal or grain have shown minimal growth. In addition, the general merchandise markets are very different from bulk commodity markets. The decision criteria for carrier selection by the general merchandise shippers are much more complex than the usual cost minimization hypothesis used in bulk commodities. Unlike bulk
commodities, general merchandises have a high value per volume and thus place low importance on transportation costs relative to value. Savings from the reduction in inventories at both origin and destination may more than offset the high freight charges required for frequent and fast delivery of smaller-sized shipments. Moreover, competition among trucking, rail and intermodal companies is highly visible. Since freight choice model considers competition explicitly, it will be more useful in general merchandise markets than in bulk markets.

3.2.1. Changes in Shippers' Business Environment

Demand for transportation is derived, i.e. shippers do not purchase transportation services in order to satisfy their own personal needs or desires but rather in order to deliver produced goods or provide services to their customers. For example, auto makers use rail services in order to deliver finished automobiles to their customers. If consumer demand for new automobiles increases, demand for railroads by auto manufacturers will also increase. That is, demand for rail services is derived from consumer demand for new automobiles. Therefore, we have to understand changes in the consumer market in order to understand changes in shippers' transportation needs. In particular, the following forces have brought about changes in the shippers' logistics systems, which in turn have resulted in changes in the transportation systems.

*Increased product variety* Intense competition has led to a proliferation of products as companies try to satisfy consumer demands by offering variety. For example, a typical supermarket used to carry a few thousand items in the 1960s, but it carried more than 16,000 items in 1990 (Progressive Grocer (April 1991). The largest supermarket can stock more than 60,000 items. Likewise, automobiles are offered in multiple colors, styles, and sizes.

*Diverse distribution channels* In addition, manufacturers now find that they have to supply diverse distribution channels such as direct marketers, catalog stores, discount
stores, warehouse clubs, department stores and specialty stores. It is almost impossible and also uneconomical to have safety stocks for all product variations in all channels.

**Shortened product life cycle** Moreover, retailers and manufacturers find that consumer tastes can change fast, and thus a product's life cycle can be very short. Consider the automobile market. While luxury sedans such as BMW and Lexus were popular in the late 1980s, compact cars like Saturn were in demand in the early 1990s and small pick-up trucks and mini-vans sold well in 1994. In this year, zeep and recreation vehicles seem to be gaining popularity. Once consumer tastes for a product change against its favor, the product's price should be cut significantly. In many cases, it is sold below the manufacturing cost in order to clear inventories. The following figure depicts the dynamic and complex nature of the transportation environment of the automobile market.

![Graph showing Diversity of Transportation Needs](image)

**Diversity of Transportation Needs**

**Proliferation of special promotions** In addition, manufacturers have difficulty in predicting accurately where and how much volume of which model will be ordered, due to numerous special promotions and/or advertising campaigns. Population is no longer concentrated in a small number of city clusters. Manufacturers have to be ready to ship orders to any place around the nation within a reasonable time interval. Therefore, many manufacturers appear willing to incur high transportation costs if they can ship small quantities to random places with speed, reliability and flexibility.
**High customer expectations**  As consumers experience high service quality, they increasingly demand it. Since expectation levels tend to increase and since competition in the market is intense with increasing international trade, consumers will continue to demand highly responsive customer service.

**Globalization**  Companies are increasingly becoming globalized. Already, many have expanded into international markets and many manufacture overseas. For instance, products are imported from the Far East to California, moved to New York through Chicago, and may be re-exported to Europe. NAFTA will also increase movements among Canada, the U.S. and Mexico. The lengthening of the supply chain increases logistics costs and risk of disruption in the supply chain.

**Information Technology**  For shippers, developments in information technology have brought many changes to their logistics systems. In particular, information technology enables firms to respond quickly to customers' demands by installing an automated order processing system. Firms such as American Hospital Supply Corporation achieved rapid growth by automating order processing which reduced their order cycle to less than 2 days, whereas the average order cycle for all manufacturing types was about 10 days. A large portion of the saving comes from increased efficiency in processing documents through information technology (Stock and Lambert 1987).

First, stock level monitoring became easy with bar-coding and with automatic calculation of inventory levels as store sales are scanned in the registers. Now, most stores aim for continuous replenishment where manufacturers automatically replenish store shelves as they are emptied. This trend will be certain to promote long-term relationships among manufacturers, retailers and carriers. In particular, frequent deliveries in small volumes have become the norm. Discounts with quantity purchase happens less frequently. Moreover, product proliferation suggests that transportation services should be differentiated, since some products require fast reliable service, whereas some do not.
Small levels of inventory at the local level suggest that transportation services should be flexible so that retailers can change the schedule and the destination of shipments as soon as they realize changes in local demand.

Second, ordering costs have been greatly reduced, and the ordering process has been decentralized. The job of ordering used to be one of the most difficult tasks in the store (Walsh 1993). Usually, one vendor or warehouse (e.g. Proctor and Gamble) supplies items to many aisles, while an aisle managed by one person contains items from several vendors. As the number of products in the mix increases, it becomes more difficult to know what model to order and how much. Moreover, department heads want to be able to take advantage of price specials by suppliers, while matching local demand. They want to take advantage of knowledge that stockers have with regard to which products sell and which do not, how often they refill the shelves, and what is piled up in the back room. Most stockers, however, do not know how to analyze sales trends and how to process the paperwork. A computerized ordering system facilitates the stockers' tasks by reducing the paperwork requirements and by automating the sales trend analysis.

As above-mentioned changes make inventory costs un-manageable, the thinking that "inventory is evil" is spreading fast among shippers. Currently, shippers are putting more emphasis on improving delivery service than on increasing safety stock. They are pushing towards just-in-time and lean distribution which requires a fast, reliable, and flexible delivery system. As a result, inventory levels relative to annual sales continue to decrease with the establishment of advanced inventory management systems. The following cases may illustrate the new thinking of logistics managers:

| Total demand for dishwashers or ranges is fairly predictable, but tastes in colors and features change like quicksilver. By 1990, GE was saddled with $160 million of obsolete inventory. Rows of big, bulky refrigerators, then out of vogue, stood in limbo. (Tully 1994). [While popular models are always out of stock, non-popular models keep being accumulated. The problem is that GE does not know in advance which model will be popular and that non-popular models will seldom become popular in the near future.] |
These parts spent only a few hours - or minutes - on the machines but sat in mountainous inventories for weeks, ... Warehouses covered enough flour space for three football fields. When an order came in, Trane needed an average of 15 days to find the components and assemble the final unit. Incredibly, despite the vast inventories, Trane often didn't have or couldn't find the parts it needed. Those pulled from stocks often turned out to be damaged. Many items rusted on the 'lac' top or took a bang from a forklift. (Tully 1994).

A leading computer company decided to close assembly lines in Korea and to open one in California. Computer lines are design-intensive and rapidly changing. While the Far East plant provided low labor costs, the long distance between product design and manufacturing led to significant problems in product quality and customer service. The new plant in California expects to track changing consumer needs fast by moving to near the market and to improve market response by reducing transportation, warehousing and duplication in inventory stocking costs. Success of local assembly companies such as Gateway prompted the decision.¹

3.2.2. Changes in Carriers' Business Environment

As much as the shippers' business environment has changed, carriers' business environment has changed as well. Deregulation of the industry lowered the barriers to entry, provided more flexibility in pricing, and relaxed many regulatory rules excluding safety-related rules. Since deregulation, shippers are shifting toward a contractual relationship, i.e. core carrier, rather than toward a transactional relationship. They are also more willing to outsource transportation services. Accordingly, carriers try to provide total solutions to shippers' transportation needs through the purchase of new technology and the establishment of a new organizational structure.

Carriers now can offer differentiated services at different rates. With regard to motor carriers, deregulation allowed independent operators to enter the industry. Since independent truckers can operate in the industry with low fixed costs, the motor carrier market has become highly competitive. In railroads, deregulation prompted mergers and acquisitions among railroads. In turn, there have been consolidations of main networks and abandonments of unprofitable branch lines. Consolidation of market power, however, ¹ Interview with a manager of Hyundai Electronics Co.
has not led to high profitability. Intermodal competition both for rate and service became fierce, and railroads had to make substantial investments on improving their services.

In addition, the fierce competition since deregulation forced carriers to adopt new technologies such as electronic data interchange, bar-coding of trailers, global positioning system, shipment tracing system, and satellite communication. For instance, full-load trucking companies began to attach sensors and communication devices to all their trucks. With satellite communication, centralized dispatchers order trucks to change routing or to refill fuel at certain stations. Moreover, a sensor attached to the engine continuously monitors engine vibrations and reports its status to central maintenance, preventing unexpected delay due to an engine problem. Such central control is known to reduce empty trailer mileage, improve operating efficiency and save fuel costs. In railroads, new equipments such as heavy-axel load cars, aluminum cars, double stack containers, and automotive trailers have improved service. A new terminal design and an electronic system control are under study and will foster another boost in productivity.

Deregulation and new technology have greatly changed the management structure of carriers as well. New management is much more likely to be finance-oriented and strategic-minded than previous managements who were more concerned about regulatory rules and operation planning (McCulloch 1994). New management is trained more rigorously in finance and is more willing to change its network design if it is necessary for maximizing shareholders' value. Customer satisfaction, strategic partnership and comparative advantage are areas that the new management puts emphasis on.

3.2.3. Changes in Shippers' Logistics System

As both the business environment and technology change, logistics systems change as well. We may classify the development of logistics systems by the following three stages.
<table>
<thead>
<tr>
<th>Issues</th>
<th>Stage I</th>
<th>Stage II</th>
<th>Stage III</th>
</tr>
</thead>
<tbody>
<tr>
<td>Role of Logistics System</td>
<td>Order Processing</td>
<td>Value Creation</td>
<td>Strategic Differentiation</td>
</tr>
<tr>
<td>Objective of Logistics System</td>
<td>1. Min transport cost 2. Min total logistics cost</td>
<td>Max customer service</td>
<td>Max comparative advantage</td>
</tr>
<tr>
<td>Key Tools</td>
<td>Economic order quantity</td>
<td>Quick response ECR</td>
<td>Business Process Redesign</td>
</tr>
<tr>
<td>Business Network Redesign</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Demand forecasting</strong></td>
<td>Long-term forecast</td>
<td>Intermediate-term forecast</td>
<td>Short-term forecast</td>
</tr>
<tr>
<td>Aggregate forecast</td>
<td>Segment forecast</td>
<td>Individualized forecast</td>
<td></td>
</tr>
<tr>
<td><strong>Manufacturing</strong></td>
<td>Made-to-stock volume production</td>
<td>Made-to-order flexible manufacturing</td>
<td></td>
</tr>
</tbody>
</table>

At the first stage, firms produce standard products in volume in order to achieve an economy of scale and to reduce facility set-up costs. Production schedules are determined based on the long-term forecast of demand for standard products. With stockfiles of inventory, the logistics departments perform the simple function of order processing. The logistics departments are treated as cost centers, and their prime objective is to minimize logistics-related costs. In order to reduce logistics costs, firms study extensively the economic order quantity model that determines optimal shipment size. Differences in logistics costs can indeed bring significant advantages. Some study shows that logistics costs occupy 5 to 45% of total cost (Ballou 1992).

Recently, firms have lowered logistics costs by reducing the working capital committed to inventories. G.E. estimates that one additional inventory turnover generates an extra $1 billion in cash, plus big savings in storage and labor. G.E. eventually wants to have 20 turns a year from the current 7.5 turns a year (Tully 1994). Connection of the production schedule to order status can also reduce product obsolescence and any need for clearance sale. Individualized delivery tracking also reduces movements of finished products among warehouses. As a result of such practices, just-in-time (JIT) management
has become standard for both in-bound materials and outbound goods. A survey of 400 companies by Purchasing magazine reveals that nearly 65% are either using JIT or planning to adopt it (Bertrand 1986). Another study looks at 131 companies and finds that 82% are using the concept. Among firms with 500 or more employees, 90% are using JIT, and the other 10% are considering using it (Celley, Clegg, Smith and Vonderembse 1986). Although JIT began as an effort to decrease the inventory of raw materials and parts waiting for production, it provided a natural transition of logistics systems towards responsive customer service.

At the second stage, firms are no longer bound by cost reduction. They realize the importance of customer service on revenue generation, and the logistics departments begin to perform the role of creating value by providing improved customer service. Customers' perceptions about a company critically depend on the context and the procedure of the company's service delivery system. Since logistics is the firms' primary contact point with customers, the logistics system can create service differentiation. Indeed, while companies can satisfy customers through quick response in product design, production schedule, and product delivery, flexibility in product delivery is the only short-term option to cope with changing consumer demands. New managerial buzzwords such as Quick Response (in the general merchandise industry) and Efficient Consumer Response (in the grocery industry) represent new interests in improving logistics systems in order to provide responsive services. At this stage, firms also understand the concept of segmented markets and a need to offer niche products. Production schedules are made based on the forecasts about segmented markets' demands.

At the third stage, firms begin to realize the importance of products and services tailored specifically to individual needs. Firms provide individually-tailored service which requires a logistics system more elaborate than the one dealing with segment-tailored services. Longterm forecasts of standard products are no longer useful. Short-term forecasts for a variety of products at the customer level becomes necessary. Developments in information technology and database support such micro-marketing and flexible
manufacturing. Many logistics departments are now undergoing major restructuring in order to make logistics a major strategic tool to compete against other companies.

In particular, more and more shippers realize that the traditional function-oriented hierarchical systems lack flexibility to cope with the rapidly changing business environment and with readily accessible market information. Bureaucracy -- the solution to industrial revolution -- now appears to be a problem to information revolution. Instead, a new system proposes a horizontal network integrated around a few major business processes. Logistics system, then, emerges as a key player in integrating the functional organizations, which we may call business process redesign.

*Business process redesign* aims to eliminate the traditional function-oriented systems and to reorganize organizations around major business processes. We may classify the major business processes as customer needs fulfillment, order-cycle management and supply-chain management (refer to figure 2.1). Customer needs fulfillment involves figuring out customers' unfulfilled needs, developing products and services that satisfy the needs, and communicating new products attributes to potential customers. Order cycle management, on the other hand, involves order generation, order receipt and entry, order selection and prioritization, scheduling, delivery, billing, returns and claim handling, and after-service (Shapiro, Rangan, and Sviokla 1992). Meanwhile, supply chain management involves production process design, capacity planning, production scheduling, material requirement planning, labor planning, supply chain integration, and facility operation.
Demand forecasts and production schedules are two major outputs that companies utilize in order to control these processes. Traditionally, management practice which emphasizes customer needs fulfillment is called a pull strategy, while the one that emphasizes supply chain management is called a push strategy. As analysis of market information becomes fast and easy, companies increasingly emphasize moving resources in accordance with pulled demands rather than pushing inventories. At this stage, logistics departments are fully integrated with other departments and play an important role in managing the order cycle and supply chain. For example, let us look at the continuous replenishment effort that Proctor and Gamble Co. and one of its major customers made jointly.

Previously, the customer used to employ stockers in order to determine shipment size and timing. With real-time market information, a computerized system might make an automatic shipment order as P&G products approaches a pre-set safety stock level. The logistics division of P&G was passively waiting for an order.

A new process will allow P&G to actively monitor inventory level of the customers and ship products as appropriate. Since P&G sells in many different categories, it is very difficult for stockers of the customer to find out what are good shipment size and timing, whereas P&G can coordinate shipments to their customers more effectively.

This new process, in turn, will give higher comparative advantage to P&G over other companies of competing products, since retailers are more interested with merchandising appropriate products than with which brands they will sell as long as P&G provides as good margins as its competitors.
As another example of business process redesign, consider the following made-to-order production system that Saturn plans to build and that Toyota has already been practicing in Japan:

When consumers want to buy a Saturn car, they will sit down at a computer terminal in a dealer's showroom. With a salesperson's help, they will select many options, e.g. tinted windows, carpet color, CD player, tires, number of doors. They can see their selection since the screen can simulate product images with computer-aided design. Once they finish making their decisions, the orders are automatically released to various departments. The computer at the showroom will immediately check the consumers' credit, arrange GMAC financing, and/or obtain insurance from GM's insurance subsidiary.

Computers at the plant will receive the orders and give it a number and an assembly date. Then orders will be released to GM's suppliers. The assembly line is organized in such a way that it can mix different options of features onto a basic frame extremely flexibly. As each car is produced and is loaded onto a truck for shipment, computers at the plant will signal the dealer that the car is on the way, instruct GMAC to start finance charges, and release payments to the suppliers for the parts. (Ballou 1992)

Ballou claims that this order cycle including production takes only a week and saves the customer $ 2,000 in the car's price by reducing working capital and space required for holding inventory. Not only does Saturn save inventory costs but it also provides its customers with the exact models they want. The customer also feels comfortable from dealing with only one salesperson for all other activities such as order status checking, order changes, after-sales claims, etc.

This approach to using customer service as a strategic marketing tool is gaining attention. Shippers, diverting from the traditional product-oriented strategy, are beginning to focus on customers' perceptions of their service quality (Albrecht 1988). In order to provide quality service that customers expect, companies are changing the way they produce and deliver products. In the above case, Saturn eliminates boundaries among different functions (e.g., marketing, distribution, finance, etc.) and requires the salesperson to handle all functions. This system, however, is a nightmare for logistics managers. Salespersons who often know little about logistics constraints handle all transportation
processing through a computer and expect fast and reliable delivery. Carriers are expected to support such individualized shipments without sacrificing system efficiency.

For carriers, business process redesign can be good news since shippers are increasingly willing to outsource their logistics functions or to make strategic partnerships with carriers in order to redesign their business processes. Eventually, firms will focus only on the processes in which they excel over others, i.e., in which company resources command core competencies. Non-core processes will be outsourced either through competitive biddings or through strategic partnerships. Core carrier programs and third-party logistics contracts for transportation and warehousing services provide such an example. The premise is that firms can control the quality of outsourced services thanks to reduction in information sharing costs.

The real comparative advantage, however, cannot be achieved until the whole logistics system is so uniquely set up that competitors can neither perform more effectively nor imitate it. For example, EDI may establish entry barriers initially, since retailers do not want to operate computer systems for all manufacturers and/or since stockers may be accustomed to the system that they learned first. The first mover advantages, however, can be abolished quickly if Congress mandates equal access of inventory information to small vendors. The first mover advantages will be sustained when a company can change its organization sufficiently enough so that other competitors cannot imitate its business process and business networks. For instance, a system that exploits the maximum high fixed costs necessary for developing an integrated EDI system may sustain entry barrier and comparative advantage.

**Business network redesign** attempts to treat customers and carriers as strategic partners so that it can overcome organizational boundaries and create efficiency by globally optimizing the whole chain (Short and Venkatraman 1992). Common mistakes that companies make during this process are the consolidation of distribution channels in order to achieve an economy of scale and the increase of safety stock in order to reduce
out-of-stock situations (Fuller, O'Connor, and Rawlinson 1993). Also, products flow
through consolidated channels at average speed and are charged out at average cost. With
this system, customers who need specialized products quickly are underserved due to slow
delivery, and customers who want commodity-like products at cheap prices are
overcharged for a delivery that is faster than necessary. By providing undifferentiated
services, companies can lose both customer segments.

Companies can completely redesign their business networks by regarding products
as integrated packages of goods and services rather than as stand alone physical goods.
For example, let us consider the case of General Electric.

Suppose that a retailer who sells G.E. products receives an order. Rather than having their
own warehouses for G.E.'s products, G.E.'s retailers now directly check the availability of
required models from G.E.'s distribution centers. If the model is available, G.E. confirms
the order and delivers directly to the purchaser's house. During this process, retailers can
completely eliminate warehousing space for G.E. products. On the other hand, G.E. saves
money by stocking only basic models and by finishing them with diverse options that
customers would like to have ("virtual warehouse").

Moreover, G.E. no longer stocks maintenance parts in many locations around the nation.
By consolidating service centers, G.E. can greatly reduce parts inventories which tend to
be overstocked and discarded long after relevant models are outdated. When G.E. got a
service call from a customer, G.E. ships required parts directly to the service
representative's residence during the same evening or next day's morning so that the
representative can go to the customer's house directly from his house ("24 hour service").
(Treacy and Wiersema 1993)

In this type of realignment, carriers are regarded as third-party logistics partners,
and shippers develop new networks that conduct business more cost-efficiently and with
higher service quality-intensity than before. This type of business network redesign is
impossible without strategic partnerships between shippers and carriers.

3.2.4. Transportation Implication

Changing Objectives In Carrier Selection
As the objective of shippers' logistics systems changes, variables that influence modal and carrier selection also change. The following table summarizes the changes:

<table>
<thead>
<tr>
<th>Stage</th>
<th>I</th>
<th>II</th>
<th>III</th>
</tr>
</thead>
<tbody>
<tr>
<td>Objective of Logistics System</td>
<td>1. Min transport cost 2. Min total logistics cost</td>
<td>Max customer service</td>
<td>Max comparative advantage</td>
</tr>
<tr>
<td>Performance Measures</td>
<td>Logistics costs</td>
<td>Customer Service</td>
<td>Sustainability</td>
</tr>
<tr>
<td></td>
<td>Operation efficiency</td>
<td>Customer Satisfaction</td>
<td>Appropriability</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total Quality</td>
<td></td>
</tr>
<tr>
<td>Structure of Logistics System</td>
<td>Functional</td>
<td>Inter-departmental</td>
<td>Inter-company</td>
</tr>
<tr>
<td></td>
<td>Hierarchical</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

At the first stage of order processing, the major focus was placed on the minimization of the total logistics costs. Initially, focus was on ensuring the availability of transportation nodes (e.g. train schedules, terminals, pick-up and delivery service) and on minimizing only the transportation costs. Accordingly, carriers competed on freight rates and terminal availability. Later, shippers began to try minimizing the total logistics costs which include ordering costs, inventory holding costs and other inventory-related costs in addition to transportation costs. The performance of the logistics system was measured in terms of operation efficiency such as order cycle length, inventory turnover, safety records, ratio of actual output over actual input (e.g. ton-mile / employee), and ratio of actual output over desired input (e.g. ton-mile / budget), etceteras (Caplice 1994). Accordingly, the performance of carriers was periodically monitored in terms of operational measures such as on-time delivery, billing accuracy, and damage percentage. Reliability of shipment delivery received much attention as a key variable that influences the amount of safety stock holding.

At the second stage of value creation, shippers focused on service competition and the objective of the logistics departments became the maximization of customer satisfaction. Initially, the flexibility in changing shipment schedules or shipment sizes
received major attention. The electronic data interchange and a real-time shipment tracing system are efforts to provide flexibility. Later, shippers began to put more focus on external measures such as customer satisfaction and continuous quality improvement. A shipper may periodically survey her customers' satisfaction level with her shipment and use the information to evaluate carriers. Or a shipper may review carriers' performance periodically through a committee whose members include relevant parties such as customers, marketing, purchasing, finance, etc. Results of such periodic surveys and/or reviews can be very important, sometimes more important than operation measures.

At the third stage of strategic differentiation, operation efficiency and customer service become less meaningful. A key focus is given to the sustainability and appropriability of a strategically-designed business process or business network. The newly designed system should provide not only a quantum leap in operation efficiency and customer service, but also should not be imitable by competitors. Sustainability means that competing firms cannot imitate a newly-designed business process and/or network and that strategic partnership between shippers and carriers can be sustained. Appropriability, on the other hand, means that new business process and network can erect entry barrier strong enough to earn extra rents over the normal economic rate of return (or the cost of capital). Specific measures have yet to be developed.

**Centralization**

<table>
<thead>
<tr>
<th>Stage</th>
<th>I</th>
<th>II</th>
<th>III</th>
</tr>
</thead>
<tbody>
<tr>
<td>Core carrier selection</td>
<td>Decentralized</td>
<td>Centralized</td>
<td>Centralized</td>
</tr>
<tr>
<td>Dispatching</td>
<td>Decentralized</td>
<td>Decentralized</td>
<td>Centralized</td>
</tr>
<tr>
<td>Carrier Relationship</td>
<td>Local partnership</td>
<td>Competitive bidding</td>
<td>Strategic partnership</td>
</tr>
</tbody>
</table>

As shippers view the role of logistics system as more and more important, they increasingly centralize in carrier relationship (Caplice 1994). Previously, carrier selection
and dispatching were done at the local level with more focus on local considerations. Current trends are to select core carriers centrally and to allow local traffic managers to choose among primary and secondary carriers. Most shippers already select carriers in their transportation headquarters through competitive bidding in order to ensure minimum freight rate as well as service quality control. The trend is towards further centralization of shipment dispatching. Transportation headquarters can now access information about local inventory status. With decision support systems that figure out the best network flow, central dispatchers can generate transportation schedules that are coordinated with local warehousing schedules. Advanced companies use developments in information technology and mathematical models (e.g. a large-scale optimization for network design, vehicle routing, and continuous move) and try to reduce empty trailer movements or trailer stopping time. In particular, large shippers are pioneering this centralization trend by engaging in strategic partnership programs with carriers (e.g. Du Pont and Roadways).

3.3. Case Study

To test the validity of the three-stage model, we conduct case studies. Case study selects one or two companies and analyzes in depth how they actually evaluate carriers. Since it does not survey a large number of shippers, the case study approach is useful for obtaining a detailed understanding of the internal process rather than for getting a broad general view of the process. Two case studies confirm that shippers are indeed changing their logistics systems as hypothesized by the three-stage development model. Both companies are putting emphasis on customer service and on strategic partnership.

The first study was exploratory. We contacted Gillette, a company that specializes in personal care products and razors, and tried to understand the role of the logistics system and the basic process of carrier selection. The summary of the first case study is provided in Appendix 1 and we learned the following four lessons:
1) The company is reducing the number of core carriers they use and is increasing centralized control over the carrier selection of local traffic managers. This implies that marketing efforts of railroads should be directed more toward transportation headquarters and that strategic performance measurements should be developed.

2) The company has a separate department for inventory planning, which is heavily oriented towards supporting production schedules in sacrifice of inventory costs. Distribution and transportation costs are evaluated based on annual budget and are fully allocated to product divisions. Although annual costs are compared to historical trends and competitors' costs, costs are not usually a big concern in day-to-day operations. The biggest concern is how to handle unexpected contingencies, e.g. production schedule change, special order, etc. Prompt response to requests from other departments such as purchasing, production or marketing managers is emphasized.

3) Network design almost exclusively determines what mode to use. The company's headquarters assigns primary, secondary and back-up carriers, and D.C. managers choose a carrier and schedule shipment. This suggests that a good modal share model can be estimated by using the data of network structure (e.g. distance, annual tonnage, etc.), although carrier selection would be more subjective to D.C. managers’ judgments.

4) Based on past experience, the company thinks that rail performance is satisfactory. With high level of inventory holding, the company does not need fast movements, as well. But since intermodal service is cheaper than rail, intermodal (and no rail) is used for inter-D.C. shipments. Note that two of their major plants are located in Boston and Chicago and can get good rates for movements to a west coast D.C. This suggests that coordination of rail and intermodal pricing may be desired.

In the second study, we looked at carrier selection criteria more specifically. The study shows that shippers' carrier evaluation process and performance measures are indeed changing according to the direction we discussed in this chapter. The subject company
monitors carrier performance in two ways; it monitors operating measures as well as service quality measures every month as shown in the following list:

<table>
<thead>
<tr>
<th>Measures</th>
<th>Criteria</th>
<th>weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ontime delivery</td>
<td>1-95%=0, 96%=8, 97%=12, 98%=17, 99%=23, 100%=30</td>
<td>30</td>
</tr>
<tr>
<td>Ontime pickup</td>
<td>1-95%=0, 96%=1, 97%=2, 98%=4, 99%=7, 100%=12</td>
<td>12</td>
</tr>
<tr>
<td>Billing accuracy</td>
<td>1-95%=0, 96%=0.2, 97%=0.5, 98%=1.2, 99%=2, 100%=3</td>
<td>3</td>
</tr>
<tr>
<td>Equip usability</td>
<td>clean, in good repair, proper size and type</td>
<td>3</td>
</tr>
<tr>
<td>Responsiveness</td>
<td>timely responses to inquiries and concerns, communications</td>
<td>3</td>
</tr>
<tr>
<td>Driver policy</td>
<td>driver screening, training, appearance, retention, courtesy, knowledge, each with 0.5</td>
<td>3</td>
</tr>
<tr>
<td>Loss/damage</td>
<td>0.5%+=0, 0.4%=0.2, 0.3%=0.5, 0.2%=1.2, 0.1%=2, 0%=3</td>
<td>3</td>
</tr>
<tr>
<td>Maintenance</td>
<td>A documented preventive maintenance (P.M.) exists = 0.5</td>
<td>3</td>
</tr>
<tr>
<td>Maintenance</td>
<td>Major P.M.'s are regularly scheduled = 0.5</td>
<td></td>
</tr>
<tr>
<td>Maintenance</td>
<td>Minor P.M.'s are regularly scheduled = 0.5</td>
<td></td>
</tr>
<tr>
<td>Maintenance</td>
<td>Maintenance parts usage/ supply are tracked = 0.5</td>
<td></td>
</tr>
<tr>
<td>Maintenance</td>
<td>Maintenance costs are tracked and analyzed = 0.5</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>60</td>
</tr>
</tbody>
</table>

The company also holds a strategic review committee meeting bi-annually. The committee members evaluate carriers according to the following list:

<table>
<thead>
<tr>
<th>Measures</th>
<th>weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 A documented quality process is in place</td>
<td>4</td>
</tr>
<tr>
<td>2 All employees are fully aware of procedures and policies</td>
<td>4</td>
</tr>
<tr>
<td>3 Visual signs of quality process are in place</td>
<td>4</td>
</tr>
<tr>
<td>4 All employees receive on-going quality training</td>
<td>4</td>
</tr>
<tr>
<td>5 Key indicators are monitored. (on-time delivery, safety, fuel cost, empty mileage)</td>
<td>4</td>
</tr>
<tr>
<td>6 All employees receive ongoing quality training</td>
<td>4</td>
</tr>
<tr>
<td>7 Improvement plans are developed, implemented and tracked</td>
<td>4</td>
</tr>
<tr>
<td>8 Quality techniques are utilized (brain storming, fishbone, questionnaires, interviews)</td>
<td>4</td>
</tr>
<tr>
<td>9 Quality tools are utilized (flow charts, histograms, pareto charts, run charts)</td>
<td>4</td>
</tr>
<tr>
<td>10 All levels of the organization are involved in projects for continuous process improvement</td>
<td>4</td>
</tr>
<tr>
<td>Total</td>
<td>40</td>
</tr>
</tbody>
</table>
When a carrier's sum of scores measuring its operation efficiency and quality improvement process exceeds 80, it is qualified as a primary carrier. If, however, the score exceeds 60, the carrier is a secondary carrier. In order to qualify as a back-up carrier, the carrier should receive at least a score of 40. Freight rates are negotiated after the core carriers are determined.

3.4. Freight Choice Model

Once shippers' objectives of managing logistics systems are defined, we can specify shippers' modal selection process in quantitative terms and estimate the mathematical relationships by employing the random utility model. The random utility model provides a tool to quantify behavioral theories. Parameters that determine choice probabilities can be estimated by using the information of observed choices. As can be easily seen, the validity of parameter estimates will greatly depend on the correct specification of shippers' objective function.

Random utility model says that decision makers (e.g. shippers) have a certain objective (or utility) function and choose the alternative with the highest utility. The objective function is defined in terms of a behavioral theory, as we will discuss later on. Since outside observers cannot measure decision makers' utility functions exactly, the utilities are treated by the analysts as random variables. Mathematically, consider the following utility function:

\[ U_{in} = \beta'x_{in} + \epsilon_{in}. \]

The utility that decision maker \( n \) achieves by choosing alternative \( i \) (\( U_{in} \)) is written as a linear function of alternative-specific service attributes (\( x_{in} \)), value of service attributes (\( \beta \)) and an unobservable disturbance \( \{\epsilon_{in}\} \). The definition assumes that we can construct a

---

2 For a summary of behavioral theories, refer to Chapter 3 in Discrete Choice Analysis by Ben-Akiva and Lerman (1985).

69
single index based on the notion of trade-offs that a decision maker uses explicitly or implicitly in comparing different attributes. The level of service attributes may compensate for each other.

While we cannot observe shippers’ utilities, we can observe choices which depend on the utilities. For example, we observe only whether each alternative is selected or not, i.e.

\[ y_{in} = 1 \quad \text{if} \quad U_{in} \geq U_{jn} \quad \text{for all} \quad j \neq i \]

\[ = 0 \quad \text{otherwise} \]

If we assume that decision makers try to maximize their own utilities, the probability that decision maker \( n \) will choose alternative \( i \) over other alternatives will be:

\[ P_n(i) = P(U_{in} \geq U_{jn} \quad \text{for all} \quad j \neq i) \]

\[ = P(\beta'(x_{in} - x_{jn}) \geq \varepsilon_{in} - \varepsilon_{in} \quad \text{for all} \quad j \neq i). \]

That is, the event of choosing alternative \( i \) is stochastic with choice probability depending on the distributional assumption of \( \{\varepsilon_{in}\} \).

For instance, if we assume that error terms \( \{\varepsilon_{in}\} \) are i.i.d. Gumbel-distributed, we have the following logit probabilities \( F(.) \):

\[ F(i| x_n; \beta) = \frac{e^{\beta x_n}}{\sum_j e^{\beta x_n}} \]

While the simple structure of this model gives it a computational advantage, it is still limited in that the cross-elasticities of the aggregate probabilities of choosing various modes with respect to any given mode are assumed to be equal. This property is called the
single index based on the notion of trade-offs that a decision maker uses explicitly or implicitly in comparing different attributes. The level of service attributes may compensate for each other.

While we cannot observe shippers' utilities, we can observe choices which depend on the utilities. For example, we observe only whether each alternative is selected or not, i.e.

\[ y_{in} = 1 \quad \text{if} \quad U_{in} \geq U_{jn} \quad \text{for all} \quad j \neq i \]
\[ = 0 \quad \text{otherwise} \]

If we assume that decision makers try to maximize their own utilities, the probability that decision maker \( n \) will choose alternative \( i \) over other alternatives will be:

\[ P_n(i) = P(U_{in} \geq U_{jn}) \quad , \text{for all} \quad j \neq i \]
\[ = P(\beta'(x_{in} - x_{jn}) \geq \epsilon_{jn} - \epsilon_{in}) \quad , \text{for all} \quad j \neq i). \]

That is, the event of choosing alternative \( i \) is stochastic with choice probability depending on the distributional assumption of \( \{\epsilon_{in}\} \).

For instance, if we assume that error terms \( \{\epsilon_{in}\} \) are i.i.d. Gumbel-distributed, we have the following logit probabilities \( F(.) \):

\[ F(i| x_n; \beta) = \frac{e^{\beta'x_n}}{\sum_j e^{\beta'x_j}} \]

While the simple structure of this model gives it a computational advantage, it is still limited in that the cross-elasticities of the aggregate probabilities of choosing various modes with respect to any given mode are assumed to be equal. This property is called the
Independence of Irrelevant Alternatives (IIA) assumption. If the IIA assumption is violated, parameter estimates can be inconsistent. We will come back to this issue later.

Similarly, we can assume that error terms are i.i.d. normal-distributed with zero mean and variance of one. The normalization of setting variance to one is necessary for the identification of the model since utility is ordinal and scale-invariant. This model is called probit, whereas the previous model is called logit. The assumption of independence among alternatives is also comparable to the IIA property of a logit model. With the normal assumption, we cannot derive a closed form for choice probabilities as we do with the logit model. For trinomial choices, we have the following distribution of disturbance terms:

\[
\varepsilon_n = \begin{pmatrix} \varepsilon_{1n} \\ \varepsilon_{2n} \\ \varepsilon_{3n} \end{pmatrix} \sim N(0, \Sigma)
\]

where \( \Sigma = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \)

and the following integral form of choice probabilities:

\[
F(1 | x_n; \beta) = P(U_{1n} \geq U_{2n} \text{ and } U_{1n} \geq U_{3n})
\]

\[
= P(\beta' x_{1n} + \varepsilon_{1n} \geq \beta' x_{2n} + \varepsilon_{2n} \text{ and } \beta' x_{1n} + \varepsilon_{1n} \geq \beta' x_{3n} + \varepsilon_{3n})
\]

\[
= P(\varepsilon_{2n} - \varepsilon_{1n} \leq \beta'(x_{1n} - x_{2n}) \text{ and } \varepsilon_{3n} - \varepsilon_{1n} \leq \beta'(x_{1n} - x_{3n}))
\]

\[
= \int_{-\infty}^{\beta'(x_{1n} - x_{2n})} \int_{-\infty}^{\beta'(x_{1n} - x_{3n})} f(\varepsilon_{2n} - \varepsilon_{1n}, \varepsilon_{3n} - \varepsilon_{1n}; \Omega) d(\varepsilon_{2n} - \varepsilon_{1n}) d(\varepsilon_{3n} - \varepsilon_{1n})
\]

where \( \begin{pmatrix} \varepsilon_{2n} - \varepsilon_{1n} \\ \varepsilon_{3n} - \varepsilon_{1n} \end{pmatrix} \sim N(0, \Omega = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}) \).

Although we do not have a closed form solution for the bivariate integral, we can evaluate them numerically. In the GAUSS software, numerical integration is provided for up to
trivariate normal distribution. An advantage of this approach is that Σ does not need to be an identity matrix and can have a general covariance structure.

3.5. Three Generations of Freight Choice Model

We have shown that the objective function of the logistics system can be defined in terms of total logistics costs, customer service and comparative advantage. Accordingly, we can classify three generations of freight choice models: cost minimization model, service maximization model and comparative advantage model. Previous researches have not modeled the modal selection process explicitly, relying on stated importance of service attributes. Some research that models the modal selection process explicitly contained a weakness in that shippers' utility is defined as a linear weighted sum of service attributes without any behavioral theory (See, for example, Swait, Louviere, and Willliams 1993, and Temple, Barker & Sloane 1991). By specifying shippers' objective of modal selection explicitly, we hope to improve over previous researches.

The first generation model estimates a demand response function by assuming that shippers minimize costs. If a shipper incurs higher costs from shipping through a particular mode than from shipping through other modes, she will be less inclined to use the mode. By relating cost attributes directly to actual modal shares, this approach can estimate the degree of influence that each cost item has on actual shares (Roberts et al. 1976, 1977, Chiang 1979, Vieira 1992). In addition, this approach allows researchers to model shipment mode in relation to shipment size or frequency (McFadden, Winston and Boersch-Supan 1985). We will discuss this approach in Chapter 4.

The second generation model assumes that shippers maximize customer service even if they have to sacrifice logistics costs a bit. Many shippers establish specific standards for customer service, such as the percentage of all orders that will be shipped or delivered within a stated time period following receipt (e.g. 95 % shipped within 24 hours or 95 % delivered within two weeks), and require their carriers to meet the standards.
Discussions with experts suggest that there is a consensus about what items of service quality are important but there is a wide disagreement on how important each item is. We will try to resolve this matter in building the service quality model in Chapter 5.

The third generation choice model assumes that shippers maximize comparative advantages. This model has not been developed yet, since the measures for comparative advantages have not been defined well. We cited sustainability and appropriability as two most important measures in the previous section. Unfortunately, how to define practical measures of sustainability and appropriability have not been studied. As shippers' logistics systems reach near the stage of strategic differentiation, we may see more research on such measures. This model is not pursued in this thesis.
Chapter 4.

The Total Logistics Cost Model

4.1. Introduction

In this chapter, we discuss models that specify shippers' utility function in terms of (negative) total logistics cost and that estimate from observed share data the effect of cost components on modal selection. We begin with a discussion on how to analyze share data in section 4.2. Typical data contain information about average modal shares instead of discrete choice at each shipment. In order to analyze share data, we propose the use of the minimum discrimination information (MDI) estimation. In section 4.3, we define components of total logistics costs and discuss their characteristics.

In section 4.4, we show the application of the MDI estimator based on total logistics costs. Here, we estimate a set of fixed parameters assuming that all shippers share the same set of taste parameters. In particular, we estimate a fixed discount rate assuming that all shippers employ the same level of discount rate in calculating inventory holding costs. In reality, such an assumption may be restrictive. Discount rates can differ by shipment (e.g. regular vs. emergency shipments), by corridor (e.g. important vs. small customer) and by company (e.g. good credit vs. bad credit company). Heterogeneity in discount rate is of great concern, since many components of total logistics costs are influenced by discount rate, and ignoring the effects of heterogeneity could result in inconsistent estimates. Thus, we estimate a freight choice model when discount rates are randomly-distributed over observations (section 4.5) and over shippers (section 4.6). In section 4.7, we conclude the chapter by discussing weaknesses of the total logistics cost model.
4.2. Modal Share Model

4.2.1. Revealed Preference Data

The revealed preference (RP) approach estimates demand response by analyzing actual choices in relation to actual situations. The conceptual framework of analyzing RP data is given in Figure 4.1.

![Diagram](image.png)

Figure 4.1. The Revealed Preference Model

Data on actual choices and choice situations may be achieved either through external monitoring or through shipper survey. While external monitoring is effective in collecting data on measurable service performance, a survey reveals shippers' perceptions that may be as important as actual performance. Eventually carriers will have to utilize both systems for data collection. Since the RP data explains actual choice outcomes based on the attribute characteristics of alternatives in actual choice situations, the attribute importance estimated in the RP approach is likely to predict future choices more accurately than hypothetical choices or stated importance.
On the other hand, the RP data are expensive to administer because researchers should observe or survey not only actual choices but also all alternative-specific situational variables that might influence choices. In order to secure sufficient data points, researchers should collect data from many respondents or observe the same respondent many times. In particular, if some alternatives are seldom chosen, the sample size should be enlarged in order to collect adequate data on seldom-chosen alternatives. As the sampling size increases, collected data may contain measurement errors, suffer from individual heterogeneity effects or get non-responses in important variables. In addition, latent or intangible variables such as perceptions cannot be observed from external monitoring.

4.2.2. Modal Share Data

Ideally, we would like to observe modal choices whenever a shipper sends a shipment. But such disaggregate observations are expensive and rarely available. Only average statistics such as modal shares among annual tonnage are available in most data sets. Such limitation in data requires a different estimation method from the usual discrete choice model. On the other hand, modal shares in our data are less subject to aggregation bias since they do not average shares of different shippers. The shares represent time-average statistics for a specific origin-destination pair for each shipper, and thus, reflect the effects of network characteristics uniquely.

A. Berkson’s method

A simple way to model modal shares is to use the Berkson’s method.¹ Let us employ the notation we used in the previous chapter (section 3.1) and also denote modal shares by s_in. Assuming that choice probabilities are determined by the multinomial logit model and that all shipment decisions are made under the same choice set, we can express

¹ Ben-Akiva and Lerman, p. 120
choice probabilities relative to that of a reference alternative (say, "J"). By taking a log on both sides, we have the following relationship:

\[ \ln(s_{jn}/s_{ln}) = \beta' (x_{jn} - x_{ln}) + \xi_{jn} \]

where \( \xi_{jn} \) are (possibly correlated) unobserved disturbances. We can apply the regression (GLS) to obtain an estimate of \( \beta \). While this approach is easy to implement, it is not suitable for our data set since many of observed shares are zero and the log of zero does not exist. A popular approach in such case is to replace zero shares as follows:

\[
s_{in} = \begin{cases} 
  \frac{1}{2K_n}, & \text{if } s_{in} = 0 \\
  s_{in}, & \text{if } 0 < s_{in} < 1 \\
  1 - \frac{1}{2K_n}, & \text{if } s_{in} = 1
\end{cases}
\]

where \( K_n \) is the number of shipments averaged for an observation \( n \) (Hughes and Savin 1994). Unfortunately, we do not have data on shipment frequency. Moreover, we tested this approach with many different \( K_n \)'s. Empirical results suggest that Berkson estimators are highly sensitive to the choice of \( K_n \). Therefore, we cannot use the "2n-rule".

**B. Extremum estimators**

The approach we adopt is Amemiya's extremum estimators (Amemiya 1985, Chapter 4). Suppose that modal choices occur based on random utility model. If the estimates correspond to the true parameters, then estimated shares should perfectly match actual shares. However, reported shares may contain measurement errors, as represented in the following relationship:

\[ s_{in} = F(i \mid x_n; \beta_o) + \zeta_{jn} \]
where $s_{in}$ denotes the actual modal share for mode $i$ and shipper $n$, $x_n$ denotes a vector of explanatory variables that influence actual choice behaviors, $F(i|x_n; \beta)$ is the choice probability that could be defined by either logit or probit c.d.f., for example, and $\{\zeta_{in}\}$ are i.i.d. unobservable random variables with zero mean and positive variance.

Assuming that each observation is independent of each other, extremum estimators estimate parameters by either maximizing or minimizing a certain function defined over the parameter space, i.e. of the following form:

$$Q_N(\beta) = \sum_{n=1}^{N} g(s_n; \beta)$$

where $s = (s_1, \ldots, s_N)'$ is a N-vector of random variables with $s_n = (s_{1n}, s_{2n}, s_{3n})'$ and $\beta$ is a K-vector of parameters. In our case, $s_n$ represents the shares of truck, rail and intermodal. The following estimators are examples of extremum estimators:

1) Maximum likelihood estimators (ML)

$$g(s_n; \beta) = \sum_{i=1}^{3} \log \left[ l(s_n, F(i|x_n; \beta)) \right]$$
where $l[s_{in}, F(i|x_n; \beta)]$ is a p.d.f.

2) Non-linear least squares estimators (NLLS)

$$g(s_n; \beta) = \sum_{i=1}^{3} (s_{in} - F(i|x_n; \beta))^2$$

3) Least absolute deviations estimator (LAD)

$$g(s_n; \beta) = \sum_{i=1}^{3} |s_{in} - F(i|x_n; \beta)|$$

4) Minimum discrimination information estimator (MDI)

$$g(s_n; \beta) = \sum_{i=1}^{3} s_{in} \log \frac{s_{in}}{F(i|x_n; \beta)}$$
In all four examples, we assume that measurement errors ($\zeta_{\text{in}}$) are independent of each other and of $x_i$. NLLS estimator will be the same as ML estimator when the variance of $\zeta_{\text{in}}$ is homogeneous and known. Statistical properties of extremum estimators are derived in Amemiya (1985). Appendix 2 shows that the MDI estimators are consistent and asymptotically normal under typical regularity conditions generally satisfied in the present case. Asymptotic efficiency differs from each other. Among different types of extremum estimators, we focus on MDI and ML throughout this paper.

**C. ML estimator**

The ML estimation approach is the most desirable since it provides a consistent, asymptotically efficient and asymptotically normal estimator. If we know the distribution of disturbance terms, we can write down the likelihood of observing market shares. Note that we have two sources of disturbance terms: one in the choice model and the other in the measurement of modal shares. Let us denote that the covariance matrix associated with $F(i \mid x_n; \beta_o)$ due to the definition of random utility (i.e. $U_{\text{in}} = \beta^\prime x_n + \varepsilon_{\text{in}}$) is $\Sigma$ and that the covariance matrix associated with $I[s_{\text{in}}, F(i \mid x_n; \beta)]$ (i.e. $s_{\text{in}} = F(i \mid x_n; \beta_o) + \zeta_{\text{in}}$) is $\Xi$. Both $\Sigma$ and $\Xi$ are $(J^*J)$ matrices where $J$ is the number of alternatives. In order to specify MLE, we assume the followings:

1) the conditional expectation of $s_{\text{in}}$ given $x_n$ is $F(i \mid x_n; \beta)$

2) the conditional covariance matrix ($\Xi$) of $s_{\text{in}}$ given $x_n$ exists for any $x_n$

3) the model is identifiable, i.e. $F(i \mid x_n; \beta_1) = F(i \mid x_n; \beta_2)$ almost surely $\Rightarrow \beta_1 = \beta_2$

Gourieroux et al. (1984) showed that if $I[s_{\text{in}}, F(i \mid x_n; \beta)]$ belongs to a linear exponential family and if the above assumptions hold, the MLE approach would provide a strongly consistent estimator. Furthermore, if $I[s_{\text{in}}, F(i \mid x_n; \beta)]$ belongs to a linear exponential family, the MLE satisfies asymptotic normality, i.e.
\[ \sqrt{N}(\hat{\beta}_N - \beta_o) \xrightarrow{d} N(0, H^{-1}JH^{-1}) \]

where

\[ H = \mathbb{E}_x\left( \frac{\partial F'}{\partial \beta} \Sigma^{-1}_o \frac{\partial F}{\partial \beta'} \right) \]

\[ J = \mathbb{E}_x\left( \frac{\partial F'}{\partial \beta} \Sigma^{-1}_o \Sigma^{-1}_o \frac{\partial F}{\partial \beta'} \right). \]

where \( \Sigma_o \) and \( \Xi_o \) are \( \Sigma \) and \( \Xi \) matrices evaluated at \( \beta_o \). Since normal and multinomial distributions belong to a linear exponential family, the MLE estimates are strongly consistent and asymptotically normal.

If \( \{\xi_{in}\} \) is i.i.d. normally distributed (with \( \sigma \) known), the objective function, ignoring constant terms, becomes:

\[ S(\beta) = -\frac{1}{\sigma^2} \sum_{n=1}^{N} \sum_{i=1}^{3} \{s_{in} - F(|x_n; \beta)\}^2. \]

This approach with normal distribution assumption is the same as non-linear least squares estimation (NLLS) which tries to achieve the statistical objective of maximizing the data fit or minimizing deviations of estimated shares from actual shares. Non-linear least square-estimators belong to extreme estimators. Amemiya show that NLLS is consistent, asymptotically efficient, and asymptotically normal under very general conditions (Amemiya 1985, Ch. 4).

When \( \{\xi_{in}\} \) is not i.i.d. normal but if we can assume that the same \( \Xi \) applies for all shippers, we will have the following log-likelihood:

\[ L(\beta, \Xi) = \log(2\pi) - \frac{1}{2} \log(\text{det}(\Xi)) - \frac{1}{2} \sum_{n=1}^{N} \left[ s_n - F(|x_n; \beta) \right] \Xi^{-1} \left[ s_n - F(|x_n; \beta) \right]. \]
for a given positive definite matrix $\Xi$. If the covariance matrix $\Xi_o$ associated
with $l(s_{in}, F(i|x_n; \beta_o))$ evaluated at true parameters $\beta_o$ is known, we can focus only on the last
term and parameter estimates become minimum distance estimators that minimize the
following distance (Malinvaud 1978):

$$D(\beta) = \sum_{n=1}^{N} [s_n - F(.|x_n; \beta)]' \Xi^{-1} [s_n - F(.|x_n; \beta)].$$

The MLE approach, however, is not directly applicable to the modal share model.
Since shares are neither negative nor greater than 1, $\{\zeta_{in}\}$ should lie at most between -1
and 2. If $\{\zeta_{in}\}$ are normally and independently distributed normally distributed $\{\zeta_{in}\}$ can
take any value between negative and positive infinity. In addition, $\{\zeta_{in}\}$ cannot be
independent of each other due to the sum-to-one constraint. Observations may also be
correlated with each other when multiple choices are observed from the same shipper.
Dependency arises due to shipper-specific attitudes, discount rates, shipment size,
company policies for risk avoidance, and etc. When distributional assumptions are
violated, the likelihood function is misspecified and the MLE will not be consistent. We
need an estimation method which is robust to the distributional assumption. For this
reason, we adopt the MDI specification.

**D. MDI estimator**

Consider an arbitrary event $E$ that occurs with probability $p$; the nature of $E$ is
completely irrelevant. Suppose that we receive a reliable message at some point in time
that $E$ has in fact occurred. The question is how much information is contained in this
message. To answer this, consider a case where $p$ is close to 1. Then, we would have
known even before the message arrives that $E$ is quite likely to occur. That is, the message
is not very informative. Suppose that $p$ is close to 0, say, $p=0.1$. If a message arrives
stating that $E$ occurred, then we are greatly surprised and the message is quite informative.
The examples suggest that if we want to measure the information content of a message in
terms of the probability of the event, we should select a decreasing (and possibly with decreasing rate) function of the probability. The function used in statistical information theory is that

\[ v(p) = - \ln(p). \]

The reason for choosing this particular form is its additivity for stochastically independent events. Suppose that events \( E \) and \( E' \) occur independently with probabilities \( p \) and \( p' \), respectively. The probability of their joint occurrence is \( pp' \). The value of the information that both \( E \) and \( E' \) occurred is equal to \( v(pp') = v(p) + v(p') \). In fact, this form is derived as the unique solution of certain axioms (unique up to the choice of the logarithmic base), the main axiom being that of additivity (Theil and Fiebig 1984).

Now, consider a complete system of mutually exclusive events \( E_1, \ldots, E_j \) (say, the choice of truck, rail or intermodal) with probabilities \( p_1, \ldots, p_j \). If a message arrives stating that \( E_i \) occurred, the value of the information is \( v(p_i) = -\ln(p_i) \). Before the message arrives, we do not know which event will occur, but we know its expected probability. Therefore, the value of expected information prior to the arrival of the message equals:

\[ \sum_{i=1}^{j} p_i v(p_i) = -\sum_{i=1}^{j} p_i \ln(p_i) \]

which is known as the entropy of the discrete distribution with probabilities \( p_1, \ldots, p_j \).

Given the above specification of information value, the Kullback-Leibler discrimination information function that measures the discrepancy between the two distributions of the observed shares \( (s_n) \) and the underlying shares \( (F(i | x_n, \beta)) \) is defined as follows:

\[ g(s_n; \beta) = \sum_{i=1}^{3} s_n \ln \frac{s_n}{F(i | x_n, \beta)} \]
The MDI approach tries to find parameters that minimize the discrimination information function subject to \( \sum_{i=1}^{3} F(i|x_n; \beta) = 1 \). Thus, parameters will be estimated such that the function \( F(i|x_n; \beta) \) of the attributes results in the final market shares as similar as possible to the observed market shares \( s_n \). In addition, since \( s_n \)'s are fixed, MDI is the same as maximizing the following function:

\[
g(s_n; \beta) = \sum_{i=1}^{3} s_n \ln F(i|x_n; \beta).
\]

We will ignore the sum-to-one constraint from now on, since the choice probabilities \( F(i|x_n; \beta) \) sum to one by construction.

In terms of statistical properties of MDI estimators, MDI estimates are consistent and asymptotically normal (refer to Appendix 2). In practice, Gensch and Soofi (1992) reported that MDI performs better than least squares estimation (i.e. Berkson's method) in terms of the information discrepancy and absolute error in the holdout samples. MDI estimators are also more robust than MLE. MDI does not presuppose any distributional form for \( \{z_{in}\} \) and uses only the information of the total number of choices regardless of whether the choices were correlated or not. Its statistical properties derived in Appendix 2 requires only the differentiability of choice probability, \( F(i|x_n; \beta) \).

Another nice property of MDI is that it can be interpreted as the ML estimator where each shipper send frequent shipments. For instance, suppose that we observe shipment frequency of each shipper via each mode \( (b_n) \) and suppose that \( b_n \) are exogenous (i.e., shipment frequency is independent of the selection process of shipment mode). The likelihood of observing the frequencies can be written as follows:
\[ L^*(\beta) = \prod_{n=1}^{N} \left( \frac{b_n!}{\prod_{i \in A} b_n!} \prod_{i \in A} F(i \mid x_n; \beta)^{D_n} \right) \]

where \( b_n = \sum_i b_{in} \) is the total number of shipments. Parameters that maximize the above likelihood will also maximize the following log-likelihood (ignoring constant terms):

\[ L(\beta) = \sum_{n=1}^{N} \sum_{i \in A} b_n \ln F(i \mid x_n; \beta) \]

or equivalently,

\[ L(\beta) = \sum_{n=1}^{N} b_n \sum_{i \in A} \frac{b_{in}}{b_n} \ln F(i \mid x_n; \beta) \].

If the total number of shipments were the same for all shippers, i.e. \( b_n = B \) for all \( n \), we can rewrite the likelihood as follows:

\[ L(\beta) = B \sum_{n=1}^{N} \sum_{i \in A} s_{in} \ln F(i \mid x_n; \beta) \].

In this case, MDI would be exactly the same as the ML estimation. Even if the total shipment frequency are not the same for all shippers, we can treat them as asymptotically equivalent if they are large for all shippers. In particular, since all shippers reported modal shares of their two largest corridors and since the reported shares might be the average shares over a long period of time, this assumption is not unrealistic with our data. The number of total shipments during a long period in large corridors would be large for all shippers. Thus, we believe that the MDI estimation applied to average shares have similar properties to the ML estimation.

If some shippers are small and seldom use rail or intermodal service due to the small size of shipment volume, whereas some shippers are large and ship a large volume
frequently through rail or intermodal service, then the MDI estimators may not be consistent. In particular, those who use intermodal service frequently are likely to establish work procedure for calling intermodal service and to have good perceptions about intermodal service. In such case, selectivity bias has to be addressed. If such heterogeneity is of great magnitude, we recommend collecting data on shipment frequency and performing the ML estimation with the multinomial specification. We believe that the heterogeneity is less of a concern with our data, since shippers of similar size in five homogeneous product types were selected as a sample.

In addition, note that MDI estimates are not efficient. If the $s_{in}$'s were outcomes of $b_{n}$ independent discrete choices, and if $b_{n}$'s were correctly modeled, the MDI coefficients would be exactly the same as the usual MLE of choice model. But, since MDI is equivalent to the assumption that each individual makes only one choice (i.e. $b_{n} = 1$ for all $n$), the standard error of MDI would be greater than that of MLE estimate by a factor of $\sqrt{b_{n}}$. Thus, if data on shipment frequency is available, MLE should be used.

4.3. Total Logistics Cost Model

4.3.1. Total Logistics Costs

Now, let us discuss the modal share model where shares are determined based on total logistics costs. We can classify total logistics costs as in Table 4.1, using the following notations:

<table>
<thead>
<tr>
<th>Dependent Variables:</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s_{in}$ share of a transportation mode i among annual corridor shipments (%)</td>
</tr>
</tbody>
</table>

---

2 Note that the MLE that we mention here is different from the MLE that we listed as one of extremum estimators. The MLE here maximizes only the choice likelihood. The MLE among extreme estimators both maximizes the choice likelihood and minimizes measurement errors.

3 The table is constructed based on the lecture note of Martland in the Freight Transportation Seminar and on Vieira’s thesis (1992). Similar constructs were already in use in Roberts (1976). I added a few more items as indirect logistics costs.
Independent Variables (Engineering Attributes):

Generic:
- \( p_n \): average price per ton ($)
- \( Q_n \): average annual tonnage in the corridor (ton)
- \( L_n \): latest acceptable delivery hours behind the schedule (hours)
- \( E_n \): earliest acceptable delivery hours before the schedule (hours)

Mode-specific:
- \( f_{in} \): average freight rate per ton per mile (cents)
- \( h_{in} \): average line-haul length between origin and destination (mile)
- \( t_{in} \): average transit time (days)
- \( q_{in} \): average shipment size in the corridor when a shipper \( n \) uses mode \( i \) (tons)
- \( d_{in} \): percentage of shipments being lost or damaged (%)\)
- \( \tau_{in} \): the percentage of times that a carrier delivered shipments within acceptable time frame (%)\)

Since the table is self-explanatory, we skip detailed explanations except out-of-stock costs. Out-of-stock costs are discussed in Appendix 3. The fourth column in the table represents the mathematical expression of each cost component to send one ton of shipment. Items within the parenthesis in the fourth columns are independent variables that can be constructed by using the RP data; and items outside the parentheses are parameters that need to be estimated from the choice model.

The table broadly classifies total logistics costs into direct costs and indirect costs. Direct costs include items that were regarded as important traditionally: transportation costs, inventory holding costs (for both in-transit stocks and safety stocks), loss and damage costs, and ordering costs. Indirect costs are seldom accounted as logistics costs but are becoming more and more important. They include bargaining power against labor union, suppliers or buyers when a shipper has a large inventory; cost decrease from smoothing production process; product obsolescence (or clearance) costs; and long-term effects of customer loss and brand image. When indirect effects are incorporated into total logistics costs, the signs and magnitudes of coefficients can be different from what were traditionally believed. For instance, the effect of in-transit stock holding costs may be different from that of safety stocks due to the difference in risk associated with different types of inventory.
TABLE 4.1. TOTAL LOGISTICS COST TABLE

DIRECT LOGISTICS COSTS

<table>
<thead>
<tr>
<th>Major costs</th>
<th>Major source of costs</th>
<th>Notation</th>
<th>Trend</th>
<th>Major developments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistics</td>
<td>order/handling cost</td>
<td>$OHC_{in} = o \left( \frac{Q_n}{q_{in}} \right)$</td>
<td>↓↓↓</td>
<td>information technology (e.g. EDI)</td>
</tr>
<tr>
<td></td>
<td>transportation cost</td>
<td>$TC_{in} = (f_{in} h_{in} Q_n)$</td>
<td>↓</td>
<td>deregulation</td>
</tr>
<tr>
<td></td>
<td>storage cost</td>
<td>$SC_{in} = s_1 \left( \frac{q_{in}}{2} \right)$</td>
<td>↑</td>
<td>specialized storage device</td>
</tr>
<tr>
<td></td>
<td>loss and damage cost</td>
<td>$LDC_{in} = r \left( \frac{m}{365} d_{in} p_n Q_n \right)$</td>
<td>↔</td>
<td>$m = 100$ (if we assumed that shippers are fully refunded within 100 days)</td>
</tr>
<tr>
<td>Finance</td>
<td>capital cost (for inventory holding)</td>
<td>$OSC_{in} = r \left( p_n \frac{q_{in}}{2} \right)$</td>
<td>↔</td>
<td></td>
</tr>
<tr>
<td></td>
<td>in-transit stock</td>
<td>$TSC_{in} = r \left( \frac{t_{in}}{365} p_n Q_n \right)$</td>
<td>↔</td>
<td></td>
</tr>
<tr>
<td></td>
<td>safety stock</td>
<td>$REL_{in} = r \left( \frac{L_n p_n Q_n}{365} \right)$</td>
<td>↓↓↓</td>
<td>regular / recurrent increased product variety costs of late arrivals</td>
</tr>
</tbody>
</table>
## INDIRECT LOGISTICS COSTS

<table>
<thead>
<tr>
<th>Marketing</th>
<th>stockout cost</th>
<th>demand fluctuation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>- seasonality</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- pulsing prom./adv.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- competitive response</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- speculation</td>
</tr>
<tr>
<td>Marketing</td>
<td>stockout cost</td>
<td>demand fluctuation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- seasonality</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- pulsing prom./adv.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- competitive response</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- speculation</td>
</tr>
<tr>
<td>Obsolescence</td>
<td>new product</td>
<td>new product</td>
</tr>
<tr>
<td></td>
<td>new fashion</td>
<td>new fashion</td>
</tr>
</tbody>
</table>

\[
SS_{in} = r \left( \frac{1 - \tau_{in}}{\Delta \tau_{in}} \right) \frac{p_{n} Q_{n}}{365}
\]

\[
OBS_{in} = v_1 \left( p_n \frac{q_{in}}{2} \right)
+ v_2 \left( \frac{t_{in}}{365} p_n Q_{n} \right)
+ v_3 \left( L_{in} + \frac{1 - \tau_{in}}{\Delta \tau_{in}} \right) \frac{p_{n} Q_{n}}{365}
\]

<table>
<thead>
<tr>
<th>Marketing</th>
<th>stockout cost</th>
<th>demand fluctuation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>- seasonality</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- pulsing prom./adv.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- competitive response</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- speculation</td>
</tr>
<tr>
<td>Marketing</td>
<td>stockout cost</td>
<td>demand fluctuation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- seasonality</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- pulsing prom./adv.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- competitive response</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- speculation</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Marketing</th>
<th>stockout cost</th>
<th>demand fluctuation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>- seasonality</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- pulsing prom./adv.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- competitive response</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- speculation</td>
</tr>
</tbody>
</table>

\[
BP_{in} = s_2 \left( q_{in} \right)
\]

\[
SP_{in} = s_3 \left( q_{in} \right)
\]

\[
EC_{in} = y \left( 1 - u_{eq_{in}} \right) \left( \frac{Q_{n}}{q_{in}} \right) q_{in}
\]

### Marketing
- sporadic / non-recurrent increased competition
- easier brand switching
- emphasis on faster service

### Obsolescence
- shortened product life cycle

### Marketing
- bargaining power
- retailer negotiation
- supplier negotiation
- labor union negotiation

### Production
- smooth production
- process shift
economy of scale
- learning curve

### Equipment Costs
- production scheduling
- quality control

### Marketing
- bargaining power
- retailer negotiation
- supplier negotiation
- labor union negotiation

### Production
- smooth production
- process shift
economy of scale
- learning curve

### Equipment Costs
- production scheduling
- quality control

### Marketing
- bargaining power
- retailer negotiation
- supplier negotiation
- labor union negotiation

### Production
- smooth production
- process shift
economy of scale
- learning curve

### Equipment Costs
- production scheduling
- quality control

### Marketing
- bargaining power
- retailer negotiation
- supplier negotiation
- labor union negotiation

### Production
- smooth production
- process shift
economy of scale
- learning curve

### Equipment Costs
- production scheduling
- quality control

### Marketing
- bargaining power
- retailer negotiation
- supplier negotiation
- labor union negotiation

### Production
- smooth production
- process shift
economy of scale
- learning curve

### Equipment Costs
- production scheduling
- quality control

### Marketing
- bargaining power
- retailer negotiation
- supplier negotiation
- labor union negotiation

### Production
- smooth production
- process shift
economy of scale
- learning curve

### Equipment Costs
- production scheduling
- quality control

### Marketing
- bargaining power
- retailer negotiation
- supplier negotiation
- labor union negotiation

### Production
- smooth production
- process shift
economy of scale
- learning curve

### Equipment Costs
- production scheduling
- quality control

### Marketing
- bargaining power
- retailer negotiation
- supplier negotiation
- labor union negotiation

### Production
- smooth production
- process shift
economy of scale
- learning curve

### Equipment Costs
- production scheduling
- quality control

### Marketing
- bargaining power
- retailer negotiation
- supplier negotiation
- labor union negotiation

### Production
- smooth production
- process shift
economy of scale
- learning curve

### Equipment Costs
- production scheduling
- quality control

### Marketing
- bargaining power
- retailer negotiation
- supplier negotiation
- labor union negotiation

### Production
- smooth production
- process shift
economy of scale
- learning curve

### Equipment Costs
- production scheduling
- quality control

### Marketing
- bargaining power
- retailer negotiation
- supplier negotiation
- labor union negotiation

### Production
- smooth production
- process shift
economy of scale
- learning curve

### Equipment Costs
- production scheduling
- quality control

### Marketing
- bargaining power
- retailer negotiation
- supplier negotiation
- labor union negotiation

### Production
- smooth production
- process shift
economy of scale
- learning curve

### Equipment Costs
- production scheduling
- quality control

### Marketing
- bargaining power
- retailer negotiation
- supplier negotiation
- labor union negotiation

### Production
- smooth production
- process shift
economy of scale
- learning curve

### Equipment Costs
- production scheduling
- quality control

### Marketing
- bargaining power
- retailer negotiation
- supplier negotiation
- labor union negotiation

### Production
- smooth production
- process shift
economy of scale
- learning curve

### Equipment Costs
- production scheduling
- quality control

### Marketing
- bargaining power
- retailer negotiation
- supplier negotiation
- labor union negotiation

### Production
- smooth production
- process shift
economy of scale
- learning curve

### Equipment Costs
- production scheduling
- quality control
We can also see that some items such as order handling costs are shipment frequency \((Q_n / q_{in})\) related; that some items such as storage costs, operating stock holding costs, bargaining power, and smooth production effects are shipment size \((q_{in})\) related; that some items such as transportation costs, reliability costs, and equipment costs are annual tonnage \((Q_n)\) related; and finally that some items such as in-transit stock holding costs, safety stock holding costs and stock-out costs are value-of-time \((r \cdot p_n Q_n/365)\) related. For instance, in-transit holding costs are in-transit time (in days) times inventory holding costs per day \((r \cdot p_n Q_n/365)\). Similarly, safety stock costs are the level of safety stock required (in days) times inventory holding costs per day. Loss-and-damage costs are calculated assuming that all damage claims are fully refunded in \(m\) days and that shippers incur only interests for the loss of opportunity during the claim period. Thus, damage costs are damage probability \((d_{in})\) times the value of product \((p_n Q_n)\) times the discount rate during the claim period \((r \cdot m/365)\).

Summarizing all the variables, we may write the total logistics costs as follows:

\[
U^*_{in} = C_{in} + v'W_{in} + r(T_{in} + \gamma'Z_{in}) + \varepsilon_{in}
\]

where \(C\) is the transportation costs, \(W\) denotes discount rate-independent variables such as on-time delivery and equipment availability performance, \(T\) is capital requirements for in-transit stock, and \(Z\) denotes discount-rate related variables such as loss and damage costs and stock-out costs. Variables in \(Z\) are assumed to follow the same distribution as in-transit stock holding costs. All variables in \(C\), \(W\), \(T\) and \(Z\) are calculated from shipment characteristics and from service attributes of available modes, by using the formula in Table 4.1.

Variables related with shipment size are omitted since shipment size is not available in our data set. We tried to approximate the shipment size by the trailer capacity (e.g. 6 kg for truck, 12 kg for intermodal and 20 kg for rail), assuming that shippers will send only one trailer each time. The ordering and storage costs calculated
with such assumption did not improve the choice model significantly. The result suggests that trailer capacity may not be a reasonable approximation of shipment size, because shippers can send many trailers and because many sizes of trailers are available. For truck shipments, the assumption that a shipper will send only one trailer each time may be reasonable. Truck rates are predominantly a function of the length of haul, shippers would prefer sending a trailer frequently rather than sending multiple trailers. But for rail and intermodal shipments, the assumption may not be reasonable since shippers would prefer sending multiple boxcars or containers. We expect that the effect of ordering costs becomes negligible with the use of EDI (electronic data interchange), particularly for rail shipments. We also expect that the effects of storage costs are low with truck shipments and high with rail shipments.

Instead of including ordering and storage cost variables, we include mode-specific variables. Vieira showed, in his Appendix C, that the effect of optimal shipment size can be reduced to $\delta_i Q_n + \xi_i Q_n h_n + \gamma_i p_n$ where he assumed that shippers select a mode based on total logistics costs for the whole annual tonnage. Yet, as will be demonstrated later, it makes more sense to assume that shippers select a mode based on total logistics costs per ton. The effect of optimal shipment size is then written as $\delta_i \cdot \xi_i h_n + \gamma_i (p_n / Q_n)$. Regarding $\delta_i$, we set rail-specific constant to zero and estimate two constants: one for truck and one for intermodal service. Regarding $\xi_i$ and $\gamma_i$, we estimate only truck-specific parameters. Due to the small number of observations, we need to maintain parsimony. In addition, it was difficult to estimate parameters for intermodal, since intermodal shares are small and have very small variations. For railroad, the omission of parameters is not too restrictive, since the ordering costs are negligible for most rail shipments. Overall, we expect that mode-specific constants and truck-specific variables will compensate for the effects of shipment size.
4.3.2. Transit time Distribution and Stock-out costs

In order to estimate stock-out costs, we need to infer the transit time distribution that shippers perceive for each mode. This section discusses how to estimate the perceived distribution of transit time. Vieira proposed that the transit time be distributed according to an exponential distribution. Yet, the exponential distribution cannot fit the transit time distribution well. The mode of exponential distribution is zero independently of parameter values and independent of transportation distance, whereas typical transit time is positive and dependent on distance. Moreover, exponential distribution has a behavioral assumption of memorylessness, i.e. no matter how long a shipper waited for a delivery, the chance of getting shipments on that day stays the same as on the previous day. Also, the variance of transit time is the same as the mean transit time in exponential distribution. We, instead, propose the use of Weibull-distribution which has the following form:

![Weibull Distribution](image)

**Figure 4.2. The Distribution of Perceived Transit Time**

Weibull distribution has two parameters, α_{in} and β_{in} which may differ by mode and by shipper. When α_{in} is greater than 1, the transit time of mode i as perceived by shipper n (t_{in}) has a left-skewed unimode at a non-zero point. Also, the mean and variance of transit time depends on transit distance (h_{in}). We can derive the parameters (i.e. α_{in} and β_{in}) from the following relationships:
1) Assuming that typical transit time \( T_{in} \) indicates the mode of Weibull-distributed transit time, we have:

\[
T_{in} = \beta_{in} h_{in} \left( \frac{\alpha_{in} - 1}{\alpha_{in}} \right)^{1/\alpha_{in}}
\]

2) From the relationship between the maximum acceptable latest delivery time \( (T_{in} + L_n) \) and the percentage of shipments delivered when wanted \( (\tau_{in}) \), we can derive:

\[
\tau_{in} = F_{in} (T_{in} + L_n) = 1 - \exp\left\{ -\left( \frac{T_{in} + L_n}{\beta_{in} h_{in}} \right)^{\alpha_{in}} \right\}
\]

where \( F_{in}(T_{in} + L_n) \) is the probability of a stock-out with \( L_n \)-days of safety stock when shipper \( n \) uses mode \( i \), i.e. \( F_{in}(T_{in} + L_n) = \int_{T_{in} + L_n}^{\infty} f_{in}(t) \, dt \). From the two equations, we can estimate two parameters, \( \alpha_{in} \) and \( \beta_{in} \) for each shipper and for each mode. The transit time of mode \( i \) perceived by shipper \( n \) is then distributed according to the following density function:

\[
f_{in} (t) = \frac{\alpha_{in} \left( \frac{t}{\beta_{in} h_{in}} \right)^{\alpha_{in} - 1} \exp\left\{ -\left( \frac{t}{\beta_{in} h_{in}} \right)^{\alpha_{in}} \right\}}{eta_{in} h_{in} \Gamma \left( \frac{1}{\alpha_{in}} \right)}
\]

with the following mean and variance:

\[
E(t_{in}) = \frac{\beta_{in} h_{in}}{\alpha_{in}} \Gamma \left( \frac{1}{\alpha_{in}} \right)
\]

\[
V(t_{in}) = \frac{(\beta_{in} h_{in})^2}{\alpha_{in}} \left\{ 2 \Gamma \left( \frac{2}{\alpha_{in}} \right) - \frac{1}{\alpha_{in}} \left[ \Gamma \left( \frac{1}{\alpha_{in}} \right) \right]^2 \right\}
\]

where \( \Gamma(.) \) is the gamma function defined by \( \Gamma(z) = \int_0^\infty t^{z-1} e^{-t} \, dt \) for \( z > 0 \).
Assuming that shippers' perception of transit time follow the Weibull distribution, we could estimate the following statistics:

<table>
<thead>
<tr>
<th></th>
<th>Truck</th>
<th>Rail</th>
<th>Intermodal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean shape parameter ($\alpha_{in}$)</td>
<td>1.629</td>
<td>1.599</td>
<td>1.686</td>
</tr>
<tr>
<td>Mean shape &amp; location para. ($\beta_{in}$)</td>
<td>0.010</td>
<td>0.027</td>
<td>0.018</td>
</tr>
<tr>
<td>Typical (Mode of) transit time (days)</td>
<td>2.184</td>
<td>8.556</td>
<td>5.774</td>
</tr>
<tr>
<td>Average (Mean of) transit time (days)</td>
<td>4.295</td>
<td>15.336</td>
<td>9.093</td>
</tr>
<tr>
<td>Std. dev. of transit time (days)</td>
<td>9.318</td>
<td>12.136</td>
<td>6.426</td>
</tr>
</tbody>
</table>

Given the transit time distribution, we can estimate the level of safety stock that shippers incur in order to prevent stock-out situations. The probability of stock-out situation depends on the randomness of transit time and daily product demands. Since our data set does not have data on the variance of daily product demand, we assumed that demand rate is constant and that only transit time is variable and calculated the size of the safety stock necessary to cope with the variability in transit time. By using the formula in Appendix 3, we found that each mode requires the following inventory level on average:

<table>
<thead>
<tr>
<th></th>
<th>Truck</th>
<th>Rail</th>
<th>Intermodal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean $d\tau_{in} = [F_{in}(T_n + L_n) - F_{in}(T_n + L_n + 1)]$</td>
<td>0.194</td>
<td>0.055</td>
<td>0.084</td>
</tr>
<tr>
<td>Mean safety stock holding ($L_n$ days)</td>
<td>0.833</td>
<td>0.833</td>
<td>0.833</td>
</tr>
<tr>
<td>Expected stockout costs $\left(\frac{1-\tau_{in}}{d\tau_{in} \text{ days}}\right)$</td>
<td>1.042</td>
<td>5.539</td>
<td>1.905</td>
</tr>
</tbody>
</table>

The result shows that when shippers use trucking, they hold inventory of only 6.1 days (e.g. 4.3 days for in-transit stock and 0.8 days for safety stock and 1.0 days for stockout costs); that when they use rail, they hold inventory of 21.6 days (e.g. 15.3 and 0.8 days and 5.5 days, respectively); and that when they use intermodal service, they hold inventory of 11.8 days (e.g. 9.1 and 0.8 days and 1.9 days, respectively). Considering that the surveyed products have relatively high prices, the big differences in the inventory requirements suggest that trucking will be favored over rail.
4.3.3. Scaling of Total Logistics Costs

We discussed that a shipper’s random utility can be interpreted as total logistics costs and that the random utility model specifies the stochastic choice system which can be compared with actual shares. In this section, we will motivate the freight choice model with a simple example. In addition, we explain why the unit cost approach is more reasonable than Vieira’s aggregate cost approach (1992). Although Vieira was one of the few who tried to model shippers’ objectives by total logistics costs; in defining total logistics costs he used aggregate costs that cover the whole annual tonnage. We show that such approach can be erroneous and argue that the total logistics cost per ton should instead be used.

Suppose that a shipper sends a shipment to where only truck and rail are available. Also suppose that product price is $2,000/ton, his discount rate is 15 %, and the annual tonnage is 10,000 tons/year. Suppose that we have information only about transportation costs, in-transit time, safety-stock level, and loss and damage probability which are given as follow:

<table>
<thead>
<tr>
<th></th>
<th>Truck</th>
<th>Rail</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transportation costs ($)</td>
<td>60</td>
<td>50</td>
</tr>
<tr>
<td>In-transit time (days)</td>
<td>3</td>
<td>10</td>
</tr>
<tr>
<td>Safety stock required (days)</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>Loss and damage probability (%)</td>
<td>0.5</td>
<td>1</td>
</tr>
</tbody>
</table>

With the information, we can calculate the total logistics costs that the shipper will incur when he uses each mode. That is,

<table>
<thead>
<tr>
<th></th>
<th>Truck</th>
<th>Rail</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transportation costs ($)</td>
<td>600,000</td>
<td>500,000</td>
</tr>
<tr>
<td>In-transit stock holding costs ($)</td>
<td>24,658</td>
<td>82,192</td>
</tr>
<tr>
<td>Safety stock holding costs ($)</td>
<td>16,438</td>
<td>41,096</td>
</tr>
<tr>
<td>Loss and damage costs ($)</td>
<td>4,110</td>
<td>8,219</td>
</tr>
<tr>
<td>TOTAL AGGREGATE COSTS ($)</td>
<td>645,206</td>
<td>631,507</td>
</tr>
</tbody>
</table>
Based on the costs calculation alone, a normative model would prescribe that he should send all shipments through rail. But suppose that we do not know his total logistics costs accurately. Due to such uncertainty, suppose that we predict modal share according to the following logit formula:

\[ p_i = \frac{\exp(-L_i/1000)}{\exp(-L_t/1000) + \exp(-L_r/1000)} = \frac{1}{1 + \exp((L_r - L_t)/1000)} \text{ for } i = \text{ truck, rail} \]

where \( L_i \) represent the total logistics costs of using mode \( i \). Our model predicts that the truck share is near 0 and the rail share is near one with the above formula.

Now, let us consider the same shipper with only one difference: the annual tonnage is 100 tons/year. His total logistics costs will be as follow:

<table>
<thead>
<tr>
<th></th>
<th>Truck</th>
<th>Rail</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transportation costs ($)</td>
<td>6,000</td>
<td>5,000</td>
</tr>
<tr>
<td>In-transit stock holding costs ($)</td>
<td>247</td>
<td>822</td>
</tr>
<tr>
<td>Safety stock holding costs ($)</td>
<td>164</td>
<td>411</td>
</tr>
<tr>
<td>Loss and damage costs ($)</td>
<td>41</td>
<td>82</td>
</tr>
<tr>
<td><strong>TOTAL AGGREGATE COSTS ($)</strong></td>
<td>6,452</td>
<td>6,315</td>
</tr>
</tbody>
</table>

In this case, our model predicts that the truck share is 0.466 and the rail share is 0.534.

Furthermore, if we assume that the shipper sends only 1 ton/year, he will have the following costs:

<table>
<thead>
<tr>
<th></th>
<th>Truck</th>
<th>Rail</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transportation costs ($)</td>
<td>60.00</td>
<td>50.00</td>
</tr>
<tr>
<td>In-transit stock holding costs ($)</td>
<td>2.47</td>
<td>8.22</td>
</tr>
<tr>
<td>Safety stock holding costs ($)</td>
<td>1.64</td>
<td>4.11</td>
</tr>
<tr>
<td>Loss and damage costs ($)</td>
<td>0.41</td>
<td>0.82</td>
</tr>
<tr>
<td><strong>TOTAL UNIT COSTS ($)</strong></td>
<td>64.52</td>
<td>63.15</td>
</tr>
</tbody>
</table>
In this case, our model predicts that the truck share will be 0.5 and the rail share is 0.5.

These examples cast a big question on what should be the unit of cost calculation in estimating shippers' modal selection. The truck share will be different (ranging from 0 to 0.5 in this example) depending on the annual tonnage, even if the service levels are the same. That is, predicted choice probabilities can differ greatly depending on the absolute magnitude of the annual tonnage, (and thus of total logistics costs). In other aspects, the truck share will be the same as long as the difference in logistics costs are the same, i.e. the logistics costs of (603 and 600 with truck and rail, respectively) will predict the same modal shares that those of (6 and 3) predict. This can create a problem if survey data are composed of diverse shippers who have different shipment size.

One way to avoid this problem is to make the scale (in the above example, 1000 in $\exp(-L_i/1000)$) change proportionally to the absolute level of total logistics costs. The principle of scaling adjustment may be shown more clearly with the assumption of normal disturbances. Suppose that the total logistics cost estimates are normally distributed with the mean being the above expected costs and with the standard error being 10% of the expected costs (i.e. $64,521$ for truck and $63,151$ for rail). Assuming that the disturbances of each cost estimates are independent of each other, the difference between total logistics costs of each mode will be normally distributed with mean of $13,699$ ($= 645,206 - 631,507$) and standard error of $90,283$ ($= \sqrt{64,521^2 + 63,151^2}$). The estimated truck share will be about 0.44 ($= \Phi((0-13,699)/90,283$). One can easily check that the truck share will be 0.44 when the expected costs of using truck and rail are $6,452$ and $6,315$, respectively, as long as we assume that standard errors of the total logistics costs are 10% of the expected costs. This approach of specifying a scale relative to annualized total logistics, however, is difficult to apply to actual estimation, since we do not know a priori the size of total logistics costs.
Another way to handle the scaling problem is to work on the total logistics costs per ton. In this thesis, we recommend the use of this unit cost approach rather than the aggregate cost approach for the following reasons:

1) It makes more sense behaviorally to assume that when a shipper selects shipment mode, he will evaluate total logistics costs per unit weight (e.g. ton) instead of total aggregate costs for annual tonnage. If shippers think in terms of unit costs, unit costs may explain average shares better.

2) It makes shippers comparable. Shippers are very diverse in terms of annual tonnage ranging from 15,000 to 1 ton/year. The total aggregate costs for a shipper who ships 15,000 tons are very different from those for shipping 100 tons. However, the unit costs per ton will be similar independently of annual tonnage. Thus, we can assume that

\[ \mu_n = \mu \text{ for all } n \]

with the unit cost approach.

3) Large shippers have a lot greater variance in their annual total logistics costs than small shippers do. For example, if a shipper sends large shipments annually, a small change in the freight rate can mean a big change in her annual budget. Yet, the variance can be similar if we compare them in terms of logistics costs per ton.

4) When service variables are included in a choice model, unit costs are more comparable to service variables. For example, suppose we want to estimate the effects of service reliability by using the variable "the percentage of times that shipments are delivered when wanted". Comparing the reliability attribute to unit cost makes much more sense than comparing it to aggregate costs. The same is true when we do not
have a directly observable engineering attribute and have to construct service quality as a latent variable.

Now, suppose that a shipper's logistics costs per ton can be written as follows:

\[
U_{in}^* = C_{in} + \nu'W_{in} + r(T_{in} + \gamma'Z_{in}) + \varepsilon_{in}^*
\]

where \( \varepsilon^{**} \) represents the uncertainty in the total logistics cost calculation. Since we will estimate alternative-specific constants, we can assume that \( \varepsilon^{**} \) has zero mean and positive variance. Also, since we employ a measure of unit logistics cost, its variance among shippers would be relatively homoschedastic. Let us denote its standard error by \(|1/\mu| \) (where \( \mu<0 \)).

If we assume that utility is an ordinal function that simply represents a preference ranking of alternatives, the function is unique only up to order-preserving transformation. The absolute level of utility would not matter as long as the ordering of utilities is preserved. Thus, we can rewrite shippers' objective function as follows:

\[
U_{in} = \mu(C_{in} + \nu'W_{in} + r(T_{in} + \gamma'Z_{in})) + \varepsilon_{in}
\]

where \( \varepsilon_{in} = \mu \varepsilon_{in}^* \sim N(0,1) \). This utility function may be interpreted as a scaled negative total logistics cost. The scaling factor can be interpreted as the uncertainty in the total logistics cost calculation.

As \( \mu \) goes to \(-\infty\), this model collapses to a deterministic model where a shipper always chooses the mode with the smallest generalized cost. For a binary case, choice probability behaves in the following way as the difference between generalized costs increases:
As $\mu$ goes to 0, this model becomes a binomial process where shippers choose randomly (i.e., with equal probability) among two alternatives. The difference between generalized costs does not influence choice probabilities at all, as shown below:

When $\mu$ is neither 0 nor $-\infty$, the choice probability is specified in-between the two extremes with choice probabilities given as follow:

$$F(i \mid x_n; \beta) = P(U_{in} \geq U_{jn}, \text{ for all } j \neq i)$$

$$= P(\mu(V_{in} - V_{jn}) \geq \epsilon_{jn} - \epsilon_{in}, \text{ for all } j \neq i)$$

which can be evaluated through either logit or probit.

4.4. The Fixed Rate Model

4.4.1. Empirical Results

In section 4.3, we specified the following function as the shippers' objective:
\[ U_{in} = \mu(C_{in} + \nu'W_{in} + r(T_{in} + \gamma'Z_{in})) + \varepsilon_{in} \]

where \( \mu < 0 \), \( C \) is the transportation costs. \( W \) include mode-specific constants, truck-specific variables that capture the effects of ordering costs, on-time delivery and equipment availability performance, \( T \) is the daily value of in-transit stocks, and \( Z \) include discount rate dependent variables such as loss and damage costs and stock-out costs. We assume that loss and damage claims will be fully refunded in 100 days from the pick-up day and that shippers incur only the interests that they could have earned from the shipment value if the damaged shipment were sold in full. Note that the coefficient of the loss and damage costs is subject to the assumption that claims are fully refunded in 100 days from the pick-up day. Therefore, its coefficient (say, \( \gamma_2 \)) can be interpreted as \( (r \text{ discount rate and } \gamma_2 \cdot 100 \text{ days for refund}) \).

Using the MDI estimation, we estimated a choice model based on unit total logistics costs (i.e. total logistics costs per ton). We chose to apply a probit specification in calculating choice probabilities, since probit is more flexible than logit in combining with structural equations in the next chapter. Logit results have similar likelihood and estimators (except the scale coefficient) at convergence. Table 4.2 shows its results. All coefficients have the correct signs.

The t-statistics reported in the above are based on White's heteroscedasticity-consistent covariance matrix. Note that standard errors should be calculated carefully when we apply the MDI estimation. If the model is well-specified (e.g. with a MLE), either the inverse of the Hessian (the matrix of second-order derivatives, say \( H \)) or the inverse of the cross-product of the first derivatives (say, \( J \)) would yield a consistent estimate of the covariance matrix of the parameters. However, when we apply the MDI to heteroscedastic data with no restriction on the distributional form of measurement errors, the two matrices can diverge. White (1980) proposed that the heteroscedastic-consistent covariance matrix (which equals \( H^{-1}JH^{-1} \)) should be calculated in such case by employing both matrices.
Table 4.2. The RP share model with i.i.d. probit probability

<table>
<thead>
<tr>
<th>Attributes</th>
<th>Notation</th>
<th>Fixed rate model</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Estimates</td>
<td>t-statistics</td>
<td></td>
</tr>
<tr>
<td>Scale (Transportation Cost)</td>
<td>( f_{\text{in}} h_{\text{in}} )</td>
<td>-1.036</td>
<td>-4.831</td>
<td></td>
</tr>
<tr>
<td>Truck-specific constant</td>
<td>1<del>0</del>0</td>
<td>-0.138</td>
<td>-0.478</td>
<td></td>
</tr>
<tr>
<td>Intermodal-specific constant</td>
<td>0<del>0</del>1</td>
<td>1.635</td>
<td>2.924</td>
<td></td>
</tr>
<tr>
<td>Product value / corridor ton (truck-specific)</td>
<td>((p_n/Q_n)\sim0~0)</td>
<td>-0.098</td>
<td>-2.131</td>
<td></td>
</tr>
<tr>
<td>Distance(truck-specific)</td>
<td>( h_{\text{TN}} \sim0~0 )</td>
<td>0.372</td>
<td>1.945</td>
<td></td>
</tr>
<tr>
<td>Delivery time reliability</td>
<td>( \tau_{\text{in}} )</td>
<td>-0.811</td>
<td>-0.918</td>
<td></td>
</tr>
<tr>
<td>Equipment usability</td>
<td>( u_{\text{eqin}} )</td>
<td>-4.346</td>
<td>-4.032</td>
<td></td>
</tr>
<tr>
<td>Discount rate</td>
<td>(%)</td>
<td>0.456</td>
<td>0.782</td>
<td></td>
</tr>
<tr>
<td>In-transit stock holding costs</td>
<td>( t_{\text{in}} p_n/365 )</td>
<td>1</td>
<td>n.a.</td>
<td></td>
</tr>
<tr>
<td>Stockout costs</td>
<td>([(1-\tau_{\text{in}})/d_{\text{in}}] p_n/365 )</td>
<td>0.169</td>
<td>1.948</td>
<td></td>
</tr>
<tr>
<td>Loss and damage costs</td>
<td>( d_{\text{in}} p_n/365 )</td>
<td>1.645</td>
<td>0.306</td>
<td></td>
</tr>
<tr>
<td>Log-likelihood at convergence</td>
<td></td>
<td>-142.77</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log-likelihood at origin</td>
<td></td>
<td>-318.60</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rho-square</td>
<td></td>
<td>0.552</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adjusted Rho-square</td>
<td></td>
<td>0.520</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Discount Rate (%)</th>
<th>Notation</th>
<th>Fixed Model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Estimates</td>
</tr>
<tr>
<td>Mean (%)</td>
<td></td>
<td>45.6</td>
</tr>
<tr>
<td>Mode</td>
<td></td>
<td>45.6</td>
</tr>
</tbody>
</table>
All parameters are scaled by the coefficient of transportation costs, and thus, need to be interpreted relative to transportation costs. Transportation costs have the following basic statistics:

<table>
<thead>
<tr>
<th></th>
<th>Notation</th>
<th>Unit</th>
<th>Mean</th>
<th>Std.dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transportation Costs</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Truck</td>
<td>$C_{in}$</td>
<td>$100/ton$</td>
<td>0.66</td>
<td>0.64</td>
</tr>
<tr>
<td>Rail</td>
<td>$C_{in}$</td>
<td>$100/ton$</td>
<td>0.57</td>
<td>0.34</td>
</tr>
<tr>
<td>Intermodal</td>
<td>$C_{in}$</td>
<td>$100/ton$</td>
<td>0.65</td>
<td>0.45</td>
</tr>
<tr>
<td>Price</td>
<td>$p_n$</td>
<td>$1000/ton$</td>
<td>7.44</td>
<td>3.92</td>
</tr>
<tr>
<td>Corridor tonnage</td>
<td>$Q_n$</td>
<td>1000 ton</td>
<td>0.70</td>
<td>1.41</td>
</tr>
<tr>
<td>Distance</td>
<td>$h_{Tn}$</td>
<td>1000 mile</td>
<td>0.86</td>
<td>0.53</td>
</tr>
</tbody>
</table>

The truck-specific constant of -0.138 suggests that if all conditions are equal, shipper will perceive smaller total logistics costs (by $13.8/ton) than those of rail, and thus will prefer truck. The intermodal-specific constant of 1.635 suggests that if all conditions are equal, shipper will perceive higher total logistics costs (by $163.5/ton) than those of rail and will shun intermodal service. This high penalty on intermodal usage may be due to the effects of omitted variables such as perceptions on ease-of-use or riskiness of using unfamiliar service. The data set was collected in 1988 when intermodal service was still unfamiliar to most shippers and when railroads did not market the service aggressively. More recent data should show much smaller penalty.

The ratio of product value to corridor ton was specified only for truck to capture the effects of order costs. The coefficient of intermodal-specific variable was not significant. The estimate of -0.098 suggests that shippers will prefer truck when product value is high or when corridor tonnage is small. The truck-specific variable of shipment distance has an estimate of 0.372. It suggests that shippers will shun truck when distance is long. Every 100 miles has an effect of increasing total logistics costs of using truck by $37.2/ton relative to rail or intermodal service.

We also employed two variables that represent quantifiable service quality: on-time performance and equipment usability. They constitute two indirect cost items in
Table 4.1. We treated them as engineering variables in this specification, but one may treat them as perceptual indicators and apply the latent factor methodology presented in Chapter 5. The coefficients suggest that shippers are willing to pay $81.1/ton for perfect on-time delivery and $434.6/ton for perfect equipment availability. After incorporating the difference between current performances of truck and rail, the coefficients suggest that shippers are willing to pay $6.5/ton (=81.1*(0.91-0.83)) more for truck than for rail because of truck's better performance in on-time delivery, and that shippers are willing to pay $30.4/ton (=434.6*(0.95-0.88)) more for truck than for rail because of truck's better performance in equipment condition.

<table>
<thead>
<tr>
<th>Attributes</th>
<th>Notation</th>
<th>Unit</th>
<th>Mean</th>
<th>Std.dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>Delivery time reliability</td>
<td>τ_in</td>
<td>[0,1]</td>
<td>0.91</td>
<td>0.08</td>
</tr>
<tr>
<td>Truck</td>
<td></td>
<td></td>
<td>0.83</td>
<td>0.13</td>
</tr>
<tr>
<td>Rail</td>
<td></td>
<td></td>
<td>0.90</td>
<td>0.08</td>
</tr>
<tr>
<td>Intermodal</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equipment usability</td>
<td>ued_in</td>
<td>[0,1]</td>
<td>0.95</td>
<td>0.05</td>
</tr>
<tr>
<td>Truck</td>
<td></td>
<td></td>
<td>0.88</td>
<td>0.10</td>
</tr>
<tr>
<td>Rail</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intermodal</td>
<td></td>
<td></td>
<td>0.92</td>
<td>0.04</td>
</tr>
</tbody>
</table>

This model also estimates the discount rate that shippers use implicitly when they evaluate the costs of holding in-transit stocks, stockout, and loss and damage. The basic statistics of the inventory holding costs are as follows:

<table>
<thead>
<tr>
<th>Attributes</th>
<th>Notation</th>
<th>Unit</th>
<th>Mean</th>
<th>Std.dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>In-transit stock costs</td>
<td>T_in=t_inPn/365</td>
<td>$100/year</td>
<td>0.46</td>
<td>0.40</td>
</tr>
<tr>
<td>Truck</td>
<td></td>
<td></td>
<td>1.78</td>
<td>1.15</td>
</tr>
<tr>
<td>Rail</td>
<td></td>
<td></td>
<td>1.21</td>
<td>0.80</td>
</tr>
<tr>
<td>Intermodal</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stockout costs</td>
<td>[(1-τ_in)/dτ_in]Pn/365</td>
<td>$100/year</td>
<td>0.47</td>
<td>2.26</td>
</tr>
<tr>
<td>Truck</td>
<td></td>
<td></td>
<td>1.32</td>
<td>1.84</td>
</tr>
<tr>
<td>Rail</td>
<td></td>
<td></td>
<td>0.62</td>
<td>1.09</td>
</tr>
<tr>
<td>Intermodal</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Loss and damage costs</td>
<td>m d_inPn/365(m=100)</td>
<td>$100/year</td>
<td>0.06</td>
<td>0.14</td>
</tr>
<tr>
<td>Truck</td>
<td></td>
<td></td>
<td>0.16</td>
<td>0.21</td>
</tr>
<tr>
<td>Rail</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intermodal</td>
<td></td>
<td></td>
<td>0.23</td>
<td>0.25</td>
</tr>
</tbody>
</table>
Since all coefficients of in-transit stock holding costs, stockout costs, and loss and damage costs represent the implicit discount rate, it is difficult to judge what is the correct level of discount rate. The coefficient of in-transit stock suggests 45.6%, while that of safety-stock suggests 7.7% and that of damage costs suggest 75%. We believe that the coefficient of in-transit stock is the correct rate. The level of stockout costs vary widely among shippers and its coefficient is not significant. Moreover, its coefficient may represent indirect effects (e.g. smoothing production schedule, volume purchasing to exploit special promotion, etc.) which cannot be included into total logistics costs. The coefficient of loss and damage costs should be interpreted as indicating that shippers may get full refund after a much longer waiting period (e.g. 164.5 days) than 100 days from the shipping date.

It is troublesome that our estimate is much higher (e.g. 45%) than a typical cost of capital (e.g. 15 to 20%). Three explanations are feasible. First, the discrepancy may arise since shippers try to control inventory costs by setting internal discount rate much higher than external costs. If this is the case, most normative approaches that calculate total logistics costs based on the surveyed discount rate and recommend the mode with minimum total logistics costs would not predict actual choices correctly. We have to find true rates by treating surveyed rates as indicators and estimating a structural equation. Secondly, discount rate may be over-estimated because we did not properly incorporate the heterogeneity of discount rates among shippers. If our data include a few shippers who have very high rates, the overall mean rate can go up. Thirdly, discount rate may be over-estimated because some important variables are omitted. For instance, shippers may prefer trucking because of its highly flexible service. Without the proper incorporation of the effects of flexibility, trucking may appear preferred due to a high discount rate. In Chapter 4 and 5, we will discuss these issues further.
4.4.2. The Distribution of Residuals

To evaluate the adequacy of the independent probit specification, we define standardized residuals as follows:

\[
e_{in}^s = \frac{s_n - F(i| x_n; \hat{\beta}_N)}{\sqrt{F(i| x_n; \hat{\beta}_N)(1 - F(i| x_n; \hat{\beta}_N))}}.
\]

We used the fact that measured shares \( s \) would be distributed around true shares \( p \) with variance \( p(1-p) \). McFadden (1987) proposed to normalize residuals as

\[
e_{in}^* = \frac{s_n - F(i| x_n; \hat{\beta}_N)}{\sqrt{F(i| x_n; \hat{\beta}_N)}}
\]

where \( d \) denotes actual discrete choices. We believe that our normalization is more reasonable conceptually. With our data, we have the following means and covariance matrix of standardized residuals:

**Mean of standardized residuals**

<table>
<thead>
<tr>
<th></th>
<th>Our normalization</th>
<th>McFadden's way</th>
</tr>
</thead>
<tbody>
<tr>
<td>Truck</td>
<td>0.063</td>
<td>0.022</td>
</tr>
<tr>
<td>Rail</td>
<td>-0.057</td>
<td>-0.052</td>
</tr>
<tr>
<td>Intermodal</td>
<td>0.014</td>
<td>0.016</td>
</tr>
</tbody>
</table>

**Covariance matrix with our normalization**

<table>
<thead>
<tr>
<th></th>
<th>Truck</th>
<th>Rail</th>
<th>Intermodal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Truck</td>
<td>0.391</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rail</td>
<td>-0.349</td>
<td>0.375</td>
<td></td>
</tr>
<tr>
<td>Intermodal</td>
<td>-0.120</td>
<td>0.007</td>
<td>0.404</td>
</tr>
</tbody>
</table>

**Covariance matrix with McFadden's approach**

<table>
<thead>
<tr>
<th></th>
<th>Truck</th>
<th>Rail</th>
<th>Intermodal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Truck</td>
<td>0.140</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rail</td>
<td>-0.158</td>
<td>0.253</td>
<td></td>
</tr>
<tr>
<td>Intermodal</td>
<td>-0.057</td>
<td>0.003</td>
<td>0.391</td>
</tr>
</tbody>
</table>

First, when we look at the standardized residuals for each mode, their means cast serious doubts on the adequacy of the choice probability specification. Truck
residuals have a strictly positive mean, while rail residuals have strictly negative means. Second, we assumed that a shipper's utility from using a freight mode is independent of that from using different modes and that the utilities are homogeneous of each other, i.e. similar size of variance. We can test the independence assumption by looking at the covariance matrix. With our way of normalization, heteroscedasticity does not seem to be a problem. But negative correlation between standardized residuals of truck and rail is very high (e.g. -0.91), raising a serious question about the reliability of the simple choice model. We need to extend the fixed rate model into a more elaborate model.

4.5. RP Model with Randomly-distributed Discount Rate

4.5.1. Sources of Heterogeneity

Suppose that we observe a population whose modal share is split 50-50 between truck and rail. At one extreme, this might be interpreted as implying that each shipper in the homogeneous population has a 50% chance of using truck and a 50% chance of using a rail. At the other extreme, this might imply that the population is composed of two groups where one group always uses truck and the other group always uses rail. A discrete choice model with fixed coefficients accepts the former explanation, assuming that the utility of an individual varies randomly from one choice to the next. Yet, one may find it difficult to accept the concept of identical shippers.

Heterogeneity can occur in two types: inter-shipper and intra-shipper heterogeneity. Inter-shipper heterogeneity occurs when shipper-specific characteristics such as attitudinal variations, unobserved perceptions on service quality, choice set generation process and decision protocol influence modal choices. On the other hand, intra-shipper (or inter-shipment) heterogeneity occurs when shipment-specific situational constraints influence modal choices.
**Attitude**  
Attitude denotes a generic latent factor that influences decision makers' tastes. For instance, shippers who believe that customer service is very important may keep using truck even when total logistics costs are cheaper with rail. On the other hand, some shippers may deliberately wish to do some business with each carrier each year. Such tendency may persist for inter-organizational reasons or in order to secure a contingency plan.

**Perceptions**  
Perceptions indicate alternative-specific latent factors that compose the systematic utility of a decision maker but cannot be observed quantitatively. For instance, beliefs that trucking provides convenient service or that rail service cannot be trusted are such examples. For further discussion, refer to Chapter 5.

**Choice set**  
Suppose that a shipper thinks that intermodal service is not available or suitable to him. Then, intermodal service would never be considered and never be chosen in all his responses. Such censoring of a choice set causes observations from a same respondent to be correlated. If the market is composed of several segments who employ choice sets similarly within the segment and differently across segments. The heterogeneity in choice sets can create inconsistency in parameter estimates.

**Decision Protocol**  
Many shippers rely on decision protocols. For instance, a shipper may use truck always (the deterministic rule); use truck if shipment distance is short, if delivery deadline is short, or if a failure to meet the delivery deadline involves huge loss (the lexicographic rule); or calculate total logistics costs and choose the least costly mode (the random utility maximization rule). Such difference in decision protocols can cause heterogeneity in choice responses.

**Situational Constraints**  
In terms of discount rate, heterogeneity can come from the difference both among shippers and among shipments.
Inter-shipper heterogeneity comes from the differences in the financial structures and credit ratings of companies. Some shippers have high discount rates and some don’t. Companies with good credit ratings and with small long-term debts would want to finance mostly by issuing long-term bonds, keeping its cost of capital close to the prime rate. In such cases, discount rates will reflect interest rates on long-term bonds. For companies who raise money through issuance of new stocks, the cost of capital includes a risk-free rate plus a risk-premium which is typically higher than interest rates on bonds. Risk premium compensates for the riskiness of the project and is usually approximated from the volatility of stock prices of companies that conduct business similar to the project. In any case, a company would set internal discount rates considering the costs and financial aspects of capital from all sources.

On the other hand, intra-shipper (or inter-shipment) heterogeneity comes from the differences in shipment. For instance, if a corridor represents a major customer base, shipments in the corridor may receive special treatment relative to other corridors. A shipper may want to use trucking always for the corridor, even if it would involve higher freight rate. The use of a premium transportation service in favor of small inventory will indicate a high discount rate in calculating inventory-holding costs. In order to justify the high transportation cost in terms of inventory holding costs, the discount rate that the shipper assumed should be very high.

4.5.2. Treatment of Heterogeneity

Ignoring parameter heterogeneity can lead to inconsistent estimates. For instance, suppose that we ignore the heterogeneity in alternative-specific constants even though the population is composed of several market segments who have different preferences for trucking. The estimated slope of the common taste model will be inconsistent as shown in Figure 4.3, even if all slope parameters are homogeneous.
Figure 4.3. cases of inconsistent estimates when intercepts are heterogeneous

The inconsistency can also occur when the slope (e.g. discount rate) parameter is heterogeneous (Figure 4.4).

Figure 4.4. cases of inconsistent estimates when slope is heterogeneous

Two major approaches have been proposed to cope with the heterogeneity parameters. One approach is to estimate fixed parameters for each shipper, e.g. alternative-specific constants for each shipper or the shipper-specific discount rate. This approach is referred to as the fixed effect model. This approach has a strength in that the
conditional density of a shipper-specific effect (given shipper characteristics) is different from its unconditional density (Hausman, Hall, and Griliches 1984), and thus, can provide more segment-specific prediction. However, this approach involves the incidental parameter problem since the number of shipper-specific parameters grows linearly with the number of shippers. Moreover, when there are only a few observations for each household, conventional maximum likelihood estimators will lead to inconsistent estimates not only of the fixed terms but also of the effects of other variables (Hsiao 1986).

In order to overcome this problem, Chamberlain proposed the conditional maximum likelihood method that estimates consistent parameters by conditioning the likelihood function on sufficient statistics for the fixed shipper-specific terms (Chamberlain 1980). Jones and Landwehr (1988) applied the method for the case of binary choices. But the conditional MLE becomes extremely complex as the conditioning involves a large number of the possible choice sequences which are permutations of various modes. Moreover, shipper-specific parameters are not estimated in the conditional MLE, and thus, we cannot use the estimation results for prediction. A recent effort includes the use of a Bayesian approach to estimating shipper-specific fixed parameters when multiple observations are available for each shipper (Rossi and Allenby 1993).

A more tractable approach than estimating a parameter for each shipper is to estimate the distribution of parameters, assuming that the parameters vary across the population according to some probability distribution. This approach is called the random effect model (McFadden 1989, Gonul and Srinivasan 1990). For our study, we adopt the random effect model since it ensures parsimony, while providing information (e.g. a distributional form) about the random effect. In particular, discount rate is assumed as random and its distribution can be estimated. In order to see the effects of the random rate model graphically, consider the following four hypothetical shipper populations:
Suppose that the distributions in all four figures have the same expectations. Figure (a) is the base case where only a fixed coefficient is estimated. Figure (b) shows a case where discount rate is symmetrically (e.g. normal) distributed. If discount rate is left-skewed (e.g. gamma, log-normal, Weibull) as shown in figure (c), elasticity will become smaller. On the other hand, elasticity will become larger when right-skewed (e.g. (d) beta distribution with $\alpha_1 > \alpha_2 > 1$). From the perspective of railroads, figure (c) is the most plausible. Shippers who are located to the left from the mean are stayers (i.e. loyal customers) who are willing to accept larger inventory requirements since their discount rates are low. Those who are located to the right are movers (i.e. switchers) who have high discount rates and quickly move to a fast service mode (even if its rate is high) if inventory requirements become large. As a result, aggregate rail share will become less elastic to transit time.

One problem of the random effect model is that the model assumes that the conditional density of the discount rate given shipper characteristics and its unconditional density are identical. This implies that the model can produce inconsistent estimates if self-selection bias occurs (Chamberlain 1980). For example, shippers who set quick response as a company policy for unobserved reasons (e.g perishable product) may make huge investments on establishing efficient trucking connections (e.g. EDI link with trucking
companies, dedicated truck transit pass, or 24-hour docking station). Then, high truck share implies small inventory costs with truck rather than small costs with truck induce high truck share. We assume that such self-selection bias is negligible in our data, since shippers are rather homogeneous. Even if selection bias did occur, we will be able to model the mean and variance of discount rate to vary systematically according to shipper characteristics.

4.5.3. Model Specification with Random Rate

In order to simplify the notation, let us denote that:

\[ U_{in} = \mu(w_{in} + r \cdot z_{in}) + \epsilon_{in} \]

where \( w_{in} = C_{in} + \nu \cdot W_{in} \) represents generalized cost variables such as freight rate and service quality variables and \( z_{in} = T_{in} + \gamma \cdot Z_{in} \) represents generalized value of time that are related to discount rate-related variables. As before, the coefficient of unit transportation cost (\( \mu \)) would indicate the degree of randomness in shippers' utility functions. Variables in \( z_{in} \) (e.g. in-transit stock holding costs, safety stock costs and loss and damage costs) share the common coefficient (i.e. discount rate). We scale the coefficients in \( z_{in} \) by the discount rate such that in-transit stock holding costs has a fixed coefficient of one and the other variables follow the same distribution as in-transit stock holding costs. Now, we allow the discount rates, \( r_n \), to take a random value, i.e.

\[ U_{in} = \mu(w_{in} + r_n \cdot z_{in}) + \epsilon_{in} \cdot \]

If discount rates are independent of \( w_{in} \) and \( z_{in} \), and are a random sampling from a univariate distribution, we can estimate consistently the distribution of discount rates as well as other parameters that determine the utility function. If \( r_n \) is correlated with \( w_{in} \) or \( z_{in} \), this approach would not work unless it is econometrically treated properly.
For the distributional form of discount rate, we assume that the discount rate is log-normally distributed for the following reasons: 1) the log-normal distribution is defined only in the positive range, 2) the log-normal distribution has a left-skewed uni-mode, and 3) a choice model with log-normally-distributed discount rates nests the fixed rate model since the two models are the same when the variance of discount rate is zero. Another distribution we experimented is the Gamma distribution. The Gamma distribution was less desired since its distribution is too compact and does not allow much variation. It also does not nest the fixed rate model since when the variance of a gamma variate goes to zero, its mean also goes to zero. If discount rate is log-normally distributed across shippers, its natural logarithm is normally distributed. That is, if

\[
\ln r \sim N(\phi, \psi^2)
\]

where \( \phi \) is \( E(\ln r) \) is the expected value of the log of discount rate, and \( \psi \) is the standard error of the log of discount rate, then it has the following density function:

\[
f(r) = \frac{1}{\psi r \sqrt{2\pi}} \exp\left(-\frac{1}{2\psi^2} \left(\frac{\ln r - \phi}{\psi}\right)^2\right)
\]

With this assumption, the discount rate has the following statistics:

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mean</strong></td>
<td>( \exp(\phi + \psi^2/2) )</td>
</tr>
<tr>
<td><strong>Mode</strong></td>
<td>( \exp(\phi - \psi^2) )</td>
</tr>
<tr>
<td><strong>Median</strong></td>
<td>( \exp(\phi) )</td>
</tr>
<tr>
<td><strong>Variance</strong></td>
<td>( \exp(2\phi + \psi^2) (\exp(\psi^2)-1) )</td>
</tr>
</tbody>
</table>

This is an asymmetric distribution skewed to the left with minimum value 0 and a tail to the right, i.e
Figure 4.5. Heterogeneity of Discount Rate

As before, choice probabilities conditional on the discount rate can be written either as logit or probit. If we assume that error terms \( \{\varepsilon_{in}\} \) are i.i.d. Gumbel-distributed, we have the following logit probabilities \( F(.) \):

\[
F(i|w_n, z_n, r) = \frac{e^{\mu(w_n + rz_n)}}{\sum_j e^{\mu(w_j + rz_j)}},
\]

If we assume that error terms are i.i.d. normal-distributed with mean zero and variance of one, the conditional choice probability for alternative 1, for example, is:

\[
F(1|w_n, z_n, r) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(\varepsilon_{2n} - \varepsilon_{1n}, \varepsilon_{3n} - \varepsilon_{1n}; \Omega) \, d(\varepsilon_{2n} - \varepsilon_{1n}) \, d(\varepsilon_{3n} - \varepsilon_{1n})
\]

where \( \begin{bmatrix} \varepsilon_{2n} - \varepsilon_{1n} \\ \varepsilon_{3n} - \varepsilon_{1n} \end{bmatrix} \sim N(0, \Omega = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}) \)

Given choice probabilities, we can estimate parameters \((\mu, \nu, \gamma, \phi \text{ and } \psi)\) via the MDI estimation. Unconditional choice probabilities are obtained by integrating over the distribution of the discount rate. The MDI approach then minimizes the following discrepancy of information between actual and expected shares.
Min \[ Q_N(s) = \sum_{n=1}^{N} \sum_{i=1}^{3} s_{in} \ln \frac{s_{in}}{\int_{0}^{\infty} F(i|w_n, z_n, r, \mu) f(r; \phi, \psi) \, dr} \]

or maximizes the following information content:

Max \[ Q^*_N(s) = \sum_{n=1}^{N} \sum_{i=1}^{3} s_{in} \ln \int_{0}^{\infty} F(i|w_n, z_n, r, \mu) f(r; \phi, \psi) \, dr. \]

### 4.5.4. Empirical Results

Numerical integration was performed by using the GAUSS-HERMITE quadrature. Its basic idea is that the integral of a function can be approximated by the weighted sum of its functional values evaluated at a set of points, i.e.

\[ \int_{-\infty}^{\infty} H(y) \exp(-y^2) \, dy \equiv \sum_{r=1}^{R} g_r H(y_r) \]

where \( g_r \) are the weights and \( y_r \) are the quadrature points where the function \( H(y) \) is evaluated. The weights and evaluation points are reported in Appendix 4. The GAUSS-HERMITE quadrature is preferred to the GAUSS-LEGENDRE quadrature (INTQUAD) provided in the GAUSS software, since the GAUSS-HERMITE quadrature is especially designed to evaluate an unbounded integral of the above form (Ben-Akiva, Bolduc and Bradley 1993). We originally used the GAUSS-LEGENDRE quadrature and the results were sensitive the specification of boundaries. Also note that a large number of quadrature points translates to high accuracy only when the integrand is very smooth, in the sense of being well approximated by a polynomial. In our case, 4 points and 12 points generally produced quite similar results.

With choice probabilities specified by i.i.d. probit, we get the result in Table 4.3. The random rate model improved the fixed rate model only by a negligible unit. All parameters are similar to the estimates of the fixed rate model. But we at least showed the feasibility of the random rate approach. The results suggest that the natural logarithm of
discount rate is normally distributed with mean of -0.783 and variance of 0.003². In terms of discount rate, the results suggest that the mean rate is 45.7% and the standard deviation is 4.7%. The small variance suggests that perceived discount rates are high and do not differ much among shippers. The standard errors of the mean, median and mode are estimated by using the delta method which is described in Appendix 5. The result of small variance is disappointing since we expected that the mean discount rate would go down to the usual cost of capital (e.g. 15-20%) after incorporating the heterogeneity. One may try different distributional assumptions (e.g. Gamma).

Table 4.3. The RP Model with Randomly-distributed Discount Rate

<table>
<thead>
<tr>
<th>Attributes</th>
<th>Fixed rate Estimates</th>
<th>Model t-statistics</th>
<th>Random rate Estimates</th>
<th>Model t-statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scale (Transportation Cost)</td>
<td>-1.036</td>
<td>-4.831</td>
<td>-1.707</td>
<td>-1.381</td>
</tr>
<tr>
<td>Truck-specific constant</td>
<td>-0.138</td>
<td>-0.478</td>
<td>-0.181</td>
<td>-0.923</td>
</tr>
<tr>
<td>Intermodal-specific constant</td>
<td>1.635</td>
<td>2.924</td>
<td>1.437</td>
<td>1.359</td>
</tr>
<tr>
<td>Value / Corridor ton (truck)</td>
<td>-0.098</td>
<td>-2.131</td>
<td>-0.368</td>
<td>-1.082</td>
</tr>
<tr>
<td>Distance(truck-specific)</td>
<td>0.372</td>
<td>1.945</td>
<td>0.438</td>
<td>1.306</td>
</tr>
<tr>
<td>Delivery time reliability</td>
<td>-0.811</td>
<td>-0.918</td>
<td>-0.848</td>
<td>-0.795</td>
</tr>
<tr>
<td>Equipment usability</td>
<td>-4.346</td>
<td>-4.032</td>
<td>-4.723</td>
<td>-1.460</td>
</tr>
<tr>
<td>Discount rate</td>
<td>0.456</td>
<td>0.782</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\phi$</td>
<td>-0.783</td>
<td>-1.060</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\ln(\psi)$</td>
<td>0.003</td>
<td>0.102</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\psi$</td>
<td>0.103</td>
<td>n.a.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>In-transit stock holding costs</td>
<td>1</td>
<td>n.a.</td>
<td>1</td>
<td>n.a.</td>
</tr>
<tr>
<td>Safety stock holding costs</td>
<td>0.169</td>
<td>1.948</td>
<td>0.033</td>
<td>2.951</td>
</tr>
<tr>
<td>Loss and damage costs</td>
<td>1.645</td>
<td>0.306</td>
<td>1.537</td>
<td>0.667</td>
</tr>
<tr>
<td>Log-likelihood at convergence</td>
<td>-142.77</td>
<td>-142.62</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log-likelihood at origin</td>
<td>-318.60</td>
<td>-318.60</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rho-square</td>
<td>0.552</td>
<td>0.552</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adjusted Rho-square</td>
<td>0.520</td>
<td>0.518</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Discount Rate (%)</th>
<th>Fixed rate Estimates</th>
<th>Model std. error</th>
<th>Random rate Estimates</th>
<th>Model std. error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean (%)</td>
<td>45.6</td>
<td>13.8</td>
<td>45.5</td>
<td>33.8</td>
</tr>
<tr>
<td>Mode</td>
<td>45.6</td>
<td></td>
<td>45.2</td>
<td>13.2</td>
</tr>
<tr>
<td>Median</td>
<td>45.6</td>
<td></td>
<td>45.7</td>
<td>54.0</td>
</tr>
<tr>
<td>standard deviation</td>
<td>n.a.</td>
<td></td>
<td>4.7</td>
<td>n.a.</td>
</tr>
</tbody>
</table>
4.5.5. Hypothesis Testing

We estimated both the fixed rate model and the random rate model. It would be interesting if we could compare the two results and test whether the random rate model is much different from the fixed rate model, i.e. explains the data generation process better. The null hypothesis is that $\sigma^2 = 0$, and the alternative hypothesis is that $\sigma^2 > 0$. Such test is possible with the lognormally-distributed rate model since it nests the fixed rate model. When the variance of random rate goes to zero, the random rate model converges to the fixed rate model. Such test is infeasible with a gamma-distributed rate model. With the gamma distribution, the mean and variance are functions of the same parameters, i.e. $E(r) = \phi \psi$ and $V(r) = \phi \psi^2$. Thus, if the variance goes to zero, its mean also has to go to zero.

Since the fixed rate model is nested under the random rate model with the log-normal distribution, three tests may easily be devised. The simplest test would be the likelihood ratio test which considers the following statistics:

$$C = -2[L_F(\hat{\beta}_F) - L_R(\hat{\beta}_R)]$$

which is asymptotically $\chi^2$ distributed with $K^*$ degrees of freedom. $K^*$ is the number of restrictions (e.g. one in our case), $\beta_F$ is the estimated coefficient from the fixed rate model, and $\beta_R$ is the estimates of the random rate model. This test is not directly applicable to our model since the information content in the MDI estimation is different from the log-likelihood in the ML estimation. If all shippers sent the asymptotically equivalent number of shipments ($B$), the log-likelihood of the ML estimation will have the following form:

---

1 The level of significance should be properly adjusted. When one variable nests a model, the significance level has to be multiplied by two in order to find the correct critical value. For details, refer to Gourieroux and Monfort (1989, p. 587).
\[ L_F(\beta) = B \sum_{n=1}^{N} \sum_{i=1}^{3} s_{ni} \ln p_m(\beta) \quad \text{and} \quad L_R(\beta) = B \sum_{n=1}^{N} \sum_{i=1}^{3} s_{ni} \int_0^\infty p_r(i, r; \beta) f(r) dr \]

Thus, the correct test statistic is \( C^* = B \cdot C \). Since we do not know the exact value of \( B \), this test is not applicable.

Secondly, we can test whether estimates are the same by comparing estimated coefficients. This is called the Hausman test. If our model is correctly specified, both the fixed rate model and the random rate model should provide consistent coefficient estimates for the same subvector of parameters. If we denote the covariance matrices of parameters as \( \Sigma_F \) and \( \Sigma_R \), then we can construct the following statistic:

\[ C = (\hat{\beta}_F - \hat{\beta}_R)(\hat{\Sigma}_F - \hat{\Sigma}_R)^{-1}(\hat{\beta}_F - \hat{\beta}_R) \]

which is asymptotically \( \chi^2 \) distributed with \( K^{**} \) degrees of freedom where \( K^{**} \) is the number of shared coefficients. The statistic calculated from our results was 7.89. Since \( \chi^2 \) (d.f. = 9) is 16.92 with significance of 5%, the Hausman test does not reject the null hypothesis of a model structure with a fixed rate. This result should be interpreted with care. First, all parameters in our model are scaled by the coefficient of transportation cost. A correct way would be to estimate parameters without scaling and to apply the Hausman test to the unscaled parameters. Secondly, in calculating the test statistic, we did not include discount rate into shared coefficients. While the Hausman test is still valid, the omission of the most important variable may reduce the reliability of the test.

The third test is to utilize standardized residuals which we defined in the previous section. Assuming that measured shares (s) would be distributed around true shares (p) with variance \( \sigma^2 \), standardized residuals should be distributed according to the standard normal distribution. This test can be done by running regression of standardized residuals of the fixed rate model on those of the random rate model. The \( R^2 \) obtained from the regression gives the following Lagrange multiplier test statistic:
\[ C = \frac{JN}{JN - 1} R^2 \]

where \( J=3 \) and \( N=290 \)

which is asymptotically \( \chi^2 \) distributed with one degree of freedom.\(^2\) Our data shows a test statistic of 0.83 which does not reject the null hypothesis that the effect of random discount rate is negligible. McFadden (1987) and Newey (1985) argued that the regression on standardized residuals can be used to test misspecification for nested models such as omitted variables, non-normality, errors-in-variables, and heteroscedasticity.

4.6. The RP Model with Agent Effects

In the previous section, we only assumed that the discount rate will differ from observation to observation, and have not differentiated inter-shipper and intra-shipper heterogeneity. But, since we have multiple observations from the same shipper, the observations may be correlated with each other due to the persistence of shipper-specific discount rates. The event that a shipper-specific discount rate persistently influences all observations of the shipper is called agent effects. When agent effects exist, the maximum likelihood estimation will still produce consistent estimates for random rate models but do not provide correct standard errors. Since the estimation does not use the information that observations from a same shipper share the same discount rate, the estimated parameters are inefficient.

In our RP data, we observe modal shares of two corridors from each shipper, i.e. \( s_n = (s_{n1}, s_{n2}) \). In order to model such multiple choices from a shipper, let us add a subscript to our notation, i.e. consider the following utility function:

\[ U_{int} = \mu (w_{int} + r_{int}^* z_{int}) + \varepsilon_{int}. \]

\(^2\) Here, we do not differentiate alternatives, since the residuals are standardized. For more accurate analysis, a test that analyze multivariate (alternative-specific) residuals may be desired.
where \( t \) denotes observations from each shipper, and

\[
R_n^* = R_n
\]

for all \( t \)

where \( R_n \) denotes inter-shipper heterogeneity which is log-normal distributed with \((\phi, \psi^2)\).

With shipper-persistent discount rates, observations from a shipper is no longer independent of each other. We have to consider the probability of observing the whole series of choices for each shipper.

In order to utilize the information of joint shares, we propose the Expected MDI estimation which maximizes the expected information value of joint share. Given a discount rate \( r \) which is log-normally distributed with parameters \((\phi, \psi)\), the conditional information value of joint shares is

\[
I_n(\mu, \nu, \gamma | r) = \sum_{t=1}^{2} \sum_{s=1}^{3} \ln F(i | w_{nt}, z_{nt}, r ; \mu, \nu, \gamma)
\]

The unconditional expected information value, then becomes:

\[
I(\mu, \nu, \gamma, \phi, \psi) = \sum_{n=1}^{N/2} \int_0^\infty \left[ \sum_{t=1}^{2} \sum_{s=1}^{3} \ln F(i | w_{nt}, z_{nt}, r ; \mu, \nu, \gamma) \right] f(r; \phi, \psi) dr
\]

The integral is again evaluated numerically by using the GAUSS-HERMITE quadrature. After specifying choice probabilities by probit, we got the result in Table 4.4.

All coefficients have expected signs and comparable magnitudes as those of fixed rate and random rate models. While the likelihood at convergence improved greatly, the Hessian failed to invert and standard errors based on the outer product of gradients exploded. A further study is required in order to determine mathematical conditions where minimizing the expected information content will bring consistent estimators.\(^3\)

\(^3\) The estimate that achieves a maximum of \( E[\ln F_{in}(\theta)] \) will be different from the estimate that achieves a maximum of \( \ln E[F_{in}(\theta)] \) due to the non-linearity of the log-function. Similarly, the estimate that achieves a maximum of \( E[s_{in} \ln F_{in}(\theta)] \) will be different from the estimate that achieves a maximum of \( s_{in} \ln E[F_{in}(\theta)] \).
Table 4.4. The RP Model with Agent Effects

<table>
<thead>
<tr>
<th>Attributes</th>
<th>Fixed rate Estimates</th>
<th>Model t-statistics</th>
<th>Agent Estimates</th>
<th>Effect model t-statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scale (Transportation Cost)</td>
<td>-1.036</td>
<td>-0.656</td>
<td>-1.318</td>
<td>-0.442</td>
</tr>
<tr>
<td>Truck-specific constant</td>
<td>-0.138</td>
<td>-0.690</td>
<td>-0.563</td>
<td>-0.435</td>
</tr>
<tr>
<td>Intermodal-specific constant</td>
<td>1.635</td>
<td>7.068</td>
<td>1.264</td>
<td>0.433</td>
</tr>
<tr>
<td>Value / Corridor ton (truck)</td>
<td>-0.098</td>
<td>-1.363</td>
<td>-0.404</td>
<td>-0.438</td>
</tr>
<tr>
<td>Distance(truck-specific)</td>
<td>0.372</td>
<td>2.908</td>
<td>0.361</td>
<td>0.416</td>
</tr>
<tr>
<td>Delivery time reliability</td>
<td>-0.811</td>
<td>-0.690</td>
<td>-1.180</td>
<td>-0.408</td>
</tr>
<tr>
<td>Equipment usability</td>
<td>-4.346</td>
<td>-3.583</td>
<td>-4.874</td>
<td>-0.442</td>
</tr>
<tr>
<td>Discount rate</td>
<td>0.456</td>
<td>3.294</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\phi$</td>
<td></td>
<td>-0.715</td>
<td>-0.001</td>
<td></td>
</tr>
<tr>
<td>$\psi$</td>
<td></td>
<td>0.481</td>
<td>-0.001</td>
<td></td>
</tr>
<tr>
<td>In-transit stock holding costs</td>
<td>1</td>
<td>n.a.</td>
<td>1</td>
<td>n.a.</td>
</tr>
<tr>
<td>Safety stock holding costs</td>
<td>0.169</td>
<td>0.228</td>
<td>0.038</td>
<td>0.009</td>
</tr>
<tr>
<td>Loss and damage costs</td>
<td>1.645</td>
<td>0.517</td>
<td>0.574</td>
<td>0.008</td>
</tr>
<tr>
<td>Log-likelihood at convergence</td>
<td>-142.77</td>
<td></td>
<td>-122.42</td>
<td></td>
</tr>
<tr>
<td>Log-likelihood at origin</td>
<td>-318.60</td>
<td></td>
<td>-318.60</td>
<td></td>
</tr>
<tr>
<td>Rho-square</td>
<td>0.552</td>
<td></td>
<td>0.616</td>
<td></td>
</tr>
<tr>
<td>Adjusted Rho-square</td>
<td>0.520</td>
<td></td>
<td>0.584</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Discount Rate (%)</th>
<th>Fixed rate Estimates</th>
<th>model std. error</th>
<th>Agent Estimates</th>
<th>Effect model std. error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean (%)</td>
<td>45.6</td>
<td>13.8</td>
<td>43.6</td>
<td>236.1</td>
</tr>
<tr>
<td>Mode</td>
<td>45.6</td>
<td></td>
<td>38.8</td>
<td>65.9</td>
</tr>
<tr>
<td>Median</td>
<td>45.6</td>
<td></td>
<td>48.9</td>
<td>351.2</td>
</tr>
<tr>
<td>standard deviation</td>
<td>n.a.</td>
<td></td>
<td>28.0</td>
<td>n.a.</td>
</tr>
</tbody>
</table>
4.7. Weakness of the Total Logistics Cost Model

In this chapter, we have discussed models that assume that shippers minimize only total logistics costs. The assumption is rather strong and we need to consider other objectives of modal selection. First, the total logistics cost approach does not incorporate opportunity costs that can occur from loss of sales due to poor customer service. If railroad managers make decisions based on this model, they are likely to view logistics as cost center rather than as customer service center and to consider cost reduction as the major dimensions of effectiveness. Naturally, railroad managers will be inclined to compete on freight rate rather than on service differentiation. Yet, if railroads compete on freight rate, they lose not only the profitability of their business but also potential sales that can come from providing quality customer service. Somehow, customer service should enter into the shippers’ objective function.

In addition, the total logistics cost model is losing external validity since the Economic Order Quantity (EOQ) formula that the model uses in determining shipment size and inventory costs is less utilized due to the following reasons:

1) EOQ assumes linear costs that do not consider trailer and container size, quantity discount and longhaul transportation rate discount.
2) No one knows what the cost of holding inventory is. The discount rate is only part of the cost. Many other costs such as product pilferage, insurance, space, and more vary widely according to the product type and are not well known.
3) Ordering costs are declining rapidly along with new developments in computer technology. With small ordering costs, EOQ prescribes very frequent deliveries with very small shipment sizes which may not make sense.
4) Inventory monitoring costs are falling fast so that more and more companies are adopting the practice of Economic Order Timing rather than EOQ.
5) EOQ assumes steady demand and production rate. If the rates are fluctuating due to uncontrollable variables such as seasonality, fads and local promotions, the ECQ formula does not work well.
6) Manufacturers prescribe shipment requirements based on production schedule (i.e. material requirement planning or just-in-time planning) or based on marketing schedule (i.e. distribution requirement planning) rather than based on EOQ. Shippers would like to achieve no interruptions in manufacturing process rather than wait until shipments reach a certain size that EOQ prescribed.
Chapter 5.

The Service Quality Perception Model

5.1. Introduction

Shippers do not make transportation choices solely based upon a rational assessment of total logistics costs. Indeed, the lack of information about some of relevant choices means that customers' perceptions about service quality will influence their choice of shipment mode. It would be interesting if we could estimate the effects of service quality perceptions in addition to total logistics costs. In section 5.2, we describe the behavioral framework of the process in which shippers select shipment mode based on both total logistics costs and service quality perceptions.

Section 5.3 discusses how to extract a perception from perceptual indicators and how to apply the technique to our case. A major problem is that we cannot measure service quality directly. We overcome this problem by asking shippers to rate their perceptions about service quality and by extracting a latent factor by utilizing correlation among the ratings. Section 5.4 extends the perceptual model into a case where alternative-specific perceptions are modeled. In section 5.5, we present a conceptual framework that combines factor extraction with modal selection and apply the combined model to our case. In section 5.6, we complete the chapter by discussing miscellaneous issues.

5.2. Behavioral Framework

5.2.1. A Behavioral Framework

In this chapter, we build on the behavioral theory discussed in Chapter 2 and extend it into the following framework of the modal choice process:
The framework assumes that engineering attributes (e.g. freight rate, transit time, on-time delivery, etc.) of freight mode determine total logistics costs and service quality perceptions, which in turn, influence modal choices. Process I and II represent processes in which engineering attributes influence shippers' modal choices through total logistics costs. In Chapter 4, we discussed how to quantify total logistics costs and estimate their effects on modal selection. Process III and IV represent processes in which engineering attributes influence shippers' modal choices through service quality perceptions. We will adopt the structural equation approach for modeling process III and the random utility approach for modeling process IV. Section 5.3 discuss process III and section 5.4 discuss the combined model of all processes of I to IV.

Note that engineering attributes affect both the total logistics costs and service quality perceptions. For example, short transit time not only reduces capital costs committed to in-transit stocks in terms of total logistics costs but also allows shippers to provide customer-responsive service. For another example, on-time delivery not only reduces inventory holding costs necessary for meeting safety stock requirements, but also eliminates a need for shippers to prepare for emergency shipments and to deal with all the liability issues. Irregular shipments may cause congestion, confusion, or poor sequencing
at shipper's receiving docks. While the exact cost of such irregular shipments cannot be calculated at every incidence, these incidences somehow will be recorded and reflected when shippers evaluate the performance of carriers, and eventually select a mode.

5.2.2. Important Service Perceptions

In addition to the reliability we discussed above, many other perceptions may also influence modal selection. The followings are examples of such perceptions:

Flexibility Traffic managers face contingencies all the time. Demand suddenly peak at unexpected places due to unexpected publicity or promotional campaigns. Production schedules are suddenly delayed due to material supply problems, labor disputes, or bad weather, and so forth. In such cases, shippers need to expedite or delay shipments in order to fulfill their sales commitments or to prevent production disruptions. Naturally, they prefer carriers who provide shipment status readily and who are easy to work with in changing shipment sizes, contents, schedules, or routing. In addition, attributes such as carriers’ responsiveness to inquiries, ease of shipment rescheduling or route change, an EDI link, ready provision of shipment tracing information, or good training of sales representatives will improve perception of flexibility of a carrier.

Riskiness Logistics managers do not want to risk a failure of the logistics system due to their decisions of modal selection. Furthermore, the size of perceived risk would be determined by both uncertainty about the outcome of a carrier selection decision and the magnitude of consequences of the wrong choice. For instance, unreliable delivery not only increases inventory holding costs for safety stock requirements, but also increases shippers' liability to delayed shipments. While the exact cost of liability can not be calculated, shippers’ perceptions about the liability will influence modal selection. Such perception can also be formed from the accuracy of billing statements, the history of loss and damage, the availability of usable equipment or the reliability of on-time delivery.
Organizational culture

Individuals are more risk-averse than a company would like them to be. Individuals do not have many options for their career development, and thus, are naturally risk-averse. On the other hand, a company has a portfolio of options and can be risk-neutral. The tendency to avoid risk by traffic managers will become stronger when multiple departments are involved with carrier selection. As many departments evaluate the decision, traffic managers perceive great risk about their carrier selection. The more rigid organizational environments are, the stronger risk avoidance will occur. For instance, if an evaluation of a traffic manager is influenced heavily by complaints of purchasing or marketing managers, the manager will choose reliable service even if it involves high costs.

Familiarity

When delivery were not made as scheduled, the size or reputation of a carrier can help ameliorate criticism of outside managers. As non-experts, outside managers have less background materials to blame a selection of a carrier or a mode that has good reputation. But with a lesser-known mode, they can easily blame transportation-related problems to the selection of a mode. For instance, shippers may not be using intermodal services, since they are afraid of using untried services. If railroads develop programs to familiarize intermodal service, intermodal revenue may grow rapidly.

Ego enhancement

Traffic managers have desires for promotion, salary increase, status within the organization or job security. Thus, if they believe that a change can prove their foresight or analytical capability, they are motivated to propose a new way of doing business. When shippers evaluate core carriers, such situations may occur. The tendency to enhance one's visibility by changing existing practice is called the desire for ego enhancement. This tendency interacts with risk aversion and influences how much individual managers will attempt to change current work process.

Convenience

Shippers may prefer trucking service because of its convenience (e.g. single source of contact, real-time shipment tracing or local pickup and delivery) even though it incurs higher total logistics costs than rail service.
**Strategic fit** From the organizational perspective, shippers will prefer a mode or a carrier who provides commitment for customer service, business-to-business communication process, and comparative advantage from strategic relationship. Corporate name, history of long-term relationship, and financial stability may also be listed.

### 5.2.3. Structural Equation

A recommended approach to modeling Process III of perception generation is to use a structural equation framework. Structural equation models are useful in order both to support behavioral theories and to avoid data problems. The modeling framework can be used when important explanatory variables have not been observed and/or when we want to model a inter-relationship between unobserved perceptions. When we are not interested in the structural relationship among latent factors as in our case, we can focus on the reduced form which is called confirmatory factor analysis. The factor analysis is applied when observed indicators contain measurement errors and when the reduction of dimensions is desirable. By extracting a small number of perceptions from many indicators, the model saves a researcher from a need to collect data on many indicators. A researcher can focus on primary indicators and generalize the finding to minor indicators. Accordingly, a survey form can be greatly reduced.

We assume that railroads can collect data of shippers' perceptions. For instance, although we cannot directly measure the degree of flexibility that shippers perceive of each transportation mode, we can ask shippers how satisfied they are with the responsiveness of a carrier to their inquiries, how easy it is to work with a carrier, and how flexible a carrier is to schedule changes or shipment size changes. Responses to such questions indicate a level of flexibility that shippers perceive of certain alternatives. We call them indicators of true perceptions, since they contain large measurement errors. If several indicators represent the same underlying perception, there will be high collinearity among them. We can extract the latent factor (say, flexibility) by utilizing the collinearity.
5.3. Modeling Perception Generation

5.3.1. A Framework for Modeling Single Perceptual Factor

In order to introduce the structural equation framework, let us consider a case of single perceptual factor. Define \( x_n^* \) as the latent perception of a given individual \( n \). We assume that \( x_n^* \) is a function of the engineering attributes \( z_n \) that influence the formation of the perception on service quality through the following form:

\[
  x_n^* = \pi' z_n + \xi_n
\]  

where \( \pi \) is a \((G*1)\) vector of parameters, \( z_n \) is a \((G*1)\) vector of engineering attributes, and \( \xi_n \) is an error term that is normally distributed with mean zero and variance \( \sigma^2 \). We cannot observe this structural equation directly, but we can observe indicators which measure the latent perception \( x_n^* \) with some measurement errors. We postulate the relationship to be as follows:

\[
  h_{mn}^* = \alpha_m x_n^* + \eta_{mn}, \quad m = 1, \ldots, M
\]  

and

\[
  h_{mn} = h_{mn}^* \quad \text{if } h_{mn} \text{ is continuous}
\]

\[
  = f(h_{mn}^*) \quad \text{if } h_{mn} \text{ is discrete}
\]

where \( h_{mn} \) is shipper \( n \)'s response to a question \( m \) such as how shippers perceive service quality, \( h_{mn}^* \) plays the role of the continuous representation of the indicator and \( \eta_{mn} \) represents errors in collecting data on perceptions. We assume that measurement errors are normally-distributed and independent of other indicators, i.e. \( \eta_n = (\eta_{1n}, \ldots, \eta_{Mn})' \sim N(0, \Theta) \) where \( \Theta \) is a \((M*M)\) diagonal matrix whose diagonal elements represent the degree of measurement errors. Also, \( \alpha = (\alpha_1, \ldots, \alpha_M)' \) is a \((M*1)\) vector of parameters.
that represents the degree that each indicator reflects a latent factor. For identification purpose, one of $\alpha_m$'s should be set to one. In the following, we will set $\alpha_1$ to 1.

5.3.2. *A Case With Two Indicators*

An example of the model system developed so far can be depicted in the below figure where $M = 2$ indicators are postulated and $G = 3$ engineering attributes are assumed to affect the perception about the flexibility of the service:

![Diagram showing the relationship between indicators and satisfaction levels](image)

Arrows on the left side indicate a structural relationship where engineering attributes influence perception formation (Process III) and arrows on the right side indicate a measurement process where respondents reveal their perceptions through indicators (Process IV). Arrows without origin represent structural or measurement error terms. In the survey at hand, satisfaction with carrier response ($rsp_n$) is measured on a continuous scale of 0 to 100, while level of efforts ($lef_n$) is measured on a discrete scale (i.e. 1 = easy, 2 = so-and-so, and 3 = difficult). If level of efforts were measured on a continuous scale ($lef_n^*$), we could write the relationships mathematically as follows:

**Structural Equation:**

$$flex_n = \pi' z_n + \xi_n$$  \hspace{1cm} (4)
Measurement Equation:

\[
\text{rsp}_n = \text{flex}_n + \eta_{1n} = \pi' z_n + \xi_n + \eta_{1n} \quad (5)
\]

\[
\text{lef}_n^* = \alpha_2 \text{flex}_n + \eta_{2n} = \alpha_2 \pi' z_n + \alpha_2 \xi_n + \eta_{2n} \quad (6)
\]

where \( \alpha_2 > 0, \xi_n \sim N(0, \sigma^2), \eta_{1n} \sim N(0, s_1^2), \eta_{2n} \sim N(0, s_2^2) \), and \( \eta_{1n} \) and \( \eta_{2n} \) are independent of each other. The measurement of the level of efforts implies that

\[
\text{lef}_n = 1 \quad \text{(i.e. easy),} \quad \text{if} \quad \text{lef}_1 \leq \text{lef}_n^* \\
= 2 \quad \text{(i.e. so-and-so),} \quad \text{if} \quad \text{lef}_2 \leq \text{lef}_n^* \leq \text{lef}_1 \\
= 3 \quad \text{(i.e. difficult),} \quad \text{if} \quad \text{lef}_n^* \leq \text{lef}_2
\]

for some thresholds \( \text{lef}_1 \) and \( \text{lef}_2 \). Since \( \text{lef}_n^* \) is latent and scale-invariant, we need to impose a normalization restriction such as \( s_2^2 = 1 \). With the assumptions, we have the following distributional relationship:

\[
\begin{pmatrix}
\text{rsp}_n^* \\
\text{lef}_n^*
\end{pmatrix}
\sim N
\begin{pmatrix}
\pi' z_n \\
\alpha_2 \pi' z_n
\end{pmatrix},
\Sigma =
\begin{bmatrix}
\sigma^2 + s_1^2 & \alpha_2 \sigma^2 \\
\alpha_2 \sigma^2 & \alpha_2^2 \sigma^2 + 1
\end{bmatrix}
\]

(7)

We can estimate parameters by maximizing the likelihood of observing \( \text{lef}_n \) and \( \text{rsp}_n \) jointly. The joint likelihood can be written in terms of conditioning on \( \text{lef}_n, \text{rsp}_n \), or a latent perception (\( \xi_n \)). We will discuss all three approaches in the below.

**Sequential estimator conditional on lef**

Current practice (Heckman 1979) suggests that the equation (6) can be estimated by running ordered probit. In order to simplify the relationship, let us divide \( \text{lef}_n^* \) by \( \sqrt{\alpha_2^2 \sigma^2 + 1} \). The newly transformed variable, \( \text{lef}_n^{**} \), implies the following relationship:

\[
\text{lef}_n^{**} = \alpha_2^{**} \pi' z_n + \alpha_2^{**} \xi_n + \eta_{2n}^{**}
\]

(8)
where $\alpha_2^{**} = \alpha_2 / \sqrt{\alpha_2^3 \sigma^2 + 1}$, $\eta_2^{**} \sim N(0, 1 / (\alpha_2^3 \sigma^2 + 1))$ and

$$
\begin{bmatrix}
\text{rsp}_n^{**} \\
\text{lef}_n^{**}
\end{bmatrix}
\sim
\mathcal{N}
\begin{bmatrix}
\pi' z_n^{**} \\
\alpha_2^{**} \pi' z_n^{**}
\end{bmatrix}
\begin{bmatrix}
\sigma^2 + s_i^2 \\
\alpha_2^{**} \sigma^2
\end{bmatrix}
$$

(9)

In addition, the measurement of the level of efforts implies that

$$
\begin{align*}
\text{lef}_n &= 1 \quad \text{(i.e. easy),} \quad \text{if lef}_n^{**} \geq \text{mlef}_1 \\
&= 2 \quad \text{(i.e. so-and-so),} \quad \text{if mlef}_2 \leq \text{lef}_n^{**} \leq \text{mlef}_1 \\
&= 3 \quad \text{(i.e. difficult),} \quad \text{if lef}_n^{**} \leq \text{mlef}_2
\end{align*}
$$

where $\text{mlef}_1 = \text{lef}_1 / \sqrt{\alpha_2^3 \sigma^2 + 1}$ and $\text{mlef}_2 = \text{lef}_2 / \sqrt{\alpha_2^3 \sigma^2 + 1}$. For identification, we assume mlef$_2$ to be zero and find the maximum likelihood estimator of mlef$_1$.

Conditional on the observation of lef$_n$, we can write down the likelihood and expectation of observing rsp$_n$. This says that if a shipper perceive a mode easy to deal with, he is likely to perceive the mode responsive.

$$
E[\text{rsp}_n \mid \text{lef}_n=1] = \pi' z_n + E[\xi_n + \eta_{1n} \mid \text{lef}_n^{**} \geq \text{mlef}_1] \\
= \pi' z_n + E[\sigma_{12} (\alpha_2^{**} \xi_n + \eta_{2n}^{**}) \mid \alpha_2^{**} \xi_n + \eta_{2n}^{**} \geq \Delta_{1n}] \\
= \pi' z_n + \sigma_{12} \frac{\phi(\Delta_{1n})}{1 - \Phi(\Delta_{1n})}
$$

where $\Delta_{1n} = \text{mlef}_1 - \alpha_2^{**} \pi' z_n$, and

$$
\sigma_{12} = \text{Cov} (\text{rsp}_n, \text{lef}_n^{**}) = \text{Cov} (\xi_n + \eta_{1n}, \alpha_2^{**} \xi_n + \eta_{2n}^{**}) = \alpha_2^{**} \sigma^2.
$$

We get consistent estimates of parameters by using least squares and probit as follow:

i) estimate (mlef$_1$) and ($\alpha_2^{**} \pi$) by running ordered probit using all observations,

ii) calculate $q_n = \phi(\Delta_{1n}) / [1-\Phi(\Delta_{1n})]$. 

131
iii) choose observations with \( \text{lef}_n = 1 \) and

\[ \text{estimate } \alpha_{2^{**}}, \, \hat{\pi} \text{ and } \hat{\sigma}_{12} \text{ by running regression of } \text{rsp}_n \text{ on } z_n \text{ and } q_n. \]

[Optional iv) estimate \( \alpha_{2^{**}} \) by running probit on \( \hat{\pi}' z_n \)

v) iterate probit and regression to get \( \alpha_{2^{**}}, \hat{\pi} \text{ and } \hat{\sigma}_{12} \).]

vi) calculate \( \sigma^2 = \sigma_{12}/\alpha_{2^{**}} \) and \( s_1^2 = (\sigma^2+s_1^2) - \sigma^2 \) where \( (\sigma^2+s_1^2) \) is estimated from the residuals of regression of \( \text{rsp}_n \) on all observations.

With the above the procedure and White’s standard errors, the estimation results are as follow:

Table 5.1. A Sequential Model Conditional on lef

<table>
<thead>
<tr>
<th>From Ordered Probit (( \alpha_{2^{**}} \pi ))</th>
<th>Estimates</th>
<th>t-statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>EDI</td>
<td>-0.313</td>
<td>-3.820</td>
</tr>
<tr>
<td>Transit time</td>
<td>-0.065</td>
<td>-3.519</td>
</tr>
<tr>
<td>Equipment availability</td>
<td>2.297</td>
<td>4.211</td>
</tr>
<tr>
<td>( \text{mlef}_1 )</td>
<td>0.421</td>
<td>13.247</td>
</tr>
</tbody>
</table>

| From Regression |
|-----------------|-----------|-------------|
| \( 1/\alpha_{2^{**}} \)   | 0.685    | 7.539       |
| \( \sigma_{12} = \alpha_{2^{**}} \sigma^2 \) | 0.018 | 4.298 |
| \( \sigma^2 = \sigma_{12}/\alpha_{2^{**}} \) | 0.013 | n.a. |
| \( \sigma^2 + s_1^2 \) (from residuals) | 0.023 | n.a. |
| \( s_1^2 = (\sigma^2 + s_1^2) - \sigma^2 \) | 0.010 | n.a. |
| \( \alpha_2 \)                  | 1.481    | n.a.       |

| The rsp equation (\( \pi \)) |
|-------------------------------|-----------|-------------|
| EDI                           | -0.214    | n.a.       |
| Transit time                  | -0.045    | n.a.       |
| Equipment availability        | 1.573     | n.a.       |
| \( \text{lef}_1 \)            | 0.288     | n.a.       |

A caution is needed in calculating t-statistics of the regression since \( q_n \) is not a fixed exogeneous variable but an estimate. Our recommendation is the use of White’s
heteroscedastic-consistent standard errors.\(^1\) \(\alpha_2\) is calculated from \(\alpha_2^{**}\). Note that \(\sigma_{12}\) is significantly greater than zero. This shows that the Heckman step is useful with our data.

**Full MLE through conditioning on \(\text{rsp}\)**

We can write the likelihood of observing both \(\text{rsp}\) and \(\text{lef}\) by conditioning on \(\text{rsp}\), i.e. \(P(\text{rsp}, \text{lef}) = P(\text{lef} | \text{rsp}) f(\text{rsp})\). Note that

\[
\begin{pmatrix} \text{rsp}_n^* \\ \text{lef}_n^* \end{pmatrix} \sim N\left( \begin{pmatrix} \pi' z_n \\ \alpha_2 \pi' z_n \end{pmatrix}, \Sigma = \begin{pmatrix} \sigma^2 + s_i^2 \\ \alpha_2 \sigma^2 & \alpha_2^2 \sigma^2 + 1 \end{pmatrix} \right)
\]

which implies that

\[
\begin{pmatrix} \text{lef}_n^* | \text{rsp}_n \end{pmatrix} \sim N\left( \alpha_2 \pi' z_n + \frac{\alpha_2 \sigma^2}{\sigma^2 + s_i^2} (\text{rsp}_n - \pi' z_n), \alpha_2^2 \sigma^2 + 1 - \frac{(\alpha_2 \sigma^2)^2}{\sigma^2 + s_i^2} \right).
\]

Since \(\text{lef}_n^*\) is scale invariant, we have to normalize the variable again by dividing it by

\[
\sqrt{\alpha_2^2 \sigma^2 + 1 - \frac{(\alpha_2 \sigma^2)^2}{\sigma^2 + s_i^2}}.
\]

If we denote it by \(\text{lef}_n^{***}\), it has the following distribution:

\[
\begin{pmatrix} \text{lef}_n^{***} | \text{rsp}_n \end{pmatrix} \sim N\left( \alpha_2^{***} \pi' z_n + \alpha_2^{***} \frac{\sigma^2}{\sigma^2 + s_i^2} (\text{rsp}_n - \pi' z_n), 1 \right)
\]

where \(\alpha_2^{***} = \alpha_2 / \sqrt{\alpha_2^2 \sigma^2 + 1 - \frac{(\alpha_2 \sigma^2)^2}{\sigma^2 + s_i^2}}\). The conditional probability of rating 1 for \(\text{lef}\) given \(\text{rsp}\) is then:

\[
P(\text{lef}_n = 1 | \text{rsp}_n) = P(\text{lef}_n^{***} \geq \text{nlef}_1 | \text{rsp}_n)
\]

\(^1\) Note that White's heteroscedastic-consistent variance in the regression is different from that in the MDI. In the MDI, it was \(H^{-1}JH^{-1}\) where \(H\) is the inverse of the Hessian and \(J\) is the inverse of the cross-product of the first derivatives of the objective function. In the regression, White's variance has the following form:

\[
V = N(z' z)^{-1} \left[ \sum_{n=1}^{N} \hat{\omega}^{-2} \text{z}_n' \text{z}_n (z' z)^{-1} \right]
\]

where \(\hat{\omega} = \text{rsp}_n - \hat{\pi}' \text{z}_n\).
\[ 1 - \Phi \left( n_{\text{leff}} - \alpha_2^{**} \left( 1 - \frac{\sigma^2}{\sigma^2 + s_i^2} \right) \pi' z_n - \alpha_2^{**} \frac{\sigma^2}{\sigma^2 + s_i^2} \right. \right. \]

where \( n_{\text{leff}} = \text{leff} / \sqrt{\alpha_2^2 \sigma^2 + 1 - \left( \frac{\alpha_2^2 \sigma^2}{\sigma^2 + s_i^2} \right)^2} \) and \( n_{\text{leff}} = 0 \). Probabilities of rating 2 or 3 for \( \text{leff} \) conditional on \( \text{rsp}_n \) can be similarly written. Parameters are estimated by maximizing the joint likelihood which has the following form:

\[ L(\text{rsp}, \text{leff}) = \sum_{n=1}^{N} \ln \left( \sum_{k=1}^{3} I(\text{leff}_n = k)P(\text{leff}_n = k|\text{rsp}_n)\psi(\text{rsp}_n) \right) \]  

(10)

where \( I(\cdot) \) is an indicator function. The estimates of \( \pi \) and \( (\sigma^2 + s_i^2) \) are identified from the "rsp" equation. Given their estimates, the estimates of \( (\alpha_2^{**}, \sigma^2 \text{ and } n_{\text{leff}}) \) are identified from the "leff" equation. With this approach, we have the following estimates:

<table>
<thead>
<tr>
<th></th>
<th>Estimates</th>
<th>t-statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>EDI</td>
<td>-0.016</td>
<td>-1.598</td>
</tr>
<tr>
<td>Transit time</td>
<td>-0.014</td>
<td>-9.198</td>
</tr>
<tr>
<td>Equipment availability</td>
<td>0.681</td>
<td>12.643</td>
</tr>
<tr>
<td>n_{leff}</td>
<td>0.424</td>
<td>2.104</td>
</tr>
<tr>
<td>( \alpha_2^{**} )</td>
<td>1.784</td>
<td>13.288</td>
</tr>
<tr>
<td>( \sigma^2 )</td>
<td>0.013</td>
<td>6.370</td>
</tr>
<tr>
<td>( s_i^2 )</td>
<td>0.006</td>
<td>3.427</td>
</tr>
<tr>
<td>( \alpha_2 )</td>
<td>1.796</td>
<td>n.a.</td>
</tr>
<tr>
<td>\text{leff}</td>
<td>0.427</td>
<td>n.a.</td>
</tr>
</tbody>
</table>

In the above table, \( \alpha_2 \) is calculated from \( \alpha_2^{**} \), and \( \text{leff} \) is calculated from \( n_{\text{leff}} \). The likelihood at convergence was - 6005.3 and the \( R^2 \) of the "rsp" equation was 0.112. Similar to the previous sequential approach, EDI and transit time have shown negative coefficients, and equipment availability showed a strongly positive coefficient. But numerical estimates seem different from what we had with the sequential estimation. We need more study of the data with different approaches.
Full MLE through conditioning on a latent factor

A third method is to directly integrate the probabilities of $r isp$ and $lef$ over the space of latent factors. Note that given latent factors, indicators are independent of each other and are normally distributed. Thus, conditional on latent factors, the likelihood of observing indicators is just a multiplication of the probabilities of observing all indicators, i.e.

$$P(risp_n, lef_n | z_n) = \int_{\xi} f(risp_n | \xi, z_n) P(lef_n | \xi, z_n) f(\xi) d\xi$$

where

$$r isp_n | \xi \sim N(\pi' z_n + \xi, \sigma^2)$$

$$lef_n^* | \xi \sim N(\alpha_2 \pi' z_n + \alpha_2 \xi, 1)$$

$$\xi \sim N(0, \sigma^2)$$

Unfortunately, this model has trouble in convergence. The decomposition of $\sigma^2$ and $\sigma^2$ from the variance of $r isp$, i.e. $(\sigma^2 + \sigma_1^2)$ seems to be troublesome. An empirical identification problem can occur if the variance of indicators is too small. Additional indicators may be required to estimate parameters. For instance, this approach worked well with three indicators, as we will show later.

Comparison

We have discussed three different approaches to extract shippers' perception of modal flexibility from indicators of $r isp$ (continous) and $lef$ (discrete). The decision as to which approach to use depends on the pertinent behavioral theory and computational considerations. The sequential approach conditional on $lef$ is convenient since it can be estimated by using standard statistical packages. It finds consistent estimates with the least computational efforts. But its estimates are not asymptotically efficient since it uses only partial information. In addition, its behavioral assumption that shippers first evaluate the required level of efforts and then evaluate a carrier's responsiveness may be questionable.
Therefore, the full MLE is more desired than the sequential estimation since the full MLE makes no such behavioral assumptions and provides consistent and asymptotically efficient estimates. Asymptotically, the full MLE through conditioning rsp and conditioning a latent factor should provide exactly the same estimates. The full MLE through conditioning a latent factor is more desired, since the full MLE conditional on an indicator becomes difficult to specify as the number of indicators increases.

5.3.3. A Case With Three Indicators

If we have more indicators, we can obtain greater information about latent factors and more accurate estimates of parameters. Additional indicators need not exactly reflect a latent factor of interest. As long as the additional indicator has a monotone relationship with the latent factor of interest, the indicator will help improve estimation. For example, in our data, another indicator, satisfaction with payment terms and billing (ptb), is available. Even if it has a low correlation with perception of flexibility, we can still think that the more satisfied a shipper is with payment terms, the more likely she is to have a positive perception on a carrier's flexibility. That is, we can treat ptb as another indicator of flex. The path diagram with ptb can then be drawn as follows:

Mathematically, we can denote the relationship as follows (refer to equations 8 and 9):

\[
\begin{align*}
\text{rsp}_n &= \pi' z_n + \xi_n + \eta_{1n} \\
\text{lef}_n &= \alpha_2 \pi' z_n + \alpha_2 \xi_n + \eta_{2n} \\
\text{ptb}_n &= \alpha_3 \pi' z_n + \alpha_3 \xi_n + \eta_{3n}
\end{align*}
\]
where
\[
\begin{pmatrix}
\text{rsp}_n \\
\text{ptb}_n \\
\text{lef}_n
\end{pmatrix} 
\sim \mathcal{N}
\left(
\begin{pmatrix}
\pi' Z_n \\
\alpha_2 \pi' Z_n \\
\alpha_3 \pi' Z_n
\end{pmatrix},
\begin{pmatrix}
\sigma^2 + s_i^2 & \alpha_2 \sigma^2 & \alpha_2 \sigma^2 + 1 \\
\alpha_2 \sigma^2 & \alpha_3 \sigma^2 & \alpha_3 \sigma^2 \\
\alpha_2 \sigma^2 & \alpha_3 \sigma^2 & \alpha_3 \sigma^2 + s_i^2
\end{pmatrix}
\right)
\]

The additional indicator not only reduces standard errors of parameters, but also can solve an identification problem. When we add one more indicator, we get three more equations (linear slope, one variance and one covariance) and two more parameters to estimate ($\alpha_3$ and $s_3^2$). $\sigma^2$ can be estimated from the covariance between rsp and ptb. Moreover, the full likelihood as conditional on a certain perception becomes difficult to specify as the number of indicators increases. Thus, the full MLE conditional on latent factors is more attractive. For instance, the rsp, ptb and lef are independent conditional on a latent factor. Thus, the joint likelihood of three indicators is written as follows:

\[
P(\text{rsp}_n, \text{ptb}_n, \text{lef}_n | z_n) = \int_\xi f(\text{rsp}_n | \xi, z_n) f(\text{ptb}_n | \xi, z_n) P(\text{lef}_n | \xi, z_n) f(\xi) d\xi
\]  

The integral is approximated through numerical integration by using the GAUSS-HERMITE quadrature discussed in the previous chapter and Appendix 4. Since we are using the ML estimation, t-statistics are calculated by using the inverse of the Hessian matrix. The results are as follow:

Table 5.3. A Full MLE with Three Perceptual Indicators

<table>
<thead>
<tr>
<th></th>
<th>Estimates</th>
<th>t-statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>EDI</td>
<td>-0.104</td>
<td>-4.422</td>
</tr>
<tr>
<td>Transit time</td>
<td>-0.094</td>
<td>-25.163</td>
</tr>
<tr>
<td>Equipment availability</td>
<td>0.977</td>
<td>7.016</td>
</tr>
<tr>
<td>lef$_1$</td>
<td>0.435</td>
<td>13.645</td>
</tr>
<tr>
<td>$\alpha_2$</td>
<td>1.647</td>
<td>7.327</td>
</tr>
<tr>
<td>$\alpha_3$</td>
<td>0.254</td>
<td>12.416</td>
</tr>
<tr>
<td>$\sigma^2$</td>
<td>0.118</td>
<td>8.106</td>
</tr>
<tr>
<td>$s_1^2$</td>
<td>0.049</td>
<td>39.767</td>
</tr>
<tr>
<td>$s_2^2$</td>
<td>0.016</td>
<td>88.573</td>
</tr>
</tbody>
</table>
The log-likelihood at convergence is -8805.61 with 12 quadrature points. The R$^2$ in the rsp equation was 0.377 and the R$^2$ in the ptb equation was 0.258. The t-statistics are based on the inverse of the Hessian. The estimates are asymptotically efficient since they are ML estimators. In addition, note that the inclusion of ptb reduced standard errors and increased efficiency. Even if the correlation between an additional indicator (ptb) and a latent factor (flex) is small, the inclusion of the indicator increases efficiency. This implies that it is better to have many indicators even if some indicators are not directly related with latent factors. If, however, the extra equation was erroneously specified (i.e. omitted variables, mis-specified functional form, etc.), then the whole parameter estimates can be biased. Therefore, researchers should carefully consider the possibility of misspecification when they add additional indicators.

As before, EDI and transit time have negative coefficients and equipment usability has a positive coefficient. The negative effect of transportation time and the positive effect of equipment usability are as expected. For railroad managers, it may be surprising or disappointing to know that EDI (Electronic Data Interchange) influences flexibility perception negatively. Railroad managers have been supporting EDI strongly since they believe that EDI can reduce time, costs and errors involved in handling documents manually. The negative effect of EDI can be interpreted in two ways. First, usually large carriers, including most railroad companies and national trucking companies, provide EDI, while many small trucking companies are slow to adopt EDI. Shippers may perceive more bureaucracy from large carriers and more flexibility from small trucking companies. Without the size of carriers, we cannot estimate this effect from our data set. If this were the case, carriers who offer EDI should change their organization structure so that their service can respond to local transportation needs more flexibly. Secondly, currently available EDI softwares may not be friendly to small shippers. Many shippers who are not accustomed to using a computer system may feel uncomfortable in doing business without human interaction. Some of them may even process twice, manually and through EDI. For them, delivery request via EDI is much less flexible than via paper, fax or telephone. If this were the case, carriers should improve their EDI software to a more user-friendly one.
5.4. Modeling Alternative-specific Perceptions

5.4.1. A General Framework to Model Single Perception With Multiple Alternatives

So far, we have not considered the effects of alternatives. Yet, perceptions are alternative-specific. That is, a perception on the flexibility of trucking service is different from that of rail service. In order to combine factor extraction with modal selection, we need to model alternative-specific perceptions and allow interactions among them. The alternative-specific aspect of perceptions makes a perceptual modeling quite different from an attitudinal modeling, since attitudes are generic to all alternatives. The modeling of alternative-specific perception generation process can be quite laborious and unparsimonious. Suppose that shippers have multiple alternatives and they generate perceptions of service quality on all available alternatives. We can write the following relationships for each alternative \((i)\), for each individual \((n)\) and for each indicator \((m)\):

\[
x_{in}^* = \pi_i \cdot z_{in} + \xi_{in} \quad \text{where } \xi_{in} \sim \mathcal{N}(0, \sigma_i^2)
\]

\[
h_{imn} = \alpha_{im} \cdot x_{in}^* + \eta_{imn} \quad \text{where } \eta_{imn} \sim \mathcal{N}(0, \sigma_{im}^2)
\]

Estimating \(\pi_i, \alpha_{im}, \sigma_i^2,\) and \(\sigma_{im}^2\)'s separately for each alternative increases the number of parameters by a factor of the number of alternatives.

It may not be unreasonable to assume that the way that engineering attributes generate perceptions about a mode is similar to the way that they do about other modes and that the way that a decision maker respond to measurement surveys for a mode is similar to the way that he does for other modes. With the two assumptions, we can impose common structures across alternatives. For instance, consider the following simplification:

\[
\pi_1 = \pi_2 = \pi_3 = \pi
\]

\[
\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma^2
\]

\[
\alpha_{1m} = \alpha_{2m} = \alpha_{3m} = \alpha_m
\]

and

\[
\xi_n = (\xi_{1n}, \ldots, \xi_{3n}) \sim \mathcal{N}(0, \Sigma_2)
\]

\[
\eta_n = (\eta_{11n}, \eta_{21n}, \eta_{31n}, \ldots, \eta_{1Mn}, \eta_{2Mn}, \eta_{3Mn})' \sim \mathcal{N}(0, S)
\]
where $S$ is a diagonal matrix whose diagonal element is a corresponding $s_{im}^2$.

Assumptions (1) to (3) basically say that the structure of latent factors is independent from alternative characteristics. (1) says that the way in which perceptions on alternative $i$ are generated is the same as the way in which perceptions on alternative $j$ are generated. (2) says that latent factors have the same covariance structure, no matter what alternative they belong to. (3) says that the way in which latent factors are measured via indicators are the same for all alternatives. Assumptions (4)-(5) follows from assumptions (1) to (3) and the normality of disturbances. Assumption (4) says that latent factors ($\xi_{in}$ and $\xi_{jn}$) are correlated among alternatives through the $(i,j)$-th element of $\Sigma_2$. Assumption (5) says that measurement errors ($\eta_{imn}$ and $\eta_{jmn}$) are not correlated among alternatives. It implies that all correlations among indicators occur only through latent factors.

If all indicators are continuously scaled, standard packages such as LISREL provide estimators with restrictions that some coefficients share the same value. If indicators are discrete, the likelihood becomes complicated. We can either approximate normal scores and run a structural equation model, or write down our own estimation program. If indicators are discrete and if the number of alternatives is large, we may want to write down a likelihood function and estimate maximum likelihood estimators with assuming that $\Sigma_2 = I$. If we assume that $\Sigma_2 = I$, we no longer need to consider alternative-specific heteroscedasticity or covariance terms. For a case of two alternatives, a direct implication of the assumption is that perceptual differences maintain the same structure, i.e.

$$x_{1n}^* - x_{2n}^* = \pi'(z_{1n} - z_{2n}) + \xi_{1n} - \xi_{2n}$$

$$h_{1mn} - h_{2mn} = \alpha_m (x_{1n}^* - x_{2n}^*) + \eta_{1mn} - \eta_{2mn}$$

where $\xi_{1n} - \xi_{2n} \sim N(0, 2\sigma^2)$ and $\eta_{1mn} - \eta_{2mn} \sim N(0, s_{1m}^2 + s_{2m}^2)$. Thus, we only need to apply a structural equation model to the difference of indicators and to the difference of engineering attributes (Morikawa 1989). Similarly, for a case of multiple alternatives, we
just need to stack indicators and engineering attributes of different modes, and run structural equation models on the stacked data without considering alternative-specific effects, as we did in section 5.3. In this section, we discuss the extension of the model into a case of a general covariance matrix among alternative-specific latent factors, $\Sigma_2$.

5.4.2. A Case with Single Perception, Three Indicators and Three Alternatives

Let us consider the case of our data whose path diagram can be drawn as follows:
Assuming that the effect of perception is generic across alternatives, i.e. share the same parameter, the mathematical relationship of variables can be denoted as follows:

\[
\begin{align*}
\mathrm{rsp}_{Tn} &= \pi' z_{Tn} + \xi_{Tn} + \eta_{T1n} \\
\mathrm{rsp}_{Rn} &= \pi' z_{Rn} + \xi_{Rn} + \eta_{R1n} \\
\mathrm{rsp}_{In} &= \pi' z_{In} + \xi_{In} + \eta_{I1n} \\
\mathrm{lef}_{Tn}^* &= \alpha_2 \pi' z_{Tn} + \alpha_2 \xi_{Tn} + \eta_{T2n} \\
\mathrm{lef}_{Rn}^* &= \alpha_2 \pi' z_{Rn} + \alpha_2 \xi_{Rn} + \eta_{R2n} \\
\mathrm{lef}_{In}^* &= \alpha_2 \pi' z_{In} + \alpha_2 \xi_{In} + \eta_{I2n} \\
\mathrm{ptb}_{Tn} &= \alpha_3 \pi' z_{Tn} + \alpha_3 \xi_{Tn} + \eta_{T3n} \\
\mathrm{ptb}_{Rn} &= \alpha_3 \pi' z_{Rn} + \alpha_3 \xi_{Rn} + \eta_{R3n} \\
\mathrm{ptb}_{In} &= \alpha_3 \pi' z_{In} + \alpha_3 \xi_{In} + \eta_{I3n}
\end{align*}
\]

where $\xi_{n}$ is now a vector of three latent factors, $(\xi_{Tn}, \xi_{Rn}, \xi_{In})$ with $f(\xi_{n}) \sim N(0, \sigma^2 \Sigma_2)$ and $\Sigma_2 = \begin{pmatrix}
1 & (sym) \\
\omega_2 & \omega_2 \\
\omega_3 & \omega_3
\end{pmatrix}$. This is the most direct extension of the model system described in section 5.3.3. Again, the arrows without origins represent error terms. Each indicator is now composed of three alternative-specific indicators which are correlated according to the following distribution:

\[
\begin{pmatrix}
\mathrm{rsp}_{Tn} \\
\mathrm{rsp}_{Rn} \\
\mathrm{rsp}_{In}
\end{pmatrix} \sim N\left(\begin{pmatrix}
\pi' z_{Tn} \\
\pi' z_{Rn} \\
\pi' z_{In}
\end{pmatrix}, \sigma^2 \begin{pmatrix}
1 & (sym) \\
\omega_2 & \omega_2 \\
\omega_3 & \omega_3
\end{pmatrix} + \begin{pmatrix}
\hat{s}_{T1}^2 & 0 & 0 \\
0 & \hat{s}_{R1}^2 & 0 \\
0 & 0 & \hat{s}_{I1}^2
\end{pmatrix}\right)
\]

and

\[
\begin{pmatrix}
\mathrm{ptb}_{Tn} \\
\mathrm{ptb}_{Rn} \\
\mathrm{ptb}_{In}
\end{pmatrix} \sim N\left(\begin{pmatrix}
\alpha_3 \pi' z_{Tn} \\
\alpha_3 \pi' z_{Rn} \\
\alpha_3 \pi' z_{In}
\end{pmatrix}, \alpha_3^2 \sigma^2 \begin{pmatrix}
1 & (sym) \\
\omega_2 & \omega_2 \\
\omega_3 & \omega_3
\end{pmatrix} + \begin{pmatrix}
\hat{s}_{T3}^2 & 0 & 0 \\
0 & \hat{s}_{R3}^2 & 0 \\
0 & 0 & \hat{s}_{I3}^2
\end{pmatrix}\right)
\]
Indicators of rsp and ptb are also correlated through the covariance matrix $\alpha_3 \sigma^2 \Sigma_2$. For the continuous representation of the level of efforts, the disturbance terms of measurement errors are normalized to have variances of one, i.e.

$$
\begin{bmatrix}
\text{lef}_{\text{Tn}}^* \\
\text{lef}_{\text{Rn}}^* \\
\text{lef}_{\text{Ln}}^*
\end{bmatrix}
\sim N
\begin{bmatrix}
\alpha_2 \pi' z_{Tn} \\
\alpha_2 \pi' z_{Rn} \\
\alpha_2 \pi' z_{Ln}
\end{bmatrix}
\begin{bmatrix}
\alpha_2^2 \sigma^2 & 1 \\
\sigma^2 & \mathbf{w}_{21} & \mathbf{w}_{22} \\
1 & \mathbf{w}_{31} & \mathbf{w}_{32} & \mathbf{w}_{33}
\end{bmatrix}
+ 
\begin{bmatrix}
1 \\
0 \\
0
\end{bmatrix}
.$$  

We can summarize all the above relationships as follows:

$$
\begin{bmatrix}
\text{rsp}_n \\
\text{lef}_n^* \\
\text{ptb}_n
\end{bmatrix}
\sim N
\begin{bmatrix}
\pi' z_n \\
\alpha_2 \pi' z_n \\
\alpha_3 \pi' z_n
\end{bmatrix}
\begin{bmatrix}
\sigma^2 & 1 \\
\sigma^2 & \mathbf{w}_{21} & \mathbf{w}_{22} \\
\sigma^2 & \mathbf{w}_{31} & \mathbf{w}_{32} & \mathbf{w}_{33}
\end{bmatrix}
\otimes
\begin{bmatrix}
1 \\
0 \\
0
\end{bmatrix}
+ S
$$

where $S$ is a diagonal matrix whose diagonal element represents the variance of measurement errors and is one for $\text{lef}^*$, i.e.

$$
S = 
\begin{bmatrix}
\text{s}_{T1}^2 & & & & & & 0 \\
0 & \text{s}_{R1}^2 & & & & & & \\
0 & 0 & \text{s}_{H1}^2 & & & & & \\
0 & 0 & 0 & 1 & & & & \\
0 & 0 & 0 & 0 & 1 & & & \\
0 & 0 & 0 & 0 & 0 & \text{s}_{T3}^2 & & \\
0 & 0 & 0 & 0 & 0 & 0 & \text{s}_{R3}^2 & \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \text{s}_{H3}^2
\end{bmatrix}
$$

In order to ensure that $\Sigma_2$ is positive definite, we estimate the Cholesky-decomposed lower triangular matrix ($\Gamma$) of $\Sigma_2$ instead of estimating $\Sigma_2$ directly, i.e. $\Sigma_2=\Gamma \Gamma'$, where $\gamma_{11}$ is set to one in order to ensure that $\mathbf{w}_1^2 = 1$, i.e.
\[
\Gamma = \begin{bmatrix}
1 & 0 & 0 \\
\gamma_{21} & \gamma_{22} & 0 \\
\gamma_{31} & \gamma_{32} & \gamma_{33}
\end{bmatrix}
\]

The approach taken in the previous section is equivalent to assuming that \( \omega_2^2 = \omega_3^2 = 1 \) and \( \omega_{12} = \omega_{13} = \omega_{23} = 0 \) in \( \Sigma_2 \) or that \( \gamma_{22} = \gamma_{33} = 1 \) and \( \gamma_{12} = \gamma_{13} = \gamma_{23} = 0 \) in \( \Gamma \). Given the normalization, the model is identifiable since we have 14 covariance parameters (\( \sigma^2, \alpha_2, \alpha_3, \Gamma \) and six diagonal elements of \( S \)) and 21 equations in the covariance matrix of 6 indicators of rsp and ptb. It is critical that three indicators of lef do not add information to the covariance matrix since their continuous representations are not observable. We recommend that railroads use continuous indicators for future research.

With the specification of alternative-specific perceptions, we estimate parameters that maximize the following log-likelihood:

\[
L(\theta) = \sum_n \ln P(rsp_n, lef_n, ptb_n | z_n)
\]

\[
P(rsp_n, lef_n, ptb_n | z_n) = \int_{\xi} f(rsp_n | \xi, z_n) P(lef_n | \xi, z_n) f(ptb_n | \xi, z_n) f(\xi) d\xi
\]

we requires the integration of a trivariate normal distribution. Right now, GAUSS allows us to evaluate an integral with arguments of up to three dimensions (INTQUAD3). We found that even for three-dimensional integration, the speed of numerical integration drops noticeably and the estimates are very sensitive to the boundary of integration. For example, if we select 12 points in each dimension, the procedure evaluates 1,728 (= 12^3) points for each observation. The GAUSS program for a trivariate case ran more than two days on an IBM-compatible 486 PC and was still far from a desired point, i.e. showing non-zero gradients.
In order to overcome the numerical problem, we instead employed simulation. The simulation approach estimates parameters by maximizing the following approximation of the log-likelihood:

\[
L^*(\theta) = \sum_n \ln \mathbb{E}(r_{sp_n}, lef_n, ptb_n | z_n)
\]

where

\[
\mathbb{E}(r_{sp_n}, lef_n, ptb_n | z_n) = \frac{1}{R} \sum_{r=1}^{R} f(r_{sp_n}|\xi^{(r)}, z_n) f(ptb_n|\xi^{(r)}, z_n) P(lef_n|\xi^{(r)}, z_n).
\]

where \(\xi\) is simulated from the normal distribution, \(N(0, \sigma^2 \Sigma_2)\), \(R\) times. Actual simulation was performed using the Cholesky-decomposed lower triangular matrix \(\Gamma\). Instead of simulating correlated variates \(\xi_n\) directly, we simulate \(\zeta_n\) from i.i.d. \(N(0, I)\) and multiply it by \(\sigma \Gamma\) to get \(\xi_n\). Since \(\Sigma_2 = \Gamma \Gamma^*\), \(\sigma \Gamma \zeta_n\) has the exactly same distribution as \(\xi_n\).

\[
\mathbb{E}(r_{sp_n}, lef_n, ptb_n | z_n) = \frac{1}{R} \sum_{r=1}^{R} f(r_{sp_n}|\zeta^{(r)}, z_n) f(ptb_n|\zeta^{(r)}, z_n) P(lef_n|\zeta^{(r)}, z_n).
\]

The basic concept of using simulation for the purpose of parameter estimation is summarized in Appendix 6. If we have more than three dimensions, simulation is the only option. Using simulation of 100 draws, we achieved Table 5.4. Most estimates are similar to what we had with the assumption that flexibility perception is not correlated among each mode. The estimated covariance matrix of latent factors shows that intra-mode correlations are indeed very small:

\[
\Sigma_2 = \begin{bmatrix}
1 & 0.048 & 0.167 \\
0.048 & 0.764 & 0.094 \\
0.167 & 0.094 & 0.735
\end{bmatrix}
\]

The log-likelihood was -8,510.48, the \(R^2\) of the \(r_{sp}\) equation was 0.327 and the \(ptb\) equation was 0.206. The log-likelihood improved, whereas the \(R^2\)'s worsened.\(^1\) Factors extracted under the most general assumption has the following statistics:

\(^1\) We estimated a model with restricting \(\Gamma = I\), allowing only the variations in measurement errors. The estimates were similar to the above with the log-likelihood -8,549.65, \(R^2\) in the \(r_{sp}\) equation 0.342 and \(R^2\) the \(ptb\) equation 0.191. The results indicate that most
<table>
<thead>
<tr>
<th></th>
<th>mean</th>
<th>std. dev.</th>
<th>max</th>
<th>min</th>
</tr>
</thead>
<tbody>
<tr>
<td>truck</td>
<td>0.341</td>
<td>0.138</td>
<td>0.562</td>
<td>-0.429</td>
</tr>
<tr>
<td>rail</td>
<td>-0.284</td>
<td>0.284</td>
<td>0.474</td>
<td>-1.321</td>
</tr>
<tr>
<td>intermodal</td>
<td>-0.057</td>
<td>0.164</td>
<td>0.410</td>
<td>-0.647</td>
</tr>
<tr>
<td>total</td>
<td>0.000</td>
<td>0.195</td>
<td>0.562</td>
<td>-1.321</td>
</tr>
</tbody>
</table>

Table 5.4. A Simulated ML of Alternative-specific Perceptions

<table>
<thead>
<tr>
<th></th>
<th>Estimates</th>
<th>t-statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>EDI</td>
<td>-0.100</td>
<td>-0.138</td>
</tr>
<tr>
<td>Transit time</td>
<td>-0.149</td>
<td>-1.214</td>
</tr>
<tr>
<td>Equipment condition</td>
<td>1.377</td>
<td>25.462</td>
</tr>
<tr>
<td>( \gamma_1 )</td>
<td>0.409</td>
<td>12.716</td>
</tr>
<tr>
<td>( \alpha_2 )</td>
<td>1.962</td>
<td>18.165</td>
</tr>
<tr>
<td>( \sigma )  (not ( \sigma^2 ))</td>
<td>0.386</td>
<td>0.229</td>
</tr>
<tr>
<td>( \gamma_{22} ) in ( \Gamma )</td>
<td>0.873</td>
<td>0.248</td>
</tr>
<tr>
<td>( \gamma_{33} ) in ( \Gamma )</td>
<td>0.835</td>
<td>0.256</td>
</tr>
<tr>
<td>( \gamma_{21} ) in ( \Gamma )</td>
<td>0.048</td>
<td>0.003</td>
</tr>
<tr>
<td>( \gamma_{31} ) in ( \Gamma )</td>
<td>0.167</td>
<td>0.012</td>
</tr>
<tr>
<td>( \gamma_{32} ) in ( \Gamma )</td>
<td>0.098</td>
<td>0.027</td>
</tr>
<tr>
<td>( s_{T1} )</td>
<td>0.144</td>
<td>0.963</td>
</tr>
<tr>
<td>( s_{R1} )</td>
<td>0.236</td>
<td>1.092</td>
</tr>
<tr>
<td>( s_{I1} )</td>
<td>0.120</td>
<td>0.681</td>
</tr>
<tr>
<td>( s_{T3} )</td>
<td>0.135</td>
<td>0.883</td>
</tr>
<tr>
<td>( s_{R3} )</td>
<td>0.141</td>
<td>1.022</td>
</tr>
<tr>
<td>( s_{I3} )</td>
<td>0.103</td>
<td>0.627</td>
</tr>
</tbody>
</table>

Improvements in the log-likelihood came from the measurement equations. Since we employ 2,610 data points (= 290 observations * 9 indicators), improvements in the log-likelihood is large, although the elements of \( \Gamma \) are not significant.

<table>
<thead>
<tr>
<th>( \Gamma = I, s_{Tm}^2 = s_{Rm}^2 = s_{Im}^2 \forall m )</th>
<th>( L(b) )</th>
<th>( R^2 ) (rsp)</th>
<th>( R^2 ) (ptb)</th>
</tr>
</thead>
<tbody>
<tr>
<td>-8,805.6</td>
<td>0.377</td>
<td>0.258</td>
<td></td>
</tr>
<tr>
<td>( \Gamma = I )</td>
<td>-8,549.7</td>
<td>0.342</td>
<td>0.191</td>
</tr>
<tr>
<td>no restriction</td>
<td>-8,510.5</td>
<td>0.327</td>
<td>0.206</td>
</tr>
</tbody>
</table>
5.5. Combination of Perception generation on Modal selection

Once we extracted latent factors from indicators, we want to estimate their effects on modal selection. In terms of a freight choice model, we want to know how perceptions of service quality influence modal selection in addition to total logistics costs. The conceptual framework of such combination of perception generation and modal selection can be drawn in Figure 5.1. Prototype models for estimating the above framework were proposed by McFadden (1986) and by Ben-Akiva and Boccara (1987). Morikawa (1989) and Morikawa and Sasaki (1994) applied the method to a binary choice case in his thesis. This research may be the first application of the method to multinomial choice. In the below, we will discuss two estimation approaches: sequential and joint.

Figure 5.1. The Modal selection model with Perception generation

5.5.1. Sequential Estimation

Consider the following relationships that represent the effects of a single perception with the simplifying assumptions which we discussed in the previous section:
\[ u_{in}^* = \beta' x_{in} + \delta x_{in}^* + \epsilon_{in} \quad \text{where } \epsilon_{i} \sim N(0, \sigma_{\epsilon_i}^2) : \text{Random utility} \]

\[ y_{in} = F_{in}(u^*) \quad : \text{Observed choice} \]

\[ x_{in}^* = \pi' z_{in} + \xi_{in} \quad \text{where } \xi_{in} \sim N(0, \sigma^2) \]

\[ h_{imn} = \alpha_m x_{in}^* + \eta_{imn} \quad \text{where } \eta_{imn} \sim N(0, s_{im}^2) \]

where \( \beta \) is a \((K*1)\) vector of parameters, \( K \) is the number of total logistics cost variables, and \( \delta \) is a scalar. Stacking variables over alternatives and changing notations accordingly, we have the following relationship:

\[ u_n^* = x_n \beta + x_n^* \delta + \epsilon_n \quad \text{where } \epsilon_n \sim N(0, \Sigma_1) : \text{Random utility (J*1)} \]

\[ y_n = F(u_n^*) \quad : \text{Observed choice (J*1)} \]

\[ x_n^* = z_n \pi + \xi_n \quad \text{where } \xi_n \sim N(0, \sigma^2 \Sigma_2) : \text{Latent factors (J*1)} \]

\[ h_n = x_n^* \otimes \alpha + \eta_{mn} \quad \text{where } \eta_n \sim N(0, S) : \text{Indicators (JM*1)} \]

where \( \Sigma_1 \) denotes the correlation among \( \epsilon_{in}'s \) (we have assumed until now that \( \Sigma_1 = I \)), \( z_n \) is now a \((J*G)\) vector of engineering attributes that influence the formation of single perception \( x_n^* \), \( \pi \) is a \((G*1)\) vector of its parameter, and \( \alpha \) is a \((M*1)\) vector of \( (\alpha_1, \ldots, \alpha_M)' \) that represent measurement equations. With the above assumptions, we can write down the following joint distributions:

\[
\begin{pmatrix}
  x_n^* \\
  u_n^* \\
  h_n
\end{pmatrix} 
\sim N
\begin{pmatrix}
  z_n \pi \\
  x_n \beta + z_n \pi \delta \\
  z_n \pi \otimes \alpha
\end{pmatrix}

\begin{pmatrix}
  \sigma^2 \Sigma_2 \\
  \delta \sigma^2 \Sigma_2 \\
  \sigma^2 \Sigma_2 \otimes \alpha \delta
\end{pmatrix}

\begin{pmatrix}
  \Sigma_1 \\
  \delta^2 \Sigma_2 + \Sigma_1 \\
  \sigma^2 \Sigma_2 \otimes \alpha + S
\end{pmatrix}

\]
where

\[ m_{1n} = z_n \pi + (\sigma^2 \Sigma_2 \otimes \alpha)'(\sigma^2 \Sigma_2 \otimes \alpha\alpha' + S)'(h_n - z_n \pi \otimes \alpha), \]

\[ m_{2n} = x_n \beta + m_{1n} \delta, \]

\[ \varphi_{11} = (\sigma^2 \Sigma_2) - (\sigma^2 \Sigma_2 \otimes \alpha)'(\sigma^2 \Sigma_2 \otimes \alpha\alpha' + S)'(\sigma^2 \Sigma_2 \otimes \alpha), \]

\[ \varphi_{22} = (\delta^2 \sigma^2 \Sigma_2 + \Sigma_1) - (\sigma^2 \Sigma_2 \otimes \alpha\delta)'(\sigma^2 \Sigma_2 \otimes \alpha\alpha' + S)'(\sigma^2 \Sigma_2 \otimes \alpha\delta) \]

for some \( \varphi_{12} \) which we do not define here for simplicity.

A sequential approach proposes to estimate latent factor models in multinomial choice cases as follow:

1) Use the ML estimation to estimate parameters in the perception generation model (i.e. \( \pi, \alpha, \sigma^2, S, \) and \( \Sigma_2 \)).

2) Calculate \( m_{1n} \) and \( \varphi_{11} \)

3) Run the discrete choice model on \( x_n \) and \( m_{1n} \) to estimate \( \beta, \delta \) and \( \varphi_{22} \).

This approach can be interpreted as an instrumental variable (IV) estimation where the extracted factors (\( m_{1n} \)) work as instrumental variables in the discrete choice model. Note that the covariance matrix of utilities (\( u_n^* \)) is now \( \varphi_{22} \) which will not be an identity matrix even if \( \Sigma_1 \) is an identity matrix. The use of probit that allows a general covariance structure is recommended.

The application of the sequential approach to our data is summarized in Table 5.5. For this result, we used an i.i.d. probit. A program that allows the estimation of a general covariance structure is under development. Standard errors are calculated by using White’s heteroscedasticity-consistent covariance matrix. As we expected, flexibility helped improve log-likelihood. Many coefficients show different estimates from those of the model without flexibility. In particular, discount rate goes down to 35.6 %, which is lower than 45.6 % estimated without flexibility. The Hausman test statistic was 5.761. Since chi-square with ten degrees of freedom (with ten common coefficients) at the 5 % significance
level is 18.31, the Hausman test do not reject the null hypothesis that common parameters in the models with and without flexibility are the same. Also, the log-likelihood ratio (times 2) is 10.38. Since chi-square with one degrees of freedom (with one restriction) at the 5% significance level is 3.84, the Likelihood ratio test reject the null hypothesis that the coefficient of flexibility is zero.

Table 5.5. The Sequential Model

<table>
<thead>
<tr>
<th>Attributes</th>
<th>without Estimates</th>
<th>flexibility t-statistics</th>
<th>with Estimates</th>
<th>flexibility t-statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scale (Transportation costs)</td>
<td>-1.036</td>
<td>-0.656</td>
<td>-1.018</td>
<td>-2.679</td>
</tr>
<tr>
<td>Truck-specific constant</td>
<td>-0.138</td>
<td>-0.690</td>
<td>-0.266</td>
<td>-0.844</td>
</tr>
<tr>
<td>Intermodal-specific constant</td>
<td>1.635</td>
<td>7.068</td>
<td>1.648</td>
<td>2.944</td>
</tr>
<tr>
<td>Value / Corridor ton (truck)</td>
<td>-0.098</td>
<td>-1.363</td>
<td>-0.466</td>
<td>-1.768</td>
</tr>
<tr>
<td>Distance (truck-specific)</td>
<td>0.372</td>
<td>2.908</td>
<td>0.627</td>
<td>1.882</td>
</tr>
<tr>
<td>Delivery time reliability</td>
<td>-0.811</td>
<td>-0.690</td>
<td>-0.905</td>
<td>-1.002</td>
</tr>
<tr>
<td>Flexibility</td>
<td></td>
<td></td>
<td>-0.769</td>
<td>-1.154</td>
</tr>
<tr>
<td>Discount Rate</td>
<td>0.456</td>
<td>3.294</td>
<td>0.356</td>
<td>1.921</td>
</tr>
<tr>
<td>In-transit stock holding costs</td>
<td>1</td>
<td>n.a.</td>
<td>1</td>
<td>n.a.</td>
</tr>
<tr>
<td>Safety stock holding costs</td>
<td>0.169</td>
<td>0.228</td>
<td>0.097</td>
<td>2.854</td>
</tr>
<tr>
<td>Loss and damage costs</td>
<td>1.645</td>
<td>0.517</td>
<td>2.277</td>
<td>0.849</td>
</tr>
<tr>
<td>Log-likelihood at convergence</td>
<td>-142.77</td>
<td></td>
<td>-137.58</td>
<td></td>
</tr>
<tr>
<td>Log-likelihood at origin</td>
<td>-318.60</td>
<td></td>
<td>-318.60</td>
<td></td>
</tr>
<tr>
<td>Rho-square</td>
<td>0.552</td>
<td></td>
<td>0.568</td>
<td></td>
</tr>
<tr>
<td>Adjusted rho-square</td>
<td>0.520</td>
<td></td>
<td>0.534</td>
<td></td>
</tr>
</tbody>
</table>

5.5.2. Joint Estimation

A more desired way than sequential estimation is to combine the processes of perception generation and modal selection. This approach maximizes the following likelihood of observing choices and indicators jointly:

\[ L(\theta) = \sum_{n=1}^{N} ln f(y_n^{RP}, h_n^{RP} | x_n^{RP}, z_n^{RP}) \]

where
\[
    f(y, h|x, z) = \int_* f(y|x, x^*) f(h|x^*) f(x^*) dx^* \\
    \quad = \int_\xi f(y|x, z, \xi) f(h|z, \xi) f(\xi) d\xi
\]

A problem with our data is that we do not observe disaggregate choices. When average shares \(s\) are observed instead of discrete choices \(y\), we can not write down the likelihood for the shares. Instead of maximizing the joint likelihood, we maximize the information value of shares (MDI) and the likelihood of indicators (MLE) integrated over the distribution of latent factors. In this case, our objective function is a combination of MDI and MLE. Although the combined value is difficult to interpret, we define it to be the pseudo-likelihood of share and indicators. We then estimate parameters that maximize the pseudo-likelihood:

\[
    f(s, h|x, z) = \int_\xi f(s|x, z, \xi) f(h|z, \xi) f(\xi) d\xi
\]

We claim that the maximization of the above pseudo-log-likelihood will achieve consistent and asymptotically normal estimators, since it is a convex combination of consistent and asymptotically normal estimators. Both MLE and MDI estimators are consistent and asymptotically normal and their convex combination achieves a maximum of the above pseudo-log-likelihood. Thus, our claim appears to be justified. Again, MDI has a weakness when viewed as an ML method in that it assumes that all shippers make the same number of shipments. Yet, MDI is more robust than MLE. By combining MLE and MDI, we complement weaknesses of both methods and get reliable estimates.

The integral can be evaluated either through numerical integration or through simulation. Numerical integration is truly justifiable for a single latent factor, i.e. when we have to integrate over one dimension. Also, when we have more than three factors, i.e. when we have to integrate over a trivariate distribution, simulation is much faster. Simulation speed can also depend on the distribution of the latent factor. When factors can be simulated in a vector form, simulation is fast due to the characteristic of the GAUSS software as a vector language. On the other hand, if factors have to be simulated element-
by-element (e.g. the acceptance-rejection method), simulation takes long time. In terms of accuracy of estimators, numerical integration is subject to the order of integration, i.e. at how many points a function is evaluated, and simulation is subject to the number of simulation draws, i.e. how many times a function is simulated. Simulation is the only feasible way for alternative-specific perceptual modeling where the dimension of integration increases proportionally to the number of alternatives.

A strength of jointly maximizing the likelihood of observing choices and perceptual indicators is that it increases efficiency by employing all the available data. Moreover, we can exploit the strengths of different data sets by combining multiple sources of data. In addition, certain parameters that could not be identified empirically due to low variability in one set of data may be identifiable when we employ all data. This system approach, of course, has weaknesses as well. The model system becomes too big to be handled in a personal computer. And an error in one of the equations may cause the inconsistency of all parameters in all equations when we estimate the whole system simultaneously. A proper modeling of structural equation is necessary.

The results of our joint estimation is provided in Table 5.6. In order to simplify the estimation, we assumed that $f(\xi_n) \sim N(0, \sigma^2 I)$ where $\xi_n = (\xi_{Tn}, \xi_{Rn}, \xi_{In})$, i.e. we assumed that $\Sigma_2 = I$ (or $\Gamma = I$). We also did not model the heterogeneity of discount rate due to the complexity of the model. The combined log-likelihood at convergence was 8,662.62. Most estimates are similar to the estimates of sequential estimation except the discount rate. The estimate of discount rate becomes 28.5 % which is lower than 35.6 % estimated with the sequential approach, and 45.6 % estimated in the model without flexibility. The effect of flexibility increased significantly, whereas the effect of equipment usability decreased. This is reasonable since equipment usability influences the formation of flexibility perception greatly. Significantly positive effect of EDI in perception generation is also interesting. Standard errors become smaller with joint estimation than with sequential estimation and than with estimating only the choice model.
Table 5.6. The Joint Model

<table>
<thead>
<tr>
<th>Attributes</th>
<th>Sequential Estimates</th>
<th>Estimation t-statistics</th>
<th>Joint Estimates</th>
<th>Estimation t-statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>EDI</td>
<td>-0.104</td>
<td>-4.422</td>
<td>0.454</td>
<td>3.623</td>
</tr>
<tr>
<td>Transit time</td>
<td>-0.094</td>
<td>-25.163</td>
<td>-0.109</td>
<td>-8.675</td>
</tr>
<tr>
<td>Equipment condition</td>
<td>0.977</td>
<td>7.016</td>
<td>0.473</td>
<td>5.818</td>
</tr>
<tr>
<td>$l_1$</td>
<td>0.435</td>
<td>13.645</td>
<td>0.443</td>
<td>12.083</td>
</tr>
<tr>
<td>$\alpha_2$</td>
<td>1.647</td>
<td>7.327</td>
<td>2.271</td>
<td>5.855</td>
</tr>
<tr>
<td>$\alpha_3$</td>
<td>0.254</td>
<td>12.416</td>
<td>0.374</td>
<td>12.031</td>
</tr>
<tr>
<td>$\sigma^2$</td>
<td>0.118</td>
<td>8.106</td>
<td>0.085</td>
<td>19.916</td>
</tr>
<tr>
<td>$s_1^2$</td>
<td>0.049</td>
<td>39.767</td>
<td>0.024</td>
<td>46.821</td>
</tr>
<tr>
<td>$s_3^2$</td>
<td>0.016</td>
<td>88.573</td>
<td>0.017</td>
<td>85.972</td>
</tr>
<tr>
<td>Scale (Transportation costs)</td>
<td>-1.018</td>
<td>-2.679</td>
<td>-0.913</td>
<td>-0.464</td>
</tr>
<tr>
<td>Truck-specific constant</td>
<td>-0.266</td>
<td>-0.844</td>
<td>-0.121</td>
<td>-0.835</td>
</tr>
<tr>
<td>Intermodal-specific constant</td>
<td>1.648</td>
<td>2.944</td>
<td>1.534</td>
<td>3.737</td>
</tr>
<tr>
<td>Value / Corridor ton (truck)</td>
<td>-0.466</td>
<td>-1.768</td>
<td>-0.119</td>
<td>-1.033</td>
</tr>
<tr>
<td>Distance (truck-specific)</td>
<td>0.627</td>
<td>1.882</td>
<td>0.642</td>
<td>3.060</td>
</tr>
<tr>
<td>Delivery time reliability</td>
<td>-0.905</td>
<td>-1.002</td>
<td>-1.955</td>
<td>-1.540</td>
</tr>
<tr>
<td>Equipment Usability</td>
<td>-4.665</td>
<td>-3.285</td>
<td>-2.843</td>
<td>-1.710</td>
</tr>
<tr>
<td>Flexibility</td>
<td>-0.769</td>
<td>-1.154</td>
<td>-2.477</td>
<td>-4.850</td>
</tr>
<tr>
<td>Discount rate</td>
<td>0.356</td>
<td>1.921</td>
<td>0.285</td>
<td>1.798</td>
</tr>
<tr>
<td>In-transit stock holding costs</td>
<td>1</td>
<td>n.a.</td>
<td>1</td>
<td>n.a.</td>
</tr>
<tr>
<td>Safety stock holding costs</td>
<td>0.097</td>
<td>2.854</td>
<td>0.228</td>
<td>0.537</td>
</tr>
<tr>
<td>Loss and damage costs</td>
<td>2.277</td>
<td>0.849</td>
<td>2.114</td>
<td>1.560</td>
</tr>
</tbody>
</table>

Log-likelihood at convergence | -8,648.06            | -8,662.52               |
RP data                       | -137.58              | -149.42                 | a posteriori

The likelihood of RP data is evaluated a posteriori by using the estimated coefficients. In order to accommodate indicators, the choice likelihood in the joint estimation becomes a little worse than that in the sequential estimation. We also found that the full joint estimation is sensitive to the model specification. In our model, $f(s|x,z,\xi)$ is a two-dimensional probability (with a trinomial choice), and $f(h|z,\xi)$ is a multiplication of nine conditional probabilities (with three indicators for each of three alternatives), and thus, the multiplication of nine density functions go to zero quickly, causing a high sensitivity to the model specification.
In order to compare the effects of total logistics costs and service quality, we calculated the average of all variables in the RP model with flexibility (Table 5.7). We further calculate the average of the averages and obtained relative performance by subtracting the global mean from the averages. The relative performance is multiplied by the coefficients estimated from the freight model. All variables are then classified into three groups: (1) mode-inherent preference that sums the effects of truck and intermodal-specific constants, price over corridor tonnage, and distance, (2) total logistics costs that adds freight costs and inventory holding costs (which are discount rate times three inventory holding costs), and (3) service quality that sums the effects of delivery time reliability, equipment usability and flexibility.

Table 5.7. Average Statistics of Total Logistics Costs and Service Quality

<table>
<thead>
<tr>
<th></th>
<th>Truck</th>
<th>Rail</th>
<th>Int’m</th>
<th>Avg.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Price</td>
<td>$1000/ton</td>
<td>7.44</td>
<td>7.44</td>
<td>7.44</td>
</tr>
<tr>
<td>Corridor tonnage</td>
<td>1000 ton</td>
<td>0.70</td>
<td>0.70</td>
<td>0.70</td>
</tr>
<tr>
<td>Distance</td>
<td>1000 mile</td>
<td>0.86</td>
<td>0.86</td>
<td>0.86</td>
</tr>
<tr>
<td>On-time delivery</td>
<td>[0,1]</td>
<td>0.91</td>
<td>0.83</td>
<td>0.90</td>
</tr>
<tr>
<td>Equipment usability</td>
<td>[0,1]</td>
<td>0.95</td>
<td>0.91</td>
<td>0.94</td>
</tr>
<tr>
<td>Flexibility</td>
<td>[-1,1]</td>
<td>0.34</td>
<td>-0.28</td>
<td>-0.06</td>
</tr>
<tr>
<td>Transportation Costs</td>
<td>$100/ton</td>
<td>0.66</td>
<td>0.57</td>
<td>0.65</td>
</tr>
<tr>
<td>In-transit stock holding costs</td>
<td>$100/year</td>
<td>0.46</td>
<td>1.78</td>
<td>1.21</td>
</tr>
<tr>
<td>Safety stock holding costs</td>
<td>$100/year</td>
<td>0.47</td>
<td>1.32</td>
<td>0.62</td>
</tr>
<tr>
<td>Loss and damage costs</td>
<td>$100/year</td>
<td>0.06</td>
<td>0.16</td>
<td>0.23</td>
</tr>
</tbody>
</table>

Table 5.8 shows that truck has disadvantage in terms of total logistics costs but recovers more than its disadvantage in truck-inherent preference and service quality. Rail shows advantage in total logistics costs. But the negative effect from low service quality is more than three times of the cost advantage. Intermodal does okay in both total logistics costs and service quality but suffers from negative inherent preference. This may be due to the fact that intermodal service was relatively unfamiliar to shippers and received negative words-of-mouth in 1986 where the survey for our data was performed. If similar survey were done now, the results might be different.
Table 5.8. Comparison of the Effects of Total Logistics Costs and Service Quality

<table>
<thead>
<tr>
<th></th>
<th>RP Model</th>
<th>Flexibility</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coefficients</td>
<td>truck</td>
</tr>
<tr>
<td>Truck-specific constant</td>
<td>-0.121</td>
<td>0.121</td>
</tr>
<tr>
<td>Intermodal-specific constant</td>
<td>1.534</td>
<td>0.000</td>
</tr>
<tr>
<td>Value / Corridor ton (truck)</td>
<td>-0.119</td>
<td>1.265</td>
</tr>
<tr>
<td>Distance (truck-specific)</td>
<td>0.642</td>
<td>-0.552</td>
</tr>
<tr>
<td>Delivery time reliability</td>
<td>-1.965</td>
<td>0.059</td>
</tr>
<tr>
<td>Equipment usability</td>
<td>-2.843</td>
<td>0.047</td>
</tr>
<tr>
<td>Flexibility</td>
<td>-2.477</td>
<td>0.846</td>
</tr>
<tr>
<td>Freight costs</td>
<td></td>
<td>-0.033</td>
</tr>
<tr>
<td>In-transit stock holding costs</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Safety stock holding costs</td>
<td></td>
<td>0.228</td>
</tr>
<tr>
<td>Loss and damage costs</td>
<td></td>
<td>2.114</td>
</tr>
<tr>
<td>mode-inherent preference</td>
<td></td>
<td></td>
</tr>
<tr>
<td>total logistics costs (r = 0.285)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>service quality</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

5.6. Issues related to Perceptual Modeling

5.6.1. Common Mistakes

A. Asking Importance Rating Directly

Most industry-sponsored market research directly ask respondents to rate the importance of service quality. As discussed in section 2.3.5, this stated importance approach has a lot of weaknesses. In particular, naming of service quality can cause big differences in survey responses. We recommend against the use of this approach.

B. Using Indicators directly as Explanatory variables

Even when shippers are asked to rate the performance of service quality, many practitioners directly employ perceptual indicators in a choice model as if they were

Such practice is not desirable, since it can cause inconsistency of estimates in many different ways. First, if indicators (\( h_n \)) measure the same latent factor, they will be highly correlated. If a choice model include several variables that are highly correlated, a small change in the value of a variable can cause a big change in the estimates (similar to the multi-collinearity problem in the linear model). Second, indicators contain measurement errors. If measurement errors are correlated with explanatory variables, the estimates are inconsistent. This fact can be proved in the linear OLS estimation and is also true in non-linear models. Furthermore, if measurement errors are correlated across alternatives, logit or i.i.d. probit estimates would not be consistent due to the violation of the "Independence of Irrelevant Alternatives" assumption. The magnitude and direction of inconsistency cannot be described in a closed form, since a freight choice model involves a non-linear relationship. The bias will differ situation by situation. The inconsistency will become more severe with a multiple of poorly measured variables than with only one poorly measured variable. Also note that a poorly measured variable can cause inconsistency not only in the estimate of the variable but also in the estimates of all parameters.

In order to compare the results of this approach and our correct approach, we estimated the choice model by including indicators directly (similar to Vieira 1992).\(^2\) This approach showed a little better fit than our approach (-135.35 vs. -137.58). But, the coefficients are not intuitive, as we expected. First, discount rate is estimated 113 %. Similarly, Vieira reported 123 % using RP data, 379 % using SP data, and 241 % using combined data in his estimation. Secondly, the coefficient of equipment usability kept increasing to -8.42. This result indicates that when highly correlated indicators are

\(^2\) The model is not exactly the same as Luiz’s model. I did not include ordering costs and storage costs. I do not have their data and Luiz found them to be insignificant.
employed directly, the Hessian of the likelihood with respect to coefficients is ill-conditioned and can fail to invert. Thirdly, two dummy variables that represent the level of effort to deal with carriers (easy and so-and-so) showed positive coefficients, where they should be negative considering the negative scale.

Finally, if we estimate choice models by employing indicators directly as explanatory variables, it is not clear how one can predict next period's demand. Indicators do not influence choices. Only latent perceptions do. Responses to indicators can vary widely depending on survey design. Prediction becomes difficult since we cannot pre-specify the levels of indicators for the next period.

C. Using Extracted Factors directly into a Choice Model

The method widely used in marketing is to insert the extracted factors directly into a choice model as explanatory variables (for example, Koppleman, Hauser and Tybout 1977). This approach can cause an inconsistency of estimates (McFadden 1986), since it ignores stochasticity inherent in the extracted factors and treats them as if they were fixed variables. Section 5.5.1 shows that an instrumental variable for a latent factor \( m_{1n} \) should be different from the factor extracted from the structural equations \( z_n \pi \). By explicitly writing down stochasticity inherent in a latent factor and conditioning over observed indicators, we could derive a term that corrects the bias.

D. Factor Analysis on Importance Ratings

Some researchers try to extract primary perceptions by running factor analysis on importance ratings rather than on perceptual ratings of attributes. This can bring erroneous results. For example, assume that on-time delivery and billing accuracy are two important factors to most shippers, and satisfaction with payment terms is less important since payment terms are largely determined exogeneously in the market. Also assume that those who are satisfied with billing accuracy are likely to be satisfied with payment terms,
since they both are perceptions on cost-efficiency. If factor analysis were applied on importance ratings, we will get counter-intuitive results that billing accuracy is combined into one factor with on-time delivery, and separate from satisfaction with payment terms.

**E. Exploratory Factor Analysis**

In addition, many researchers rely on exploratory factor analysis that impose no constraints on the factor loading matrix. When the loading matrix is difficult to interpret, they would rotate the loading so that it is easy to interpret the extracted factors. While such a procedure may be useful for obtaining insights at the initial stage, the final analysis should be based on confirmatory factor analysis. There can be an infinite number of factor loading matrices that generate the same covariance matrix of indicators. In order to find desired factors uniquely, researchers should impose a structure.

**5.6.2. Normalization of Indicators**

The estimates of a perceptual model depend only on the covariance matrix of indicators and the expected value of perceptions are set to zero. Thus, the absolute levels of perceptions do not enter the covariance matrix of indicators. The effects of differences in the absolute levels of perceptions on modal selection should be captured by alternative-specific constants. Therefore, before estimating a perceptual model, we normalize all variables (both indicators and explanatory variables) in terms of differences from their means so that the normalized variables have zero means. Such normalization eliminates constant terms and simplifies equations.

A problem is that there are several ways to normalize perceptual indicators and we need to determine the best way. Let us decompose the variation in an indicator as follows:

\[ h_{in} = h_{..} + \delta_n + \delta_i + \delta_{in} \]
where \( \mu \) denotes a mean with respect to the corresponding dimension, where each of \( \delta_{\cdot,n} \), \( \delta_{\cdot,i} \), and \( \delta_{\cdot,in} \) has mean zero. First, we can subtract the individual mean from each response:

(\text{Approach 1}) \quad h_{\cdot,n}^* = h_{\cdot,n} - (h_{\cdot,n} + \delta_{\cdot,n}) = \delta_{\cdot,i} + \delta_{\cdot,in}

or subtract the modal mean from each response:

(\text{Approach 2}) \quad h_{\cdot,n}^* = h_{\cdot,n} - (h_{\cdot,i} + \delta_{\cdot,i}) = \delta_{\cdot,n} + \delta_{\cdot,in}

or subtract the global mean from each response:

(\text{Approach 3}) \quad h_{\cdot,n}^* = h_{\cdot,n} - h_{\cdot,i} = \delta_{\cdot,n} + \delta_{\cdot,i} + \delta_{\cdot,in} \quad \text{and} \quad h_n = \text{vec}(h_{\cdot,n}^*)

After a series of estimation, we concluded that the first approach performs better than the other approaches. The first approach is recommendable since it reduces heterogeneity. We found that a respondent who gives a high rating for a mode is very likely to give a high rating for other modes. That is, there is response bias such that some shippers are very generous and some shippers are not in rating their perceptions. By differentiating with respect to individual mean, we can remove heterogeneity in responses. Following statistics support our hypothesis:

**Means of normalized indicators**

<table>
<thead>
<tr>
<th>Mean ((\text{rsp}_{\cdot,n}^*))</th>
<th>Approach (1)</th>
<th>Approach (2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean ((\text{rsp}_{\cdot,i}^*))</td>
<td>6.004</td>
<td>0.000</td>
</tr>
<tr>
<td>Mean ((\text{rsp}_{\cdot,in}^*))</td>
<td>-4.892</td>
<td>0.000</td>
</tr>
<tr>
<td>Mean ((\text{ptb}_{\cdot,i}^*))</td>
<td>-1.111</td>
<td>0.000</td>
</tr>
<tr>
<td>Mean ((\text{ptb}_{\cdot,in}^*))</td>
<td>8.989</td>
<td>0.000</td>
</tr>
<tr>
<td>Mean ((\text{ptb}_{\cdot,in}^*))</td>
<td>-8.318</td>
<td>0.000</td>
</tr>
<tr>
<td>Mean ((\text{ptb}_{\cdot,in}^*))</td>
<td>-0.671</td>
<td>0.000</td>
</tr>
</tbody>
</table>

**Correlations of normalized indicators with Approach (1)**

<table>
<thead>
<tr>
<th>(\text{r}^*)</th>
<th>(\text{r}^*_n)</th>
<th>(\text{r}^*_i)</th>
<th>(\text{r}^*_in)</th>
<th>(\text{ptb}^*_n)</th>
<th>(\text{ptb}^*_i)</th>
<th>(\text{ptb}^*_in)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\text{r}^*_n)</td>
<td>1.000</td>
<td>-0.4498</td>
<td>-0.4555</td>
<td>0.1664</td>
<td>-0.0163</td>
<td>-0.1925</td>
</tr>
<tr>
<td>(\text{r}^*_i)</td>
<td>1.000</td>
<td>0.5902</td>
<td>-0.5902</td>
<td>-0.1088</td>
<td>0.3624</td>
<td>-0.3153</td>
</tr>
<tr>
<td>(\text{r}^*_in)</td>
<td>1.000</td>
<td>0.0419</td>
<td>0.0419</td>
<td>-0.0419</td>
<td>0.4883</td>
<td>0.4883</td>
</tr>
<tr>
<td>(\text{ptb}^*_n)</td>
<td>-0.0163</td>
<td>-0.6021</td>
<td>-0.6021</td>
<td>1.000</td>
<td>-0.6021</td>
<td>-0.6021</td>
</tr>
<tr>
<td>(\text{ptb}^*_i)</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>(\text{ptb}^*_in)</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
</tbody>
</table>
Correlation of normalized indicators with Approach (2)

<table>
<thead>
<tr>
<th></th>
<th>rsp\textsubscript{Tn}*</th>
<th>rsp\textsubscript{Rn}*</th>
<th>rsp\textsubscript{In}*</th>
<th>ptb\textsubscript{Tn}*</th>
<th>ptb\textsubscript{Rn}*</th>
<th>ptb\textsubscript{In}*</th>
</tr>
</thead>
<tbody>
<tr>
<td>rsp\textsubscript{Tn}*</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rsp\textsubscript{Rn}*</td>
<td>0.3524</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rsp\textsubscript{In}*</td>
<td>0.2422</td>
<td>0.1096</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ptb\textsubscript{Tn}*</td>
<td>0.1824</td>
<td>0.0261</td>
<td>0.1022</td>
<td>1.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ptb\textsubscript{Rn}*</td>
<td>0.2602</td>
<td>0.4126</td>
<td>0.0857</td>
<td>0.1569</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>ptb\textsubscript{In}*</td>
<td>0.2008</td>
<td>0.1002</td>
<td>0.6920</td>
<td>0.1670</td>
<td>0.1473</td>
<td>1.000</td>
</tr>
</tbody>
</table>

Note the positive correlations among all indicators in Approach (2). They are counter-intuitive since they do not capture trade-offs among perceptions about the performance of available modes. In addition, if rsp and ptb reflect one latent factor, their covariance matrix should be symmetric. The covariance matrix between rsp and ptb are more symmetric with (1) than with (2). Thus, we chose to adopt the first approach.

5.6.3. Normal Approximation of Ordinal Indicators

When indicators are ordinal, we may approximate a linear relationship by using normal scores and apply the standard package programmed for continuous indicators. Suppose that we have an indicator rated on a 1 to 3 scale, and suppose that we denote h to be the indicator reported by shippers, h* the continuous representation of latent ratings and x* the latent factor (perception) that influences shippers' ratings. Then, we have:

\[ h^* = \alpha x^* + \eta \]

\[ h = \begin{cases} 1 & \text{if } h_1 \leq h^* \\ 2 & \text{if } h_2 \leq h^* \leq h_1 \\ 3 & \text{if } h^* \leq h_2 \end{cases} \]
Now, denote $n_j$ to be the number of observations in the j-th category. The threshold values are approximated from the marginal distribution of $h^*$ by

$$h_k = \Phi^{-1}\left(\frac{\sum_{j=k+1}^{3} n_j}{N}\right)$$

where $\Phi^{-1}$ is the inverse standard normal distribution function and $N$ is the total number of observations. The normal score $z_k$ is the mean of $h^*$ in the interval corresponding to $h=k$.

For the case where $h^*$ is censored above $h_1$,

$$z_1 = \frac{\phi(h_1)}{1-\Phi(h_1)} = \frac{\phi(h_1)}{n_1/N}$$

where $\phi$ and $\Phi$ are the standard normal density and distribution function, respectively. For the case where $h^*$ lies within the interval $h_2 \leq h^* \leq h_1$,

$$z_2 = \frac{\phi(h_2) \cdot \phi(h_1)}{\Phi(h_1) \cdot \Phi(h_2)} = \frac{\phi(h_2) \cdot \phi(h_1)}{n_2/N}$$

and for the case where $h^*$ is censored below $h_2$,

---

3 Johnson and Kotz, 1970, pp. 81-82
\[ z_3 = -\frac{\phi(h_2)}{\Phi(h_2)} = -\frac{\phi(h_2)}{n_3/N} \]

We can easily see that the mean of the normal scores is zero and its variance is:

\[ \text{Var}(z) = \frac{[\phi(h_1)^2]}{n_1/N} + \frac{[\phi(h_2)\cdot \phi(h_1)]^2}{n_2/N} + \frac{[\phi(h_3)]^2}{n_3/N}, \]

i.e. we avoid the identification problem by imposing zero mean and a fixed variance.

This approach replaces observations of discrete indicators \((h_k)\) by continuous normal scores \((z_k)\) and applies structural equation models to the variance-covariance matrix of normal scores and other indicators. While simple, this approach is not recommended except for obtaining starting values of parameters for the joint ML estimation or for the SP data analysis which we will discuss in the next chapter. Note that the solution is derived separately for each indicator without considering correlation among indicators through \(x^*\). This approach is inefficient since it does not use information about \(h^*\) from the other indicators of \(x^*\). Nor it provides standard errors of the thresholds.

5.6.4. General Model of Multiple Perception With Multinomial Alternatives

Shippers’ perceptions of service quality can be defined in many dimensions (e.g. riskiness, strategic fit, convenience, etc). Our proposed approach can be easily generalized into such a case of multiple perceptions. For the sequential estimation, we can specify the conditional distribution of shippers’ utility given on observed indicators and derive instrumental variables from the conditional distribution. For the joint estimation, we can specify a likelihood as an integral that integrates the multiplication of the conditional probabilities of indicators and choices over the dimension of latent factors. The same principles are applied to a case of multiple perceptions.
Chapter 6.
Stated Preference Data

6.1. Introduction

While actual choice data are difficult to obtain, railroads can collect a large data set by asking shippers to make choices in hypothetical situations. Data obtained in this way is called the stated preference (SP) data. SP data is different from stated importance data in that it does not ask for the importance of attributes directly. It rather asks stated preferences (or hypothetical choices) and tries to infer the importance of attributes by relating preference to pre-specified values of attributes. Appendix 7 discusses different forms to solicit stated preference and their estimation methods, strengths and weaknesses.

Section 6.2 discusses empirical results with SP data, assuming homogeneous shippers and a fixed discount rate. Section 6.3 discusses a model with randomly distributed rates. We note that since SP data is composed of multiple observations from each respondent, the persistency of respondent-specific heterogeneity may occur more significantly in SP data than in RP data. The persistency is called agent effects. Section 6.4 discusses how to estimate a model with such agent effects in SP data. Section 6.5 discusses a model with both random rates and agent effects. Since respondents don’t have to behave as they responded to SP experiments, SP data can contain large response bias. One way to ensure that both the SP and RP data measure the true preference function is to specify that the utility function shares common coefficients in SP and RP data with scale adjustments (Ben-Akiva and Morikawa 1990). We estimate demand response by adopting this approach. The results of the combined model are discussed in section 6.6.
6.2. The SP Model with Fixed Rate

In this section, we discuss how to estimate parameters using SP data. Similarly to the analysis of RP data, we assume that shippers determine total logistics costs and service quality perceptions from attribute levels and that shippers make choices based on the preference formed by logistics costs and service quality perceptions. We estimate parameters that maximize the likelihood of observing stated preferences conditional on pre-specified attribute levels. The conceptual framework can be drawn as follows:

![Diagram of the Stated Preference Model]

Figure 6.1. The Stated Preference Model

Shippers were presented with a pair of alternatives consisting of partial profiles and asked to indicate which one they prefer on a 1 to 9 scale. Figure 6.2 shows an example of SP experiment. In terms of the classification we presented in Appendix 7, the experiment is rating-based, making partial-list (pair-wise) comparison of partially profiled alternatives. Each experiment present trade-offs between two attributes out of nine attributes (transportation mode, freight rate, transit time, on-time performance, loss or damage, usability of equipment, EDI, payment terms, responsiveness, and level of efforts).
Since each pair presents only partial profiles, we assume that the alternatives share the same values for all omitted attributes.

<table>
<thead>
<tr>
<th>Which service offering would you prefer?</th>
<th>Which service offering would you prefer?</th>
</tr>
</thead>
<tbody>
<tr>
<td>10% lower than average rate I pay now</td>
<td>10% higher than average rate I pay now</td>
</tr>
<tr>
<td>80% of shipments arrive when I want them to</td>
<td>90% of shipments arrive when I want them to</td>
</tr>
<tr>
<td>1  2  3  4  5</td>
<td>6  7  8  9</td>
</tr>
<tr>
<td>Strongly prefer left</td>
<td>Strongly prefer right</td>
</tr>
</tbody>
</table>

Fig. 6.2. Example of SP Experiment

If we assume that all shippers have the same fixed discount rate, shippers' utility function at each experiment can be written as follows,

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

where \( \varepsilon_{in}^{SP} \) is i.i.d. normal-distributed, \( i \) denotes either \( L \) (e.g. the left-side alternative) or \( R \) (e.g. the right-side alternative), and \( n \) denotes each observation. We put a superscript "SP" in order to denote the variables of SP data. We will put a superscript "RP" for RP data. Shippers indicated which alternative they prefer on a 1 to 9 scale. If they strongly prefer the right alternative (e.g. if \( U_{Rn} - U_{Ln} \) is greater than a certain threshold, say \( \theta_9 \)), they will answer 9. We model this response process by ordered probit, i.e.

\[ P_n(1) = P(U_{Rn} - U_{Ln} = \beta'(x_{Rn} - x_{Ln}) + (\varepsilon_{Rn} - \varepsilon_{Ln}) \leq \theta_1) = \Phi(\theta_1 - \beta'(x_{Rn} - x_{Ln})) \]

\[ P_n(2) = P(\theta_1 \leq \beta'(x_{Rn} - x_{Ln}) + (\varepsilon_{Rn} - \varepsilon_{Ln}) \leq \theta_2) = \Phi(\theta_2 - \beta'(x_{Rn} - x_{Ln})) - \Phi(\theta_1 - \beta'(x_{Rn} - x_{Ln})) \]
\[ P_n(9) = P(\theta \leq \beta'(x_{Rn} - x_{Ln}) + (\epsilon_{Rn} - \epsilon_{Ln})) \]
\[ = 1 - \Phi(\theta - \beta'(x_{Rn} - x_{Ln})) \]

Fig. 6.3. Distribution of Stated Preference

For thresholds, we can either specify them as fixed, symmetric or variable. Fixed specification employs total 8 thresholds. Symmetric thresholds reduce the number of parameters by setting \( \theta_1 = -\theta_8, \theta_2 = -\theta_7, \theta_3 = -\theta_6, \theta_4 = -\theta_5 \) where \( \theta_8 \geq \theta_7 \geq \theta_6 \geq \theta_5 \geq 0 \). We also assume for notation purpose that \( \theta_0 = -\infty \) and \( \theta_9 = \infty \). [Even if SP responses choose among 10 ratings, we will still need only 4 thresholds by setting the center to zero.]

Variable thresholds specifies that the distribution center depends on socio-characteristics of respondents, i.e. \( \theta_{kn} = \omega_k Z_n + \zeta_{kn} \). Experience shows that variable thresholds improve model fit, but the estimated utility function does not change much.\(^1\) This result implies that different segments of respondents basically share the same shape of a utility function, even if they have different distribution centers (location adjustment). Since our objective is not to show the difference among threshold specifications, we chose a simple symmetric specification. Also note that we do not employ constant in the specification of utility difference. We assume that alternatives do not get extra preference for being placed in the left (or in the right) in SP experiments.

\(^1\) From in-class discussion with Professor Ben-Akiva
If we denote $I(.)$ to be an indicator function, the probability of observing a SP choice ($d_{n}^{SP}$) is:

$$P(d_{n}^{SP} | x_{n}^{SP}, \beta, r) = \prod_{k=1}^{9} P(k | x_{n}^{SP}, \beta, r)^{I(d_{n}^{SP}=k)}$$

or equivalently (we use the fact that the sum of indicators is one for each observation),

$$P(d_{n}^{SP} | x_{n}^{SP}, \beta, r) = \sum_{k=1}^{9} I(d_{n}^{SP} = k) \cdot P(k | x_{n}^{SP}, \beta, r).$$

Assuming that responses are independent of each other, the likelihood of observing SP responses is a multiplication of the above choice probabilities. We thus estimate parameters that maximize the following log-likelihood:

$$L(\beta, \mu) = \sum_{n=1}^{N^{SP}} \ln[P(d_{n}^{SP} | x_{n}^{SP}, \beta, r)]$$

For explanatory variables, transit time, freight rate, reliability of delivery time, loss and damage ratio, and equipment usability were available. Transportation modes were included in about a fifth of the experiments. Using the variables, we constructed total logistics cost variables in such a way that they have the same measurement units as those in the RP model. Transit time, freight rate, loss and damage, and stockout costs were defined relative to the current time and rate of a pre-specified mode (if the experiment specified a mode) or of the primary mode (if the experiment did not). For most shippers, trucking is the primary mode.

In addition, we extracted a latent perception of flexibility from indicators and employed it as an explanatory variable. Unlike the RP data where respondents provide perception ratings, indicators (e.g. satisfaction with responsiveness, satisfaction with payment terms, and level of efforts) are exogenously given to respondents as choice contexts in the SP data. Thus, it is neither feasible nor reasonable to estimate structural equations from exogenously given indicators. However, we can extract a factor if we are willing to assume that the structural equations of the RP data will still hold between
indicators and a latent factor in the SP data. With the assumption, a factor can be obtained as a weighted average of indicators where weights are determined by the measurement equation of the RP data. We first transformed a discrete indicator (i.e. level of efforts) into a continuous indicator by using normal scores (refer to section 5.6.3). Perceptions (x*) are then defined by the following weighted least square method:

\[
\text{if } h = \Lambda \, x^* + \nu \quad \text{where } \nu \sim N(0, \Theta)
\]
\[
\text{then } x^* = (\Lambda' \Theta^{-1} \Lambda)^{-1} (\Lambda' \Theta^{-1} h) \quad \text{where } \Lambda \text{ and } \Theta \text{ are estimated from the RP data.}
\]

Estimation results are summarized in Table 6.1.

Table 6.1. The Fixed Discount Rate SP Model

<table>
<thead>
<tr>
<th>Attributes</th>
<th>RP Estimates</th>
<th>t-statistics</th>
<th>SP Estimates</th>
<th>t-statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Truck-specific constant</td>
<td>-0.266</td>
<td>-0.844</td>
<td>-0.250</td>
<td>-3.042</td>
</tr>
<tr>
<td>Intermodal constant</td>
<td>1.648</td>
<td>2.944</td>
<td>0.065</td>
<td>0.940</td>
</tr>
<tr>
<td>Price (truck-specific)</td>
<td>-0.466</td>
<td>-1.768</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distance(truck-specific)</td>
<td>0.627</td>
<td>1.882</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delivery time reliability</td>
<td>-0.905</td>
<td>-1.002</td>
<td>-3.731</td>
<td>-6.961</td>
</tr>
<tr>
<td>Flexibility</td>
<td>-0.769</td>
<td>-1.154</td>
<td>-0.800</td>
<td>-6.712</td>
</tr>
<tr>
<td>Scale (Transportation costs)</td>
<td>-1.018</td>
<td>-2.679</td>
<td>-1.336</td>
<td>-10.096</td>
</tr>
<tr>
<td>Discount Rate</td>
<td>0.356</td>
<td>1.921</td>
<td>0.178</td>
<td>5.560</td>
</tr>
<tr>
<td>In-transit stock holding costs</td>
<td>1</td>
<td>n.a.</td>
<td>1</td>
<td>n.a.</td>
</tr>
<tr>
<td>Stockout costs</td>
<td>0.097</td>
<td>2.854</td>
<td>0.381</td>
<td>0.571</td>
</tr>
<tr>
<td>Loss and damage costs</td>
<td>2.277</td>
<td>0.849</td>
<td>5.663</td>
<td>6.056</td>
</tr>
<tr>
<td>Threshold 5</td>
<td></td>
<td></td>
<td>0.240</td>
<td>17.421</td>
</tr>
<tr>
<td>Threshold 6</td>
<td></td>
<td></td>
<td>0.515</td>
<td>26.865</td>
</tr>
<tr>
<td>Threshold 7</td>
<td></td>
<td></td>
<td>0.840</td>
<td>35.361</td>
</tr>
<tr>
<td>Threshold 8</td>
<td></td>
<td></td>
<td>1.303</td>
<td>43.684</td>
</tr>
<tr>
<td>Log-likelihood at convergence</td>
<td>-137.58</td>
<td></td>
<td>-3156.94</td>
<td></td>
</tr>
<tr>
<td>Log-likelihood at origin</td>
<td>-318.60</td>
<td></td>
<td>-3317.81</td>
<td></td>
</tr>
<tr>
<td>Rho-square</td>
<td>0.568</td>
<td></td>
<td>0.048</td>
<td></td>
</tr>
<tr>
<td>Adjusted rho-square</td>
<td>0.534</td>
<td></td>
<td>0.045</td>
<td></td>
</tr>
</tbody>
</table>

Since we are using the ML estimation, all t-statistics (including all other tables in this chapter) are calculated by using the inverse of the Hessian matrix. Several coefficients were different from the results of RP data analysis. First, the coefficient of on-time delivery is bigger than that of RP data (3.7 vs. 0.9). On the other hand, the importance of equipment usability is smaller than that of RP data (1.4 vs. 4.7). The results suggests either that shippers may have tried to influence carriers' service design by emphasizing the importance of on-time delivery more than necessary, or that the model might omit some important latent perceptions (e.g. credibility or trustworthiness of carriers, or risk of system failure) that are correlated with on-time delivery. This type of response bias in which respondents try to influence surveyor's policy often occurs in other SP experiments. Similar thing happens with loss and damage attribute (5.7 vs. 2.3). Shippers respond that loss and damage is much more important than what their actual choices suggest. On the other hand, discount rate is estimated at a lower rate (17.8 % vs. 35.6 %) than in RP data.

6.3. The SP Model with Randomly-distributed Rates

As discussed in the analysis of RP data, discount rate may vary among observations. In order to model randomness in discount rate, we will employ the concept of the expected choice probability, as we did in the analysis of RP data. But, let us first add a subscript to our notation in order to model multiple choices from a shipper, i.e. t denotes choice events for each shipper. We also decompose x into w and z where w denotes variables independent of discount rate, and z denotes variables whose coefficients have the same distribution with discount rate. Shippers' utility function can then be written as follows:

$$U^{SP}_{int} = \mu(w^{SP}_{int} + r^{SP}_{nt} z^{SP}_{nt}) + \varepsilon^{SP}_{int}$$

where $i \in \{R, L\}$,

or in terms of a differenced form,

$$U^{SP}_{nt} = \mu(w^{SP}_{nt} + r^{SP}_{nt} z^{SP}_{nt}) + \varepsilon^{SP}_{nt}$$
where  \( U_{nt}^{\text{SP}} = U_{nt}^{\text{SP}} - U_{nt}^{\text{SP}} \),
\( W_{nt}^{\text{SP}} = W_{nt}^{\text{SP}} - W_{nt}^{\text{SP}} \),
\( Z_{nt}^{\text{SP}} = Z_{nt}^{\text{SP}} - Z_{nt}^{\text{SP}} \),
\( \varepsilon_{nt}^{\text{SP}} = \varepsilon_{nt}^{\text{SP}} - \varepsilon_{nt}^{\text{SP}} \sim N(0,1) \), and

\( r_{nt} \) are randomly distributed over the population according to the log-normal distribution, i.e. \( ln r \sim N(\phi, \psi^2) \) where \( \phi \) is \( E(ln r) \) is the expected value of the log of discount rate, and \( \psi \) is the standard error of the log of discount rate. Again, the discount rate has the following density function and statistics:

\[
f(r) = \frac{1}{\psi r \sqrt{2\pi}} \exp\left(-\frac{1}{2} \left( \frac{ln r - \phi}{\psi} \right)^2 \right)
\]

Mean = \( \exp(\phi + \psi^2/2) \)
Mode = \( \exp(\phi - \psi^2) \)
Median = \( \exp(\phi) \)
Variance = \( \exp(2\phi + \psi^2)(\exp(\psi^2)-1) \)

Given a discount rate \( r_{nt} \), the conditional probability of rating \( k \) is:

\[
F(k| W_{nt}^{\text{SP}}, Z_{nt}^{\text{SP}}, r_{nt}; \mu^{\text{SP}}) = \Phi(\theta_{k} - \mu^{\text{SP}}(W_{nt}^{\text{SP}} + r_{nt} Z_{nt}^{\text{SP}})) - \Phi(\theta_{k-1} - \mu^{\text{SP}}(W_{nt}^{\text{SP}} + r_{nt} Z_{nt}^{\text{SP}}))
\]

Choice probability conditional on shipper-specific discount rate is then:

\[
P(d_{nt}^{\text{SP}}| W_{nt}^{\text{SP}}, Z_{nt}^{\text{SP}}, r_{nt}; \mu^{\text{SP}}) = \sum_{k=1}^{g} I(d_{nt}^{\text{SP}} = k) F(k| W_{nt}^{\text{SP}}, Z_{nt}^{\text{SP}}, r_{nt}; \mu^{\text{SP}})
\]

The log-likelihood of observing choices can then be written as follows:

\[
L(\mu^{\text{SP}}, \phi, \psi) = \sum_{n=1}^{N} \sum_{t=1}^{T} \ln \int_{0}^{\infty} P(d_{nt}^{\text{SP}}| W_{nt}^{\text{SP}}, Z_{nt}^{\text{SP}}, r_{nt}; \mu^{\text{SP}}) f(r_{nt}| \phi, \psi) dr_{nt}
\]

Empirical results with random rates are summarized in Table 6.2. Again, the GAUSS-HERMITE quadrature with change of variable of \( y = \frac{ln r - \phi}{\sqrt{2\psi}} \) was used to integrate conditional choice probabilities over the log-normal density. In order to ensure
that $\psi$ is positive, we used the exponential form, $\psi = \exp(\Psi)$. Standard errors for the mean, median and mode are estimated by using the delta method provided in Appendix 5.

Table 6.2. The SP Model with Random Rates

<table>
<thead>
<tr>
<th>Attributes</th>
<th>SP (Fixed Estimates)</th>
<th>Rate t-statistics</th>
<th>SP(Random Estimates)</th>
<th>Rate t-statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Truck-specific constant</td>
<td>-0.250</td>
<td>-3.042</td>
<td>-0.234</td>
<td>2.858</td>
</tr>
<tr>
<td>Intermodal constant</td>
<td>0.065</td>
<td>0.940</td>
<td>0.083</td>
<td>1.205</td>
</tr>
<tr>
<td>Delivery time reliability</td>
<td>-3.731</td>
<td>-6.961</td>
<td>-3.735</td>
<td>-6.964</td>
</tr>
<tr>
<td>Equipment Usability</td>
<td>-1.369</td>
<td>-7.412</td>
<td>-1.345</td>
<td>-7.337</td>
</tr>
<tr>
<td>Flexibility</td>
<td>-0.800</td>
<td>-6.712</td>
<td>-0.793</td>
<td>-6.670</td>
</tr>
<tr>
<td>Scale (Transportation costs)</td>
<td>-1.336</td>
<td>-10.096</td>
<td>-1.342</td>
<td>-9.917</td>
</tr>
<tr>
<td>In-transit stock holding costs</td>
<td>1</td>
<td>n.a.</td>
<td>1</td>
<td>n.a.</td>
</tr>
<tr>
<td>Stockout costs</td>
<td>0.381</td>
<td>0.571</td>
<td>0.381</td>
<td>0.600</td>
</tr>
<tr>
<td>Loss and damage costs</td>
<td>5.663</td>
<td>6.056</td>
<td>5.664</td>
<td>5.685</td>
</tr>
<tr>
<td>Discount rate (agent effects)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fixed rate</td>
<td>0.178</td>
<td>5.560</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\phi$</td>
<td></td>
<td></td>
<td>-1.735</td>
<td>8.607</td>
</tr>
<tr>
<td>$\ln(\psi)$</td>
<td></td>
<td></td>
<td>-0.980</td>
<td>1.417</td>
</tr>
<tr>
<td>$\psi$</td>
<td></td>
<td></td>
<td>0.375</td>
<td>n.a.</td>
</tr>
<tr>
<td>Threshold 5</td>
<td>0.240</td>
<td>17.421</td>
<td>0.241</td>
<td>17.110</td>
</tr>
<tr>
<td>Threshold 6</td>
<td>0.515</td>
<td>26.865</td>
<td>0.517</td>
<td>25.910</td>
</tr>
<tr>
<td>Threshold 7</td>
<td>0.840</td>
<td>35.361</td>
<td>0.843</td>
<td>33.266</td>
</tr>
<tr>
<td>Threshold 8</td>
<td>1.303</td>
<td>43.684</td>
<td>1.310</td>
<td>40.151</td>
</tr>
<tr>
<td>Log-likelihood at convergence</td>
<td>-3156.94</td>
<td></td>
<td>-3156.51</td>
<td></td>
</tr>
<tr>
<td>Log-likelihood at origin</td>
<td>-3317.81</td>
<td></td>
<td>-3317.81</td>
<td></td>
</tr>
<tr>
<td>Rho-square</td>
<td>0.048</td>
<td></td>
<td>0.048</td>
<td></td>
</tr>
<tr>
<td>Adjusted rho-square</td>
<td>0.045</td>
<td></td>
<td>0.044</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Discount Rate (%)</th>
<th>Fixed rate Estimates</th>
<th>Fixed rate std. error</th>
<th>Random rate Estimates</th>
<th>Random rate std. error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean (%)</td>
<td>17.8</td>
<td>3.2</td>
<td>18.9</td>
<td>3.6</td>
</tr>
<tr>
<td>Median</td>
<td>17.8</td>
<td></td>
<td>17.6</td>
<td>5.1</td>
</tr>
<tr>
<td>Mode</td>
<td>17.8</td>
<td></td>
<td>15.3</td>
<td>3.5</td>
</tr>
<tr>
<td>standard deviation</td>
<td>n.a.</td>
<td></td>
<td>7.4</td>
<td>n.a.</td>
</tr>
</tbody>
</table>
The coefficient estimates and the log-likelihood at convergence are almost similar to the fixed rate model. Randomness in discount rate appears small. One possibility is that randomness in discount rate may occur with different characteristics in SP data. In the SP data, each individual shipper responds to several questions. The responses from the same shipper may be influenced by the same discount rate. We explore such possibility next.

6.4. The SP Model with Agent Effects

As discussed in the RP data analysis, if we have multiple observations from the same shipper, the observations may be correlated with each other due to the persistence of shipper-specific discount rates. Since SP data is composed of multiple observations from each shipper, this respondent-persistent heterogeneity, called agent effects, is a much bigger problem in SP data than in RP data. For instance, our data set has ten observations from each shipper. When agent effects exist, the maximum likelihood estimation will still produce consistent estimates, but the estimates will be inefficient since the estimation does not use the information that observations from a same shipper are subject to the same shipper-specific disturbance. Previous research, however, has largely ignored the agent effects in analyzing SP data.

Suppose that a shipper-specific disturbance influences the total utility function. If we denote such disturbance by \( \nu_n \), we can write shippers’ utility difference function as follows:

\[
U_{nt}^{SP} = \mu (w_{nt}^{SP} + rz_{nt}^{SP}) + \nu_n + \varepsilon_{nt}^{SP} \quad \text{where } \nu_n \sim N(0, \sigma^2)
\]

where \( r \) is a fixed discount rate. Given the shipper-specific heterogeneity \( \nu_n \), the conditional probability of each rating \( k \) is:

\[
F(k | w_{nt}^{SP}, z_{nt}^{SP}, \nu_n, \mu^{SP}) = \Phi(\theta_k - \mu^{SP} (w_{nt}^{SP} + rz_{nt}^{SP}) - \nu_n) - \Phi(\theta_{k-1} - \mu^{SP} (w_{nt}^{SP} + rz_{nt}^{SP}) - \nu_n)
\]

and choice probability is:

\[
P(d_{nt}^{SP} | w_{nt}^{SP}, z_{nt}^{SP}, \nu_n, \mu^{SP}) = \sum_{k=1}^{q} I(d_{nt}^{SP} = k) F(k | w_{nt}^{SP}, z_{nt}^{SP}, \nu_n, \mu^{SP}).
\]
The likelihood of observing the whole series of choices $d_{n}^{SP} = (d_{n1}^{SP}, d_{n2}^{SP}, ..., d_{nN}^{SP})$ from each shipper is then:

$$L(\mu^{SP}, \sigma) = \sum_{n=1}^{N^{SP}/10} ln \int_{0}^{\infty} \left[ \prod_{t=1}^{10} P(d_{nt}^{SP} | w_{nt}^{SP}, z_{nt}^{SP}, \nu, \mu^{SP}) \right] f(\nu | \sigma) d\nu$$

where $f(\nu)$ is a normal $(0, \sigma^{2})$ distribution. The GAUSS-HERMITE quadrature can be applied with change of variable of $\nu = \frac{\nu}{\sqrt{2}\sigma}$. In order to ensure that $\sigma$ is positive, we again used the exponential form, i.e. $\sigma = \exp(\lambda)$. Empirical results with agent effects are summarized in Table 6.3. Again, the improvement in fit was small. Randomness due to the shipper-specific disturbance in the utility function does not appear large.

6.5. The SP Model with Both Effects

This section estimates a model with both randomness in discount rates and agent effects. Since both the random rate model and the agent effect model fail to improve the fixed rate model, we should not expect a large improvement from the combined model. Our objective is rather to show the feasibility of the combined model. In specifying the combined model, we can employ two different approaches: an error component model and a discount rate decomposition model. We will discuss both models in the below.

A. Error Component Model

Suppose that a shipper-specific disturbance influences his or her total utility function and that discount rates also vary at each observation. If we denote the disturbance of utilities by $\nu$ and the randomness in discount rate by $r$, we can write a shipper’ utility difference function at a specific choice situation as follows:

$$U_{nt}^{SP} = \mu(w_{nt}^{SP} + r_{nt} z_{nt}^{SP}) + \nu_{n} + \epsilon_{nt}^{SP} \quad \text{where} \ ln(r) \sim N(\phi, \psi^{2}) \ and \ \nu_{n} \sim N(0, \sigma^{2}).$$
Table 6.3. The SP Model with Agent Effects (Error Component Model)

<table>
<thead>
<tr>
<th>Attributes</th>
<th>SP (Fixed Estimates)</th>
<th>Rate t-statistics</th>
<th>SP (Agent Estimates)</th>
<th>Effects t-statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Truck-specific constant</td>
<td>-0.250</td>
<td>-3.042</td>
<td>-0.263</td>
<td>3.411</td>
</tr>
<tr>
<td>Intermodal constant</td>
<td>0.065</td>
<td>0.940</td>
<td>0.068</td>
<td>0.893</td>
</tr>
<tr>
<td>Delivery time reliability</td>
<td>-3.731</td>
<td>-6.961</td>
<td>-3.720</td>
<td>-7.261</td>
</tr>
<tr>
<td>Equipment Usability</td>
<td>-1.369</td>
<td>-7.412</td>
<td>-1.373</td>
<td>-7.985</td>
</tr>
<tr>
<td>Flexibility</td>
<td>-0.800</td>
<td>-6.712</td>
<td>-0.795</td>
<td>-6.224</td>
</tr>
<tr>
<td>Scale (Transportation costs)</td>
<td>-1.336</td>
<td>-10.096</td>
<td>-1.328</td>
<td>-10.663</td>
</tr>
<tr>
<td>In-transit stock holding costs</td>
<td>1</td>
<td>n.a.</td>
<td>1</td>
<td>n.a.</td>
</tr>
<tr>
<td>Stockout costs</td>
<td>0.381</td>
<td>0.571</td>
<td>0.387</td>
<td>0.614</td>
</tr>
<tr>
<td>Loss and damage costs</td>
<td>5.663</td>
<td>6.056</td>
<td>5.188</td>
<td>6.893</td>
</tr>
<tr>
<td>Discount rate</td>
<td>0.178</td>
<td>5.560</td>
<td>0.197</td>
<td>5.971</td>
</tr>
<tr>
<td>Shipper Heterogeneity</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\lambda = \ln(\sigma_v)$</td>
<td></td>
<td></td>
<td>-3.376</td>
<td>0.774</td>
</tr>
<tr>
<td>$\sigma_v$</td>
<td></td>
<td></td>
<td>0.034</td>
<td>n.a.</td>
</tr>
<tr>
<td>Threshold 5</td>
<td>0.240</td>
<td>17.421</td>
<td>0.241</td>
<td>26.617</td>
</tr>
<tr>
<td>Threshold 6</td>
<td>0.515</td>
<td>26.865</td>
<td>0.516</td>
<td>36.684</td>
</tr>
<tr>
<td>Threshold 7</td>
<td>0.840</td>
<td>35.361</td>
<td>0.840</td>
<td>49.257</td>
</tr>
<tr>
<td>Threshold 8</td>
<td>1.003</td>
<td>43.684</td>
<td>1.303</td>
<td>62.980</td>
</tr>
<tr>
<td>Log-likelihood at convergence</td>
<td>-3156.94</td>
<td></td>
<td>-3156.59</td>
<td></td>
</tr>
<tr>
<td>Log-likelihood at origin</td>
<td>-3317.81</td>
<td></td>
<td>-3317.81</td>
<td></td>
</tr>
<tr>
<td>Rho-square</td>
<td>0.048</td>
<td></td>
<td>0.049</td>
<td></td>
</tr>
<tr>
<td>Adjusted rho-square</td>
<td>0.045</td>
<td></td>
<td>0.046</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Discount Rate (%)</th>
<th>Fixed rate model Estimates</th>
<th>std. error</th>
<th>Agent effect model Estimates</th>
<th>std. error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean (%)</td>
<td>17.8</td>
<td>3.2</td>
<td>19.7</td>
<td>3.3</td>
</tr>
<tr>
<td>Median</td>
<td>17.8</td>
<td></td>
<td>19.7</td>
<td></td>
</tr>
<tr>
<td>Mode</td>
<td>17.8</td>
<td></td>
<td>19.7</td>
<td></td>
</tr>
<tr>
<td>standard deviation</td>
<td>n.a.</td>
<td></td>
<td>n.a.</td>
<td></td>
</tr>
</tbody>
</table>
Given the heterogeneity of shipper-specific utility $\nu_n$ and the observation-specific discount rate $r_{nt}$, the conditional probability of each rating $k$ is:

$$F(k \mid w_{nt}^{SP}, z_{nt}^{SP}, r_{nt}, \nu_n, \mu^{SP}) = \Phi(\theta_k - \mu^{SP}(w_{nt}^{SP} + r_{nt}z_{nt}^{SP}) - \nu_n)) - \Phi(\theta_{k-1} - \mu^{SP}(w_{nt}^{SP} + r_{nt}z_{nt}^{SP}) - \nu_n)$$

and choice probability is:

$$P(d_{nt}^{SP} \mid w_{nt}^{SP}, z_{nt}^{SP}, r_{nt}, \nu_n, \mu^{SP}) = \sum_{k=1}^{9} I(d_{nt}^{SP} = k) F(k \mid w_{nt}^{SP}, z_{nt}^{SP}, r_{nt}, \nu_n, \mu^{SP})$$

Conditional on the heterogeneity of shipper-specific utility $\nu_n$, the likelihood of observing a choice $d_{nt}^{SP}$ is then:

$$l_{nt}(d_{nt}^{SP} \mid w_{nt}^{SP}, z_{nt}^{SP}, \nu_n, \mu^{SP}, \phi, \psi) = \int_{0}^{\infty} P(d_{nt}^{SP} \mid w_{nt}^{SP}, z_{nt}^{SP}, \nu_n, r, \mu^{SP}) f(r \mid \phi, \psi) dr$$

where $f(r)$ is a log-normal ($\phi, \psi$) distribution. The likelihood of observing the whole series of choices $d_{n}^{SP} = (d_{n1}^{SP}, d_{n2}^{SP}, \ldots, d_{n10}^{SP})$ from each shipper is then:

$$l_{n}(d_{n}^{SP} \mid \mu^{SP}, \phi, \psi, \sigma_{v}) = \int_{0}^{\infty} \prod_{i=1}^{10} l_{nt}(d_{nt}^{SP} \mid w_{nt}^{SP}, z_{nt}^{SP}, \nu, \mu^{SP}, \phi, \psi) f(v \mid \sigma_{v}) dv$$

where $f(v)$ is a normal ($0, \sigma_{v}^2$) distribution. The likelihood of observing the whole series of choices $d_{n}^{SP} = (d_{n1}^{SP}, d_{n2}^{SP}, \ldots, d_{n10}^{SP})$ from each shipper is then:

$$L(\mu^{SP}, \phi, \psi, \sigma_{v}) = \sum_{n=1}^{N_{SP}/10} ln l_{n}(d_{n}^{SP} \mid \mu^{SP}, \phi, \psi, \sigma_{v})$$

which has the following double integral form:

$$\sum_{n=1}^{N_{SP}/10} ln \int_{0}^{\infty} \left[ \prod_{i=1}^{10} \left( \int_{0}^{\infty} P(d_{nt}^{SP} \mid w_{nt}^{SP}, z_{nt}^{SP}, \nu, r, \mu^{SP}, \phi, \psi) f(r \mid \phi, \psi) dr \right) \right] f(v \mid \sigma_{v}) dv.$$
Table 6.4. The SP Model with Combined Effects (Error Component Model)

<table>
<thead>
<tr>
<th>Attributes</th>
<th>SP (Fixed Estimates)</th>
<th>Rate t-statistics</th>
<th>SP (Both Estimates)</th>
<th>Effects t-statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Truck-specific constant</td>
<td>-0.250</td>
<td>-3.042</td>
<td>-0.221</td>
<td>2.907</td>
</tr>
<tr>
<td>Intermodal constant</td>
<td>0.065</td>
<td>0.940</td>
<td>0.093</td>
<td>1.218</td>
</tr>
<tr>
<td>Delivery time reliability</td>
<td>-3.731</td>
<td>-6.961</td>
<td>-3.796</td>
<td>-7.454</td>
</tr>
<tr>
<td>Equipment Usability</td>
<td>-1.369</td>
<td>-7.412</td>
<td>-1.353</td>
<td>-7.966</td>
</tr>
<tr>
<td>Flexibility</td>
<td>-0.800</td>
<td>-6.712</td>
<td>-0.778</td>
<td>-6.176</td>
</tr>
<tr>
<td>Scale (Transportation costs)</td>
<td>-1.336</td>
<td>-10.096</td>
<td>-1.343</td>
<td>-10.591</td>
</tr>
<tr>
<td>In-transit stock holding costs</td>
<td>1</td>
<td>n.a.</td>
<td>1</td>
<td>n.a.</td>
</tr>
<tr>
<td>Stockout costs</td>
<td>0.381</td>
<td>0.571</td>
<td>0.315</td>
<td>0.463</td>
</tr>
<tr>
<td>Loss and damage costs</td>
<td>5.663</td>
<td>6.056</td>
<td>5.568</td>
<td>6.026</td>
</tr>
<tr>
<td>Discount rate</td>
<td>0.178</td>
<td>5.560</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\phi$</td>
<td></td>
<td>-1.705</td>
<td></td>
<td>8.612</td>
</tr>
<tr>
<td>$ln(\psi_1)$</td>
<td></td>
<td>-1.468</td>
<td></td>
<td>0.973</td>
</tr>
<tr>
<td>$(\psi_1)$</td>
<td></td>
<td>0.230</td>
<td></td>
<td>n.a.</td>
</tr>
<tr>
<td>Shipper Heterogeneity</td>
<td></td>
<td>$ln(\sigma_v)$</td>
<td>-3.174</td>
<td>1.058</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$(\sigma_v)$</td>
<td>0.034</td>
<td>n.a.</td>
</tr>
<tr>
<td>Threshold 5</td>
<td>0.240</td>
<td>17.421</td>
<td>0.240</td>
<td>26.070</td>
</tr>
<tr>
<td>Threshold 6</td>
<td>0.515</td>
<td>26.865</td>
<td>0.516</td>
<td>33.905</td>
</tr>
<tr>
<td>Threshold 7</td>
<td>0.840</td>
<td>35.361</td>
<td>0.841</td>
<td>45.980</td>
</tr>
<tr>
<td>Threshold 8</td>
<td>1.303</td>
<td>43.684</td>
<td>1.305</td>
<td>54.905</td>
</tr>
<tr>
<td>Log-likelihood at convergence</td>
<td>-3156.94</td>
<td></td>
<td>-3156.54</td>
<td></td>
</tr>
<tr>
<td>Log-likelihood at origin</td>
<td>-3317.81</td>
<td></td>
<td>-3317.81</td>
<td></td>
</tr>
<tr>
<td>Rho-square</td>
<td>0.048</td>
<td></td>
<td>0.049</td>
<td></td>
</tr>
<tr>
<td>Adjusted rho-square</td>
<td>0.045</td>
<td></td>
<td></td>
<td>0.046</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Discount Rate (%)</th>
<th>Fixed rate Estimates</th>
<th>std. error</th>
<th>Error component Estimates</th>
<th>std. error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean (%)</td>
<td>17.8</td>
<td>3.2</td>
<td>18.7</td>
<td>3.6</td>
</tr>
<tr>
<td>Median</td>
<td>17.8</td>
<td></td>
<td>18.2</td>
<td>5.3</td>
</tr>
<tr>
<td>Mode</td>
<td>17.8</td>
<td></td>
<td>17.2</td>
<td>3.3</td>
</tr>
<tr>
<td>standard deviation</td>
<td>n.a.</td>
<td></td>
<td>4.4</td>
<td>n.a.</td>
</tr>
</tbody>
</table>
Again, the GAUSS-HERMITE quadrature is applied with proper changes of variables. With the double integral, computation took much longer time than with a case of single integral. We also estimated the model with using simulation for the evaluation of the double integral. Computation time of simulation depends on the number of simulation draws, whereas computation time of numerical integration depends on the number of evaluation points. Overall, it was not clear which method takes shorter time. Numerical integration with 4 quadrature points achieved as good estimates as with 12 points. Simulation with 20 draws resulted in the same estimates as one with 100 draws. In order to ensure that $\psi$ and $\sigma_v$ are positive, exponential forms are again used. Empirical results are summarized in Table 6.4. As expected, the improvement in fit is small.

\section*{B. Discount rate Decomposition Model}

Instead of assuming the heterogeneity in shippers' total utilities, we may assume that only discount rate is heterogeneous. With this assumption, discount rate at each choice occasion is decomposed into two terms: shipper-specific discount rate and occasion-specific discount rate. Thus, let us assume that

\[ r_{nt} = \exp(\rho_n + \rho_{nt}) \]

where $\rho_n \sim N(0, \psi_1^2)$ and $\rho_{nt} \sim N(\phi, \psi_2^2)$ are independent of each other. Since the sum of two normal variates is also normally distributed, $r_{nt}$ is log-normally distributed. We have to set the mean (i.e., the location parameter) of $\rho_n$ to zero in order to identify the parameters.

Given the shipper-specific discount rate $\rho_n$ and choice occasion-specific discount rate $\rho_{nt}$, the conditional probability of each rating $k$ is:

\[ F(k \mid w_{nt}^{sp}, z_{nt}^{sp}, \rho_n, \rho_{nt}; \mu^{sp}) = \Phi(\theta_k - \mu^{sp}(w_{nt}^{sp} + \exp(\rho_n + \rho_{nt})z_{nt}^{sp})) \]

\[ - \Phi(\theta_{k-1} - \mu^{sp}(w_{nt}^{sp} + \exp(\rho_n + \rho_{nt})z_{nt}^{sp})) \]
and choice probability is:

\[ P(d_{nt}^{SP} \mid w_{nt}^{SP}, z_{nt}^{SP}, \rho_{nt}; \mu^{SP}) = \sum_{k=1}^{9} I(d_{nt}^{SP} = k) F(k \mid w_{nt}^{SP}, z_{nt}^{SP}, \rho_{nt}; \rho_{n}, \mu^{SP}) \].

Conditional on the heterogeneity of shipper-specific utility \( \rho_{n} \), the likelihood of observing a choice \( d_{nt}^{SP} \) is then:

\[ l_{nt}(d_{nt}^{SP} \mid w_{nt}^{SP}, z_{nt}^{SP}, \rho_{n}; \mu^{SP}, \phi, \psi_{1}) = \int_{-\infty}^{\infty} P(d_{nt}^{SP} \mid w_{nt}^{SP}, z_{nt}^{SP}, \rho_{n}; \rho_{1}; \mu^{SP}) f(\rho_{1} \mid \phi, \psi_{1}) d\rho_{1} \]

where \( f(\rho_{1}) \) is a normal (\( \phi, \psi_{1} \)) distribution. The likelihood of observing the whole series of choices \( d_{n}^{SP} = (d_{n1}^{SP}, d_{n2}^{SP}, ..., d_{n10}^{SP}) \) from each shipper is then:

\[ l_{n}(d_{n}^{SP} \mid \mu^{SP}, \phi, \psi_{1}, \psi_{2}) = \int_{-\infty}^{\infty} \left[ \prod_{i=1}^{10} l_{nt}(d_{nt}^{SP} \mid w_{nt}^{SP}, z_{nt}^{SP}, \rho_{2}; \mu^{SP}, \phi, \psi_{1}) \right] f(\rho_{2} \mid \psi_{2}) d\rho_{2} \]

where \( f(\rho_{2}) \) is a normal (0, \( \psi_{2} \)) distribution. The likelihood of observing the whole series of choices \( d_{n}^{SP} = (d_{n1}^{SP}, d_{n2}^{SP}, ..., d_{n10}^{SP}) \) from each shipper is then:

\[ L(\mu^{SP}, \phi, \psi_{1}, \psi_{2}) = \sum_{n=1}^{N_{SP}/10} ln l_{n}(d_{n}^{SP} \mid \mu^{SP}, \phi, \psi_{1}, \psi_{2}). \]

Empirical results with this approach are summarized in Table 6.5. The t-statistics are based on the inverse of the Hessian matrix. The GAUSS-HERMITE quadrature is called twice to evaluate the double integral. In order to ensure that \( \psi_{1} \) and \( \psi_{2} \) are positive, we again used the exponential forms. As before, the improvement in fit was small.
Table 6.5. The SP Model with Combined Effects (Discount rate Decomposition Model)

<table>
<thead>
<tr>
<th>Attributes</th>
<th>SP (Fixed Estimates)</th>
<th>Rate t-statistics</th>
<th>SP(Both Estimates)</th>
<th>Effects t-statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Truck-specific constant</td>
<td>-0.250</td>
<td>-3.042</td>
<td>-0.232</td>
<td>3.055</td>
</tr>
<tr>
<td>Intermodal constant</td>
<td>0.065</td>
<td>0.940</td>
<td>0.086</td>
<td>1.149</td>
</tr>
<tr>
<td>Delivery time reliability</td>
<td>-3.731</td>
<td>-6.961</td>
<td>-3.782</td>
<td>-7.443</td>
</tr>
<tr>
<td>Equipment Usability</td>
<td>-1.369</td>
<td>-7.412</td>
<td>-1.357</td>
<td>-7.983</td>
</tr>
<tr>
<td>Flexibility</td>
<td>-0.800</td>
<td>-6.712</td>
<td>-0.809</td>
<td>-6.414</td>
</tr>
<tr>
<td>Scale (Transportation costs)</td>
<td>-1.336</td>
<td>-10.096</td>
<td>-1.346</td>
<td>-10.727</td>
</tr>
<tr>
<td>In-transit stock holding costs</td>
<td>1</td>
<td>n.a.</td>
<td>1</td>
<td>n.a.</td>
</tr>
<tr>
<td>Stockout costs</td>
<td>0.381</td>
<td>0.571</td>
<td>0.041</td>
<td>0.062</td>
</tr>
<tr>
<td>Loss and damage costs</td>
<td>5.663</td>
<td>6.056</td>
<td>5.263</td>
<td>6.251</td>
</tr>
<tr>
<td>Discount rate</td>
<td>0.178</td>
<td>5.560</td>
<td>-1.643</td>
<td>8.383</td>
</tr>
<tr>
<td>( \phi )</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \ln(\psi_1) )</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \psi_1 )</td>
<td>-5.730</td>
<td>0.003</td>
<td>0.003</td>
<td>n.a.</td>
</tr>
<tr>
<td>Shipper Heterogeneity</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \ln(\psi_2) )</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \psi_2 )</td>
<td>-1.092</td>
<td>0.336</td>
<td>0.336</td>
<td>n.a.</td>
</tr>
<tr>
<td>Threshold 5</td>
<td>0.240</td>
<td>17.421</td>
<td>0.241</td>
<td>26.671</td>
</tr>
<tr>
<td>Threshold 6</td>
<td>0.515</td>
<td>26.865</td>
<td>0.517</td>
<td>35.358</td>
</tr>
<tr>
<td>Threshold 7</td>
<td>0.840</td>
<td>35.361</td>
<td>0.843</td>
<td>48.410</td>
</tr>
<tr>
<td>Threshold 8</td>
<td>1.303</td>
<td>43.684</td>
<td>1.309</td>
<td>61.439</td>
</tr>
<tr>
<td>Log-likelihood at convergence</td>
<td>-3156.94</td>
<td></td>
<td>-3155.76</td>
<td></td>
</tr>
<tr>
<td>Log-likelihood at origin</td>
<td>-3317.81</td>
<td></td>
<td>-3317.81</td>
<td></td>
</tr>
<tr>
<td>Rho-square</td>
<td>0.048</td>
<td></td>
<td>0.049</td>
<td></td>
</tr>
<tr>
<td>Adjusted rho-square</td>
<td>0.045</td>
<td></td>
<td>0.046</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Discount Rate (%)</th>
<th>Fixed rate model Estimates</th>
<th>std. error</th>
<th>Error component Estimates</th>
<th>std. error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean (%)</td>
<td>17.8</td>
<td>3.2</td>
<td>20.5</td>
<td>n.a.</td>
</tr>
<tr>
<td>Median</td>
<td>17.8</td>
<td></td>
<td>19.3</td>
<td>n.a.</td>
</tr>
<tr>
<td>Mode</td>
<td>17.8</td>
<td></td>
<td>17.3</td>
<td>n.a.</td>
</tr>
<tr>
<td>standard deviation</td>
<td>n.a.</td>
<td></td>
<td>7.1</td>
<td>n.a.</td>
</tr>
</tbody>
</table>

179
6.6. The Combined Model of RP and SP

So far, we have analyzed RP and SP data separately. Since both data are collected from the same shippers, they will reflect the same preference function. In order to ensure such property, we can estimate parameters by maximizing the likelihood of observing RP and SP data jointly. By combining both data, we may be able to correct response variability inherent in the SP data, while ensuring external validity in the RP data. The conceptual framework is provided below:

![Diagram](image)

Figure 6.4. The Combined Model of RP and SP

We adopt the MDI estimation for analyzing RP data and the ML estimation for analyzing SP data. This practice is defended in that both estimation methods produce consistent and asymptotically normal estimators, that a convex combination of consistent parameters is also consistent, and that a convex combination of asymptotically normal parameters is also asymptotically normal. Note that the same claim cannot be made for asymptotic efficiency. The MDI estimator is not efficient and we sacrifice the efficiency for the case of estimation.
Let us first consider the case where choice probabilities of multiple observations from the same shipper are independent of each other. As before, we decompose $x$ into $w$ and $z$ depending on whether its coefficients show the same distribution as discount rate. We also denote $\pi$ to be explanatory variables that influence actual choice behaviors but is not included in the SP data, and $g$ to be variables that affect SP responses but do not affect actual choices. Then, we can write the utility of an individual $n$ for an alternative $i$ at the $t$-th observation as follows:

$$U_{nt}^{RP} = \mu(w_{nt}^{RP} + r z_{nt}^{RP} + \alpha \pi_{nt}^{RP}) + \epsilon_{nt}^{RP}$$

$$U_{nt}^{SP} = \mu(w_{nt}^{SP} + r z_{nt}^{SP} + \gamma g_{nt}^{SP}) + \epsilon_{nt}^{SP}$$

where $\alpha$ and $\gamma$ are coefficients of $\pi$ and $g$, respectively. We will also continue to denote coefficients associated with $x$ (and thus with $w$ and $z$) by $\beta$.

In order to ensure that both data reflect the same true preference function of shippers, we force RP and SP data to share the same coefficients. In forcing the same coefficients, we have to consider that SP data have different variability from RP data, i.e.,

$$\text{Var}(\epsilon_{nt}^{SP}) = \kappa^2 \text{Var}(\epsilon_{nt}^{RP})$$

If we denote that $\mu^{RP} = \mu$, then $\mu^{SP} = \mu/\kappa$. It is likely that $\kappa \geq 1$ due to the response bias discussed in Appendix 7. In addition, we normalize the variance of $\epsilon^{RP}$ to be 1, since utility functions are scale invariant. We also assume that the disturbance terms of RP and SP data, e.g. $\epsilon^{RP}$ and $\epsilon^{SP}$, are independent of each other, i.e.

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

As we have seen in previous chapters, $P(s^{RF})$ and $P(d^{SP})$ are cumulative distribution functions of normal or Gumbel disturbances that represent random utility components. In
this thesis, we assume that they are normally distributed. With the normal assumption, choice probability in the RP data is defined by:

\[
F(l| w_{nt}^{RP}, z_{nt}^{RP}, \pi_{nt}^{RP}) = P(U_{1nt}^{RP} \geq U_{2nt}^{RP} \text{ and } U_{1nt}^{RP} \geq U_{3nt}^{RP})
\]

\[
= \int_{-\infty}^{\mu_k(x_{nt}^{RP} - \mu_{nt}^{RP} + \alpha_{nt}^{RP} - \xi_{nt}^{RP})} \int_{-\infty}^{\mu_k(x_{nt}^{RP} - \mu_{nt}^{RP} + \alpha_{nt}^{RP} - \xi_{nt}^{RP})} f(e_{2nt}^{RP} - e_{1nt}^{RP}, e_{3nt}^{RP} - e_{1nt}^{RP}, \Omega) d(e_{2nt}^{RP} - e_{1nt}^{RP}) d(e_{3nt}^{RP} - e_{1nt}^{RP})
\]

where \( \begin{bmatrix} e_{2nt}^{RP} - e_{1nt}^{RP} \\ e_{3nt}^{RP} - e_{1nt}^{RP} \end{bmatrix} \sim N(0, \Omega = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}) \)

and choice probability in the SP data is defined by:

\[
P(k| w_{nt}^{SP}, z_{nt}^{SP}, g_{nt}^{SP}) = \Phi(\theta_k - \frac{1}{\kappa}(w_{nt}^{SP} + r z_{nt}^{SP} + \gamma g_{nt}^{SP})) - \Phi(\theta_{k-1} - \frac{1}{\kappa}(w_{nt}^{SP} + r z_{nt}^{SP} + \gamma g_{nt}^{SP}))
\]

where \( x_{nt}^{SP} = x_{nt}^{SP} - x_{nt}^{SP} \) and \( g_{nt}^{SP} = g_{nt}^{SP} - g_{nt}^{SP} \). We also impose the symmetric assumption of thresholds. The likelihood of observing actual shares and SP responses can be written as:

\[
L(\mu, \alpha, r, \gamma, \kappa) = \sum_{n=1}^{N_{RP}/2} \sum_{i=1}^{2} \sum_{t=1}^{3} s_{nt}^{RP} \ln F(i| \mu, r, \alpha) + \sum_{n=1}^{N_{SP}/10} \sum_{t=1}^{10} d_{nt}^{SP} \ln P(d_{nt}^{SP} | \mu, r, \gamma, \kappa)
\]

where \( N_{RP} \) and \( N_{SP} \) are the number of observations; \( s_{nt}^{RP} \) is the observed shares in RP; \( d_{nt}^{RP} \in A_{RP} = \{ \text{truck, intermodal, and rail} \} \) denote alternatives available in actual shipments and \( d_{nt}^{SP} \in A_{SP} = \{ 1, 2, \ldots, 9 \} \) denote alternatives available in hypothetical SP experiments.

Parameters can be estimated either sequentially or simultaneously. A sequential approach allows us to use existing discrete choice softwares by estimating parameters sequentially (Ben-Akiva and Morikawa 1990). That is,
1) estimate $\beta, \alpha$ by using only the RP data  
2) construct instrumental variable $\beta'x^{sp}$ by using $\beta$ estimated from step 1  
3) estimate $\gamma, \kappa$ by using the SP data and the instrumental variable $\beta'x^{sp}$  
4) repeat step 1, 2 and 3 with different instrumental variables, if necessary

Several measures are taken to reduce response bias of SP. First, earlier results (Ben-Akiva and Morikawa 1990) show that alternative-specific constants estimated from RP data are very different from those estimated from SP data. We, thus, estimated mode-specific constants separately for RP and SP data.

Secondly, SP responses often have a bias that frequently-used modes in actual shipments often receive high preference over infrequently-used modes. In order to capture the degree of such response bias, we included modal shares in RP data as an explanatory variable of SP data so that SP utility increases monotonically with modal shares, i.e. we include $s_{int}$ of the largest corridor into $g_{int}$. Since this bias correction term was not included in previous estimations of SP data, we re-estimate the fixed rate SP model with the term. Its results are provided in the second and third columns of Table 6.6. The coefficient of the bias correction term is significant and brought an improvement in the log-likelihood.

Thirdly, for SP data, latent perception of flexibility is constructed by applying the weighted least square method to exogeneously given indicators. For RP data, latent perception of flexibility is constructed by employing structural equations on engineering attributes and reported perceptual indicators.

The results are summarized in Table 6.6. The t-statistics are based on the Hessian matrix. As parameter estimates try to fit both RP and SP data, the likelihood at convergence decreased somewhat with both data. In addition, $\kappa$ is greater than 1 with high statistical significance, indicating that SP data show larger response bias than RP data. In addition, by employing the Likelihood ratio test, we can test whether the RP and SP models share the same coefficients subject to the scale variation. The unrestricted likelihood is the sum of the convergence likelihoods with each data (in our case, -3292.42
The restricted likelihood is the convergence likelihood of the combined model (in our case, $-3296.41 = -141.32 -3155.09$). The likelihood ratio, which is twice the difference (e.g. 7.98), is distributed according to the chi-square distribution with six degrees of freedom (since the number of shared coefficients are six). Since the test statistic is 12.59 with 5% of significance level, the Likelihood ratio test does not reject the null hypothesis that RP and SP data share the same coefficients.

Table 6.6. The Combined Model with a Fixed Discount Rate

<table>
<thead>
<tr>
<th>Attributes</th>
<th>SP Estimates</th>
<th>Model t-statistics</th>
<th>Combined Estimates</th>
<th>Model t-statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Truck-specific constant (RP)</td>
<td>-0.256</td>
<td>-0.776</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intermodal-specific (RP)</td>
<td>1.669</td>
<td>4.588</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Truck-specific constant (SP)</td>
<td>-0.221</td>
<td>-2.682</td>
<td>-0.201</td>
<td>-2.557</td>
</tr>
<tr>
<td>Intermodal-specific (SP)</td>
<td>0.077</td>
<td>1.101</td>
<td>0.088</td>
<td>1.380</td>
</tr>
<tr>
<td>Value / Corridor ton (truck)</td>
<td></td>
<td>-0.466</td>
<td>-1.865</td>
<td></td>
</tr>
<tr>
<td>Distance(truck-specific)</td>
<td></td>
<td>0.602</td>
<td>2.464</td>
<td></td>
</tr>
<tr>
<td>Delivery time reliability</td>
<td>-3.777</td>
<td>-6.926</td>
<td>-3.615</td>
<td>-7.803</td>
</tr>
<tr>
<td>Equipment usability</td>
<td>-1.364</td>
<td>-7.315</td>
<td>-1.3**</td>
<td>-7.795</td>
</tr>
<tr>
<td>Flexibility</td>
<td>-0.805</td>
<td>-6.666</td>
<td>-0.820</td>
<td>-7.430</td>
</tr>
<tr>
<td>RP Modal share (in SP)</td>
<td>-0.122</td>
<td>-1.884</td>
<td>-0.118</td>
<td>-1.735</td>
</tr>
<tr>
<td>Variability ($\kappa$ in SP response)</td>
<td>1.297</td>
<td>4.737</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scale (Transportation costs)</td>
<td>-1.324</td>
<td>-9.998</td>
<td>-1.029</td>
<td>-4.592</td>
</tr>
<tr>
<td>Discount rate</td>
<td>0.197</td>
<td>5.946</td>
<td>0.202</td>
<td>6.601</td>
</tr>
<tr>
<td>In-transit stock holding costs</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>n.a</td>
</tr>
<tr>
<td>Safety stock holding costs</td>
<td>0.003</td>
<td>0.004</td>
<td>0.029</td>
<td>0.116</td>
</tr>
<tr>
<td>Loss and damage costs</td>
<td>5.194</td>
<td>6.522</td>
<td>5.049</td>
<td>6.431</td>
</tr>
<tr>
<td>Threshold 5</td>
<td>0.239</td>
<td>17.406</td>
<td>0.240</td>
<td>17.466</td>
</tr>
<tr>
<td>Threshold 6</td>
<td>0.513</td>
<td>26.835</td>
<td>0.514</td>
<td>26.139</td>
</tr>
<tr>
<td>Threshold 7</td>
<td>0.837</td>
<td>35.294</td>
<td>0.838</td>
<td>35.652</td>
</tr>
<tr>
<td>Threshold 8</td>
<td>1.300</td>
<td>43.439</td>
<td>1.301</td>
<td>43.014</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>RP</th>
<th>SP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log-likelihood at convergence</td>
<td>-3154.84</td>
<td>-141.32   -3155.09</td>
</tr>
<tr>
<td>Log-likelihood at origin</td>
<td>-3317.81</td>
<td>-318.60   -3317.81</td>
</tr>
<tr>
<td>Rho-square</td>
<td>0.049</td>
<td>0.556      0.049</td>
</tr>
<tr>
<td>Adjusted rho-square</td>
<td>0.046</td>
<td>n.a.       n.a.</td>
</tr>
</tbody>
</table>
Chapter 7.
Applications of Freight Choice Models

7.1. Introduction

There are many activities that railroad marketing departments perform. They have
to identify the principal market segments that might be attracted to railroad service, and at
each specific time, recommend the ideal marketing mix that can maximize revenues or net
profits. They also have to forecast the volumes of business that might be obtained from
each segment, and provide guidelines about what price to charge for each segment. This
chapter discusses the use of demand models for such decisions.

In section 7.2, we introduce the concept of own-elasticity and analyze own-
elasticities of the shippers who participated in our survey. In section 7.3, we analyze own-
elasticities further in terms of market segments. Such analysis will help railroads design
customer-tailored service. In section 7.4, we propose a measure of shippers’ willingness to
pay for service improvements. The measure will help railroads determine proper price for
service improvements. We also discuss a case study of using the measure for an actual
investment decision. In section 7.5, we introduce the concept of cross-elasticity and
propose the market strength map. The map will allow a railroad to perform comparative
analysis of the freight industry and to devise an effective marketing strategy.

7.2. Own Elasticity to Service Attribute

Definition

Demand elasticity can provide a useful guideline about how the market responds to
service performance. Let us start with defining the concept. Demand elasticity can be
defined in the both disaggregate and aggregate level and for both own and cross effects (Ben-Akiva and Lerman 1985, p.111). A disaggregate elasticity represents the responsiveness of an individual's choice probability to a change in the value of some attribute. An aggregate elasticity represents the responsiveness of a group of decision makers rather than that of an individual. An own-elasticity represents how an individual's (or a group's) probability of choosing an alternative is influenced by a change in the value of some attributes of the alternative. A cross-elasticity represents how the probability of choosing an alternative is influenced by a change in the attribute value of other alternatives. We will discuss the own-elasticity in the sections 7.2 and 7.3 and the cross-elasticity in section 7.5.

The own-elasticity ($\eta_{ni}$) is defined as the percentage change in the share of mode $i$ corresponding to a percentage change in the attribute value of mode $i$. Suppose that an individual $n$'s probability of choosing an alternative $i$ is $P_n(i)$ when an attribute $k$ has a value of $x_{nk}$. Also suppose that his probability of choosing $i$ changes by $\partial P_n(i)$ as the value of the attribute changes by $\partial x_{nk}$. The elasticity is the ratio of a percentage change in the choice probability ($\frac{\partial P_n(i)}{P_n(i)}$) over a percentage change in the attribute value ($\frac{\partial x_{nk}}{x_{nk}}$). By comparing percentage changes in choice probability and attribute value, the elasticity measure ensures that it is unit-free, i.e. attribute value can be measured in any units. Mathematically, disaggregate own-elasticity is defined as follows:

$$E_{x_{nk}}^{P_n(i)} = \frac{\partial P_n(i)}{\partial x_{nk}} \frac{x_{nk}}{P_n(i)}$$

Now, suppose that we want to know how much an incremental change in an attribute influences the expected share of alternative $i$ by a group of shippers. Aggregate elasticity provides an answer by utilizing disaggregate elasticities. Suppose that a value of an attribute is changed by $\partial x_{nk}$. Each individual's choice probability will be changed by
\( \partial P_n(i) \). If we sum up changes of all individuals' choice probabilities by using individual choice probabilities as weights, we will obtain the elasticity of the expected share of the group. The aggregate own elasticity is thus defined as:

\[
\eta_i = E_{x_{ia}}^{P(i)} = \frac{\sum_{n=1}^{N} P_n(i) E_{x_{ia}}^{P_n(i)}}{\sum_{n=1}^{N} P_n(i)}
\]

where

\[
x_{jk} = \frac{1}{N} \sum_{n=1}^{N} x_{jn_k} \quad \text{and} \quad \bar{P}(i) = \frac{1}{N} \sum_{n=1}^{N} P_n(i).
\]

We use the survey data and analyze aggregate elasticities to five service attributes: freight rate, on-time delivery, transit time, loss and damage, and equipment availability. Elasticities are calculated using the results of our models: a model based on total logistics costs, a model with flexibility perception, a model based on stated preference data, and a model using the combined data of RP and SP. Models with randomly-distributed rates failed to improve the fixed rate model and we do not include them here. But their estimates should be similar to the value in the following table. In addition, we will make all the discussions from now on based on the result of the RP model with flexibility.

We also found that four other elasticity estimates are available: Vieira, John Morton, Kansas State, and Marketing manager survey. Vieira’s results are reported in his thesis (Vieira 1992). The other estimates are reported in Smith and Resor (1991). John Morton is a market research company. They ran preference regression on hypothetical choices (i.e. SP data). Their results can be erroneous since they assumed that the preference scale of 1 to 9 is continuous. A correct approach is to acknowledge its discreteness and to use ordered probit which we employed in Chapter 6. The results of a Kansas State University study are based on the regression of actual share on “actual car-miles per car-year”. Since the results were based on data of the 1970s and their surrogate for service is questionable, Smith and Resor expressed caution on using their results. In
addition, marketing managers of Burlington Northern were interviewed and asked to forecast the market gain that could be achieved from improved service. The interview results represent expert opinions but could be subjective. We will include all of their results as benchmarks.

To most attributes, own elasticities of truck share are smaller than those of rail or intermodal shares. Such results should not be surprising since own elasticity is reverse-proportional to its own share by the definition. If a mode's share is large already, the share is difficult to increase. If it is small, the share can be increased quickly with small improvements in service. The Kansas State University study ignored this fact and estimated the same elasticity for both truck and rail, which gives another reason to take caution in interpreting its results.

Own-Elasticities to Service Attributes

Table 7.1 summarizes our estimation results in terms of own-elasticities to freight rate. The table reads such that in the third row, the RP model with flexibility estimated that if truck service increased its rate by 1%, it would lose its share by 0.136%, and that if rail service increased its rate by 1%, it would lose its share by 0.419%, and so forth. The table can also be interpreted in the opposite way. That is, if truck service decreased its rate by 1%, it would gain its share by 0.136%.

<table>
<thead>
<tr>
<th>Estimate Type</th>
<th>Truck</th>
<th>Rail</th>
<th>Intermodal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marketing managers' estimates</td>
<td>n.a.</td>
<td>-1.2 to ∞</td>
<td>n.a.</td>
</tr>
<tr>
<td>Kansas State's estimates</td>
<td>n.a.</td>
<td>-1.3 to -3.0</td>
<td>n.a.</td>
</tr>
<tr>
<td>John Morton's estimates</td>
<td>n.a.</td>
<td>-1.3 (-0.9 to -1.6)</td>
<td>n.a.</td>
</tr>
<tr>
<td>Vieira's Estimation</td>
<td>-1.150</td>
<td>-0.448</td>
<td>-0.324</td>
</tr>
<tr>
<td>RP data with Fixed rate</td>
<td>-0.173</td>
<td>-0.532</td>
<td>-0.729</td>
</tr>
<tr>
<td>RP data with Flexibility</td>
<td>-0.136</td>
<td>-0.419</td>
<td>-0.574</td>
</tr>
<tr>
<td>SP data</td>
<td>-0.214</td>
<td>-0.658</td>
<td>-0.903</td>
</tr>
<tr>
<td>Combined data</td>
<td>-0.209</td>
<td>-0.645</td>
<td>-0.885</td>
</tr>
</tbody>
</table>
Overall, the table shows that shares of rail or intermodal service are more sensitive to freight rate than truck shares. Our elasticity measures are close to Viera’s result and lower than the benchmark estimates. Our estimates of shippers’ willingness to pay for service quality, which we will discuss later, are close to the benchmark estimates.

Elasticity to on-time delivery is rather complex since it influences choice model through two variables: safety stock holding costs and reliability costs. Since safety stock holding cost is a non-linear function of on-time delivery ($\tau_{in}$) and the probability of incremental changes in on-time delivery ($d\tau_{in}$), we need careful evaluation. Table 7.2 shows the changes in choice probabilities due to a 1% change in on-time delivery.

<table>
<thead>
<tr>
<th>Table 7.2. Own Elasticity of Choice Probability to On-time Delivery</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Marketing managers’ estimates</strong></td>
</tr>
<tr>
<td>n.a.</td>
</tr>
<tr>
<td><strong>Kansas State’s estimates</strong></td>
</tr>
<tr>
<td><strong>John Morton’s estimates</strong></td>
</tr>
<tr>
<td><strong>Vieira’s Estimation</strong></td>
</tr>
<tr>
<td><strong>RP data with Fixed rate</strong></td>
</tr>
<tr>
<td><strong>RP data with Flexibility</strong></td>
</tr>
<tr>
<td><strong>SP data</strong></td>
</tr>
<tr>
<td><strong>Combined data</strong></td>
</tr>
</tbody>
</table>

Overall, modal shares are very elastic to service reliability. Trucks would not gain share much from improvements in on-time performance since trucks are already providing reliable service. But, as shown below, rail and intermodal lagged in on-time delivery in 1988, and their shippers wanted an improvement in on-time performance. The elasticity of rail share to on-time delivery is much larger than its elasticity to freight rate or transit time, which suggests the importance of providing reliable rail service.

<table>
<thead>
<tr>
<th>Actual On-time Performance</th>
<th>Notation</th>
<th>Unit</th>
<th>Mean</th>
<th>Std.dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Truck</td>
<td>ctt_in</td>
<td>%</td>
<td>91.10</td>
<td>8.15</td>
</tr>
<tr>
<td>Rail</td>
<td>ctt_Rn</td>
<td>%</td>
<td>82.99</td>
<td>13.10</td>
</tr>
<tr>
<td>Intermodal</td>
<td>ctt_ipn</td>
<td>%</td>
<td>89.85</td>
<td>8.41</td>
</tr>
</tbody>
</table>
Table 7.3 shows the changes in choice probabilities due to a 1% change in transit time. The table shows that trucking companies would not gain much share from faster service than the current operation. This is reasonable since trucking companies may be already providing fast service. On the other hand, if rail or intermodal service improves its transit time, it can gain large share. The table suggests that with 1% increase in transit time, its share would increase by about 0.34%.

Table 7.3. Own Elasticity of Choice Probability to Transit time

<table>
<thead>
<tr>
<th></th>
<th>Truck</th>
<th>Rail</th>
<th>Intermodal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marketing managers’ est.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td>Kansas State’s est.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td>John Morton’s est.</td>
<td>n.a.</td>
<td>-1.2 (-0.9 to -1.6)</td>
<td>n.a.</td>
</tr>
<tr>
<td>Vieira’s Estimation</td>
<td>-0.059</td>
<td>-0.865</td>
<td>-0.291</td>
</tr>
<tr>
<td>RP data with Fixed rate</td>
<td>-0.059</td>
<td>-0.455</td>
<td>-0.447</td>
</tr>
<tr>
<td>RP data with Flexibility</td>
<td>-0.024</td>
<td>-0.336</td>
<td>-0.274</td>
</tr>
<tr>
<td>SP data</td>
<td>-0.053</td>
<td>-0.435</td>
<td>-0.430</td>
</tr>
<tr>
<td>Combined data</td>
<td>-0.053</td>
<td>-0.438</td>
<td>-0.432</td>
</tr>
</tbody>
</table>

Table 7.4 shows the changes in choice probabilities due to a 1% change in loss and damage. Overall, shippers are less sensitive to loss and damage than to other attributes. This may suggest that all modes provide satisfactory service in terms of loss and damage, as shown in the below table. In addition, truck shippers are even less sensitive relative to rail shippers. The actual performance table shows that only 0.31 percentage of the total value shipped through truck was lost or damaged.

Table 7.4. Own Elasticity of Choice Probability to Loss and damage

<table>
<thead>
<tr>
<th></th>
<th>Truck</th>
<th>Rail</th>
<th>Intermodal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marketing managers’ est.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td>Kansas State’s est.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td>John Morton’s est.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td>Vieira’s Estimation</td>
<td>-0.057</td>
<td>-0.286</td>
<td>-0.279</td>
</tr>
<tr>
<td>RP data with Fixed rate</td>
<td>-0.012</td>
<td>-0.061</td>
<td>-0.105</td>
</tr>
<tr>
<td>RP data with Flexibility</td>
<td>-0.012</td>
<td>-0.065</td>
<td>-0.113</td>
</tr>
<tr>
<td>SP data</td>
<td>-0.021</td>
<td>-0.206</td>
<td>-0.358</td>
</tr>
<tr>
<td>Combined data</td>
<td>-0.017</td>
<td>-0.171</td>
<td>-0.297</td>
</tr>
</tbody>
</table>
Table 7.5 shows the changes in choice probabilities due to a 1% change in equipment availability. For truck shippers, this attribute would mean the availability of a truck and a driver and the equipment dimension and size. For rail and intermodal shippers, it's about the availability, cleanliness, door opening and maintenance condition of trailers and cars. Since it influences choice probability both directly and through the formation of flexibility perception, its elasticity is carefully calculated. Overall, shippers are very sensitive to equipment availability than to other attributes. The high sensitivity is a direct result of the large coefficient of equipment availability in the choice model. Conventional stated-importance approach found equipment availability as important, but did not find this high level of sensitivity. Shippers may value equipment availability much higher than what they state as its value. In addition, rail shippers are more sensitive to truck shippers.

Table 7.5. Own Elasticity of Choice Probability to Equipment usability

<table>
<thead>
<tr>
<th>Attributes</th>
<th>Notation</th>
<th>Unit</th>
<th>Mean</th>
<th>Std.dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loss and damage percentage</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Truck</td>
<td>$d_{in}$</td>
<td>%</td>
<td>0.31</td>
<td>0.62</td>
</tr>
<tr>
<td>Rail</td>
<td>$d_{in}$</td>
<td>%</td>
<td>0.85</td>
<td>0.82</td>
</tr>
<tr>
<td>Intermodal</td>
<td>$d_{in}$</td>
<td>%</td>
<td>1.16</td>
<td>1.05</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Truck</th>
<th>Rail</th>
<th>Intermodal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marketing managers’ estimates</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td>Kansas State’s estimates</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td>John Morton’s estimates</td>
<td>n.a.</td>
<td>2.5 (1.9 to 3.3)</td>
<td>n.a.</td>
</tr>
<tr>
<td>Vieira’s Estimation</td>
<td>0.977</td>
<td>0.502</td>
<td>3.737</td>
</tr>
<tr>
<td>RP data with Fixed rate</td>
<td>0.773</td>
<td>2.542</td>
<td>3.210</td>
</tr>
<tr>
<td>RP data with Flexibility</td>
<td>1.238</td>
<td>2.688</td>
<td>3.610</td>
</tr>
<tr>
<td>SP data</td>
<td>0.674</td>
<td>2.214</td>
<td>2.795</td>
</tr>
<tr>
<td>Combined data</td>
<td>0.819</td>
<td>2.692</td>
<td>3.399</td>
</tr>
<tr>
<td>Attributes</td>
<td>Notation</td>
<td>Unit</td>
<td>Mean</td>
</tr>
<tr>
<td>----------------------------------</td>
<td>----------</td>
<td>------</td>
<td>-------</td>
</tr>
<tr>
<td>Equipment availability percentage</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Truck</td>
<td>ueq_{in}</td>
<td>%</td>
<td>94.99</td>
</tr>
<tr>
<td>Rail</td>
<td>ueq_{in}</td>
<td>%</td>
<td>87.56</td>
</tr>
<tr>
<td>Intermodal</td>
<td>ueq_{in}</td>
<td>%</td>
<td>92.06</td>
</tr>
</tbody>
</table>

7.3. Market Segmentation

In order to analyze the difference of demand response among different market segments, we categorized shippers in terms of distance, annual tonnage, and commodity type and compared their differences of own-elasticities. Again, all the analysis is based on the result of the RP model with flexibility. The analysis illustrates the different responses of different segments and the importance of designing customer-tailored service.

*Short-vs. Long-distance Shippers*

We categorized shippers into three categories depending on whether the distance between origin and destination of a corridor is less than 400 mile, greater than 1000 miles, or in-between. They have the following statistics:

<table>
<thead>
<tr>
<th>Distance</th>
<th>mean (mile)</th>
<th>std. dev. (mile)</th>
<th>sample size</th>
</tr>
</thead>
<tbody>
<tr>
<td>less than 400 miles</td>
<td>273</td>
<td>38</td>
<td>120</td>
</tr>
<tr>
<td>in-between</td>
<td>610</td>
<td>122</td>
<td>100</td>
</tr>
<tr>
<td>greater than 1000 miles</td>
<td>1576</td>
<td>645</td>
<td>70</td>
</tr>
</tbody>
</table>

Their own elasticities to freight rate are shown in the below table. Overall, shippers who ship a long distance are more sensitive to freight rate than short-distance shippers. The tendency is clearer in the intermodal service. For corridors with distance less than 400 miles, trucking is more convenient and has natural advantages since trucking does not need to go through drayage (pick-up and delivery) and terminal operations. Thus, short-distance shippers are captive to trucking. Their own elasticity shows that they are indeed insensitive to freight rate.
<table>
<thead>
<tr>
<th>Elasticity to freight rate</th>
<th>truck</th>
<th>rail</th>
<th>intermodal</th>
</tr>
</thead>
<tbody>
<tr>
<td>less than 400 miles</td>
<td>-0.042</td>
<td>-0.142</td>
<td>-0.108</td>
</tr>
<tr>
<td>in-between</td>
<td>-0.114</td>
<td>-0.357</td>
<td>-0.348</td>
</tr>
<tr>
<td>greater than 1000 miles</td>
<td>-0.434</td>
<td>-0.673</td>
<td>-1.306</td>
</tr>
</tbody>
</table>

The own-elasticity to on-time delivery is shown in the below table. Overall, long distance shippers are more sensitive to reliable service than short distance shippers. For a corridor whose distance is less than 400 miles, trucks would be on-time almost 100%. Thus, reliability is not a big concern for such truck users. As transportation distance increases, possible sources of unexpected delay increase (e.g. engine breakdown, highway accidents, or a driver’s off-duty activities for trucking, and missed train connections for rail and intermodal service). Thus, it is understandable that long-distance shippers are more worried about on-time delivery than short-distance shippers. If a carrier provides a guaranty of on-time delivery to long-distance shippers, they might be willing to pay more for such service (e.g. a high-priority service). We will later discuss how much shippers are willing to pay for such service.

<table>
<thead>
<tr>
<th>Elasticity to on-time delivery</th>
<th>truck</th>
<th>rail</th>
<th>intermodal</th>
</tr>
</thead>
<tbody>
<tr>
<td>less than 400 miles</td>
<td>0.245</td>
<td>1.385</td>
<td>1.429</td>
</tr>
<tr>
<td>in-between</td>
<td>0.379</td>
<td>1.345</td>
<td>1.504</td>
</tr>
<tr>
<td>greater than 1000 miles</td>
<td>0.633</td>
<td>0.982</td>
<td>1.747</td>
</tr>
</tbody>
</table>

For the same reason as the own-elasticity to on-time delivery, long-distance shippers are more sensitive to transit time when they send freight via truck. A shipper would not be really worried about transit time of trucking service for a short distance. But if he sent a shipment for a long distance via truck, he would be worried about transit time (and reliability). For rail and intermodal service, no discernible trend is shown between short- and long-distance shippers.

<table>
<thead>
<tr>
<th>Elasticity to transit time</th>
<th>truck</th>
<th>rail</th>
<th>intermodal</th>
</tr>
</thead>
<tbody>
<tr>
<td>less than 400 miles</td>
<td>-0.007</td>
<td>-0.359</td>
<td>-0.262</td>
</tr>
<tr>
<td>in-between</td>
<td>-0.024</td>
<td>-0.396</td>
<td>-0.270</td>
</tr>
<tr>
<td>greater than 1000 miles</td>
<td>-0.063</td>
<td>-0.269</td>
<td>-0.290</td>
</tr>
</tbody>
</table>
In terms of loss and damage, no discernible trend is shown with distance.

<table>
<thead>
<tr>
<th>Elasticity to loss and damage</th>
<th>truck</th>
<th>rail</th>
<th>intermodal</th>
</tr>
</thead>
<tbody>
<tr>
<td>less than 400 miles</td>
<td>-0.004</td>
<td>-0.057</td>
<td>-0.132</td>
</tr>
<tr>
<td>in-between</td>
<td>-0.020</td>
<td>-0.075</td>
<td>-0.111</td>
</tr>
<tr>
<td>greater than 1000 miles</td>
<td>-0.015</td>
<td>-0.062</td>
<td>-0.096</td>
</tr>
</tbody>
</table>

In terms of equipment availability, long distance shippers are more worried about the shortage of truck driver and are very sensitive to equipment availability. No such concern appears for rail and intermodal.

<table>
<thead>
<tr>
<th>Elasticity to equipment availability</th>
<th>truck</th>
<th>rail</th>
<th>intermodal</th>
</tr>
</thead>
<tbody>
<tr>
<td>less than 400 miles</td>
<td>0.834</td>
<td>3.131</td>
<td>3.259</td>
</tr>
<tr>
<td>in-between</td>
<td>1.255</td>
<td>3.034</td>
<td>3.470</td>
</tr>
<tr>
<td>greater than 1000 miles</td>
<td>2.107</td>
<td>2.173</td>
<td>4.094</td>
</tr>
</tbody>
</table>

The above analysis shows that for intermodal service, short-distance and long-distance shippers are not really different in terms of their sensitivity to transit time, loss and damage and equipment availability. Long-distance shippers are, however, sensitive to freight rate and on-time delivery. Therefore, when railroad companies design long-distance intermodal service, they have to focusing on improving these two attributes.

**Small vs. Large Shippers**

Secondly, we classified shippers into three groups depending on whether the annual corridor tonnage is less than 100 tons, greater than 1,000 tons, or in-between. Their statistics are as follow:

<table>
<thead>
<tr>
<th></th>
<th>mean (ton)</th>
<th>std. dev. (ton)</th>
<th>sample size</th>
</tr>
</thead>
<tbody>
<tr>
<td>less than 100 tons / year</td>
<td>34</td>
<td>30</td>
<td>119</td>
</tr>
<tr>
<td>in-between</td>
<td>456</td>
<td>266</td>
<td>116</td>
</tr>
<tr>
<td>greater than 1000 tons/yr</td>
<td>2662</td>
<td>2346</td>
<td>55</td>
</tr>
</tbody>
</table>
Shippers' own-elasticities to freight rate are provided in the below table. The table shows that small shippers are more sensitive to freight rate than large shippers. For small shipments, shippers can choose among truck, rail and intermodal services. Their modal shares can change greatly due to the high competition. On the other hand, shippers with large shipments are captive to rail and their modal shares are less sensitive to freight rate.

<table>
<thead>
<tr>
<th>Elasticity to freight rate</th>
<th>truck</th>
<th>rail</th>
<th>intermodal</th>
</tr>
</thead>
<tbody>
<tr>
<td>less than 100 tons / year</td>
<td>-0.184</td>
<td>-0.634</td>
<td>-0.630</td>
</tr>
<tr>
<td>in-between</td>
<td>-0.088</td>
<td>-0.356</td>
<td>-0.534</td>
</tr>
<tr>
<td>greater than 1000 tons/yr.</td>
<td>-0.121</td>
<td>-0.293</td>
<td>-0.495</td>
</tr>
</tbody>
</table>

In terms of on-time delivery, large shippers prefer reliable service than small shippers. Small shippers may be able to cope with transportation contingencies more flexibly. But in large organizations, a small deviation from the original plan can cause a series of adjustments in many departments. Thus, it is reasonable that large shippers are more sensitive to reliable service.

<table>
<thead>
<tr>
<th>Elasticity to on-time delivery</th>
<th>truck</th>
<th>rail</th>
<th>intermodal</th>
</tr>
</thead>
<tbody>
<tr>
<td>less than 100 tons / year</td>
<td>0.270</td>
<td>1.087</td>
<td>1.127</td>
</tr>
<tr>
<td>in-between</td>
<td>0.417</td>
<td>1.302</td>
<td>1.947</td>
</tr>
<tr>
<td>greater than 1000 tons/yr.</td>
<td>0.502</td>
<td>1.164</td>
<td>1.988</td>
</tr>
</tbody>
</table>

In terms of transit time, small shippers are more sensitive than large shippers. Large shippers can plan ahead and may be less sensitive to transit time, as long as reliability is satisfied. Moreover, large shippers are likely to have lower discount rates than small shippers. Thus, large shippers incur lower inventory holding costs during in-transit time. Note that the elasticity of large shippers to on-time performance is 1.988, whereas that to transit time is -0.210. This implies that railroads are better off by offering a service with a longer transit time but with a guaranteed reliability. Such service would require railroads to hold trailers at rail terminals for some days in certain cases. But large shippers
would respond more favorably to a slow but reliable (and thus predictable) service than to a fast but unreliable service.

<table>
<thead>
<tr>
<th>Elasticity to transit time</th>
<th>truck</th>
<th>rail</th>
<th>intermodal</th>
</tr>
</thead>
<tbody>
<tr>
<td>less than 100 tons / year</td>
<td>-0.026</td>
<td>-0.457</td>
<td>-0.283</td>
</tr>
<tr>
<td>in-between</td>
<td>-0.022</td>
<td>-0.321</td>
<td>-0.293</td>
</tr>
<tr>
<td>greater than 1000 tons/yr.</td>
<td>-0.024</td>
<td>-0.231</td>
<td>-0.210</td>
</tr>
</tbody>
</table>

In terms of loss and damage, small and large shippers appears to be more sensitive than middle-size shippers. But no major trend is shown.

<table>
<thead>
<tr>
<th>Elasticity to loss and damage</th>
<th>truck</th>
<th>rail</th>
<th>intermodal</th>
</tr>
</thead>
<tbody>
<tr>
<td>less than 100 tons / year</td>
<td>-0.013</td>
<td>-0.105</td>
<td>-0.134</td>
</tr>
<tr>
<td>in-between</td>
<td>-0.012</td>
<td>-0.046</td>
<td>-0.084</td>
</tr>
<tr>
<td>greater than 1000 tons/yr.</td>
<td>-0.008</td>
<td>-0.053</td>
<td>-0.112</td>
</tr>
</tbody>
</table>

In terms of equipment availability, large shippers are more sensitive than small shippers. This result can also be interpreted in terms of large shippers’ tendency to adhere to pre-established plans and to avoid any delays due to equipment shortage.

<table>
<thead>
<tr>
<th>Elasticity to equipment availability</th>
<th>truck</th>
<th>rail</th>
<th>intermodal</th>
</tr>
</thead>
<tbody>
<tr>
<td>less than 100 tons / year</td>
<td>0.922</td>
<td>2.540</td>
<td>2.710</td>
</tr>
<tr>
<td>in-between</td>
<td>1.384</td>
<td>2.771</td>
<td>4.410</td>
</tr>
<tr>
<td>greater than 1000 tons/yr.</td>
<td>1.664</td>
<td>2.729</td>
<td>4.501</td>
</tr>
</tbody>
</table>

Overall, small shippers are more sensitive to freight rate and transit time. Large shippers are more sensitive to on-time delivery and equipment availability. No difference between small and large shippers is found for loss and damage. The results suggest that when railroad companies contact small shippers, they would better emphasize cost advantage, and that when they contact large shippers, they would better emphasize plan integrity and no risk of system failure.

**Commodity type**

Thirdly, we analyze own elasticities for each commodity type. As shown below, we have five commodity types. Pet foods are the least expensive. Plastics, paper and
aluminum are more expensive in that order. Tires are the most expensive commodities. They represent a wide range of the carload freight market and none of them are perishable.

<table>
<thead>
<tr>
<th>Commodity</th>
<th>mean price ($/ton)</th>
<th>sample size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pet food</td>
<td>530</td>
<td>40</td>
</tr>
<tr>
<td>Plastics</td>
<td>5,986</td>
<td>48</td>
</tr>
<tr>
<td>Paper</td>
<td>7,201</td>
<td>97</td>
</tr>
<tr>
<td>Aluminum</td>
<td>9,440</td>
<td>69</td>
</tr>
<tr>
<td>Tires</td>
<td>13,872</td>
<td>36</td>
</tr>
</tbody>
</table>

The below table shows the own-elasticities to freight rate by commodity type. Overall, high-value commodities are less sensitive than low-value commodities. This is reasonable since transportation costs are a small portion of the total costs for high-value commodities, and a large portion for low-value commodities.

<table>
<thead>
<tr>
<th>Elasticity to freight rate</th>
<th>truck</th>
<th>rail</th>
<th>intermodal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pet food</td>
<td>-0.193</td>
<td>-0.469</td>
<td>-0.636</td>
</tr>
<tr>
<td>Plastics</td>
<td>-0.156</td>
<td>-0.270</td>
<td>-0.643</td>
</tr>
<tr>
<td>Paper</td>
<td>-0.139</td>
<td>-0.335</td>
<td>-0.603</td>
</tr>
<tr>
<td>Aluminum</td>
<td>-0.126</td>
<td>-0.647</td>
<td>-0.588</td>
</tr>
<tr>
<td>Tires</td>
<td>-0.064</td>
<td>-0.564</td>
<td>-0.415</td>
</tr>
</tbody>
</table>

In terms of on-time delivery, commodities (e.g. newspaper materials) whose receivers have stricter time windows for delivery are more sensitive than other commodities. In industries where retailers (e.g. discount chain or supermarket) have greater bargaining power than manufacturers, shippers may be sensitive to on-time delivery in order to deliver within a pre-specified time window.

<table>
<thead>
<tr>
<th>Elasticity to on-time delivery</th>
<th>truck</th>
<th>rail</th>
<th>intermodal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pet food</td>
<td>0.323</td>
<td>1.259</td>
<td>1.921</td>
</tr>
<tr>
<td>Plastics</td>
<td>0.434</td>
<td>1.080</td>
<td>2.035</td>
</tr>
<tr>
<td>Paper</td>
<td>0.465</td>
<td>1.356</td>
<td>2.087</td>
</tr>
<tr>
<td>Aluminum</td>
<td>0.307</td>
<td>1.068</td>
<td>1.093</td>
</tr>
<tr>
<td>Tires</td>
<td>0.211</td>
<td>1.101</td>
<td>1.104</td>
</tr>
</tbody>
</table>
In terms of transit time, high-valued commodities (e.g. tires, paper, aluminum) are sensitive to transit time, since inventory holding costs during transit time is high for high-valued items. On the other hand, pet food shows almost no sensitivity to transit time. In addition, tire shippers show an unexpectedly high sensitivity for rail. Their customers may want quick response or may incur large inventory holding costs during in-transit time.

<table>
<thead>
<tr>
<th>Elasticity to transit time</th>
<th>truck</th>
<th>rail</th>
<th>intermodal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pet food</td>
<td>0.000</td>
<td>-0.002</td>
<td>-0.002</td>
</tr>
<tr>
<td>Plastics</td>
<td>-0.014</td>
<td>-0.191</td>
<td>-0.141</td>
</tr>
<tr>
<td>Paper</td>
<td>-0.022</td>
<td>-0.262</td>
<td>-0.212</td>
</tr>
<tr>
<td>Aluminum</td>
<td>-0.025</td>
<td>-0.459</td>
<td>-0.228</td>
</tr>
<tr>
<td>Tires</td>
<td>-0.065</td>
<td>-1.364</td>
<td>-0.747</td>
</tr>
</tbody>
</table>

In terms of loss and damage, high value commodities are more sensitive than low-value commodities. For instance, pet foods show nearly zero sensitivity.

<table>
<thead>
<tr>
<th>Elasticity to loss and damage</th>
<th>truck</th>
<th>rail</th>
<th>intermodal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pet food</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Plastics</td>
<td>-0.013</td>
<td>-0.016</td>
<td>-0.206</td>
</tr>
<tr>
<td>Paper</td>
<td>-0.005</td>
<td>-0.119</td>
<td>-0.077</td>
</tr>
<tr>
<td>Aluminum</td>
<td>-0.030</td>
<td>-0.063</td>
<td>-0.056</td>
</tr>
<tr>
<td>Tires</td>
<td>-0.010</td>
<td>-0.042</td>
<td>-0.307</td>
</tr>
</tbody>
</table>

In terms of equipment availability, shippers of high value commodities (tire, paper, aluminum) are less sensitive than low value commodities. It is feasible that they pay higher freight rate than shippers of low value commodities, and have little worry about equipment availability. On the other hand, shippers of low value commodities may have to bargain hard in order to ensure low freight rate, and then worry about equipment availability and equipment condition.
<table>
<thead>
<tr>
<th>Elasticity to equipment availability</th>
<th>truck</th>
<th>rail</th>
<th>intermodal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pet food</td>
<td>1.117</td>
<td>2.641</td>
<td>4.220</td>
</tr>
<tr>
<td>Plastics</td>
<td>1.487</td>
<td>2.608</td>
<td>4.852</td>
</tr>
<tr>
<td>Paper</td>
<td>1.501</td>
<td>3.010</td>
<td>4.746</td>
</tr>
<tr>
<td>Aluminum</td>
<td>1.004</td>
<td>2.362</td>
<td>2.580</td>
</tr>
<tr>
<td>Tires</td>
<td>0.810</td>
<td>2.414</td>
<td>2.625</td>
</tr>
</tbody>
</table>

Overall, shippers of high value commodities are sensitive to transit time and loss and damage. Shippers of low value commodities, on the other hand, are sensitive to freight rate and equipment availability. An interesting finding is that tire shippers are very sensitive to transit time. This results shows the importance of market segmentation and a need for developing customer-tailored service.

7.4. Shippers' Willingness to Pay for Service Improvements

Definition

Railroad managers may be interested in knowing how much shippers are willing to pay for improved service. Improvements in service quality can increase revenue in many ways. First, shippers may engage in more transportation activities than currently after the service improvement (i.e. an increase in total freight volume). Secondly, a shipper may use more rail than truck service (i.e. an increase in modal share). Thirdly, a shipper may be willing to pay more for the improved service (i.e. an increase in unit freight rate). The three effects are mixed with each other and difficult to decompose. Our approach focuses on estimating the size of rate increase that shippers are willing to pay without changing modal selection assuming the same level of freight volume.

We define shippers' willingness to pay for service quality by the absolute ratio of the own-elasticity of choice probability to service quality and that to freight rate.

\[
\text{Willingness}_{k}^{i} = \left| \frac{\text{Own - elasticity}_{k}^{i}}{\text{Own - elasticity}_{\text{freight rate}}^{i}} \right|
\]
where k denotes service attribute and i denotes a freight mode. This concept is similar to the marginal rate of substitution. Suppose that a shipper would increase rail share by 1% when on-time performance is improved by 1% or when freight rate is decreased by 2%. Then we can say that the shipper would be willing to pay 2% higher freight rate for an 1% improvement in on-time delivery. The 1% increase in on-time performance has the same effect to rail share as the 2% decrease in freight rate. We estimated shippers’ willingness based on the above formula and the results of the RP model with flexibility.

We have to note that whereas own elasticities of an attribute are reversely proportional to current shares of modes, shippers’ willingness to pay for an attribute should be similar for all modes. Shippers’ willingness represents shippers’ taste or attitude which is generic rather than mode-specific. On the other hand, small differences of willingness among modes may come from the heterogeneity of shippers who patronize a mode or their desire for improvements in the current service level. For instance, if those who are sensitive to on-time delivery use trucks rather than rail, truck users would be willing to pay higher rates for improvements in on-time delivery than rail users. On the other hand, if rail users were really suffering from low performance in on-time delivery, they would be willing to pay higher rates than truck users.

**Willingness to pay for On-time Delivery**

Let us first consider shippers’ willingness to pay for on-time delivery. The RP model with flexibility indicates that truck shares will be decreased by 0.136% with an 1% increase in truck rate, but will be increased by 0.370% with an 1% increase in on-time delivery. This implies that shippers are willing to pay an 2.73% (=0.370/0.136) higher rate in order to get an 1% improvements in on-time delivery of trucking. As expected, shippers’ willingness to pay are very high for all modes. Note that the high willingness does not necessarily means the high elasticity of shares to on-time performance. In terms of elasticity, rail and intermodal service show much higher levels of elasticities than truck. But in terms of willingness to pay, shippers show similar levels of willingness for all
modes. Shippers’ willingness is estimated as much higher in SP data (and thus in combined dataset) than in RP data. This result may arise as shippers try to emphasize reliability in SP experiments more than what they are willing to pay for in actual choice.

Table 7.6. Shippers’ Willingness to Pay for an 1% Improvement in On-time Delivery

<table>
<thead>
<tr>
<th></th>
<th>Truck</th>
<th>Rail</th>
<th>Intermodal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marketing managers’</td>
<td>n.a.</td>
<td>0 to 0.4%</td>
<td>n.a.</td>
</tr>
<tr>
<td>Kansas State’s estimates</td>
<td>n.a.</td>
<td>0.2 to 3.3%</td>
<td>n.a.</td>
</tr>
<tr>
<td>John Morton’s estimates</td>
<td>n.a.</td>
<td>4.1 (2.9 to 6.9)%</td>
<td>n.a.</td>
</tr>
<tr>
<td>Vieira’s Estimation</td>
<td>0.501%</td>
<td>6.663%</td>
<td>7.049%</td>
</tr>
<tr>
<td>RP data with Fixed rate</td>
<td>1.451%</td>
<td>1.139%</td>
<td>1.077%</td>
</tr>
<tr>
<td>RP data with Flexibility</td>
<td>2.721%</td>
<td>2.862%</td>
<td>2.720%</td>
</tr>
<tr>
<td>SP data</td>
<td>4.157%</td>
<td>3.315%</td>
<td>3.128%</td>
</tr>
<tr>
<td>Combined data</td>
<td>3.328%</td>
<td>2.596%</td>
<td>2.452%</td>
</tr>
</tbody>
</table>

We can adopt the same approach to estimate the willingness to pay for on-time delivery in different market segments. The results are provided in the below tables. Long-distance shippers are less willing to pay for improvements in on-time delivery than short-distance shippers. Large shippers are more willing than small shippers, and shippers of plastics and paper are more willing than other commodity shippers. When short-haul shipments are shipped through rail or intermodal service, they might undergo the same number of terminal operations as long-haul shipments. Thus, the sensitivity of shippers to on-time reliability is higher with short-haul shipments than with long-haul shipments.

**In terms of shipment distance**

<table>
<thead>
<tr>
<th></th>
<th>truck</th>
<th>rail</th>
<th>intermodal</th>
</tr>
</thead>
<tbody>
<tr>
<td>less than 400 miles</td>
<td>5.833%</td>
<td>9.754%</td>
<td>13.231%</td>
</tr>
<tr>
<td>in-between</td>
<td>3.325%</td>
<td>3.768%</td>
<td>4.322%</td>
</tr>
<tr>
<td>greater than 1000 miles</td>
<td>1.459%</td>
<td>1.459%</td>
<td>1.338%</td>
</tr>
</tbody>
</table>

**In terms of shipment size**

<table>
<thead>
<tr>
<th></th>
<th>truck</th>
<th>rail</th>
<th>intermodal</th>
</tr>
</thead>
<tbody>
<tr>
<td>less than 100 tons / year</td>
<td>1.467%</td>
<td>1.715%</td>
<td>1.789%</td>
</tr>
<tr>
<td>in-between</td>
<td>4.739%</td>
<td>3.657%</td>
<td>3.646%</td>
</tr>
<tr>
<td>greater than 1000 tons/yr.</td>
<td>4.149%</td>
<td>3.973%</td>
<td>4.016%</td>
</tr>
</tbody>
</table>
In terms of commodity type

<table>
<thead>
<tr>
<th></th>
<th>truck</th>
<th>rail</th>
<th>intermodal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pet food</td>
<td>1.674 %</td>
<td>2.684 %</td>
<td>3.020 %</td>
</tr>
<tr>
<td>Plastics</td>
<td>2.782 %</td>
<td>4.000 %</td>
<td>3.165 %</td>
</tr>
<tr>
<td>Paper</td>
<td>3.345 %</td>
<td>4.048 %</td>
<td>3.461 %</td>
</tr>
<tr>
<td>Aluminum</td>
<td>2.437 %</td>
<td>1.651 %</td>
<td>1.859 %</td>
</tr>
<tr>
<td>Tire</td>
<td>3.297 %</td>
<td>1.952 %</td>
<td>2.660 %</td>
</tr>
</tbody>
</table>

Willingness to pay for Transit time

Compared to shippers’ willing to pay in order to ensure on-time delivery, shippers’ willingness to pay for transit time are relatively low. The above table shows that shippers’ willingness is low for truck, and high for rail service. This result may arise since shippers are already enjoying fast service with truck but not with rail. Shippers show desire to have fast rail service through a high willingness to pay.

Table 7.7. Shippers’ Willingness to Pay for an 1% Improvement in Transit Time

<table>
<thead>
<tr>
<th></th>
<th>Truck</th>
<th>Rail</th>
<th>Intermodal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marketing managers’ estimates</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td>Kansas State’s estimates</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td>John Morton’s estimates</td>
<td>n.a.</td>
<td>0.9 (0.6 to 1.8) %</td>
<td>n.a.</td>
</tr>
<tr>
<td>Vieira’s Estimation</td>
<td>0.051 %</td>
<td>1.931 %</td>
<td>0.898 %</td>
</tr>
<tr>
<td>RP data with Fixed rate</td>
<td>0.341 %</td>
<td>0.855 %</td>
<td>0.613 %</td>
</tr>
<tr>
<td>RP data with Flexibility</td>
<td>0.176 %</td>
<td>0.802 %</td>
<td>0.477 %</td>
</tr>
<tr>
<td>SP data</td>
<td>0.247 %</td>
<td>0.661 %</td>
<td>0.477 %</td>
</tr>
<tr>
<td>Combined data</td>
<td>0.254 %</td>
<td>0.678 %</td>
<td>0.488 %</td>
</tr>
</tbody>
</table>

Willingness to pay for Loss and damage

Shippers’ willingness to pay for an improvement in loss and damage is low, relative to their willingness to other attributes. This may be due to the fact that the percentage of shipment value lost or damaged are low and shippers do not perceive loss and damage as an important attribute.
Table 7.8. Shippers’ Willingness to Pay for an 1% Improvement in Loss and damage

<table>
<thead>
<tr>
<th></th>
<th>Truck</th>
<th>Rail</th>
<th>Intermodal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marketing managers’ estimates</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td>Kansas State’s estimates</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td>John Morton’s estimates</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td>Vieira’s Estimation</td>
<td>0.050 %</td>
<td>0.638 %</td>
<td>0.861 %</td>
</tr>
<tr>
<td>RP data with Fixed rate</td>
<td>0.069 %</td>
<td>0.115 %</td>
<td>0.144 %</td>
</tr>
<tr>
<td>RP data with Flexibility</td>
<td>0.088 %</td>
<td>0.155 %</td>
<td>0.197 %</td>
</tr>
<tr>
<td>SP data</td>
<td>0.100 %</td>
<td>0.313 %</td>
<td>0.396 %</td>
</tr>
<tr>
<td>Combined data</td>
<td>0.084 %</td>
<td>0.265 %</td>
<td>0.336 %</td>
</tr>
</tbody>
</table>

Willingness to pay for Equipment usability

Shippers’ willingness to pay for an improvement in equipment usability was very high, exceeding those for all other attributes. This result indicates the importance of securing qualified drivers and usable trailers in a desired time frame. The perception of flexibility depends on equipment availability and shippers really reward a mode or a carrier who provides flexible service and thus usable equipment readily in the desired time.

Table 7.9. Shippers’ Willingness to Pay for an .1% Improvement in Equipment usability

<table>
<thead>
<tr>
<th></th>
<th>Truck</th>
<th>Rail</th>
<th>Intermodal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marketing managers’ estimates</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td>Kansas State’s estimates</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td>John Morton’s estimates</td>
<td>n.a.</td>
<td>1.9 (1.3 to 3.0) %</td>
<td>n.a.</td>
</tr>
<tr>
<td>Vieira’s Estimation</td>
<td>0.850 %</td>
<td>1.121 %</td>
<td>11.534 %</td>
</tr>
<tr>
<td>RP data with Fixed rate</td>
<td>4.468 %</td>
<td>4.778 %</td>
<td>4.403 %</td>
</tr>
<tr>
<td>RP data with Flexibility</td>
<td>9.103 %</td>
<td>6.415 %</td>
<td>6.289 %</td>
</tr>
<tr>
<td>SP data</td>
<td>3.229 %</td>
<td>3.430 %</td>
<td>3.158 %</td>
</tr>
<tr>
<td>Combined data</td>
<td>3.926 %</td>
<td>4.171 %</td>
<td>3.842 %</td>
</tr>
</tbody>
</table>

Application of the Willingness measure to Service Design

The measure of shippers’ willingness to pay for service improvements can help railroad managers determine investment priorities. Firms can fail to evaluate market response, while concentrating on technological developments and engineering aspects of service improvements. A problem is that customers may not appreciate a new service as
the engineers do, or may not want to take the time to understand new technology. The new service can succeed only when customers appreciate the improved service. Thus, a big issue in developing a new marketing and/or service program is how to forecast demand response to the new program. While costs of new programs can be projected accurately based on engineering assessment, a projection of revenue growths is typically controversial. Without knowing what benefits each option will provide to shippers and how shippers will change their modal selections with new benefits, revenue potentials are uncertain. Such uncertainty of demand forecast made railroad managers favor projects that provide cost reductions with certainty. With proper use of freight choice models, railroad may be able to evaluate projects more effectively.

The ARES project is an example of a large-scale investment project for improving service performance.\(^1\) Burlington Northern Railroad initiated the ARES (Advanced Railroad Electronics System) project with an objective of using a computerized control system to improve operations and service quality. The project required an investment of about 350 million dollars for purchasing hardware such as vehicle identification (global positioning), radio and satellite communication, and locomotive monitoring system, and for developing software such as automated train scheduling/dispatching, locomotive analysis/control system and energy management. After years of study, the ARES team concluded that the primary benefits would be accrued in the form of reduced expenditures on fuel, equipment, labor, track-side equipment and damage prevention, and that the cost reduction alone was not enough to justify the investment (Burlington Northern: The ARES Decision (A), 1991, p.121). In order to forecast revenue increase, Burlington Northern properly adopted the following procedure:

\(^1\) \textit{defining and predicting improvements in service performance}

The following lists the primary benefits that ARES was expected to bring in:

\(^1\) refer to Harvard Case Study: The ARES Project (1991) and Smith and Resor (1991)
1) Operational Safety
   - constant monitoring of wayside signals and locomotive and track condition
   - constant monitoring of real-time position, speed of locomotives and trains.
2) Line/Terminal Capacity
   - Automatic train move authorization
   - Terminal activity control
3) Operating Efficiency: fuel, yard
   - Adjustment of train speed
   - Train / Terminal Interface
4) Dispatcher Productivity (automating routine activities)
   - Threat monitoring, warrant generation
   - Traffic planning, train sheet documentation

2) translating the operational improvements in terms of shippers’ perceived benefits

From the shippers' point of view, operational benefits are not interesting unless they are transformed into better customer service. Shippers’ primary concerns would include low freight rate, fast transit time, transit time reliability, fast shipment tracing, easy shipment expediting / rescheduling, early notification of shipment delays, provision of well-maintained equipment, accurate billing statements, or small loss and damages. The ARES team found that they can bring the following benefits to shippers:

<table>
<thead>
<tr>
<th>Benefit</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>total trip time</td>
<td>reduction by 17 % (Zeta-Tech Associates)</td>
</tr>
<tr>
<td>on-time reliability</td>
<td></td>
</tr>
<tr>
<td>terminal time</td>
<td>reduction in missed connections by 15-17 % (Martland)</td>
</tr>
<tr>
<td>trip time</td>
<td>reduction by 7-8 % (A &amp; L Associates)</td>
</tr>
<tr>
<td>shipment rescheduling</td>
<td>easier (no quantification)</td>
</tr>
</tbody>
</table>

3) predicting shippers’ willingness to pay for the improved service

A railroad can estimate shippers’ willingness to pay for the improved service through expert opinions (e.g. Delphi method), data analysis (e.g. disaggregate choice model or aggregate time series), or managerial judgments. This process is the most important but the least researched. Most current researches, including those by leading consulting companies, depend heavily on the stated importance survey. Yet, this approach is not reliable as we reviewed in Chapter 2, since stated responses are usually very
different from actual choices. Shippers have no risk of expressing that the value for service quality is higher than what they are actually willing to pay for. Indeed, Burlington Northern marketing managers could not agree on how much more shippers would be willing to pay as a result of service improvements or how much rail share it will gain from the improved. The followings were estimated as the percentage increase in customers' willingness to pay for one percent increase in service reliability (Smith and Resor 1991).

<table>
<thead>
<tr>
<th>Percentage</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 - 0.4 %</td>
<td>BN marketing managers' perception</td>
</tr>
<tr>
<td>0.2 - 3.3 %</td>
<td>Kansas State University study of actual choices</td>
</tr>
<tr>
<td>2.9 - 6.9 %</td>
<td>John Morton research: regression on hypothetical preference</td>
</tr>
<tr>
<td>1.1 - 3.3 %</td>
<td>Our estimates</td>
</tr>
</tbody>
</table>

4) performing the cost-benefit analysis

The wide range of estimates makes it difficult for railroad managers to draw any conclusion about how much revenue can be increased by improving service. In addition, marketing managers have no incentive to agree on a higher willingness to pay, since if they had agreed, they would have to be responsible for revenue increase after the project. On the other hand, market researches can be biased towards favorable recommendations for the project, since the researches are likely to be funded by the project team. A lack of trust on the demand analysis and the wide range of estimates of willingness-to-pay discouraged railroad managers from supporting the project. Although the project was justified at the 2 % willingness, the 2 % was much higher than marketing managers' assessment. Without consensus on demand response, Burlington Northern eventually had to drop the ARES project. Even though the ARES project was not implemented, its use of a freight choice model to evaluate revenue potentials of service improvement was well-conceived. We believe that railroads can benefit by adopting similar procedures for evaluating service improvement projects.
7.5. Cross-elasticity and the Market Strength Map

**Definition**

Similarly to own-elasticity, a disaggregate cross-elasticity represents the responsiveness of an individual's probability of choosing an alternative to a change in the value of some attribute of other alternatives. We can also calculate aggregate cross-elasticities by utilizing disaggregate elasticities. Mathematically, they are defined as follow:

\[
E_{x_{jnk}}^{P(i)} = \frac{\partial P_n(i)}{\partial x_{jnk}} \frac{x_{jnk}}{P_n(i)}
\]

and

\[
\eta_{ij} \equiv E_{x_{jnk}}^{P(i)} = \frac{\sum_{n=1}^{N} P_n(i) E_{x_{jnk}}^{P(i)}}{\sum_{n=1}^{N} P_n(i)}
\]

where

\[
x_{jk} = \frac{1}{N} \sum_{m=1}^{N} x_{jnk} \quad \text{and} \quad \bar{P}(i) = \frac{1}{N} \sum_{n=1}^{N} P_n(i).
\]

Cross-elasticities can be categorized into two groups. Cross elasticity of the form \((\eta_{ij})\) reports the percentage change in the share of mode \(i\) corresponding to a percentage change in the attribute value of a competing mode \(j\). Therefore, it represents the vulnerability of mode \(i\) to its competing modes. On the other hand, cross elasticity of the form \((\eta_{ji})\) is the percentage change in \(j\)'s share with a percentage change in \(i\)'s attribute, and represents the power of mode \(i\) to take share away from competing modes. Based on this notion, we define market power and market vulnerability as follow:

\[
\text{market power}_i = \frac{1}{N} \sum_{n=1}^{N} \sum_{j \neq i} P_n(j) E_{x_{jnk}}^{P(j)}
\]

\[
\text{market vulnerability}_i = \frac{1}{N} \sum_{n=1}^{N} \sum_{j \neq i} P_n(j) E_{x_{jnk}}^{P(i)}
\]
The two measures can be approximated with aggregate statistics as follow:

\[ \text{market power}_i \approx \left| \sum_{j \neq i} \bar{P}(j) \ \eta_j \right| \]

\[ \text{market vulnerability}_i \approx \left| \sum_{j \neq i} \bar{P}(j) \ \eta_j \right| \]

where the summation runs over all competing modes of mode \( i \).

The two measures capture the asymmetric effects of cross elasticities after considering market shares of all available alternatives. For negative attributes such as freight rate, transit time and damage percentage, market power and vulnerability indicate how choice probabilities would react to the decrease of the attribute value. For positive attributes such as equipment availability, on-time delivery, they indicate how choice probabilities would react to the increase of the attribute value. By taking the form of absolute value, the two measures represent how choice probabilities would react to the improvement of attributes. Clearly, high power and low vulnerability are desired.

In addition, we define the market strength map by drawing each alternative by a circle whose center is located according to the two axes of market power and market vulnerability, and whose area represents its average choice probability, \( \bar{P}(i) \). Railroad companies may draw the map for each product category or for origin-destination pairs, and can easily monitor their comparative strength in each market segment.\(^2\)

\(^2\) The two measures and the market strength map are first proposed in this paper. Cooper (1989) proposed the following similar measures:

\[ \text{Clout}_i = \sum_{j \neq i} \eta_j^2 + \eta_i^2 \]

\[ \text{Receptivity}_i = \sum_{j \neq i} \eta_j^2 + \eta_i^2 \]
The Market Strength Map of Freight Rate

Table 7.10. Cross-elasticities of Freight rate

<table>
<thead>
<tr>
<th>Changes in Choice Probability of</th>
<th>Changes in rate of</th>
<th>Truck</th>
<th>Rail</th>
<th>Intermodal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Truck</td>
<td>-0.139</td>
<td>0.123</td>
<td>0.048</td>
<td></td>
</tr>
<tr>
<td>Rail</td>
<td>0.385</td>
<td>-0.419</td>
<td>0.065</td>
<td></td>
</tr>
<tr>
<td>Intermodal</td>
<td>0.386</td>
<td>0.159</td>
<td>-0.574</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Truck</th>
<th>Rail</th>
<th>Intermodal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Market Power</td>
<td>0.087</td>
<td>0.101</td>
<td>0.049</td>
</tr>
<tr>
<td>Market Vulnerability</td>
<td>0.025</td>
<td>0.300</td>
<td>0.329</td>
</tr>
</tbody>
</table>

Table 7.10 reads such that the first column says that if trucking increases its rate by one percent, trucking will lose its share by 0.139 %, while rail will gain by 0.385 % and intermodal will gain 0.386 %. The cross-elasticities can also be interpreted in the opposite direction. If trucking decreases its rate by one percent, it will gain share by 0.139 % and reduce rail share by 0.385 %. While a change in truck rate has a major impact on other modes' shares, a rate change in the intermodal service does not influence shares of other modes except its own share. If intermodal service decrease its rates by 1 %, its share will rise by 0.57 % without hurting other modes' shares much. Thus, trucking has a large ability to take shares of other modes by reducing its rate, while intermodal service does not. Intermodal share is rather vulnerable to the rate changes of other modes.

What has happened in reality is that intermodal service was able to decrease its rates by 30 % or more since 1988 thanks to the use of double stack trailers, whereas trucking companies could not decrease their rates much since they was already operating

and Kamakura and Russell (1989) proposed the following measures:

\[
\text{Competitive Clout}_i = \sum_{j \neq i} \eta_{ij}^2
\]

\[
\text{Vulnerability}_i = \sum_{j \neq i} \eta_{ij}^2.
\]
at rock bottom rates. If this trend continues and if improvements in on-time performance of intermodal service continues, we can expect continued growth of intermodal revenue.

The market map in Figure 7.1 (a) show a summarized relationship of market power and market vulnerability. The map shows that trucking has high power to lead freight rate policy in the industry and a low vulnerability to other modes' pricing policies. Rail has high power but high vulnerability as well. Intermodal service has both small power and high vulnerability. Small values of market power indicate that freight rate does not influence the dynamics of modal shares much. Thus, railroad companies need not focus on freight rate as a strategic differentiation area. In fact, we believe that railroads’ practice of reducing rail rates less than reductions in trucking rates has ensured profitability of railroads in the 1990s without hurting revenue growth much.

**The Market Strength Map of On-time Delivery**

<table>
<thead>
<tr>
<th>Changes in Performance of Changes in Choice Probability of</th>
<th>Truck</th>
<th>Rail</th>
<th>Intermodal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Truck</td>
<td>0.370</td>
<td>-0.198</td>
<td>-0.069</td>
</tr>
<tr>
<td>Rail</td>
<td>-0.488</td>
<td>1.199</td>
<td>-0.082</td>
</tr>
<tr>
<td>Intermodal</td>
<td>-0.507</td>
<td>-0.246</td>
<td>1.561</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Truck</th>
<th>Rail</th>
<th>Intermodal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Market Power</td>
<td>0.112</td>
<td>0.161</td>
<td>0.069</td>
</tr>
<tr>
<td>Market Vulnerability</td>
<td>0.041</td>
<td>0.381</td>
<td>0.440</td>
</tr>
</tbody>
</table>

In terms of cross-elasticity, the market strength map, Figure 7.1 (b), shows that truck is not vulnerable to other modes' improvements in on-time delivery, whereas rail and intermodal are very vulnerable to trucking's improvement in on-time delivery. On the other hand, trucking already provides 91% of on-time performance, whereas rail is only 83% on-time. What could happen in reality is that the reliability of rail service could improve at a much faster rate than that of trucking. The market strength map could then shift greatly to the favor of railroad companies.
**The Market Strength Map of Transit Time**

Table 7.12. Cross-Elasticity of Choice Probability to Transit Time

<table>
<thead>
<tr>
<th>Changes in Performance of Changes in Choice Probability of</th>
<th>Truck</th>
<th>Rail</th>
<th>Intermodal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Truck</td>
<td>-0.024</td>
<td>0.147</td>
<td>0.040</td>
</tr>
<tr>
<td>Rail</td>
<td>0.113</td>
<td>-0.336</td>
<td>0.047</td>
</tr>
<tr>
<td>Intermodal</td>
<td>0.123</td>
<td>0.195</td>
<td>-0.274</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Truck</th>
<th>Rail</th>
<th>Intermodal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Market Power</td>
<td>0.026</td>
<td>0.121</td>
<td>0.039</td>
</tr>
<tr>
<td>Market Vulnerability</td>
<td>0.030</td>
<td>0.089</td>
<td>0.133</td>
</tr>
</tbody>
</table>

The market strength map in Figure 7.1 (c) confirms the fact that rail has the high power to change the market structure by improving transit time. While truck has low vulnerability, it has also low power. Rail and intermodal services will not lose shares much due to faster trucking service, since most shippers who need fast service are already using truck. On the other hand, rail has both high power and high vulnerability. The map indicates that If rail service can improve its transit time, truck shippers will switch to rail quickly. truck share is the most vulnerable.

**The Market Strength Map of Loss and Damage**

Table 7.13. Cross-Elasticity of Choice Probability to Loss and Damage

<table>
<thead>
<tr>
<th>Changes in Performance of Changes in Choice Probability of</th>
<th>Truck</th>
<th>Rail</th>
<th>Intermodal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Truck</td>
<td>-0.012</td>
<td>0.021</td>
<td>0.009</td>
</tr>
<tr>
<td>Rail</td>
<td>0.023</td>
<td>-0.065</td>
<td>0.014</td>
</tr>
<tr>
<td>Intermodal</td>
<td>0.024</td>
<td>0.027</td>
<td>-0.113</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Truck</th>
<th>Rail</th>
<th>Intermodal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Market Power</td>
<td>0.005</td>
<td>0.017</td>
<td>0.010</td>
</tr>
<tr>
<td>Market Vulnerability</td>
<td>0.004</td>
<td>0.018</td>
<td>0.024</td>
</tr>
</tbody>
</table>
Modal shares are not elastic to loss and damage. This may be explained by the fact that loss and damages are small and similar across all modes. The market strength map in Figure 7.1 (e) confirms that no mode stands out in market power and vulnerability. This implies that carriers should not try to differentiate in terms of loss and damage.

**The Market Strength Map of Equipment availability**

Table 7.14. Cross-Elasticity of Choice Probability to Equipment Availability

<table>
<thead>
<tr>
<th>Changes in Performance of Changes in Choice Probability of</th>
<th>Truck</th>
<th>Rail</th>
<th>Intermodal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Truck</td>
<td>1.238</td>
<td>-0.879</td>
<td>-0.318</td>
</tr>
<tr>
<td>Rail</td>
<td>-2.406</td>
<td>2.688</td>
<td>-0.380</td>
</tr>
<tr>
<td>Intermodal</td>
<td>-2.448</td>
<td>-1.047</td>
<td>3.610</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Truck</th>
<th>Rail</th>
<th>Intermodal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Market Power</td>
<td>0.546</td>
<td>0.714</td>
<td>0.320</td>
</tr>
<tr>
<td>Market Vulnerability</td>
<td>0.181</td>
<td>1.874</td>
<td>2.097</td>
</tr>
</tbody>
</table>

The market strength map in Figure 7.1 (d) shows that equipment availability can be a key strategic attribute that carriers should emphasize when they advertise their services or when their salesmen contact potential shippers. Rail and intermodal services are particularly vulnerable to other modes' improvements in equipment provision. On the other hand, rail has a high power to change modal shares by improving equipment availability.

Railroad companies may be able to improve equipment availability strategically and to communicate such improvement to shippers. But, such improvements in equipment availability will necessarily involve the lower utilization of equipment. Railroad top managers need to determine whether they will pursue such a revenue growth strategy even if it increases base costs and deteriorates its cost structure towards higher fixed costs.
Figure 7.1. The Market Strength Map
Chapter 8. Conclusion

8.1. The Significance of Research

In Chapter 3, we hypothesized that the objective of the logistics system changes from processing orders to creating customer value and to sustaining comparative advantage. Such trends provide both risks and opportunities to carriers. An increasing number of shippers (e.g. Gillette) plan to reduce the number of carriers they work with in order to develop strategic partnerships. A carrier that is not selected as a core carrier faces extreme difficulty in recruiting new customers. On the other hand, a carrier chosen as a core carrier can be assured of high volumes and a long-term relationship. Carriers with creative ideas can find growth opportunities, since shippers are more willing to outsource transportation services and to try new ideas for business network redesign. In order to attract shippers, the new ideas need to be customer-oriented in such diverse areas as developing customer-tailored equipment; contracting a network of lanes; designing differentiated services, etc.

Information technology can provide tools such as databases and mathematical models that are essential for designing customer-tailored service. In this thesis, we propose an information system to assist railroad managers in making effective marketing decisions. The system is designed to perform three tasks: collect market information, analyze market response, and generate marketing recommendations. Of the three tasks, the ability to analyze market response is the most critical but least developed. Thus, this thesis focuses on developing a model that analyzes shippers’ modal selection. Our proposed system is a model-based decision-analytic system that utilizes state-of-art statistical methodologies.
The methodology proposed in this thesis would be relevant to general merchandise and intermodal service. While intermodal service is the railroad industry’s biggest growth market (Welty 1994), it is also characterized by heavy competition with trucking, scheduled operations and short-term spot market. Due to heavy competition, forecasting accurately shippers’ response to changes in price or service performance is much more difficult than in other markets. Intensive use of market information is necessary. Accordingly, top managers are paying more attention to service design and pricing than to operational planning. Currently, most information about intermodal customers and demand projections comes from intermediaries such as trucking, ocean shipping, or third party logistics companies. Although these projections can provide rough estimates of shippers’ response to railroad marketing plans, they need to be complemented by direct analysis of actual shippers.

We expect that railroad marketing managers will in future make increasing use of freight demand models. The freight demand modeling approach proposed in this thesis provides a tool to help railroad managers analyze market response, prioritize service design projects and determine an appropriate pricing policy. It can also provide estimates of demand elasticities by market segments, shippers’ willingness to pay for service improvements, and the market strength maps for analyzing industry structure. Such information will be useful initially for service design teams and marketing sales managers. Instead of focusing on administrative paperwork, sales managers will be able to analyze shipment data in order to understand evolving patterns of shippers’ demand. The system could eventually be integrated with the corporate-wide decision support system to make it accessible to other managers, e.g. service design and strategic planning. Aligning operating plans with marketing strategy will provide significant advantages in competing against other modes. Furthermore, when a railroad owns a third party logistics company and tries to be a single source of contact (e.g. CSX), the modeling approach provides a method to analyze the revenue and profit potential of customers at the individual level.
8.2. Research Contributions

We believe that railroad companies can increase their shares of inter-city freight by monitoring customer needs regularly, applying a demand analysis model and readily updating marketing decisions. We also showed the feasibility of such a decision support system based on the freight choice model. Our research advances the freight choice model in the following three areas: representation of behavioral theories, utilization of multiple sources of data, and application of advanced statistical methodologies (Ben-Akiva 1995). Each of these areas represents an on-going research effort to conceptualize complex choice processes in quantitative terms.

Representation of Behavioral Theories

For an accurate representation of actual choice process, a freight choice model should explicitly reflect all factors that influence choice decisions. After reviewing the literature, we concluded that a freight choice model should reflect shippers' three objectives: total logistics costs, service quality and comparative advantage. We developed freight choice models that reflect such behavioral theory. In particular, we estimated a model where both total logistics costs and service quality influence modal selection.

Figure 8.1 shows the framework we adopted in this thesis. In the figure, ellipses represent unobservable constructs, while rectangles represent observable variables relevant to the problem context. Thick lines represent structural relationships; thin lines represent measurement relationships. This model assumes that shippers choose the "best" mode in terms of total logistics costs and service quality perceptions, both of which are determined by characteristics of shippers, shipments and modes. The framework also indicates that although we cannot observe perceptions and preference directly, we can measure their indicators.
First, a total logistics cost model was developed both with and without heterogeneity of discount rate. In order to estimate the framework, we perform detailed analyses of total logistics cost components which include transportation costs and inventory holding costs. Since total logistics costs items are easily quantified, their influence on modal selection is easy to estimate. We first estimate a model while assuming that the discount rate which shippers employ in calculating inventory holding costs is fixed and the same for all shippers. We then estimated a model with random discount rates. Although we expected that the incorporation of the heterogeneity of discount rates would improve the estimation since the assumed discount rate will differ by shippers and by shipments, the random rate models failed to improve the fixed rate model. We also considered the possibility of correlated observations. If shipper-specific discount rates persistently influence multiple observations from the same shipper, parameter estimates can be inefficient. Our results show that such effect is small in our data set.
Secondly, a model that estimates the impacts of service quality was developed. Perceptually, many shippers may prefer truck despite the higher total logistics costs than rail. Quality perceptions are difficult to analyze since they are unobservable. Because of the measurement problem, researchers typically ask shippers directly to rate the importance of service attributes on a Likert scale (e.g., 1 to 7 depending on the level of importance). While this approach is easy to administer, the reliability of stated importance in explaining actual modal shares is unclear. We, instead, treat the ratings as indicators, extract instrumental variables of perceptions from the ratings and estimate their importance in modal selection. The incorporation of perceptions into a choice model improves the data fit and prevents the inconsistency of parameter estimates due to omitted variables.

Thirdly, we show how the estimates of demand models can be applied to actual decisions such as market segmentation, pricing and industry analysis. We proposed a measure of shippers' willingness to pay for improved service based on own-elasticities and proposed the market strength map for analyzing industry structure based on cross-elasticities.

**Utilization of Multiple Sources of Data**

In order to estimate demand response accurately, we can utilize multiple sources of data. In particular, this thesis utilizes three database: actual shipment records (revealed preference), choices under hypothetical situations (stated preference), and perceptual indicators. Stated preference data include responses to questions such as "Which mode would you prefer in a certain situation?" and "What alternative do you consider when you ship from A to B?". Perceptual indicators include responses to questions such as "How important do you think it is for a carrier to provide on-time delivery?" and "How safe from unexpected incidents do you feel when you ship through rail?". Data can be
collected either from shippers or from railroad salesforce. These indicators provided a large database that can improve a demand model.

In this thesis, we discuss ways to combine different sources of data. Since each data has different characteristics, it is not easy to analyze them together. But, since each data source has its own advantages, the combined dataset can increase the estimation efficiency and the external validity of research results. For example, data on actual choice contexts are most desired for explaining actual choices but are difficult to collect. On the other hand, stated preference data are easy to collect and easy to define choice contexts. By showing actual application, our research advances toward the utilization of multiple sources of data.

Application of Advanced Statistical Methodologies

In order to achieve the better representation of behavioral theories and to utilize multiple sources of data, we need methodologies more advanced than the popular multinomial logit. We believe that the approaches proposed in this thesis can provide a useful guideline for future research in this area.

First, we propose the use of the minimum discrimination information (MDI) estimation in order to analyze modal share data. As an extremum estimator, the MDI estimator is consistent and asymptotically normal. We further extend the concept into the MDI estimator with expected probability for the case of random rate model, the expected MDI estimator for the case of multiple observations from each shipper, and the combination of MDI and MLE to analyze RP and SP data jointly.

Secondly, we apply the concept of structural equations to capture the effects of shippers’ quality perception. Although service perceptions are unobservable, we could extract perceptions from perceptual indicators and estimate their importance by combining
structural equation models with a discrete choice model. To the best of the author's knowledge, such combination in multinomial choice setting is pioneered in this thesis.

Thirdly, we propose the use of simulation for parameter estimation in order to overcome a numerical evaluation problem inherent to complex model specifications. When we combine RP and SP data or when we combine discrete choice model and structural equations, the likelihood becomes complicated. Simulation allows us to estimate consistent parameters of the complex system even if a numerical evaluation of choice probabilities is difficult. We believe that simulation will expand the horizon of econometrics.

8.3. Future Research

There are several aspects in which the proposed model should be improved. We propose such future research in terms of the three research themes we have been pursuing; representation of elaborate behavioral theories, utilization of multiple sources of data and development of advanced statistical methodologies; and also in terms of managerial implications.

Representation of Elaborate Behavioral Theories

First, we focus on total logistics costs and service quality as shippers' objectives in logistics decision making. Yet, as we discussed in Chapter 4, strategic differentiation becomes a more and more important consideration in selecting modes and carriers. How to define comparative advantage and measure its importance by using the freight choice model is an important task. We believe that the approach of measuring service quality, i.e. using indicators, can be applied for this purpose. Proper indicators need to be found.

Second, the total logistics cost model can be improved by modeling shipment size and mode jointly. Since shippers would prefer rail service for large shipment sizes and
trucking for small shipment sizes (i.e. selectivity bias), this approach can improve the accuracy of estimation. Parameters that maximize the likelihood of observing modal share and shipment size simultaneously can be estimated by writing the likelihood as follow:

1) $L_n(\theta) = P_n(\text{modal share} \mid \text{shipment size}) \cdot P_n(\text{shipment size})$

2) $L_n(\theta) = P_n(\text{shipment size} \mid \text{modal selection}) \cdot P_n(\text{modal selection})$

The first approach derives shipment size from the profit function by using the well-known proposition of Roy’s Identity (McFadden, Winston and Boersch-Supan 1985). The second approach fits a choice model first, and fits a shipment size model after correcting selectivity bias (Heckman 1979). The first approach is preferred since industry practice seems to follow the sequence of deciding shipment size first and transportation mode second, and since it can be applied to aggregate share data. This thesis did not try either approach since shipment size data were not available.

Third, one particular characteristic of freight choice models is that the major element of the shipper's objective function, i.e. total logistics costs, is defined and measured in dollar terms. In most applications of random utility model, the utility is an ordinal function that is unique only up to order-preserving transformation. The absolute level of utility does not matter as long as the ordering of utilities is preserved. On the other hand, in the freight choice model, the absolute level of utility matters and its numerical assignment is unique and can be added and multiplied. This cardinality property suggests that not only the expectation but also the variance of utility should enter the decision process. Shippers may design their optimal portfolio of modes and carriers based on expected utilities and perceived risks, similarly to what investors do in financial markets. We tried to estimate the expected utility model assuming that the risk of total logistics costs is proportional to the variance of on-time performance. The result was not as good as we expected, and we did not pursue the implications of the cardinality property further. But these certainly are interesting topics for future research.
Fourth, further study of market segmentation is needed. The random coefficient model we showed in Chapter 4 assumes that the heterogeneity of shippers are continuously distributed. If there are discrete segments in the shipper population (e.g. freight rate-sensitive segment, transit time-sensitive segment, on-time delivery-sensitive segment, etc.), it would be more effective to estimate a separate set of coefficient for each segment. Since we do not know a priori what segment a shipper belongs to, class membership is latent. Recent developments show that if we have appropriate explanatory variables, we can estimate the membership probability (referred as the latent class model, Gopinath 1995). This area deserves more study, to promote understanding of shipper segments.

Fifth, the framework in figure 8.1 can be further elaborated by incorporating attitudes and choice set generation (Figure 8.2).

![Diagram of Static Freight Choice Model](image)

Figure 8.2. The General Framework of Static Freight Choice Model
*Attitudes* are variables that influence shippers’ tastes or importance of attributes. Attitudes are generic to alternatives, while perceptions are alternative-specific. For example, some shippers are time sensitive and some shippers are reliability sensitive. The same time-sensitive shipper may perceive truck as reliable and rail as unreliable. The incorporation of attitudes, thus, allows us to model the heterogeneity of decision makers. *Latent choice sets* differentiate decision makers who are loyal to certain alternatives and those who are switching among alternatives. It also implies that alternatives exogenously available at a choice situation and alternatives a shipper considers at a specific choice context may not be the same. Differences in shippers’ choice sets may come from differences in a shipper's attitudes (e.g. loyalty to an alternative), perceptions (e.g. riskiness of trying unfamiliar alternatives), decision protocols (e.g. reduction of choice set into a manageable size by using a lexicographic or satisfaction rule), and situational constraints (e.g. lack of time for search or evaluation).

Sixth, the static model can be extended further into the dynamic model when multiple observations for shippers are available over time. Extending the static model, the dynamic model assumes that once a shipper experiences the service and cost level of a certain mode, he forms two latent constructs: captivity and satisfaction. *Captivities* are alternative-specific variables that influence a choice set generation process for the next choice. For instance, when a shipper repeatedly uses intermodal services, he may gain administrative efficiency in shipping through intermodal service by investing in EDI, purchasing intermodal trailer, or hiring an intermodal agent. Such changes increase transaction costs to other alternatives and captivity to the repeatedly-chosen alternative. *Satisfactions* are another of the alternative-specific variables that update perceptions about the performance of each alternative. The dynamic process in which satisfactions influence perceptions is documented in Boulding et al. (1993). While the dynamic model will improve an understanding of shippers’ decision process, it requires panel data for which a group of shippers are surveyed over time.
Utilization of Multiple Sources of Data

First, the issue of how to structure existing data requires more study. Railroad managers should figure out how to handle the massive volume of data that can waste management time and distract management attention. Not only do they have to integrate the variety of data sources such as customer service records, sales call reports and market research records, but they also have to coordinate the interrelationships among various marketing reports such as sales quotas, schedules, forecasts, marketing plans and marketing messages. Quantification of customer service records and sales call reports are desired. A way of documenting salesmen's knowledge about shippers' perceptions and choice environments is necessary. If possible, multimedia surveys can also be devised.

Second, a shipper panel may be maintained. A panel is a sample of shippers whose decisions are followed over time so that they provide a choice history as well as multiple
observations. By carefully designing the composition of a panel, railroads can collect both cross-sectional and longitudinal data. Cross sectional data provides information about the response of different market segments. Longitudinal data provides information about changes of shipper behaviors responding to railroad marketing decisions. Such data allow researchers to distinguish inter-individual differences from intra-individual differences, and thus, help identify economic models and discriminate between competing economic hypotheses. Panel data can also help reduce estimation bias, and reduce problems of multicollinearity (Hsiao 1986).

**Development of Advanced Statistical Methodology**

First, more study on simulation is necessary. Simulation is useful when a likelihood is difficult to evaluate. Such cases occur frequently as we specify a large system of relationships and try to estimate the system of equations. For instance, simulation allows us to estimate a total logistics cost model with random and agent effects, a service quality perception model, a correlated probit model, a model of determining shipment size and mode jointly. Statistical properties of parameters estimated using simulation are now well studied and established. Currently a major problem is how to reduce the variance of parameters due to simulation, or how to achieve the same level of accuracy in parameter estimates with a small number of draws. Variance reduction may be feasible with a well-designed simulation scheme; more study on this subject is necessary.

Secondly, if railroads chose to maintain a shipper panel, econometric studies are necessary to understand statistical properties of a choice model using panel data. The special features of panel data create difficult estimation problems in freight choice models. Individual heterogeneity, state dependence, and serial correlation in the disturbances need to be distinguished from each other (Heckman 1981). The agent effect model that we applied in this thesis can be used for analyzing panel data. When shippers are observed many times over a certain time period, the effects of shipper-specific tastes persist. The
agent effect model provides a useful methodology to analyze inter-temporal variations after capturing the effect of heterogeneity.

Thirdly, a study of comparing the MDI estimation and the generalized IPF estimation is desired. Both uses the ratio between actual \((s_n)\) and predicted shares \((F_n(\theta))\). If the shares are exactly the same, their ratio will be one. Now, consider the following error function:

\[
\text{Error}_{\text{in}}(\theta) = \frac{F_n(\theta)}{s_n} \ln \left( \frac{s_n}{F_n(\theta)} \right) - \frac{s_n}{F_n(\theta)} + 1
\]

which has the following form:

![Error function graph]

The generalized IPF minimizes the following weighted error:

\[
Q_N(\theta) = \sum_n \sum_i s_n \text{Error}_{\text{in}}(\theta)
\]

\[
= \sum_n \sum_i \left[ F_n(\theta) \ln \left( \frac{s_n}{F_n(\theta)} \right) - s_n + F_n(\theta) \right]
\]

with the following score vector:

\[
S_N(\theta) = \sum_n \sum_i \ln \left( \frac{s_n}{s_n(\theta)} \right)
\]

Note that the MDI minimizes the following information discrepancy:

\[
Q_N(\theta) = \sum_n \sum_i s_n(\theta) \ln \left( \frac{s_n}{F_n(\theta)} \right).
\]
Promotion of Managerial Applications

First, managerial implications of RMSS should be studied. Issues such as what information railroad marketing managers desire, in what format, and how often have to be studied carefully in terms of establishing the Marketing Report System. How to structure managers' knowledge into databases and how to connect with EDI system is also important in terms of building the Marketing Information System.

Second, for policy makers, our model can be extended into a model that forecasts economic, energy and environmental impacts of various social policies (e.g. such as legislatures to impose a higher gas tax, to subsidize the construction of intermodal terminals, or to impose congestion tolls in the highway). We may extend RMSS to support such public transportation planning. A basic framework for such a model is as follows. For further discussions, refer to Roberts et al. (1976).

Figure 8.4. System Configuration of a Modal Share Model for Policy Makers
Appendix 1. Case Study

Gillette: A Strategic Transition of Transportation Department

In August 1994, John Donovan, Director of Transportation at Gillette, was busy preparing for the annual Strategic Planning Board. He has to explain the performance of the transportation division to the Board. While summarizing the logistics system of Gillette, he began to think about new trends in transportation that Gillette and its competitors are currently observing.

Among them, outsourcing transportation activities is the most notable trend. John always believed that the transportation activity is one of Gillett's core competencies that provides comparative advantage over major competitors. But, he also knew that there are some people, both within the company and among investors, who think that Gillette's core technology is to design a new product, manufacture and market the product and that distribution and transportation activities do not add value as much as the designing, manufacturing or marketing activities. He, however, had few quantitative data that showed the value created by his division, or compared the benefits and risks involved with outsourcing transportation activity.

He would also need to ask for an increase in the annual budget for the transportation division. Retailers were willing to share sales information and many of Gillette's competitors were spending heavily in integrating point-of-sales information directly into distribution planning. He wanted to study how Gillette should change the current transportation planning process in order to utilize the readily available sales data.

Company Backgrounds

In 1895, Mr. King Gillette, a salesman for the Baltimore Seal Company, originated the idea of a disposable razor blade while shaving with a dull straight razor at his home in Brook-line, Massachusetts. In 1901, MIT machinist William Nickerson joined with Gillette and perfected the safety razor. With the financial support of some wealthy friends, the 2 men formed the American Safety Razor Company in Boston. Gillette put his safety razor on the market in 1903, selling only 51 sets the first year. But the good news spread fast, and in 1904, Gillette sold 90,844 sets. In 1905, he established his first overseas operation in London and during the World War I, the company sold 3.5 million Service Set shaving kits to the US government.

---

1 This case is prepared by Jai-Kue Park under the supervision of Senior Research Associate Carl Martland for the class of Freight Transportation Seminar at the Center for Transportation Studies, M. I. T. This case is written to serve the basis for class discussion rather than to illustrate either effective or ineffective handling of an administrative situation.
In 1948, the company began diversifying by purchasing Toni (a home permanent kit), which became the Personal Care Division in 1971. During the 1950s, Gillette adopted its present name, introduced Foamy (shaving foam, 1953), and bought Paper Mate (pens, 1955). During the 1960s and 1970s, the company expanded its product line further by introducing Right Guard (deodorant, 1960), Trac II (twin-blade razors, 1971), Cricket (disposable lighters, 1972; sold to Swedish Match, 1984), Good News (disposable razors, 1975), and Eraser Mate (erasable pens, 1979).

The company branched into related markets by acquiring Braun (electric shavers and appliances, 1967), Liquid Paper (correction fluid, 1979) and Oral-B (dental products, 1984). But, the 1989-90 purchase of Swedish Match (maker of Wilkinson Sword razors) has been problematic, since the US and the UK and other EC governments opposed the high market share it gave Gillette. In 1991, Gillette was ordered to divest the UK portion of Swedish Match, and in 1992, it was ordered by antitrust officials to sell its European holdings in Wilkinson Sword.

The company believe in the new product leadership. The company launched the successful Sensor razor campaign in 1990; introduced the Gillette Series, a line of men's shaving and skin products in 1992; and launched the Cool Wave line of male grooming aids in 1993. In early 1994, The company announced plans to roll out the SensorExcel in the US later that year. This trend will continue. The company launched 20 products, its most ever, in 1994 with at least 20 more due in 1995. The company is also active in establishing global presence. In 1993, Oral-B opened a joint venture toothbrush manufacturing facility in Shanghai, China. The company introduced the SensorExcel razor in Europe and Canada.

Products, Markets and Competitors

Gillette is an international company with annual sales of $5.4 billion in 1993. Its sales have been growing 8-10% every year, whereas its profits grew at a rate of about 15%. Since it grew during the recessionary periods (1988-1992), its stock price grew at an annual 35% rate (vs. 16% for S&P 500). Refer to Exhibit 1 for recent financial data.

Gillette has 5 product lines that are all market leaders. Its razor and blades division (including Trac II, Atra, and Good News) account for about 39% of sales and about 75% of profits. The company recently became the world's top seller of writing instruments when it added Parker Pen Holdings to its stationery products division, which was already a leading international seller of pens and correction fluids (e.g. Paper Mate, Waterman, and Liquid Paper). The company's Oral-B division is also a leader in US toothbrush sales, and Braun makes a variety of small appliances and is one of the world's leading producers of electric shavers. Its toiletries division produces such well-known brand names as Right Guard, Foamy, and Dry Idea. Refer to Exhibit 2 and 3 for the performance, major brands and competitors of each product line.
Gillette sells its products in more than 200 countries and manufactures them at 62 locations in 28 countries. Gillette continues to expand its geographic reach. It is restructuring its worldwide operations to take advantage of emerging global opportunities, a move that will affect about 2,000 jobs, mostly outside the US. As part of the restructuring, Gillette took a realignment charge that reduced pretax profit from operations by $263 million in 1993. Refer to Exhibit 4 for the global presence data.

**Distribution and Logistics System Configuration**

**Organizational Structure**

Exhibit 5 shows the organizational chart of Gillette's distribution and transportation division. Gillette operates 4 distribution centers (D.C.). Gillette has been reducing the number of D.C.'s (e.g., close Dallas and Atlanta D.C.'s). Consolidation is also occurring in manufacturing. Shampoo used to be produced at two plants but has since been consolidated into a single plant. While such movements increased distribution costs and response time, they have been more than offset by smaller inventory and lower periodic costs in production and purchasing; It was also believed that the consolidation saved labor costs by eliminating personnels related with duplicated purchasing and material handling activities. Quantifiable statistics were not available about performance measures before and after the change.

Braun and Oral-B are subsidiary companies, not divisions, and have a separate distribution system. However, since their customers are similar to those of blades and toiletries, joint operation may decrease costs. For instance, customers of personal care products, blades and Oral-B are the same supermarkets, discount stores or warehouse clubs. Braun sells at department stores such as Lechmere as well as at Wal-mart. Thus, Braun and Oral-B use major customers of the Gillette distribution system. Stationeries is also a part of Gillette's U.S. and Canada distribution system. The company has a separate distribution system for the international market.

**From D.C. to Customers**

Gillette classifies customers into two categories: major and minor. Major customers are primarily large retail stores (e.g., Wal-mart, K-mart, etc.) and supermarket chains (e.g., Stop and Shop, H.E.B, Kroger, etc.). Gillette ships directly to the large buyers' D.C.'s from its own D.C.'s (e.g. Andover and St. Paul). For these shipments, trucking is the main mode. Refer to exhibit 6 for the network configuration.

In addition, Gillette has many small customers around the nation. For small buyers, Gillette divides the U.S. into 4 regions and supports them by employing pool distributors (P.D.). P.D.'s are either public warehousing companies or small trucking companies with a warehouse. Refer to exhibit 7 for the P.D. network configuration. They operate their own trucks and warehouses, and are required to establish EDI with Gillette. Most P.D.'s come to D.C. once a week to receive Gillette's products which they distribute to local retailers.
They are responsible for local shipments. They also arrange potential shipments during empty front-haul movements. It is not clear as to whether such a scheduling of operation to make up a certain shipment size (e.g., a fixed day ordering and shipment) is cost efficient (in sacrifice of fast customer service). Costs may differ depending on the sales volume and the variability due to seasonality or due to local promotion.

The overall network is summarized as follows:

<table>
<thead>
<tr>
<th>Plants</th>
<th>Products</th>
<th>D.C.</th>
<th>Regions</th>
<th>Direct shipment to customer (million lbs.)</th>
<th>Number of P.D.'s Supported</th>
</tr>
</thead>
<tbody>
<tr>
<td>Andover</td>
<td>blades, deodorants shaving</td>
<td>Andover</td>
<td>Northeast</td>
<td>39.5 (22.9%)</td>
<td>9</td>
</tr>
<tr>
<td>Boston</td>
<td></td>
<td>(MA)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Chicago</td>
<td>St. Paul</td>
<td>Midwest, Texas, Florida</td>
<td>15.0 (4.4%)</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>(IL)</td>
<td>(MN)</td>
<td>Mid- &amp; northwest</td>
<td></td>
<td></td>
</tr>
<tr>
<td>St. Paul</td>
<td>hair care</td>
<td>Ontario</td>
<td>Southwest &amp; west-coast</td>
<td>7.5 (8.7%)</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(CA)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Santa</td>
<td>stationery</td>
<td></td>
<td></td>
<td>172.1 (65.0%)</td>
<td>39</td>
</tr>
<tr>
<td>Monica</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Gillette decided to maintain Chicago D.C. since about 30% of sales are made within 100 miles from Chicago. But note that P.D.'s in Florida, Georgia and Texas are supported from Chicago, even though the distance between them are quite large. Rail or intermodal service may be cost-efficient.

Overall, Gillette utilizes core carriers as shown in the below table. Each D.C. is assigned a list of primary, secondary, and back-up carriers so that traffic managers at D.C.'s do not need to think about which carrier they will use for specific shipments (centralized core carrier and centralized carrier assignment). Carrier scheduling and dispatching are done by each D.C., based on daily requirements.

<table>
<thead>
<tr>
<th>Carrier Type</th>
<th>Number of core carriers</th>
<th>Shipment composition (volume)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TL (from D.C. to Large Customers)</td>
<td>6</td>
<td>53%</td>
</tr>
<tr>
<td>TL (from D.C. to P.D.)</td>
<td>38</td>
<td>35%</td>
</tr>
<tr>
<td>LTL / UPS</td>
<td>3</td>
<td>12%</td>
</tr>
</tbody>
</table>

From Plant (D.C.) to D.C.

For shipments between D.C.'s, Gillette utilizes the following core carriers:
Note that since the distance between Chicago and St. Paul is quite small, trucking is more efficient for this corridor. For other corridors, Gillette uses intermodal services widely. Delivery performance is of less concern. For example, 6 days is considered satisfactory for transporting a shipment from Chicago to L.A. and even boxcar performance was considered satisfactory. Trucking is used only when shipments needed to be delivered quickly. The main reason that Gillette uses the intermodal service is its cheap rate. Note that a large volume of Gillette's shipments moves from Boston to Chicago and from St. Paul to California. For those west-bound shipments with high flexibility in shipment schedule, Gillette can negotiate a good rate with intermodal carriers who have to move empty trailers to Chicago and to California. Given the fact, the intermodal rate is actually cheaper than the boxcar rate.

**From Supplier to Plant**

Total inbound tonnage is 150 million pounds per year. Gillette has a program to select and recommend its preferred carriers to vendors. It is believed that Gillette can negotiate better rates and impose stricter service standards than small vendors. Most of them are TL and LTL shipments, and the following table shows the major movements:

<table>
<thead>
<tr>
<th>Source</th>
<th>Destination</th>
</tr>
</thead>
<tbody>
<tr>
<td>New England</td>
<td>Andover, MA</td>
</tr>
<tr>
<td>Metro NY, Northern NJ</td>
<td>South Boston, MA</td>
</tr>
<tr>
<td>Erie, Corry, Milroy, Philadelphia, PA</td>
<td></td>
</tr>
<tr>
<td>Midwest</td>
<td>Evansville, IN</td>
</tr>
<tr>
<td>Chicago, IL</td>
<td>St. Paul, MN</td>
</tr>
<tr>
<td>Kettering, OH</td>
<td>Racine, WI</td>
</tr>
<tr>
<td>South</td>
<td>Gordonsville, VA</td>
</tr>
<tr>
<td>Atlanta, GA</td>
<td>Marion, SC</td>
</tr>
<tr>
<td>Charlotte, NC</td>
<td></td>
</tr>
<tr>
<td>New England</td>
<td>Newburyport, MA</td>
</tr>
<tr>
<td>All of NJ</td>
<td>St. Paul, MN</td>
</tr>
<tr>
<td>Philadelphia, Williamsport, PA</td>
<td></td>
</tr>
<tr>
<td>Midwest</td>
<td>Chicago, Vandalia, IL</td>
</tr>
<tr>
<td>Peosta, IA</td>
<td>Racine, Janesville, WI</td>
</tr>
<tr>
<td>Ligonier, IN</td>
<td></td>
</tr>
<tr>
<td>South</td>
<td>Forest Park, GA</td>
</tr>
<tr>
<td>West</td>
<td>L.A., CA</td>
</tr>
<tr>
<td></td>
<td>Dallas, TX</td>
</tr>
</tbody>
</table>
In addition to the above shipments from small vendors, Gillette has bulk tanker shipments as follows:

<table>
<thead>
<tr>
<th>Source</th>
<th>Destination</th>
</tr>
</thead>
<tbody>
<tr>
<td>I North Chicago</td>
<td>St. Paul, Andover</td>
</tr>
<tr>
<td>II Points in NJ</td>
<td>St. Paul, Andover, North Chicago</td>
</tr>
<tr>
<td>III Baltimore, MD</td>
<td>St. Paul</td>
</tr>
<tr>
<td>IV Kankakee, IL</td>
<td>St. Paul</td>
</tr>
<tr>
<td>V Janesville, WI</td>
<td>St. Paul</td>
</tr>
</tbody>
</table>

Distribution and Logistics Performance

**Customer Service Standards**

Gillette primarily measures the order fill rate in terms of lines filled/lines ordered, dollars shipped/dollars ordered, and the number of invoices shipped completely/total invoices. Order cycle lead time needs to be at most 10-12 days from receipt of order to customer (more specifically, 1-5 days for order transmission, 2 days for order processing, and 3-4 days for delivery). Gillette has an internal (but not explicit) policy on customer service standards, accuracy in order filling, rush service, product return, and response to customers' schedule changes.

**Total Logistics Costs**

Gillette estimates that approximately 4-5% of sales are spent on total logistics costs for shaving, personal care and stationery divisions. Direct costs (freight) are about 2-3% and period costs (warehousing and administration) are about 2-3%. A review of the annual report shows the following inventory level:

<table>
<thead>
<tr>
<th></th>
<th>1992</th>
<th>1991</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$ million</td>
<td>days of sales</td>
</tr>
<tr>
<td>Net sales</td>
<td>$5,162.8</td>
<td></td>
</tr>
<tr>
<td>Raw materials and supplies</td>
<td>$192.4</td>
<td>13.6</td>
</tr>
<tr>
<td>Work-in-process</td>
<td>$94.1</td>
<td>6.7</td>
</tr>
<tr>
<td>Finished goods</td>
<td>$565.9</td>
<td>40.0</td>
</tr>
</tbody>
</table>

Inventory levels vary by product line. Products like Sensor, which is in high demand, may have 2-3 weeks of inventory. On the other hand, hair care products which deplete slowly may have a few months of inventory. People tend to think that companies will maintain a high level of inventory for fast-moving items and low level for slow-moving items. But, such a presumption can be wrong depending on the flexibility of the production technology. In fact, Gillette may produce hair care products only twice a year in order to reduce production set-up costs. Whether such a practice is good or not is a
managerial decision. Certainly such a practice is not responsive to shortening product life cycle and changing consumer demand, but it may be cost-effective. With a high level of inventory, transit time may not be important. Reliability of delivery time, trust-worthiness of a carrier, and shipment tracing capability combined with cheaper freight rate may be more important.

**Core Carrier program**

As described in section 2, Gillette has established a strong core carrier network since 1985. Initial major selection criteria were carriers' experience with similar products and volume with Gillette's customers. The reasons are that if a carrier handles similar products, the carrier would understand customer requirements well and that if a carrier delivers many products to Gillette's customers, the carrier might achieve cost efficiency from achieving economy of scale.

Gillette measures performance of its core carriers continuously. Each month, Gillette evaluates carriers on the following items:

- accordance to delivery standards
  (i.e. ___ days from order receipt and within ___ hours from appointment time)
- response to special requirements
- billing accuracy
- loss and damage records

At the end of a year, Gillette selects "The Carrier of the Year" based on the performance records. In addition, Gillette monitors customers' perception of carriers' service quality through a quarterly telephone survey and semi-annual customer visit. Every quarter, Gillette surveys key account customers by telephone. Gillette has more than 100 major accounts. In addition, A joint team of marketing, sales and distribution visits major customers (usually their purchasing and operations departments) twice a year. The objective of the visit is broad, targeting the improved understanding and communication among each other. Yet, the discussion topics include customer service, palletization and shipment methods.

Gillette plans to keep reducing the number of carriers that it utilizes in its distribution system. In general, Gillette does not want to change carriers often. Its policy is that new carriers will be brought into the carrier list only when existing carriers fail to provide sufficient advantage in price and service, and that existing carriers will be given an opportunity to respond to a proposal of a new carrier. The following items are the major current criteria for considering a new carrier:

- competitive rate
- financial stability
- safety records
o service and quality commitment
o evaluation of customers
o EDI capability
o understanding of Gillette procedures and needs (e.g., P.D., palletization, etc.)
o reasons for change from current practice (e.g., route change)

**Electronic Data Exchange**

Gillette requires all core carriers to establish and communicate through EDI. The process is as follows:

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1) Gillette sends a transportation order one day before shipment.</td>
<td></td>
</tr>
<tr>
<td>2) Carrier replies with the intended delivery time</td>
<td>to both the sales and the distribution department.</td>
</tr>
<tr>
<td>3) Gillette notifies the receiver.</td>
<td></td>
</tr>
<tr>
<td>4) Carrier delivers shipments.</td>
<td></td>
</tr>
<tr>
<td>5) Carrier sends the name of the individual who received the shipments.</td>
<td></td>
</tr>
<tr>
<td>6) Gillette notifies its bank for payments.</td>
<td></td>
</tr>
</tbody>
</table>

All TL and LTL/UPS core carriers have EDI capabilities. They are willing and have the resources to work with Gillette in centralizing dispatching, vehicle routing, or scheduling. Most P.D.'s are small independent distributors. They are more reliant on the old system and tend to think that things go well with most weekly deliveries. Many of them think that EDI is not really necessary (and that a fax can do the same job with less cost) and Gillette had to press them to establish the EDI communication.

**Strategic Options for Transportation**

Gillette strongly believes in new product leadership. Challenges imposed on the transportation division are how to cope with shortened product life cycle and increased product variety. There are many thousands product models (e.g. SKU's) that are constantly changing. Their palletization, distribution and transportation requirements are diverse and changing. Moreover, Gillette's customers are increasingly demanding short time windows for delivery. If they used to ask for delivery to a store every Tuesday and were happy as long as Gillette delivered without too much discrepancy from the scheduled date, they now wants Gillette to deliver between 9:30 and 9:45 a.m. every Tuesday. Such a trend is unavoidable since retailers themselves suffer from congestion in the receiving docks due to increased product variety, increased number of suppliers and increased shelf space to fill up. Manufacturers with superior logistics capability (e.g. coordinating multiple items to the same customer) are now strongly favored by store managers. Retailers would appreciate if a manufacturer can consolidate smaller lots of larger variety and deliver to them frequently. They are now willing to provide day-to-day inventory information calculated directly from point-of-purchase data. Incorporating sales information directly into logistics planning has become a hot issue recently, as can be inferred from many
managerial buzzwords such as Continuous Replenishment, Quick Response, and Efficient Consumer Response.

In addition, John has been asked by the Strategic Planning Board how much value the transportation division is adding to the company. Consider the case of Benetton which maintains only design, marketing and purchasing teams. Once the design team in Italy predicts next year's fashion, the purchasing team in Japan secures fabric, the purchasing team in Korea arranges for production, and the marketing team contacts potential retailers in the U.S. Everything else is outsourced. By focusing on core competencies, a company may achieve the greatest value added per employee. It is very interesting that Benetton outsources distribution, because inventory management is presumably very important in the apparel industry. Apparel inventory can lose its value dramatically once a season is over or if new models do not make a hit.

Gillette tries to provide similar benefits to all of its employees. If a division does not add value as much as other divisions, it would not make sense to maintain the lagging division. Thus, the current trend towards re-engineering would trim the organization until all the remaining employees possessed their own core technology and provided similar amounts of value to the company. One target area can be the P.D. system. Should Gillette continue to use P.D.'s, even when their service schedules are long and rigid and their cost advantages are eroding? Another area is the inventory planning process. If Gillette updates inventory and transportation plans directly based on point-of-sales data, ships goods overnight, transforms warehouses into cross-docking stations that run 24 hours, the inventory level would be drastically reduced. Having reviewed such situations, John asked what alternatives are available in order to cope with such changing demands for transportation.

**Strategic Investment** The first option is to develop transportation as one of core technologies that Gillette's resources perform better than other third-party logistics companies and as one of the competitive advantages that Gillette has to maintain over its competitors. Gillette might be able to improve the efficiency of the transportation organization by consolidating local control and pursuing global optimization. Performance will improve, for example, if Gillette finds the best vehicle dispatching/routing plan that coordinates inbound and outbound movements or that enables a continuous movement of vehicles. John heard that Mobil Oil Corporation, DEC, and Abbott Labs claim savings of 10 to 15% of the annual transportation budget by centralizing control. The core of their control process was to build mathematical models that analyze the network structure and figure out the best vehicle schedule. This approach has a merit in that if built properly, the model-based system can provide a great degree of differentiation from other competitors. This approach, however, assumes that Gillette has human resources who can run the model and interpret the results. It would also require large investments for developing the model, restructuring the whole logistics system, and training all personnel. Since Gillette has not implemented a similar approach before, this approach involves a huge risk of failure. It is also not clear as to how quickly Gillette can recover the initial investments.
**Strategic Partnership** The company may achieve centralized control by building a strategic partnership with a carrier and to let the carrier optimize globally over its own network. The current system utilizes multiple carriers whose levels of employee training, local service quality, and/or price formats may be different. By building a close relationship with one or two carriers, Gillette can provide opportunities for the carriers to improve productivity. The selected carriers can provide uniform data flow, integrate into Gillette’s transportation needs, and standardize customer service. Presumably, the carriers will find a better solution than Gillette since they will optimize over a larger network. John heard about the new partnerships between DuPont and Roadways and between Kodak and Bekins. Gillette has already been moving in this direction, but the trend can be further strengthened by eliminating the P.D. system and/or by eliminating Chicago D.C. This approach, however, involves transferring a significant part of current transportation planning activities to the carrier. It is not clear as to how much Gillette can trust the service quality and commitment of the carriers.

**Strategic Obliteration** Eventually, the institutional approaches would lead to the outsourcing of transportation activities. Many companies have decided to focus on their principal activities and to hire external specialists to perform ancillary business activities. This type of activity is not new, as can be seen in outsourcing of accounting, advertising, and information technology services. Recently, the transportation industry is now seeing an evolution where contracts are based less on price and more on intangible benefits resulting from a strong relationship between the client and contractor. In many cases, the outsourcing users transfer assets and staff to the service company, which then manages the operation for a fee and operates much like a business partner. With this approach, the whole transportation department will be consolidated except the carrier control function. Whenever transportation needs occur, that department will call a pre-determined carrier directly, explaining delivery requirements. The carrier will be responsible for meeting the requirements. John agreed that outsourcing may be utilized. But it would be done in order to free up important company resources from secondary activities, rather than to reduce operating costs. For instance, Gillette outsources billing auditing to the Bank of Boston and warehousing to public warehouses. If Gillette manages a warehouse, it will have to provide salaries, health benefits and all other benefits which quickly become fixed costs. Gillette wants to maintain a lean organization by keeping only the control department and by leaving the actual operation to another company.

There are at least four reasons why outsourced partners can achieve better performance than Gillette. First, they can achieve economies of scale by combining movements of similar products over similar origins and destinations. For instance, a large carrier can purchase elaborate material handling equipment or a large-scale information system. Second, carriers can achieve economy of scope by combining movements of different products. By serving a portfolio of products, carriers can smooth seasonality peaks or reduce empty trailer movements in certain directions. Third, carriers can achieve learning effects through specialization. Researchers found that companies learn as they accumulate experience from handling similar businesses. Sending shipments all over the U.S. or even around the world responsively can be an extremely difficult job for in-house
organizations who are less equipped and less trained than specialized carriers. Fourth, carriers can achieve economies of risk pooling. By employing a third party contract or a public warehousing, shippers no longer need to maintain a large organization or give attention to asset maintenance, real estate risk, or technological changes. Shippers basically can transfer the risk of owning assets in non-core activities to carriers. Carriers, on the other hand, can pool the risks and spread them over a large base just as insurance companies provide insurance to their policy holders. The fixed costs needed to develop a large-scale optimization model can be ameliorated in a short time period over the large customer base.

On the other hand, it is not clear as to whether the outsourced company would be committed to the high service standards and customer-orientation that Gillette requires. In view of his long experience with public warehousing, John found it very difficult to believe that a third-party would be committed to Gillette's mission. A third-party may be more efficient for commodity-type products, but not for Gillette. The competitive structure of the third-party logistics industry may provide a buffer for unexpected contingencies. If there is enough competition, Gillette could switch to competing companies when the outsourced company did not commit to Gillette's mission statements. But, there are transaction costs involved with changing hands. Furthermore, Gillette can have control problems whether the whole distribution industry is constrained by capacity (due to driver shortage, intermodal equipment shortage, or rail track shortage) or not. If the industry is over-capacitated, the outsourced company might try to exploit Gillette. Gillette would have a very small bargaining leverage without its own distribution network and might end up paying high freight rates. If the industry is under-capacitated, the carrier might become financially weak and go bankrupt. Gillette would also lose a chance to negotiate favorable rates.

Mutual trust and longterm commitment seem to be the most important issues. Nobody is willing to invest heavily on assets which reduce their bargaining leverage when they negotiate with their shippers. This so-called "asset hold-up" problem occurs when a party (say, "A") invests on specialized assets that cannot be used for any other purpose except for a particular party (say, "B"). Before the investment, B has many incentives to induce A to invest. But once the investment has been made by A, B has little incentive to pay high prices to A, since A's investment has already become a sunk cost and cannot be used for anyone else except for B. Taking advantage of the situation, B will keep pushing A's prices down to its marginal costs. The burden of finding an alternative use for the asset is on A. There should be a way to institutionalize a no exploitation clause into the contract.

**Independent Company** Given that Gillette's major concerns about outsourcing are quality control, flexibility to environmental changes, and issue of continuity, one option for Gillette is to establish the distribution and transportation division as a separate company. The mutual trust necessary to outsource logistics activities may come from a shared capital base. John believed that the current Gillette system had core technology strong enough to compete against other logistics companies. The new company would serve all
five Gillette product lines and actively seek outside customers. In some sense, the Gillette logistics system already might be moving in such a direction. Braun, Oral-B and stationaries have been using the Gillette distribution system in such a way. John knew that 3M established a such system. In this way, the new company could provide the same special care that an internal organization provides to Gillette's products, while focusing all its resources on improving transportation productivity. The new company would understand Gillette's internal missions, organizations and procedures, and would tailor its services to meet the distribution needs of Gillette's customers. It would also provide the flexibility necessary to cope with such environmental changes as special requests by clients, special marketing promotion, new product introduction, product phase out, product recall, disruption in supply, and service customization. On the other hand, the new company should command sufficient efficiency in transportation planning and execution. Further elaboration of mathematical modelling may be necessary.

**Do nothing**

Obviously, the easiest option is to do nothing. The feasibility or economic benefits of the proactive logistics system as talked by outside consultants has not been proven yet. Although John heard about the joint work between Wal-Mart and P&G, a similar effort between K-Mart and Gillette is on stand-still as K-Mart found that the coordination of information flow is much more difficult than they originally expected. Also, the true value of the logistics system is likely to be underestimated since Gillette currently does not include the costs of clearance sales and the loss of sales opportunity due to out-of-stock situations. Activity-based costing system is still weak in reflecting the quality of logistics activities. If the company include such costs in estimating the value added by its logistics and transportation activities, the current operation might be producing as big a value as other divisions do.

John pondered about the benefits and risks of each option and what he should recommend to the Strategic Planning Board as the best direction that Gillette should take now.
### Exhibit 1. Summary of Financial Data

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Sales ($ mil.)</td>
<td>3,819</td>
<td>4,345</td>
<td>4,684</td>
<td>5,163</td>
<td>5,411</td>
</tr>
<tr>
<td>Net income ($ mil.)</td>
<td>285</td>
<td>368</td>
<td>427</td>
<td>513</td>
<td>427</td>
</tr>
<tr>
<td>Income as % of sales</td>
<td>7.5%</td>
<td>8.5%</td>
<td>9.1%</td>
<td>9.9%</td>
<td>7.9%</td>
</tr>
<tr>
<td>Earnings per share ($)</td>
<td>1.35</td>
<td>1.60</td>
<td>1.94</td>
<td>2.32</td>
<td>1.92</td>
</tr>
<tr>
<td>Stock price - high ($)</td>
<td>24.88</td>
<td>32.63</td>
<td>56.13</td>
<td>61.25</td>
<td>63.75</td>
</tr>
<tr>
<td>Stock price - low ($)</td>
<td>16.50</td>
<td>21.75</td>
<td>28.19</td>
<td>43.88</td>
<td>47.38</td>
</tr>
<tr>
<td>P/E - high</td>
<td>19</td>
<td>21</td>
<td>30</td>
<td>27</td>
<td>33</td>
</tr>
<tr>
<td>P/E - low</td>
<td>12</td>
<td>14</td>
<td>15</td>
<td>19</td>
<td>25</td>
</tr>
<tr>
<td>Dividends per share ($)</td>
<td>0.48</td>
<td>0.54</td>
<td>0.62</td>
<td>0.72</td>
<td>0.84</td>
</tr>
<tr>
<td>Book value per share ($)</td>
<td>0.36</td>
<td>0.85</td>
<td>4.82</td>
<td>6.64</td>
<td>6.72</td>
</tr>
<tr>
<td>Employees</td>
<td>30,400</td>
<td>30,400</td>
<td>31,200</td>
<td>30,900</td>
<td>33,400</td>
</tr>
</tbody>
</table>

### Exhibit 2. Performance of Different Product Lines

<table>
<thead>
<tr>
<th></th>
<th>1993</th>
<th>Sales</th>
<th>Operating Income</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$ mil.</td>
<td>% of total</td>
</tr>
<tr>
<td>Blades &amp; razors</td>
<td></td>
<td>2,118</td>
<td>39</td>
</tr>
<tr>
<td>Toiletries &amp; cosmetics</td>
<td></td>
<td>1,047</td>
<td>19</td>
</tr>
<tr>
<td>Stationery products</td>
<td></td>
<td>633</td>
<td>12</td>
</tr>
<tr>
<td>Braun</td>
<td></td>
<td>1,249</td>
<td>23</td>
</tr>
<tr>
<td>Oral-B &amp; other</td>
<td></td>
<td>364</td>
<td>7</td>
</tr>
<tr>
<td>Adjustments</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>5,411</td>
<td>100</td>
</tr>
</tbody>
</table>
### Exhibit 3. Major Brands and Competitors

<table>
<thead>
<tr>
<th>Blades and Razors</th>
<th>Major Brands</th>
<th>Key Competitors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atra</td>
<td>Good News</td>
<td>BIC</td>
</tr>
<tr>
<td>Sensor, Sensor for Women</td>
<td>SensorExcel</td>
<td></td>
</tr>
<tr>
<td>Trac II</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Toiletries and</strong></td>
<td><strong>Cosmetics</strong></td>
<td></td>
</tr>
<tr>
<td>Dry Idea (antiperspirant)</td>
<td>Foamy (shaving cream)</td>
<td>Amway</td>
</tr>
<tr>
<td>Gillette Series (men's toiletries)</td>
<td>Jafra (skin care)</td>
<td>Avon</td>
</tr>
<tr>
<td>Right Guard (deodorant)</td>
<td>Soft &amp; Dri (deodorant)</td>
<td>Procter &amp; Gamble</td>
</tr>
<tr>
<td>White Rain (hair care)</td>
<td></td>
<td>Dial</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L'Oreal</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mary Kay</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Unilever</td>
</tr>
<tr>
<td><strong>Stationery Products</strong></td>
<td>Helit (desk accessories)</td>
<td>Carter-Wallace</td>
</tr>
<tr>
<td>Liquid Paper (correction fluid)</td>
<td>Paper Mate (markers and pens)</td>
<td>A. T. Cross</td>
</tr>
<tr>
<td>Parker (pens)</td>
<td>Waterman (pens)</td>
<td></td>
</tr>
<tr>
<td><strong>Small Appliances</strong></td>
<td>Braun (shavers, blenders, juicers, coffee makers, oral care appliances)</td>
<td>Black &amp; Decker</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mr. Coffee</td>
</tr>
<tr>
<td><strong>Dental Products</strong></td>
<td>Oral-B (toothbrushes, toothpaste, floss)</td>
<td>Bausch &amp; Lomb</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Colgate-Palmolive</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Johnson &amp; Johnson</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Pfizer</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SmithKline Beecham</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Warner-Lambert</td>
</tr>
</tbody>
</table>

### Exhibit 4. Global Presence

<table>
<thead>
<tr>
<th></th>
<th>1993</th>
<th>Sales</th>
<th>Operating Income</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$ mil</td>
<td>% of total</td>
<td>$ mil</td>
</tr>
<tr>
<td>US</td>
<td>1,759</td>
<td>33</td>
<td>256</td>
</tr>
<tr>
<td>Europe</td>
<td>1,949</td>
<td>36</td>
<td>316</td>
</tr>
<tr>
<td>Latin America</td>
<td>762</td>
<td>14</td>
<td>178</td>
</tr>
<tr>
<td>Other regions</td>
<td>941</td>
<td>17</td>
<td>126</td>
</tr>
<tr>
<td>Adjustments</td>
<td></td>
<td>(51)</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>5,411</td>
<td>100</td>
<td>825</td>
</tr>
</tbody>
</table>

241
Reference


Teaching Note

1. Discuss changing environments of Gillette

   1) Basic trends
      - increased product variety
      - shortened product life cycle
      - multiple/hybrid distribution channels
      - proliferation of special promotions
      - high customer expectation
      - globalization

   2) Industry analysis (what's happening in the bargaining power: See M. Porter)
      - supplier
      - buyer: - growing power of retailers
          - consumers
      - competitors
      - company resources
      - technology: information technology
      - substitute products

2. Discuss strengths and weaknesses of Gillette' distribution network

   - organizational chart
   - control: strategic: network design - modal selection
     tactical: shipment size, frequency - carrier selection
     operational: vehical dispatching - carrier allocation
   - D.C. system
   - P. D. system
   - Inbound vs. outbound

3. Discuss the pros and cons of strategic options

   - develop mathematical models
   - develop strategic partnership
   - outsource completely to a third-party logistics company
   - offspring as a separate entity specializing in distribution and transportation
   - do nothing & sponser the study of the true value of transportation
Exhibit 5. Organizational Structure of Distribution and Transportation at Gillette
Appendix 2. Statistical Property of MDI Estimators

1. The MDI estimator

1.1. Consistency

As an extreme estimator, the MDI estimator is consistent and asymptotic normal. In order to prove consistency, consider the following objective function of MDI:

\[
\text{Min } Q_N(\beta) = \sum_{n=1}^{N} \sum_{i=1}^{3} s_{in} \ln \frac{s_{in}}{F(i|x_n; \beta)}.
\]

If we take a Taylor expansion on (1) in an open neighborhood of \( \beta_0 \), we have:

\[
\frac{1}{N} Q_N(\beta) = \frac{1}{N} Q_N(\beta_0) + \frac{1}{N} \frac{\partial Q_N}{\partial \beta} \bigg|_{\beta_0} (\beta - \beta_0) + \frac{1}{2} (\beta - \beta_0)' \frac{1}{N} \frac{\partial^2 Q_N}{\partial \beta \partial \beta'} \bigg|_{\beta_0} (\beta - \beta_0)
\]

where \( \beta^* \) lies between \( \beta \) and \( \beta_0 \). As \( \beta \) approaches \( \beta_0 \), \( \beta^* \) also approaches \( \beta_0 \). The first component, \( \frac{1}{N} Q_N(\beta_0) \), converges to a non-stochastic function \( Q(\beta_0) \), in this case 0, in probability uniformly in an open neighborhood of \( \beta_0 \). This is true since choice probabilities predict actual shares perfectly at the true parameter value and the log of their ratio become zero. The second component also goes to zero as \( \beta \) approaches \( \beta_0 \). For the third component, let us suppose that the following assumption holds:

\[
\text{(A1) } \lim N \frac{\partial^2 Q_N}{\partial \beta \partial \beta'} \text{ exists and is continuous in an open and convex neighborhood of } \beta_0.
\]

\[
\text{(A2) } \frac{1}{N} \frac{\partial^2 Q_N}{\partial \beta \partial \beta'} \bigg|_{\beta_n^*} \rightarrow A(\beta_0) \text{ in probability for any sequence } \beta_n^* \text{ such that } \lim \beta_n^* = \beta_0
\]

where \( A(\beta_0) = \lim E N \frac{\partial^2 Q_N}{\partial \beta \partial \beta'} \bigg|_{\beta_0} \) is a finite non-singular and positive definite matrix.

If we take the probability limit of both sides of (2), we have:

\[
Q(\beta) = Q(\beta_0) + \frac{1}{2} (\beta - \beta_0)' A^* (\beta - \beta_0) \text{ where } A^* = \lim N \frac{\partial^2 Q_N}{\partial \beta \partial \beta'} \bigg|_{\beta_0}.
\]
Since $A^*$ is positive definite, $Q(\beta) > Q(\beta_0)$ for $\beta \neq \beta_0$. That is, $Q(\beta)$ attains a strict local minimum at $\beta_0$. Then, the MDI estimator is consistent based on Amemiya (Theorem 4.1.2., p. 110).

The assumptions, (A1) to (A3), depend on the distributional form of choice probabilities. But note that the followings hold for choice probabilities defined in logit or independent probit form:

(4) $\frac{\partial F}{\partial \beta}$ exists and is continuous in an open and convex neighborhood of $\beta_0$.

(5) $\frac{\partial^2 F}{\partial \beta \partial \beta'}$ exists and is continuous in an open and convex neighborhood of $\beta_0$.

(6) $\frac{1}{N} \frac{\partial^2 F}{\partial \beta \partial \beta'} \rightarrow C(\beta_0)$ in probability for any sequence $\beta^*_\nu$ such that $\text{plim} \; \beta^*_\nu = \beta_0$,

where $C(\beta_0) = \lim E \left( \frac{1}{N} \frac{\partial^2 F}{\partial \beta \partial \beta'} \right)_{\beta_0}$ is a finite non-singular and negative definite matrix.

When choice probability is twice continuously differentiable, the score and hessian of the objective function can be written as linear combinations of the first and second derivatives of choice probabilities as follow:

(7) $\frac{1}{N} \frac{\partial Q_N}{\partial \beta} = - \frac{1}{N} \sum_{n=1}^{N} \sum_{i=1}^{3} s_{ni} \left( \frac{\partial F}{\partial \beta} \right)$

(8) $\frac{1}{N} \frac{\partial^2 Q_N}{\partial \beta \partial \beta'} = \frac{1}{N} \sum_{n=1}^{N} \sum_{i=1}^{3} s_{ni} \left[ \left( \frac{\partial F}{\partial \beta} \right) \left( \frac{\partial F}{\partial \beta} \right)' - \frac{\partial^2 F}{\partial \beta \partial \beta'} \right]$.

The ratio of continuous functions and the outer product of a continuous function are continuous. Therefore, the score and hessian of the objective function are continuous in an open neighborhood of true parameters. Moreover, the outer product of the first derivative is positive definite and the second derivative of the choice probability is negative definite. Thus, (A2) is satisfied.
Our proof of consistency is limited since it merely states that one of the local minima is consistent and since it does not give any guide about how to choose a consistent minimum. The minimum is not unique and there can be many estimators that satisfies:

\[
\sum_{i=1}^{3} s_{in} \ln \frac{s_{in}}{F(i|x_n, \beta)} = 0.
\]

Both actual and predicted shares have to sum to one. Thus, if an estimator over-predicts a share of a certain mode, it has to under-predict shares of other modes. Then, the sum of their log-ratios can be zero in several values for \( \beta \). That is, the MDI estimator is subject to the identification condition. We need to ensure the consistency of the MDI estimator by checking if the solution gives a reasonable value in terms of an economic theory and it the iteration by which the local minimum was obtained started from a consistent estimator. If both are not satisfactory, we would better find an estimator that minimizes the following absolute information discrepancy:

\[
\text{Min } Q_N(\beta) = \sum_{n=1}^{N} \sum_{i=1}^{3} s_{in} \ln \left| \frac{s_{in}}{F(i|x_n, \beta)} \right|
\]

1.2. Asymptotic Normality

Now, let us show that the MDI estimator is asymptotically normal. MDI finds an estimator that minimizes the discrepancy between actual shares and predicted shares. The first condition says that

\[
\frac{\partial Q_N}{\partial \beta} = 0 \quad \text{at the estimator}
\]

By applying a Taylor expansion in an open neighborhood of \( \beta_0 \), we have:

\[
\frac{1}{\sqrt{N}} \frac{\partial Q_N}{\partial \beta} = 0 = \frac{1}{\sqrt{N}} \frac{\partial Q_N}{\partial \beta} \bigg|_{\beta_0} + \frac{1}{N} \frac{\partial^2 Q_N}{\partial \beta \partial \beta'} \bigg|_{\beta_0} \sqrt{N} (\beta - \beta_0)
\]

where \( \beta^{**} \) lies between \( \beta \) and \( \beta_0 \). Now, suppose the following assumptions hold:

\[
(A3) \quad \frac{1}{\sqrt{N}} \left[ \frac{\partial Q_N}{\partial \beta} \right]_{\beta_0} \rightarrow \mathcal{N}(0, B(\beta_0)) \text{ where } B(\beta_0) = \lim E \left[ \frac{1}{N} \frac{\partial Q_N}{\partial \beta} \bigg| \beta_0 \right] \frac{\partial Q_N}{\partial \beta} \bigg|_{\beta_0}.
\]

If we have a sequence, \( \{ \hat{\beta}_N \} \), such that \( \frac{\partial Q_N}{\partial \beta} = 0 \) and \( \text{plim } \hat{\beta}_N = \beta_0 \), then
\[ \sqrt{N} (\hat{\beta}_N - \beta_o) = \left( \frac{1}{N} \frac{\partial^2 Q_N}{\partial \beta \partial \beta'} \right)^{-1} \frac{1}{\sqrt{N}} \frac{\partial Q_N}{\partial \beta} \bigg|_{\beta_o} \rightarrow \mathcal{N} [0, A(\beta_o)^{-1}B(\beta_o)A(\beta_o)^{-1}] \]

Note that we do not require any distributional assumption on measurement errors except the differentiability of choice probability in deriving the statistical properties of MDI estimators. In particular, this shows the robustness of MDI estimators. Since it looks only at the ratio between actual shares and predicted shares, the MDI is consistent even when shippers send different number of shipments.

2. The MDI estimator with the Expected Probability

We propose the MDI estimator with the expected probability when discount rates are randomly distributed. This specification finds an estimator that solves the following problem:

\[
\text{Min} \quad Q_N(s; \beta, \phi) = \sum_{n=1}^{N} \sum_{i=1}^{3} s_{in} \ln \frac{s_{in}}{\int_{0}^{\infty} F(i|x_{n}, r; \beta)f(r; \phi) \, dr}
\]

The statistical properties of the estimators can be shown similarly to what we have done before since choice probabilities evaluated at the true parameter value will predict actual shares perfectly and the log of their ratio will become zero. Moreover, the first derivative of the objective function with respect to \( \beta \) or \( \phi \) converges to a normal variate with zero mean and finite variance due to the following theorem (Lang 1987, p. 199):

Assume that
1. \( f \) is continuous on the rectangle \([a \leq \beta \leq b \) and \( c \leq r \leq d\)], and that \( df/d\beta \) exists and is continuous.
2. \( df/d\beta \) exists and is continuous.

Then, \( \psi(\beta) = \int_{a}^{b} f(r, \beta) \, dr \) is differentiable and \( \frac{d\psi}{d\beta} = \int_{a}^{b} \frac{\partial f(r, \beta)}{\partial \beta} \, dr \)

In addition, we need to show that the second derivatives (own and cross) exist, is continuous and is positive definite. This is more difficult since they depend on cross-derivatives of the hessian with respect to \( \beta \) and \( \phi \). But, it will be true if \( \beta \) and \( \phi \) are independently distributed of each other.
Appendix 3. Calculation of Safety Stock Holding Costs

Shippers maintain safety stocks in order to prepare for unexpected delay of shipments or for unexpected change in demands. We will discuss two approaches based on whether demand variability is considered or not.

1. Without Demand Variability

Given the distribution of transit time, we can calculate expected stockout costs to shippers. When a shipper maintains $L_n$-days of safety stock, the annual expected stockout costs is the cost per each stockout times the probability of stockout times shipment frequency, i.e.

$$SS_n = F_n(T_n + L_n) \cdot s_n \cdot \frac{Q_n}{q_n}$$

where $s_n$ is the cost that a shipper $n$ would incur for each stockout incidence such as opportunity costs of customer loss and emergency shipment costs, $Q_n$ is annual tonnage, and $q_n$ is typical shipment size with mode $i$. Anticipating the stockout costs, a shipper decides whether to maintain $L_n$-days or $(L_n+1)$ days of safety stock.

If he decides to maintain $(L_n+1)$ days of safety stock, the cost of carrying inventory for an additional day is $r \cdot \frac{p_n Q_n}{365}$. On the other hand, the expected stockout costs are reduced by $[F_n(T_n + L_n + 1) - F_n(T_n + L_n)] \cdot s_n \cdot \frac{Q_n}{q_n}$. Let us denote that

$$d \tau_n = F_n(T_n + L_n + 1) - F_n(T_n + L_n).$$

Since the shipper decided to hold $L_n$ days of safety stock, the inventory costs must be larger than the savings in stockout costs, i.e.

$$d \tau_n \cdot s_n \cdot \frac{Q_n}{q_n} \leq r \cdot \frac{p_n Q_n}{365} \quad \text{or} \quad s_n \leq r \cdot \frac{1}{d \tau_n} \cdot \frac{p_n q_n}{365}.$$ 

This formula suggests that the value of $\frac{q_n}{d \tau_n}$ should roughly be equal across modes. This restriction comes since we assumed that shippers maintain the same level of safety stock ($L_n$) no matter which mode they use. If shippers maintain larger safety stock when they use rail than when they use truck, we will have to survey the maximum acceptable lateness of delivery time for each mode. In any case, the annual stockout costs is then limited in the above by:
\[ SS_{in} \leq r \frac{F_{in}(T_{in} + L_n) \ p_n \ Q_n}{d \tau_{in}} \frac{1}{365}. \]

We can calculate the maximum stockout costs by letting \( F_{in}(T_{in} + L_n) = 1 - \tau_{in} \) and by estimating \( d \tau_{in} \) from the Weibull distribution.

2. With Demand Variability

If we know the mean and variance of the daily shipment needs (\( d_n \)), we can calculate the level of safety stocks for demand variation as well. Suppose that we denote \( X_{in} \) to be the amount of inventory required during transit time, i.e.

\[ X_{in} = \sum_{w=1}^{t_{in}} d_{wn} \]

where both transit time and daily demands are random variables. We assume that the distribution of transit time (carrier performance) is independent of the distribution of daily demands (consumer taste). The mean and the variance of the required inventory would then be (Ross 1985)\(^1\):

\[
\begin{align*}
E[X_{in}] &= E[t_{in}] \ E[d_n] \\
V[X_{in}] &= E[t_{in}] \ V[d_n] + V[t_{in}] \ E[d_n]^2
\end{align*}
\]

Safety stock would be then \( k \times \sqrt{V[X_{in}]} \) for some decision variable \( k \). The cost of holding in-transit stock and safety stock are:

\[ r \ p_n \ (E[X_{in}] + k \sqrt{V[X_{in}]}) \]

where \( r \) is the discount rate and \( p_n \) is product price. We can estimate the degree of \( k \) by including \( E[X_{in}] \) and \( \sqrt{V[X_{in}]} \) as separate explanatory variables in a choice model. Note that in order to estimate this model, we need to know the mean and the variance of the daily shipment demand. While its mean can be calculated from annual tonnage, its variance is difficult to survey. Instead, we may survey how variable a shipper's demand is relative to her mean demand, and estimate the importance of safety stock from the importance of the coefficient of variation \( (CV_{in} = \sqrt{V[X_{in}]} / E[d_n]) \).

Appendix 4. The GAUSS-HERMITE Quadrature

The following program presents the quadrature points (pt) and their weights (wt) that the GAUSS-HERMITE quadrature employs. Their value depend on the pre-specified number of quadrature points (npt).

if npt == 4;
    let pt =
        1.650680123885784 0.5246476232752904
        -0.5246476232752904 -1.650680123885784;
    let wt =
        8.1312835447245060E-02 0.8049140900055128
        0.8049140900055128 8.1312835447245060E-02;
elseif npt == 8;
    let pt =
        2.930637420257244 1.981656756695843
        1.157193712446780 0.3811896902073222
        -0.3811896902073222 -1.157193712446780
        -1.981656756695843 -2.930637420257244;
    let wt =
        1.9960407221136764E-04 1.7077983007413460E-02
        0.2078023258148917 0.6611470125582410
        0.6611470125582410 0.2078023258148917
        1.7077983007413460E-02 1.9960407221136764E-04;
else; /* npt = 12 */
    let pt =
        3.889724897869782 3.020637025120890
        2.279507080501060 1.597682635152605
        0.9477883912401637 0.3142403762543591
        -0.3142403762543591 -0.9477883912401637
        -1.597682635152605 -2.279507080501060
        -3.020637025120890 -3.889724897869782;
    let wt =
        2.6585516843563044E-07 8.573687043587822222E-05
        3.9053905846290365E-03 5.1607985615883901E-02
        0.2604923102641610 0.5701352362624796
        0.5701352362624796 0.2604923102641610
        5.1607985615883901E-02 3.9053905846290365E-03
        8.573687043587822222E-05 2.6585516843563044E-07;
endif,
Appendix 5. The Delta Method

After the distribution of random discount rates is estimated, we can calculate the median, mode and mean. We now want to know their standard errors. The delta method is used to derive standard errors of such functions of parameters.

Suppose that we are interested in the standard error of \( \gamma \) as a function of estimated parameters \((\phi, \psi)\), i.e. \( \gamma = f(\phi, \psi) \). The Taylor series expansion around the estimated values \((\hat{\phi}, \hat{\psi})\) gives:

\[
f(\phi, \psi) \approx f(\hat{\phi}, \hat{\psi}) + \left( \frac{\partial f}{\partial \phi} \right)_{(\hat{\phi}, \hat{\psi})} (\phi - \hat{\phi}) + \left( \frac{\partial f}{\partial \psi} \right)_{(\hat{\phi}, \hat{\psi})} (\psi - \hat{\psi}).
\]

The variance of \( \gamma \) is given by:

\[
Var(\gamma) \approx \left( \frac{\partial f}{\partial \phi} \right)_{(\hat{\phi}, \hat{\psi})}^2 Var(\phi) + \left( \frac{\partial f}{\partial \psi} \right)_{(\hat{\phi}, \hat{\psi})}^2 Var(\psi) + 2 \left( \frac{\partial f}{\partial \phi} \right)_{(\hat{\phi}, \hat{\psi})} \left( \frac{\partial f}{\partial \psi} \right)_{(\hat{\phi}, \hat{\psi})} Cov(\phi, \psi)
\]

We now use the above formula to the statistics of a log-normal distribution.

1) Median: \( \gamma = \exp(\phi) \)

\[
\frac{\partial f}{\partial \phi} = \exp(\phi) \quad \frac{\partial f}{\partial \psi} = 0
\]

2) Mode: \( \gamma = \exp(\phi - \exp(\psi)^2) \)

\[
\frac{\partial f}{\partial \phi} = \exp(\phi - y) \quad \frac{\partial f}{\partial \psi} = -2y \exp(\phi - y) \quad \text{where } y = \exp(\psi)^2
\]

2) Mean: \( \gamma = \exp(\phi + \exp(\psi)^2/2) \)

\[
\frac{\partial f}{\partial \phi} = \exp(\phi + \frac{y}{2}) \quad \frac{\partial f}{\partial \psi} = y \exp(\phi + \frac{y}{2}) \quad \text{where } y = \exp(\psi)^2
\]
Appendix 6. Simulation Method for Parameter Estimation

1. Introduction

Simulation (or Monte Carlo) methods have been used widely when mathematical problems are analytically intractable (or prohibitively expensive to solve) but can be solved by substituting an equivalent stochastic problem (Hendry 1984). In particular, distribution sampling is used to evaluate features of a statistical distribution by representing it numerically and by drawing observations from that the desired distribution. For example, in order to investigate the distribution of the mean of random samples of $T$ observations from a distribution which was uniform between zero and one, one could simply draw large numbers within the interval $[0,1]$ and plot the resulting distribution of the mean. Suppose that we need to evaluate choice probabilities which are written in the integral form which does not have an analytical solution and is also difficult to evaluate numerically. For example, suppose that one has to evaluate the following probability where $\theta = (\beta, f(\eta))$: \[
P_n(i|x; \theta) = \int \frac{P_n(i|x_n; \beta, \eta)f(\eta)d\eta}{P_n(i|x_n; \beta, \eta)}
\]
one can approximate the probability by averaging probabilities conditional on simulated variates $\eta$'s instead of trying to get its numerical integration directly, i.e.
\[
P_n(i|x; \theta) \approx \frac{1}{R} \sum_{r=1}^{R} P_n(i|x_n; \beta, \eta_0^{(r)})
\]
where $\eta$ is simulated from $f(\eta)$ $R$ times.

McFadden (1989) showed that as we make more simulation draws relative to the number of observations, the bias of parameter estimates resulting from simulation noise will disappear. Basically, he showed that the parameters estimated from the simulation method are consistent and asymptotically normal, and have reasonable asymptotic variance. With this finding, the application of the simulation approach to parameter estimation has received a great attention recently from econometricians. In particular, the simulated maximum likelihood (SML) estimation and the methods of simulated moments (MSM) have been widely researched. In section 5.2, we discuss various estimation methods and their properties.

As parameters estimated by using simulation can be proved to be consistent and asymptotically normal, more focus will be given on how to simulate more efficiently. In particular, algorithms to generate variates from various distributions other than normal distribution are not well documented. Section 5.3 discusses such algorithms. In addition, researchers have not understood that currently available algorithms, even for well-documented distributions, can be made more efficient by various methods of variance
reduction. Section 5.4 discusses ideas for variance reduction. Studies on statistical properties of such methods are necessary.

2. Statistical Property of Parameters Estimated Using Simulation

2.1. SML Estimators

The first estimator to which simulation was applied is the simulated maximum likelihood estimation (SML) proposed by Lerman and Manski (1981). The SML method maximizes the following log-likelihood:

$$L(\theta) = \sum_n \sum_i y_{in} \ln(p_{in}(\theta))$$

where $$L(\theta)$$, $$\theta = (\beta, V)$$, is approximated by

$$p_n(i|x; \theta) \approx \frac{1}{R} \sum_{r=1}^{R} p_n(i|x, \eta^{(r)}; \beta) \text{ where } \eta^{(r)} \sim f_\eta(V).$$

At the parameter values where the log-likelihood is maximized, the first derivative of the log-likelihood function, i.e. the associated score vector, will be zero. The score can be decomposed into the five terms as shown in the below:

$$0 = \frac{1}{\sqrt{N}} \frac{\partial L}{\partial \theta}$$

$$= \frac{1}{\sqrt{N}} \sum_n \frac{\partial l(X_n|\hat{\theta})}{\partial \theta}$$

$$= \frac{1}{\sqrt{N}} \sum_n \frac{\partial l(X_n|\theta^*)}{\partial \theta} \text{ data noise} \quad > \quad N(0, \Sigma_{nn})$$

$$+ \frac{1}{\sqrt{N}} \sum_n \left[ \frac{\partial l(X_n|\theta^*, w)}{\partial \theta} - E_w \left( \frac{\partial l(X_n|\theta^*, w)}{\partial \theta} \right) \right. \text{ simulation noise} \quad > \quad N(0, \Sigma_n)$$

$$+ \frac{1}{\sqrt{N}} \sum_n \left[ E_w \left( \frac{\partial l(X_n|\hat{\theta})}{\partial \theta} \right) - \frac{\partial l(X_n|\hat{\theta})}{\partial \theta} \right] \text{ simulation bias} \quad > \quad o\left( \frac{1}{\sqrt{N}} \right) \text{ as } R \rightarrow \infty$$

$$+ \frac{1}{\sqrt{N}} \sum_n \left[ \frac{\partial l(X_n|\hat{\theta})}{\partial \theta} - \frac{\partial l(X_n|\theta^*)}{\partial \theta} \right]$$

+ Remaining Parts \quad \rightarrow \quad o\left( \frac{1}{\sqrt{N}} \right) \text{ if } \hat{\theta} \text{ is near } \theta^*.$$

where $$l(X_n | \theta)$$ denotes the log-likelihood of an observation, $$\theta^*$$ denotes the true values of parameters, $$\hat{\theta}$$ denotes their estimated values using simulation, and $$w$$ denotes the randomness introduced by simulation. This decomposition basically shows that the bias
from data noise and simulation noise have zero means and that other terms disappear as we increase the number of simulation draws.

The last part requires that the bias from the interaction between data noise and simulation noise will be near zero if \( \hat{\theta} \) is near \( \theta^* \). In econometric jargon, the interaction bias disappears if the simulation noise is stochastically equi-continuous in \( \theta \) or if the probability that there is discontinuity in the neighborhood of the true value (\( \theta^* \)) is very small enough to ensure continuity. With this property, the variance of parameter estimates can be decomposed into variance from data generation and variance from simulation noise, and the variance due to interaction between data noise and simulation noise approaches zero as the number of simulation draws increases.

The fourth term is equivalent to \( \frac{1}{N} \sum_n \left[ \frac{\partial^2 I(X_n|\bar{\theta})}{\partial \theta \partial \theta'} \right] \sqrt{N}(\hat{\theta} - \theta^*) \) for some \( \bar{\theta} \in [\hat{\theta}, \theta^*] \). Thus, if we rearrange the above decomposition in terms of the fourth terms, the parameter estimates from the usual Newton-Raphson method will have the following property:

\[
\sqrt{N}(\hat{\theta} - \theta^*) \rightarrow N(0, H_{ML}^{-1}(\Sigma_{mm} + \Sigma_{ss})H_{ML}^{-1})
\]

if \( H_{ML} \) is non-singular where \( H = \frac{1}{N} \sum_n \left[ \frac{\partial^2 I(X_n|\bar{\theta})}{\partial \theta \partial \theta'} \right] \) is the hessian matrix evaluated at the estimated parameter \( \theta_{ML} \), \( \Sigma_{mm} \) is the variance from data noise and \( \Sigma_{ss} \) is the variance from simulation noise.

Although the above shows that bias from simulation can be controlled, the SML method involves a bias from the non-linearity of the natural log function. Take a Taylor series-expansion of the objective function around the estimated parameters and take expectation on both sides. Then we have:

\[
E[ln P_jn(\theta)] - ln E[P_jn(\theta)] \approx - \text{Var}(P_jn(\theta)) / [2 P_jn(\theta)^2] < 0
\]

This bias can be quite serious when some of the predicted choice probabilities (i.e., \( P_jn(\theta) \)) are very small or when the probabilities have high variances. Thus, this model should be used only when choice probabilities of all alternatives would be relatively large and the number of simulation draws would be big enough to have a small variance of choice probabilities (Borsch-Supan and Hajivassiliou 1993).

2.2. MSM estimators

This method of simulated moments (MSM) have been popular after McFadden proposed in 1989. The score vector of the log-likelihood is:
\[
\frac{\partial}{\partial \theta} L(\theta) = \sum_n \sum_i y_{in} W_{in}(\theta) \text{ where } W_{in}(\theta) = \frac{1}{P_{in}(\theta)} \frac{\partial}{\partial \theta} P_{in}(\theta).
\]

Note that since \( \sum_i P_{in}(\theta) = 1 \) for all \( n \), we have:

\[
\sum_i \frac{\partial P_{in}(\theta)}{\partial \theta} = \sum_i P_{in}(\theta) W_{in}(\theta) = 0.
\]

and we can rewrite the score vector as follows:

\[
\frac{\partial}{\partial \theta} L(\theta) = \sum_n \sum_i (y_{in} - P_{in}(\theta)) W_{in}(\theta)
\]

Since the score vector should be zero at the true parameter, we may search for an estimator that minimizes its quadratic forms. Partly motivated by the idea, the MSM method minimizes the following objective function (In fact, any moment estimator of \( y_{in} \) can be used instead of \( P_{in}(\theta) \) for the MSM method.):

\[
S(\theta) = \sum_n \sum_i (y_{in} - P_{in}(\theta)) W_{in} D W_{in}' (y_{in} - P_{in}(\theta))'
\]

where its gradient becomes now:

\[
\frac{\partial}{\partial \theta} S(\theta) = -2 \sum_n \sum_i \frac{\partial P_{in}(\theta)}{\partial \theta} W_{in} D W_{in}' (y_{in} - P_{in}(\theta))'
\]

where is an instrumental variable is chosen for \( W_{jn} \) and \( D \) is a weighting matrix.

Statistical properties of MSM estimators can be derived in a similar way as the SML. Let us first consider the case without simulation. From the unbiasedness of \( \hat{\theta} \), we know that \( 0 = \sum_n y' W' D W (y - P(\hat{\theta})) \). Inserting this into the first order condition gives the following relationship:

\[
0 = \sum_n (y - \hat{\theta} + \frac{\partial P(\hat{\theta})}{\partial \theta})' W' D W (y - P(\hat{\theta}))
\]

\[
= \sum_n (y - \theta^*)' W' D W (y - P(\theta^*))
\]

\[
+ \sum_n (y - \hat{\theta})' W' D W (y - \hat{\theta}) (\hat{\theta} - \theta^*) \text{ for some } \hat{\theta}.
\]
Therefore, if we assume
\[ \sum_n \left( y - \frac{\partial P(\theta \ast)}{\partial \theta} \right)' W' \rightarrow G \text{ and } \sqrt{N} \left( y - P(\theta) \right) \rightarrow N(0, A), \]
\[ \sqrt{N} (\hat{\theta} - \theta \ast) = - \left[ \sum_n \left( y - \frac{\partial P(\hat{\theta})}{\partial \theta} \right)' W' D W \left( y - \frac{\partial P(\hat{\theta})}{\partial \theta} \right) \right]^{-1} \]
\[ \times \sum_n \left( y - \frac{\partial P(\theta \ast)}{\partial \theta} \right)' W' D \sqrt{N} W \left( y - P(\theta \ast) \right) \]
\[ \rightarrow (G'DG)^{-1} G'DW \ast N(0, A) = N(0, (G'DG)^{-1} G'DW A W'(G'DG)^{-1}). \]
\[ \rightarrow N(0, (G'(W'AW)^{-1}G)^{-1}) \text{ if } D = (W'AW)^{-1}. \]

As we can see from the last equation, the variance of MSM estimators is minimized when \( D = (W'AW)^{-1} \). How we can find the optimal weight matrix \( D \) is a tough question. McFadden (1989) used an identity matrix for \( D \) in his paper. Clearly \( D = I \) is not optimal. Review of econometric literature suggests that when disturbances are not auto-correlated, Newey's non-parametric regression (1990) or Robinson's k-th nearest neighborhood methods (1991) can be used, and that with auto-correlation, Hansen's approach (1985, 1988) may be used.

Now, consider the case with simulation. Similar to the case of the SML, the decompositional approach suggests that only the variance of data noise and simulation noise will stay as we increase the number of simulation draws. That is, if we denote \( \Sigma = W'AW, \hat{\Sigma} = \hat{\Sigma}_{mm} + \hat{\Sigma}_{ss} \). The problem is then how to estimate \( (\Sigma_{mm} + \Sigma_{ss}) \). Fortunately, McFadden showed that the outer product of gradients evaluated at the parameter value estimated by using simulation can be used as the asymptotic variance, i.e.
\[ \hat{\Sigma} = \frac{1}{N} \sum_n \sum_{i,j} W_{in}'(y_{in} - P_{in}(\hat{\theta}))(y_{jn} - P_{jn}(\hat{\theta}))W_{jn} \]

This is a good news since we can use the B.H.H. method which involves only the first derivatives (Berndt, Hall, Hall and Hausman, 1974). The BHHH method approaches the optimal value iteratively by taking the following steps:
\[ \theta^{(k+1)} = \theta^{(k)} + \lambda M^{(k)} \]

where the direction matrix \( M \) can be computed as follows:
\[ M^{(k)} = \left[ \frac{1}{N} \sum_n \frac{\partial l(X_n|\theta^{(k)})}{\partial \theta} \right]' \left[ \frac{1}{N} \sum_n \frac{\partial l(X_n|\theta^{(k)})}{\partial \theta} \right] \]

Note that in the Newton-Raphson method, we use the hessian matrix \( H \) instead. Bunch (1988) compared the properties of different hill-climbing methods and found that the BHHH method performs pretty well, considering that it saves lots of computational efforts.
by evaluating only the first derivatives. Its basic intuition is that at the parameter value that achieves Cramer-Rao lower bound, the negative of the expected value of the hessian matrix is the same as the expected value of the outer product of the gradient matrix. The outer product of gradients provides more robust standard errors than the hessian. If more robust standard errors are desired, we use the White heteroscedastic-consistent matrix \( H_{\text{ML}}^{-1} M_{\text{ML}}^{-1} H_{\text{ML}}^{-1} \) as the covariance matrix.

Another way to measure standard errors would be to use the jack-knife method. Instead of estimating \( \hat{\theta} \) by using all \( R \) draws (say, the naive approach), let us divide the draws into \( G \) runs with \( B \) draws for each run (e.g. \( G \cdot B = R \)). The method takes the following five steps:

1. Estimate parameters by using all data, say \( \hat{\theta} \).
2. Estimate parameters \( G \) times by using all data except each run of simulation draws, \( \Rightarrow \) say \( \hat{\theta}^{(i)}, \ldots, \hat{\theta}^{(G)} \).
3. Calculate the pseudo-values: \( \hat{\theta}_{(g)}^* = G \hat{\theta} - (G-1) \hat{\theta}^{(g)} \) for \( g = 1, \ldots, G \)
4. Calculate the expectation of \( \hat{\theta} \) by \( \bar{\theta}^* = \frac{1}{G} \sum_{g=1}^{G} \hat{\theta}_{(g)}^* \)
5. Calculate the variance of \( \hat{\theta} \) by \( \text{ar}(\hat{\theta}) = \frac{1}{(G-1)G} \sum_{g=1}^{G} (\hat{\theta}_{(g)}^* - \bar{\theta}^*)(\hat{\theta}_{(g)}^* - \bar{\theta}^*)' \)

The strength of the jack-knife method is that it provides a non-parametric estimation of the standard error with little distribution assumption.

2.3. SMDI Estimators

Similar to the SML, we can define the simulated minimum discrimination information estimator which minimizes the discrepancy between observed share and simulated information, i.e. \( \sum_{n=1}^{N} \sum_{i=1}^{3} s_m \log \frac{s_m}{P_m(\theta)} \). Note that since shares are fixed, the MDI estimation is equivalent to maximizing \( \sum_{n=1}^{N} \sum_{i=1}^{3} s_m \log P_m(\theta) \). This objective function is no different from that of the SML, since \( \sum_{i=1}^{3} s_m = 1 \) for all \( n \). Thus, the SMDI estimator shares all the properties of the SML estimator. It is consistent and asymptotically normal, i.e.

\[
\sqrt{N} (\hat{\theta} - \theta^*) \rightarrow N(0, (G'DG)^{-1} G'D \Sigma D'G' (G'DG)^{-1})
\]

where

\[
\hat{\Sigma} = \frac{1}{N} \sum_n \sum_{ij} W_{ni}(s_m - P_m(\hat{\theta}))'(s_{jn} - P_{jn}(\hat{\theta})) W_{ji},
\]
and \( \hat{\theta} \) is the parameter estimated from the SMDI at each iteration. Also note that the SMDI estimator involves a bias from the non-linearity of the natural log function, as the SML estimator does.

2.4. Empirical Findings

Since simulations are driven by random inputs, they produce random outputs. Only when we apply proper statistical techniques to simulation output, the results can be properly analyzed, interpreted and utilized. The decompositional approach to identify statistical properties of simulated estimators shows that the bias from simulation can be controlled as we increase the number of simulation draws. This property makes simulation attractive, since other approximation methods like Clark approximation does not provide such control over bias. For greater details, refer to McFadden (1989). In addition, the decompositional approach indicates that all extremum estimators can be estimated consistently by using simulation with little loss of generality. In effect, we propose that the simulated minimum discrimination information (SMDI) estimation is feasible, and that the SMDI estimator share all the properties of the SML estimator. For the random rate model in Chapter 4, we estimated parameters by employing both numerical integration (e.g. the GAUSS-HERMITE quadrature) and simulation. Their results were similar, providing at least an empirical support for our proposal.

Between the MSM and SML (and thus SMDI), the MSM method has several advantages over the SML method. First, Ben-Akiva and Bolduc (1991) claimed by using a law of large numbers that the simulation bias averaged across individuals in MSM goes to zero at a faster rate than in SML. McFadden suggests in his lecture note that usually 20 draws are enough for the MSM method, whereas the SML method needs the number of draws to be big enough to ensure \( R/\sqrt{N} \rightarrow \infty \) where \( R \) is the number of draws and \( N \) is the number of observations. McFadden (1989) also claimed that iteration using smooth simulators (e.g., Logit-kernel) is faster than that using frequency simulators in MSM. Second, the MSM method is more robust to mis-specification than the SML method. Note that the bias-free of \( W_{jn}(\theta) \) does not affect the estimation results much in the MSM method, as long as \( W_{jn}(\theta) \) and \( P_{jn}(\theta) \) are independent of each other. If we use different sets of random variables in approximating \( W_{jn}(\theta) \) and \( P_{jn}(\theta) \), MSM estimates \( \theta \) that makes \( P_{jn}(\theta) \) unbiased. Third, the MSM method involves no bias in approximating the objective function through the approximation of \( P_{jn}(\theta) \), because the objective function goes to zero when we replace \( P_{jn}(\theta) \) by its approximation from simulation.

On the other hand, the SML method is computationally more economical than the MSM method since we need to simulate only \( W_{jn}(\theta) \) in the SML method and since we have to simulate \( P_{jn}(\theta) \) independently of \( W_{jn}(\theta) \) in the MSM method. Ben-Akiva and Bolduc (1991) found that the objective function of the SML method behaves more nicely than that of the MSM method. The reason is that kernel-smoothed frequency simulators have local flats with the objective function of the MSM method. They also report that it is more economical to use the SML method with hundreds of simulation draws than to use
the MSM method with a small number of draws, while the simulation bias get averaged-out across individuals at a faster rate in the MSM method than in the SML method.

3. Uniform or Normal Random Number Generators

The basic source needed for every method of generating random variables from any distribution or process is i.i.d. uniform (0,1) random variables. The GAUSS software provides random number generators for both uniform distribution and standard normal distribution, and we use their procedures. We can generate random variates with a very general distribution by using the i.i.d. uniform variates. Broadly, we identify four methods of generating random variates: inverse transformation, composition, acceptance and rejection, and special properties. In this section, we briefly summarize the four methods. Commonly, we use more than one method in an algorithm.

3.1. Inverse Transformation

This approach utilizes the fact that the cumulative distribution function (c.d.f.) of any random variate is continuous, monotone increasing, and uniform (0,1) distributed.

Using this fact, we can generate a random variable $x$ with c.d.f. $F(x)$, in the following way:

1) Generate $u$ from uniform[0,1] distribution
2) Set $x = F^{-1}(u)$ and return

For example, suppose that the transit time of a mode $i$ is Weibull distributed with $(\alpha_i, \beta_i h_{in})$. We know that the c.d.f. of Weibull distribution is

$$F(x) = 1 - \exp\left(-\left(x / \beta_i h_{in}\right)^{\alpha_i}\right) \equiv u \in [0,1]$$

Solving the above equation in terms of $x$, we have:

$$x = \left(\beta_i h_{in}\right) \left[-\ln(1-u) \right]^{\frac{1}{\alpha_i}}$$

Also note that if $u \sim \text{uniform}[0,1]$, then $(1-u) \sim \text{uniform}[0,1]$ as well. Thus, we can generate Weibull random variates by the following steps:
1) Generate $u$ from uniform $(0,1) \perp$

2) Set $x = (\beta, h_\infty)[-\ln(u)]^{\alpha}$ and return.

This approach has been the most popular since it is fail-proof and easy to understand. It can be applied to discrete as well as continuous random variables. On the other hand, in order to implement this approach, a researcher should know the form of $F(.)$ exactly in order to specify $F^{-1}(.)$.

### 3.2. Composition

This technique applies when the distribution function ($f$) from which we want to sample can be expressed as a convex combination of other distribution functions ($f_1, f_2, ...$). If $f_1$ and $f_2$ are easier to simulate than $f$, then this approach can be used. For instance, suppose that we want to sample from a negative binomial distribution for a $k$-th success with probability $p$. Rather than directly simulating a negative binomial distribution, we can use the following relationship:

$$h(x) = \int_0^\infty f_{x|y}(x|y)g(y)dy$$

where $f_{x|y}(.)$ is a Poisson distribution and $g(.)$ is a Gamma distribution. Thus, we can generate a negative binomial variate in the following way:

1) Generate $y$ from a Gamma distribution with parameters $(k, (1-p)/p)$.
2) Generate $x$ from a Poisson distribution with mean $y$

### 3.3. Acceptance-Rejection

This approach is not as direct as earlier approaches but can be useful when direct approaches fail to exist or are inefficient. For instance, we use this approach in generating a gamma distribution with $\alpha > 1$. This approach first specifies a function ($t$) that majorizes the density ($f$), i.e. $t(x) \geq f(x)$ for all $x$. The majorizing function $t$ is not a density since

$$c = \int_{-\infty}^{\infty} t(x)dx \geq \int_{-\infty}^{\infty} f(x)dx = 1.$$ 

But the function $r(x)$, defined by $r(x) = t(x)/c$, is a density since $\int_{-\infty}^{\infty} r(x)dx = 1$. If $r(x)$ is easy to simulate, we can generate a random variable from $f(x)$ by the following method:
1) Generate \( y \) from \( r(y) \).
2) Generate \( u \) from \( U(0,1) \), independent of \( y \).
3) If \( u \leq f(y)/t(y) \), set \( x = y \) and return.
   Otherwise, go back to step 1 and try again.

Let us first discuss why this approach is valid. Suppose that \( A \) denotes the event that acceptance occurs in step 3. Then, \( x \) is defined only when the event \( A \) occurs, which is a subset of the entire space of \( y \) and \( u \). Also note that

\[
P(A|Y = y) = P(U \leq \frac{f(y)}{t(y)}) = \frac{f(y)}{t(y)}
\]

For any value of \( x \), the random variable \( y \) accepted only when the event \( A \) occurs has the same distribution as \( x \), since

\[
P(Y \leq x | A) = \frac{P(A \text{ and } Y \leq x)}{P(A)} = \frac{\int_{-\infty}^{x} P(A|Y = y)r(y)dy}{\int_{-\infty}^{\infty} P(A|Y = y)r(y)dy} = \frac{\int_{-\infty}^{x} \frac{f(y)}{t(y)}\frac{t(y)}{c}dy}{\int_{-\infty}^{\infty} \frac{f(y)}{t(y)}\frac{t(y)}{c}dy} = \int_{-\infty}^{x} f(y)dy
\]

A random variable \( x \) that has the value \( y \) generated from \( r(y) \) and accepted randomly with probability \( f(y)/t(y) \) is distributed according to the desired distribution \( f(x) \). Therefore, if it is difficult to simulate from \( f(x) \), we would rather simulate \( y \) from \( r(y) \) and accept it with some criterion function. While this technique always works, the choice of the majorizing function \( t(.) \) influences efficiency greatly. First, we want to generate random variates from \( r(y) = t(y)/c \) quickly. Second, we want to make the rejection probability, \( 1 - f(y)/t(y) \), as small as possible, i.e. to make \( t(.) \) fit \( f(.) \) as closely as possible from the above. Using the above discussed technique, we can generate Gamma \((\alpha,\beta)\) random variates where \( \alpha > 1 \) as follows:

1) Generate \( u_1 \) and \( u_2 \) as i.i.d. \( U(0,1) \).
2) Let \( v = a \ln(u_1/(1-u_1)) \),
   \( y = \alpha \exp(v) \),
   \( z = u_1^2 u_2 \), and
   \( w = b + qv - y \).
3) If \( w+d-\theta z \geq 0 \), set \( x = y \) and return. Otherwise, proceed to step 4.
4) If \( w \geq \ln(z) \), set \( x = y \) and return. Otherwise, go back to step 1.
where \( a = (2\alpha - 1)^{-1/2} \), \( b = \alpha - \ln(4) \), \( q = \alpha + 1/a \), \( \theta = 4.5 \), and \( d = 1 + \ln(\theta) \). Step 3 is an added pre-test which avoids computing the logarithm in the regular acceptance-rejection test in step 4. Without step 3, the algorithm would still be valid, but less efficient. For a detailed explanation of the method, refer to Cheng (1977).

3.4. Special Properties

We can generate random variates by using their special properties. For example, the sum of Exponential variables is Gamma distributed. Since it is easy to inverse-transform the c.d.f. of the Exponential distribution, we can generate Exponential variates first and use their sum as our random variable. For another example, we know that the ratio of two gamma distributions \( \alpha_1 / (\alpha_1 + \alpha_2) \) is Beta \( \alpha_1, \alpha_2 \)-distributed. While Beta distribution has no closed form of c.d.f., we can generate two Gamma variates first, and then use their ratio as our Beta variates.

As another example, we can generate correlated normal variates by using a Cholesky lower triangular matrix. Suppose that the covariance matrix of desired normal variates is \( \Sigma \). Since \( \Sigma \) is a covariance matrix, \( \Sigma \) is positive definite and can be decomposed into \( \Gamma \Gamma^\prime \) where \( \Gamma \) is a Cholesky-decomposed lower triangular matrix. For a positive definite symmetric matrix, a unique Cholesky lower triangular matrix exists. [This property does not require that the first element of \( \Sigma \) be one. Normalization is required for the identification of the probit model rather than for the Cholesky-decomposition.] The GAUSS software provides a procedure for calculating a Cholesky lower triangular matrix. Then, we get correlated random variates from the following relationship:

\[
\xi_n = \Gamma \varepsilon_n \quad \text{where} \quad \varepsilon_n \sim N(0, I)
\]

1. Generate standard normal variates \( \varepsilon_n \)
2. Multiply \( \varepsilon_n \) by \( \Gamma \).

4. Variance Reduction

If we can somehow reduce the variance of the output random variable of interest without disturbing its expected value, we can obtain greater precision with the same number of simulation draws. In general, variance reduction techniques can be classified into five categories: common random numbers, antithetic variates, control variates, indirect estimation, conditional expectations, stratified sampling and importance sampling. We will briefly discuss them in this section.

Let us first note that the effectiveness of different techniques depends on the model specification and that we usually do not know in advance which technique will be the most effective. Moreover, these techniques have different computing costs at each iteration. Thus, we should consider a tradeoff among how easy to program a technique, how many
random numbers we should draw, how much variance the technique reduces, how much computational costs it takes at each iteration. Preliminary pilot runs are recommended.

4.1. Common Random Numbers

This key to using simulation for parameter estimation is to have the same set of random variates generated at each iteration. With the same set of random variates, parameter estimates do not move around, and the algorithm converges in finite steps without cycling. We can achieve common random numbers by re-setting the seed of a random number generator to a fixed value at the beginning of each iteration. The idea of common random numbers can also be used when we want to compare different specifications. By comparing their results under the same situations, we can analyze whether the differences come from specification or from simulation.

4.2. Antithetic Variates

Suppose that we plan to run $R$ simulation draws. We divide them into $R/2$ pairs of runs resulting in estimates $(\theta^{(r,1)}, \theta^{(r,2)})$ such that each pair is independent of every other pair, i.e. $\theta^{(r,1)}$ and $\theta^{(r,2)}$ are independent of each other if $r \neq s$. Let

$$\theta^{(r)} = \frac{\theta^{(r,1)} + \theta^{(r,2)}}{2}$$

and let $\hat{\theta} = E[\theta^{(r)}] = \frac{2}{R} \sum_{r=1}^{R/2} \theta^{(r)}$ be the point estimator of the parameter. Since the $\theta^{(r)}$'s are i.i.d. random variables, the variance of the parameters will be:

$$ar(\hat{\theta}) = \frac{2}{R} Var(\theta^{(r)}) = \frac{Var(\theta^{(r,1)}) + Var(\theta^{(r,1)}) + 2Cov(\theta^{(r,1)}, \theta^{(r,2)})}{2R}$$

Thus, if the random numbers used in the second run are negatively correlated with those in the first run, the variance of our estimates will be reduced.

4.3. Control Variates

If we have two random variables that are correlated with each other, we can reduce the variance without making pairs of runs. Suppose that we want to estimate $\beta$, and that $\alpha^{(r)}$ is another random variable which is believed to be correlated with $\beta^{(r)}$ and has a known expectation $\alpha$. Then, the random variable $\gamma^{(r)} = \beta^{(r)} - k(\alpha^{(r)} - \alpha)$ is also an unbiased estimator of $\beta$ for any real number $k$. Moreover,
\[ ar(\gamma^{(r)}) = Var(\beta^{(r)}) + k^2 Var(\alpha^{(r)}) - 2k Cov(\beta^{(r)}, \alpha^{(r)}) \]

which will be smaller than \( ar(\beta^{(r)}) \) if \( k^2 Var(\alpha^{(r)}) < 2k Cov(\beta^{(r)}, \alpha^{(r)}) \) where \( k \) can be either positive or negative. The random variable \( \alpha^{(r)} \) is called a control variable. Good control variables should be strongly correlated with the parameters of our interest, either positively or negatively. This approach can be useful for Gibbs sampler which performs Bayesian updating. The input variable that determines a priori distribution is usually correlated with the output variable which is our estimates for the posterior distribution.

### 4.4. Indirect Estimation

If the parameters can be decomposed into two groups - one with unknown value and the other with theoretical or a priori known value, then we can replace parameters in the second group by their fixed values and estimate parameters in the first group. In this way, the variance of parameters in the first group is greatly reduced.

### 4.5. Conditional Expectations

Suppose that \( \beta^{(r)} \) is an output random variable of interest and we want to estimate \( \beta = E(\beta^{(r)}) \). Also, suppose that there is some other random variable \( \alpha^{(r)} \) such that we can calculate the conditional expectation \( E(\beta^{(r)}|\alpha^{(r)} = a) \) for any real number \( a \). Then,

\[ \beta = E(\beta^{(r)}) = E[E(\beta^{(r)}|\alpha^{(r)})] \]

Moreover,

\[ ar[E(\beta^{(r)}|\alpha^{(r)})] = Var(\beta^{(r)}) - E[Var(\beta^{(r)}|\alpha^{(r)})] \leq Var(\beta^{(r)}) \]

Thus, \( E(\beta^{(r)}|\alpha^{(r)}) \) is unbiased for \( \beta \) and has a smaller variance than \( \beta^{(r)} \). Thus, we would rather observe \( E(\beta^{(r)}|\alpha^{(r)}) \) rather than observe \( \beta^{(r)} \) directly. Note that \( E(\beta^{(r)}|\alpha^{(r)} = a) \) is a function of the real number \( a \), whereas \( E(\beta^{(r)}|\alpha^{(r)}) \) is a random variable.

### 4.6. Stratified Sampling

Suppose that we want to simulate interest rate from a distribution that has a long tail. For example, suppose that interest rate is from uniform[0, 20\%] distribution for most shipments (90 \% of the times) and from uniform[20, 80\%] distribution for 10\% of the times. A naive approach is to generate interest rate randomly with the required probability. With 100 simulation draws, this approach would generate high interest rate (i.e. greater than 20\%) in only about 10 cases. Yet, it is usually a high interest rate (e.g. emergency shipments) that is of our interest, while emergency shipments are rare events (e.g. production stop due to a shortage of some critical parts or cancelation of local promotions...
due to a lack of sufficient inventory) under reasonable operation plans. Thus, we want to ensure enough number of emergency shipments in the simulation draws.

One way to increase simulation efficiency is to draw interest rates separately from different groups. In order to implement this stratified sampling, we need to understand the characteristics of each group. Suppose that our prime objective is to understand the probability of choosing truck for intercity freight shipment and that we got the following mean and variance of truck choice probability by running a pilot simulation of 20 draws from each interest rate.

<table>
<thead>
<tr>
<th>Interest Rate</th>
<th>P(event)</th>
<th>Number of runs</th>
<th>E[P(truck)]</th>
<th>V[P(truck)]</th>
</tr>
</thead>
<tbody>
<tr>
<td>between 0 and 20%</td>
<td>0.90</td>
<td>20</td>
<td>0.45</td>
<td>0.20</td>
</tr>
<tr>
<td>between 20% and 80%</td>
<td>0.10</td>
<td>20</td>
<td>0.85</td>
<td>0.60</td>
</tr>
</tbody>
</table>

Based on the pilot results, we may expect that with $n_1$ runs for interest rate between 0 and 50%, the variance of our estimate of truck choice probability will be approximately:

$$VR = 0.90^2 \cdot 0.20 / n_1 + 0.10^2 \cdot 0.60 / (T - n_1)$$

where $T$ is the total number of draws that simulation budget allows (say, 100 draws). The variance is minimized when $n_1$ is about 76 and $n_1$ is about 24. Since we have already run 20 draws from each group, we need to run only 56 and 4 additional draws from each group. Note that we gain in the requirement of sample size, i.e. we can run 14 ($=24-10$) more draws for the second group this way without increasing variance.

Suppose that we have $E_1$ and $E_2$ as the mean truck choice probability of each group by using both pilot and additional draws. We can calculate the overall mean truck choice probability by reweighing observations as follows:

$$ER = 1.18 \times 0.90 \times E_1 + 0.42 \times 0.10 \times E_2$$

Reweight parameters are calculated by dividing population probability by the sampling probability (i.e. $1.18=0.90/0.76$ and $0.42=0.10/0.24$). If we also denote $V_1$ and $V_2$ to be the variance of truck choice probability of each group by using both pilot and additional draws, the overall variance of truck choice probability will be as follow:

Variance from stratified sampling = $1.18^2 \times 0.90 \times V_1 + 0.42^2 \times 0.10 \times V_2$

Now, let us extend the above observation into a case where the heterogeneity of interest rate is Exponentially distributed. Its density function is:

$$f(x) = c \cdot \exp(-\alpha x), \quad x \geq 0$$

and its cumulative density function (c.d.f.) is
\[ F(x) = 1 - \exp(-\alpha x), \quad x \geq 0 \]

Since we have a closed form of the c.d.f., we can simulate interest rate easily by using the following inverse transformation:

1) Generate \( u \) from uniform \((0,1)\)
2) Set \( x = - (\ln u) / \alpha \) and return.

The problem with this approach is that since this density has a long tail, the time required to obtain draws from a tail-side may be very long. Suppose now that we collect data from two separate groups (say, \( n_1 \) observations with a c.d.f. smaller than \( t \) and \( n_2 \) observations with a c.d.f. larger than \( t \)). Using the fact that a c.d.f. is equally likely to have all values between 0 and 1, we can generate interest rates of two groups as follows:

1) Generate \( u \) from uniform[0,1] distribution
2) Set \( x_1 = - (\ln (u^*t)) / \alpha \) and return \( n_1 \) times.
3) Set \( x_2 = - (\ln (t + u^*(1-t))) / \alpha \) and return \( n_2 \) times.

With the above draws, we have the following characteristics of sample draws:

<table>
<thead>
<tr>
<th></th>
<th>Group 1</th>
<th>Group 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population Ratio</td>
<td>( t )</td>
<td>( 1 - t )</td>
</tr>
<tr>
<td>Sampling Ratio</td>
<td>( n_1 / (n_1 + n_2) )</td>
<td>( n_2 / (n_1 + n_2) )</td>
</tr>
<tr>
<td>Reweight Parameter</td>
<td>( t(n_1 + n_2) / n_1 )</td>
<td>( (1 - t)(n_1 + n_2) / n_2 )</td>
</tr>
</tbody>
</table>

We now apply pilot and additional runs in a similar way as we did with uniform distributions.

### 4.7. Importance Sampling

We often need to run many simulation draws in order to ensure that some random variates are generated from the tail side as well as from the center of a distribution. Typically, random numbers are generated easily from the center but seldom from the tail. And models sensitive to outlier behaviors need many draws to ensure outlier behaviors. In such case, the number of simulation draws can be reduced by segmenting simulation
samples into two groups, one simulated from the center and the other from the tail, and by applying a correction factor to compensate.

For instance, suppose that we want to estimate rail share as a function of interest rate which has the Gamma(2, 1) distribution, i.e.

\[ P_R(r) = \int P(R|r)f(r)dr \]

where \( P(R|r) \) is the probability of choosing rail conditional on interest rate \( r \), and

\[ f(r) = r \exp(-r) \quad x>0. \]

One way to estimate rail share is:

\[ P_R(r) = \int \left( \frac{P(R|r)f(r)}{h(r)} \right) h(r)dr \]

where \( h(r) \) is a distribution which is easy to simulate such as Exponential (\( \lambda=1 \)), e.g.

\[ h(r) = \exp(-r) \]

Note that \( h(r) \) is easy to simulate and that it has a lighter tail than \( f(r) \) so that it falls away in the tail faster than \( f(r) \).

Alternatively, we can combine the original distribution and a different distribution. Note that our major concern is the difficulty to simulate from a tail of \( f(r) \). If we decide to sample separately from a c.d.f. smaller than \( t \) and from a c.d.f. larger than \( t \), we may consider the following distribution:

\[ b(r) = \begin{cases} b_1(r) = \frac{1}{c}r\exp(-r), & 0 < r < t \\ b_2(r) = \frac{1}{c}t\exp(-r), & t \leq r \end{cases} \]

where \( c \) is a normalizing constant such that \( \int_0^\infty b(r)dr = \) (in this case, \( c=1-\exp(-t) \)) and \( t \) is an arbitrary parameter. The probability of being on \((t, \infty)\) is \( \int_t^\infty b_2(r)dr = \frac{t}{e^t - 1} \). The ratio of \( f(r)/b(r) \) is as follows:

\[ \frac{f(r)}{b(r)} = \begin{cases} c, & 0 < r < t \\ cr, & t \leq r \end{cases} \]

270
By sampling separately, we can control the number of samples from a tail relative to the number of samples from a center. In summary, we can sample from $b(r)$ by the following composition method:

1) Generate random number $u$ from uniform$[0,1]$
2) If $u$ is greater than $t/\exp(t)-1$, we will generate $r \in (0, t)$.
   Generate variates from Gamma $(2, 1)$ distribution and reject any draw if it exceeds $t$. 
3) If $u$ is less than $t/\exp(t)-1$, we will generate $r \in (t, \infty)$.
   Generates variates from an Exponential distribution and multiply by $t/c$.

5. Desired Properties of Random Number Generators

   There is no simple answer as to which approach is "best". The choice of the best algorithm depends on how many variates we need, how much variance we want to reduce, whether parameter values are fixed or changing, and whether order statistics are needed. It is up to the researcher to judge and decide which algorithm will be used in order to generate random variates from a desired distribution. As a minimum guideline, the following three properties may be checked before selecting an algorithm:

   1) exactness: the algorithm produces random variables with the exact distribution.
   2) efficiency: the algorithm is efficient in terms of storage space and execution time
   3) robustness: the algorithm behaves in a desired way for all parameter values
Appendix 7. Stated Preference Data

Conventional shipper surveys solicit shippers to evaluate the importance of service attributes. We have discussed in Chapter 2 that this stated-importance approach does not provide reliable estimates of the importance since the response is made without relation to actual choices. Instead, the stated-preference format has been recommended for shippers' survey. In marketing, the SP approach is usually called conjoint analysis, and many commercial software are available for. In this section, we review variations of the SP approach and discuss what approach would best fit the railroad's needs. We may first categorize the SP approach as follows:

1) By the response variable
   - rating vs. ranking vs. choice vs. share vs. willingness-to-pay

2) By the method of presenting alternatives
   - Line-by-line vs. Partial list vs. Full list

3) By the method of presenting alternative attributes
   - Full-profile vs. Partial-profile

4) By the method of handling large number of attributes
   - Bridging vs. Hierarchical vs. Bundling

1. Response variable

A. Rating-based SP Approach

One can give descriptions of an alternative and ask how much they would like the alternative on a continuous scale. For example:

Suppose that a carrier proposes the following service to you.

<table>
<thead>
<tr>
<th>Service Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Freight rate per cwt relative to the current truck rate</td>
<td>-20%</td>
</tr>
<tr>
<td>Door-to-door transit time</td>
<td>5 days</td>
</tr>
<tr>
<td>The percentage of shipments that arrive on the scheduled day</td>
<td>80%</td>
</tr>
<tr>
<td>The percentage of bills that have no errors</td>
<td>90%</td>
</tr>
</tbody>
</table>

How likely are you to purchase its service?

0% --- 10% --- 20% --- 30% --- 40% --- 50% --- 60% --- 70% --- 80% --- 90% --- 100%
When all data are collected, reported preferences are regressed on the attribute levels. The coefficients estimated from the preference function are the importance of attributes. While this approach has been, and still is, very popular in marketing research, there are several drawbacks, and railroad managers will be better off avoiding this approach.

First, this rating-based approach lacks validity since respondents do not need to consider any trade-offs in rating alternatives. In order to improve validity, marketing researchers have been debating about a wrong question - how to define the preference scale. Proposed ideas vary widely, including how much they would like the alternative, how satisfied they will be with the alternative, how likely they are to purchase the alternative and how likely they would recommend the purchase of the alternative to others. Secondly, preference regression assumes that consumers have deterministic preferences.\(^1\) It lacks a behavioral background which random utility models have, and it only tries to fit data by reducing the variations that are not explained by the fitted line. Thirdly and most importantly, many marketing research run a regression for each respondent, and call the estimates part-worths of attributes for the respondent. Each respondent evaluates at most 20 to 30 alternatives, and these 20 to 30 observations do not give consistent estimates of the importance of 5 to 7 attributes. That is, the analysis of individual-level part-worths is totally unreliable.

If this rating-based data were collected, we would recommend that a researcher segment respondents into homogeneous groups and estimate attribute importances by running ordered probit for the segments rather than running a regression for each respondent. Ordered probit can then be estimated with the following assumptions:

\[
U_{in} = \beta' X_{in} + \varepsilon_{in} \text{ where } \varepsilon \sim N(0, \sigma^2) \\
P_{in}(10\%) = P(U_{in} = \beta' X_{in} + \varepsilon_{in} \leq \mu_1) = \Phi((\mu_1 - \beta' X_{in})/\sigma) \\
P_{in}(20\%) = P(\mu_1 \leq \beta' X_{in} + \varepsilon_{in} \leq \mu_2) = \Phi((\mu_2 - \beta' X_{in})/\sigma) - \Phi((\mu_1 - \beta' X_{in})/\sigma) \\
... \\
P_{in}(100\%) = P(\mu_9 \leq \beta' X_{in} + \varepsilon_{in}) = 1 - \Phi((\mu_9 - \beta' X_{in})/\sigma)
\]

With the above choice probabilities, we can estimate parameters that maximize the following log-likelihood:

\[
L(\beta, \mu) = \sum_{n=1}^{N_s} \sum_{i \in A_{ni}} d_{in}^{SP} \log(p_{in}^{SP}(d_{in}^{SP}))
\]

In practice, the locations of thresholds rarely affect the systematic utility estimates, \(\beta\). Thus, we usually can reduce the number of parameters by setting \(\mu_5=0, \mu_9=-\mu_1, \mu_8=-\mu_2, \mu_7=-\mu_3, \mu_6=-\mu_4\) where \(\mu_1 \leq \mu_2 \leq \mu_3 \leq \mu_4 \leq 0\)

---

B. Ranking-based SP Approach

The ranking-based approach was proposed as a way to force respondents to compare alternatives. For example:

Suppose that a carrier proposes the following services to you in terms of price, transit time, service reliability and billing accuracy. [Variables are defined as in the above example.] Please rank them in the order you would prefer to have them.

<table>
<thead>
<tr>
<th>Alternative</th>
<th>Price</th>
<th>Transit Time</th>
<th>Service Reliability</th>
<th>Billing Accuracy</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-20%</td>
<td>5 days</td>
<td>70%</td>
<td>90%</td>
<td>_____</td>
</tr>
<tr>
<td>2</td>
<td>+10%</td>
<td>2 days</td>
<td>95%</td>
<td>99%</td>
<td>_____</td>
</tr>
<tr>
<td>:</td>
<td>:</td>
<td>:</td>
<td>:</td>
<td>:</td>
<td>:</td>
</tr>
<tr>
<td>15</td>
<td>-40%</td>
<td>7 days</td>
<td>60%</td>
<td>80%</td>
<td>_____</td>
</tr>
</tbody>
</table>

In analyzing ranking data, market researchers tend to perform erroneous analysis by treating them as if they represent utilities and by applying regression on them directly. A correct method for analyzing ranking data with the proper behavioral theory is as follows (Ben-Akiva et al.): Suppose that a person ranked alternatives in the order of A, B, . . . , Z.

\[
U_{in} = \beta' X_{in} + \epsilon_{in} \quad \text{for } i = A, B, \ldots, Z
\]

\[
P_n(A,\{A,B,\ldots,Z\}) = P(A|\{A,B,\ldots,Z\}) \cdot P(B|\{B,C,\ldots,Z\}) \ldots P(Y|\{Y,Z\})
\]

If we assume that all choice probabilities follow the same logit model with the same variance, the rank order probability is simply the product of logit choice probabilities conditional on the ordered choice sets. The rank order probability for alternatives that have different levels of variability is simply the product of logit choice probabilities with some adjustment for the scale effects, i.e.

\[
P_n(i|\{i,i+1,\ldots,Z\}; \beta, \mu_i) = \frac{\epsilon^{\beta'X_{in}/\mu_i}}{\sum_{j=i}^{Z} \epsilon^{\beta'X_{jn}/\mu_i}}
\]

Although this approach ensures the full utilization of the ranking data, the estimation quickly becomes problematic as the number of alternatives needed to be ranked increases.

---

C. Choice-based SP Approach

The choice-based approach asks respondents to choose the alternative they prefer the most. For example:

Suppose that a carrier proposes the following services to you in terms of price, transit time, service reliability and billing accuracy. [Variables are defined as in the above example.] Which alternative would you prefer to have?

<table>
<thead>
<tr>
<th>Transport Alternative</th>
<th>Price</th>
<th>Transit Time</th>
<th>Service Reliability</th>
<th>Billing Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Intermodal</td>
<td>-20%</td>
<td>5 days</td>
<td>70%</td>
<td>90%</td>
</tr>
<tr>
<td>2 Truck</td>
<td>+10%</td>
<td>2 days</td>
<td>95%</td>
<td>99%</td>
</tr>
<tr>
<td>3 Rail</td>
<td>-40%</td>
<td>7 days</td>
<td>60%</td>
<td>80%</td>
</tr>
</tbody>
</table>

Since choice is our ultimate object of study, this approach ensures external validity. Moreover, it has the advantage of allowing a researcher to experiment with a choice set provided to respondents and to measure the effects of the choice set.

This approach, however, greatly increases data requirements, since a researcher needs to provide a detailed description of choice situations. The huge size of survey sometimes scares respondents and causes non-response bias. In addition, Elrod et al. claimed that both the rating-based and the choice-based approach predict holdout shares well and that neither approach was better than the other.\(^3\)

D. Share-based SP Approach

The share-based approach asks respondents to allocate 100 points to the proposed alternatives in terms of usage likelihood. For example:

Suppose that a carrier proposes the following services to you in terms of price, transit time, service reliability and billing accuracy. [Variables are defined as in the above example.] What percentage of your annual shipments would you ship via each mode?

<table>
<thead>
<tr>
<th>Transport Alternative</th>
<th>Price</th>
<th>Transit Time</th>
<th>Service Reliability</th>
<th>Billing Accuracy</th>
<th>Share</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Intermodal</td>
<td>-20%</td>
<td>5 days</td>
<td>70%</td>
<td>90%</td>
<td>____%</td>
</tr>
<tr>
<td>2 Truck</td>
<td>+10%</td>
<td>2 days</td>
<td>95%</td>
<td>99%</td>
<td>____%</td>
</tr>
<tr>
<td>3 Rail</td>
<td>-40%</td>
<td>7 days</td>
<td>60%</td>
<td>80%</td>
<td>____%</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>100%</td>
</tr>
</tbody>
</table>

---

When survey data is collected, we can estimate the importances of attributes by applying the same estimation method which we used for the RP share data.

**E. Willingness-to-pay-based SP Approach**

Another idea that recently received attention is to ask respondents directly how much they would be willing to pay for a proposed service ("open-ended"). One may give respondents a series of prices that keep increasing (or decreasing) until respondents finally say that they are willing to pay the amount ("single referendum"). Or one may start at the mid-range price and change prices according to respondent's responses ("double referendum"). An example of the double referendum question is as follow:

<table>
<thead>
<tr>
<th>Suppose that a carrier proposes the following service to you.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transportation model is intermodal</td>
</tr>
<tr>
<td>Door-to-door transit time is 5 days</td>
</tr>
<tr>
<td>The percentage of shipments that arrive on the scheduled day is 80 %</td>
</tr>
<tr>
<td>The percentage of bills that have no errors is 90 %</td>
</tr>
</tbody>
</table>

Would you use this service for a freight rate per cwt that is 10% lower than the current truck rate?

[A researcher pre-specifies the rate range. For example, -40%, -30%, -20%, -10%, 0%, 10%, and 20%. If the response is "yes", a researcher asks the same question with the next higher price, i.e. 0% in this case. If the response is "no", a researcher asks the same question with the next lower price, i.e. -20% in this case.]

This approach tries to estimate an utility function from the reservation price that respondents report. Suppose that a respondent reports prices $P_{An}, \ldots, P_{Zn}$ for alternatives $(A, \ldots, Z)$ with non-price attributes $Z_{An}, \ldots, Z_{Zn}$. The prices represent the respondent's reservation price $P_{An}^*, \ldots, P_{Zn}^*$ with some errors, i.e.

$$P_{An}^* = P_{An} + \delta_{in}$$

where $\delta_{in}$ may be modeled as an i.i.d. normally-distributed across the population with mean zero and variance $\sigma^2_{\delta}$. Random utility $\epsilon_{in}$ is also assumed to be i.i.d. normally-distributed with mean 0 and variance $\sigma^2_{\epsilon}$. We also denote $\beta_0$ and $\beta_1$ as the importance weight for price and for other attributes, respectively. Respondents will have the same level of utility at the reported price levels, i.e.

$$\beta_0 P_{An}^* + \beta_1 Z_{An} + \epsilon_{An} = \ldots = \beta_0 P_{Zn}^* + \beta_1 Z_{Zn} + \epsilon_{Zn}$$

which can be rewritten as
\[ \beta_0 P_{An} + \beta_1' Z_{An} + \beta_0 v_{An} + \varepsilon_{An} = \ldots = \beta_0 P_{Zn} + \beta_1' Z_{zn} + \beta_0 v_{Zn} + \varepsilon_{Zn} \]

We may estimate \( \beta_0 \) and \( \beta_1 \) which minimize the measurement errors, if we ignore the random utility effects, i.e.

\[
\min_{\beta} \sum_{n} \sum_{i=A}^{Z} (\beta_0 (P_{in} - P_{in}) + \beta_1' (Z_{in} - Z_{in}))^2
\]

for a reference alternative \( J \). This approach can be elaborated further by asking for descriptions of the alternative that respondents are currently using. The current alternative can be a good reference alternative, as long as the current alternative is the one most preferred among all available alternatives. We may also elaborate the model into a generalized minimum distance estimator by incorporating the covariance matrix. The problem with this approach is that the model does not assume random utility. Moreover, respondents may find it difficult to state the maximum price they are willing to pay, increasing the possibility of measurement errors or non-response bias.

One may give a price list which respondents can choose from and ask them to choose the maximum price they are willing to pay for the proposed alternatives ("price-listed referendum"). Denoting the price list by \( (P(1), \ldots, P(K)) \) which is in ascending order and assuming that a respondent \( n \) chooses \( P_{in}^{(k_n)} \),

\[ \beta_0 P_{in}^{(k_n)} + \beta_1' Z_{in} + \mu_{in} \geq U_{in} \geq \beta_0 P_{in}^{(k_{in} + 1)} + \beta_1' Z_{in} - \mu_{in} \]

where \( \mu_{in} \) is i.i.d. normally distributed with mean zero and heteroscedastic variance \( \sigma_{\mu}^2 \). Also, suppose that \( U_{in} \) is i.i.d. normal distributed across the population with mean \( \kappa_i \), i.e. \( U_{in}^{\ast} = \kappa_i + \xi_{in} \) where \( \xi_{in} \) is i.i.d. normally distributed with mean zero and heteroscedastic variance \( \sigma_{\xi}^2 \). In such case, the likelihood of observing price choices would be:

\[ L(\beta) = \prod_{n} \prod_{i} P(\beta_0 P_{in}^{(k_{in})} + \beta_1' Z_{in} - \kappa_i \geq \xi_{in} - \mu_{in} \geq \beta_0 P_{in}^{(k_{in} + 1)} + \beta_1' Z_{in} - \kappa_i) \]

This model can be estimated with the ordered probit model, assuming that \( \xi_{in} \) and \( \mu_{in} \) are independent of each other.

Willingness-to-pay-based approaches try to find the maximum price customers are willing to pay. If respondents can reveal their maximum value correctly, this approach will allow firms to practice value-based pricing. McFadden and Leonard, however, found that both single and double referendum showed a starting point bias, due to fatigue or anchoring, that made the outcomes of repetitive surveys sensitive to the response path. After extensive experiments, they concluded that the estimates of willingness-to-pay for

---

environmental goods are highly subjective to survey design, including situational descriptions and wording and sequence of questions. No survey form provided superior power in explaining choices. The willingness-to-pay approach also focuses too much on the price attribute which may become less important relative to service quality. Moreover, a joint estimation of modal selection and shipment size becomes very difficult.

2. Method of Presenting Alternatives

Conjoint experiments can be classified in terms of how to present the choice scenario. In the line-by-line approach, respondents are given one alternative at a time and are asked to evaluate the alternative. In the partial-list approach, respondents are given a subset of alternatives to evaluate at each time. In particular, respondents compare two alternatives at a time in pair-wise comparisons. In full-list comparisons, respondents evaluate the whole list of alternatives at the same time. While the rating-based approach usually employs line-by-line evaluation, the ranking-based approach employs full-list comparison. As for the choice-based approach, it usually employs pair-wise or partial-list comparison. Meanwhile, the share-based approach employs full-list comparison, and the willingness-to-pay-based approach uses line-by-line evaluation.

3. Method of Presenting Alternative Attributes

In describing alternatives, one may want to provide profiles that are as detailed as possible. Yet, when alternatives have many, i.e. more than 10, potential attributes, presenting all of them at the same time can confuse and distract respondents. In addition, a long survey form may discourage respondents from completing the survey or may induce them to look at only one or two important variables and respond randomly. If so, presenting full profiles increases either response randomness and non-response bias. Simplification in the pair-wise approach facilitates the comparison and reduces the efforts of respondents. On the other hand, some expressed concerns about the possibility that respondents could forget where they are in the trade-off table when only partial profiles are given or about the possibility that respondents may be forced to make inferences about omitted attribute levels.⁵

Nevertheless, pairwise comparison and partial profiles seem to be much more popular than others in actual market studies. In fact, Huber and Hansen found, in an empirical study of pair-wise comparison, that presenting five attributes at the same time did not do better than presenting two (or possibly three) attributes each time.⁶ This result is attributed to the fact that simple questionnaires make it easier for respondents to judge utility differences.

---

4. Large-Number of Attributes

Depending on how deep we are interested in learning the effects of service attributes, we may define numerous dimensions that have to be analyzed. For example, freight transportation has the long list of service attributes described in the Appendix 3.1. The list includes attributes that are required as a minimum for long-term relationship; extra services that a carrier provides in order to differentiate her service from others; engineering attributes that can be measured objectively; and perceptual attributes that can only be surveyed. As the number of attributes increases, the list of hypothetical alternatives can explode. The more alternatives a respondent has to evaluate, the more likely he will not participate in the survey. How to condense the large number of potential alternatives in order to make a short survey form has been a big problem. So far, three different approaches have been proposed to handle the large-number of alternative attributes: bridging, hierarchical and bundling. We will review them below:

A. Bridging Process

The first approach employs a bridging process where separate designs are bridged by at least one common attribute, and estimates from subexperiments are rescaled to a common scale. Naturally, this approach employs sequential data collection and tries to use information collected from one stage in the next stage. As examples of this approach, we review adaptive and hybrid conjoint analysis below:

Adaptive conjoint analysis is basically based on pair-wise comparison of partial profiles. In the first stage, it asks respondents to evaluate the importance of each attribute and selects profiles of the second stage based on the self-stated importance. In the third stage, the computer constructs several full profiles based on previous responses and asks the respondents to evaluate them on a purchase likelihood scale. The procedure is claimed to be adaptive in the sense that a computerized system uses the information obtained about a respondent's partworths at each step and selects the best profiles for evaluation, namely pairs whose utility differences cannot be very large. Their selection criteria for good profiles use only the estimated utility rather than the design matrix of explanatory variables and this may be improved.\(^7\)

Hybrid conjoint analysis asks respondents to self-state importance of attributes in the first stage. In the second stage, alternatives are fully profiled with different attribute levels but with the same price. The profile is composed of seven facets with each facet having five levels. Using factorial design, the model selects 50 plans that are minimal for capturing orthogonal main effects. Then, respondents are asked ten times to choose among five alternatives in terms of the likelihood of purchase frequency, i.e. among:

---

"purchase all the time, on a regular basis, now and then, rarely, and never". It is interesting that the likelihood of purchase frequency is used interchangeably with purchase likelihood. Modeling purchase frequency jointly with choice as a function of utility should be a necessary extension. In the third stage, respondents allocate 100 points to five alternatives that are fully profiled with different prices, based on how likely they would be to purchase the alternative at the given price.  

B. Hierarchical Process

Hierarchical conjoint analysis is based on the theory of hierarchical information integration that decision makers do not evaluate all available attributes in a bottom-up fashion, and that they rather form perceptions of attributes on a high or abstract level and make choices based on the high-level perceptions. This approach utilizes choice experiments by categorizing attributes into several non-overlapping sets, running separate subexperiments for each category, and running overall experiments where respondents make choices based on their overall impression defined from the subexperiments. For instance, respondents are presented with two alternatives whose attributes differ only within a predetermined category and are asked to evaluate them on a 1 to 9 scale of preference. Then preference ratings of other categories are given to the two alternatives, and respondents are asked to allocate a budget among the two alternatives and a "none of the two" option. The following is an example:

1. Suppose that you have two modes available. How favorable do you think the services are in terms of convenience if they provide the following attributes?

<table>
<thead>
<tr>
<th></th>
<th>Alternative 1</th>
<th>Alternative 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>provide EDI (electronic data interchange)</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>provide pickup and delivery</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>a single call provides all necessary service</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>provides credit lines and electronic fund transfer</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>YOUR IMPRESSION OF CONVENIENCE</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Extremely unfavorable | Very bad | Moderately bad | Moderately So-and-so | good | Very good | Extremely favorable
1-------------------2-------------------3-------------------4-------------------5-------------------6-------------------7

2. Suppose that the above alternatives also provide the following levels of service in addition to convenience. How would you allocate your shipments between the two?

---

C. Bundling Process

Instead of asking respondents to choose among different alternatives, we may ask them to design an alternative that they desire the most. The following is an example:

Suppose that you are considering a selection of new core carriers for inter-D.C. shipments. Based on the proposals of carriers participating in a bid, you discovered that the carriers can be differentiated by the following nine features. Moreover, you also found that you have to pay for each feature in addition to the base freight rate of 500 cents per trailer-mile, as indicated in the parenthesis.

<table>
<thead>
<tr>
<th>Feature</th>
<th>Desired level (additional cost: cents per trailer-mile)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transit time</td>
<td>20% faster than now (50) same (0) 20% slower (-50)</td>
</tr>
<tr>
<td>Consistency</td>
<td>95% on-time (10) 90% (0) 80% on-time (-30)</td>
</tr>
<tr>
<td>Loss or damage</td>
<td>≤0.3% of value lost (20) ≤1% (0) ≤3% value lost (-10)</td>
</tr>
<tr>
<td>Equipment usability</td>
<td>99% of times (30) 95% (0) 90% usable (-20)</td>
</tr>
<tr>
<td>EDI</td>
<td>Offer (10) Do not offer (-10)</td>
</tr>
<tr>
<td>Billing accuracy</td>
<td>99% of times (20) 95% (0) 95% of times (-10)</td>
</tr>
<tr>
<td>Real-time tracing</td>
<td>Offer (50) Do not offer (0)</td>
</tr>
<tr>
<td>Level of efforts</td>
<td>easy to work with (20) normal (0) difficult (-10)</td>
</tr>
<tr>
<td>Responsiveness</td>
<td>always satisfactory (20) sometimes (0) seldom (-10)</td>
</tr>
</tbody>
</table>

Given the above costs for service levels, please mark inside the parenthesis what levels of service do you think are the most appropriate for your company?
<table>
<thead>
<tr>
<th>Transit time</th>
<th>20% faster than now ( )</th>
<th>same ( )</th>
<th>20% slower ( )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consistency</td>
<td>95% on-time ( )</td>
<td>90% ( )</td>
<td>80% on-time ( )</td>
</tr>
<tr>
<td>Loss or damage</td>
<td>≤0.3% of value lost ( )</td>
<td>≤1% ( )</td>
<td>≤3% value lost ( )</td>
</tr>
<tr>
<td>Equipment usability</td>
<td>99% of times ( )</td>
<td>95% ( )</td>
<td>90% usable ( )</td>
</tr>
<tr>
<td>EDI</td>
<td>Offer ( )</td>
<td></td>
<td>Do not offer ( )</td>
</tr>
<tr>
<td>Billing accuracy</td>
<td>99% of times ( )</td>
<td>95% ( )</td>
<td>90% of times ( )</td>
</tr>
<tr>
<td>Real-time tracing</td>
<td>Offer ( )</td>
<td></td>
<td>Do not offer ( )</td>
</tr>
<tr>
<td>Level of efforts</td>
<td>easy to work with ( )</td>
<td>normal ( )</td>
<td>difficult ( )</td>
</tr>
<tr>
<td>Responsiveness</td>
<td>always satisfactory ( )</td>
<td>sometimes ( )</td>
<td>seldom ( )</td>
</tr>
</tbody>
</table>

Note that if we try to construct alternatives as combinations of attributes, we can construct $8,748 (=3^7 \times 2^2)$ different alternatives. Even if we do not consider the interaction effects among attributes and focus on main effects, we still need to generate 25 alternatives with an orthogonal array. Rating or ranking 25 fully-profiled alternatives is a difficult job that respondents will be discouraged from performing. Constructing alternatives with partial profiles takes obviously a much larger number of questionnaires and is not feasible. By soliciting respondents to reveal their preferences at the attribute level rather than at the alternative level, we can greatly reduce the complexity of a survey form.

Few research, however, have studied how to analyze data collected in this way. A simple way is to treat attributes as being independent of each other (i.e. focus on main effects) and to estimate attribute choices separately as a function of price and other individual-specific characteristics. This approach treats an alternative as multi-dimensional choices and does not assume that prices can change by bundling attributes. Moreover, this approach does not provide the degree of trade-offs among attributes.

A second way is to randomly select a sample of bundled alternatives and to estimate discrete choice models as a function of attribute level dummies. By including no alternative-specific constant, this approach is identified even when no observation is reported for an alternative. This approach is cost-inefficient, since we do not use observations who chose unselected alternatives. Ideally, we may want to use all observations by including all reported alternatives. However, if we construct a choice set by using only reported alternatives and by ignoring unreported alternatives, we will estimate inconsistent coefficients by ignoring the choice set generation process. Either we should include adjusting factors for importance sampling, or we should explicitly model the choice set generation process.

A third approach is to ask respondents to choose among bundled services after the above survey. For example, suppose that we administer the following question after the above question:

Now suppose that you narrowed candidates into the following three carriers. Please indicate which carrier you would prefer.
<table>
<thead>
<tr>
<th>Feature provided</th>
<th>Carrier A</th>
<th>Carrier B</th>
<th>Carrier C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Freight rate</td>
<td>650 cents per trailer-mile</td>
<td>500 cents</td>
<td>300 cents</td>
</tr>
<tr>
<td>Transit time</td>
<td>20% faster than now</td>
<td>same</td>
<td>20% slower</td>
</tr>
<tr>
<td>Consistency</td>
<td>95% on-time</td>
<td>90%</td>
<td>80% on-time</td>
</tr>
<tr>
<td>Loss or damage</td>
<td>≤0.3% of value lost</td>
<td>≤1%</td>
<td>≤3% value lost</td>
</tr>
<tr>
<td>Equipment usability</td>
<td>99% of times</td>
<td>95%</td>
<td>90% usable</td>
</tr>
<tr>
<td>EDI</td>
<td>Offer</td>
<td>Offer</td>
<td>Do not offer</td>
</tr>
<tr>
<td>Billing accuracy</td>
<td>99% of times</td>
<td>95%</td>
<td>90% of times</td>
</tr>
<tr>
<td>Real-time tracing</td>
<td>Offer</td>
<td>Do not offer</td>
<td>Do not offer</td>
</tr>
<tr>
<td>Level of efforts</td>
<td>easy to work with</td>
<td>normal</td>
<td>difficult</td>
</tr>
<tr>
<td>Responsiveness</td>
<td>always satisfactory</td>
<td>sometimes</td>
<td>seldom</td>
</tr>
</tbody>
</table>

I would prefer the following carrier (or I would allocate shares in the following way):

<table>
<thead>
<tr>
<th>Carrier A</th>
<th>(or %)</th>
<th>Carrier B</th>
<th>(or %)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carrier C</td>
<td>(or %)</td>
<td>Neither of them</td>
<td>(or %)</td>
</tr>
</tbody>
</table>

In the above example, freight rate is calculated by multiplying the discount rate to the sum of base rate (e.g. 500 cent) and prices corresponding to all provided features. As we mentioned earlier, we can construct many alternatives by combining different levels of attributes. Instead, we use the technique of "importance sampling of alternatives" in order to reduce the choice set to a manageable size.\(^{10}\) Using responses to this questionnaire, we can estimate the importance of attributes as well as the complementarity or substitutability of attributes when bundled in an alternative.\(^{11}\) Responses to the first questionnaire provide an indicator for determining how much respondents are willing to pay for desired services.

5. *Strengths and Weaknesses of SP Approach*

The stated preference approach provides opportunities to estimate attributes' importances without observing actual choices.\(^{12}\) Since researchers provide descriptions of choice environments, researchers can include intangible attributes such as perceptions and service quality, avoid multi-collinearity among attributes, provide attributes without measurement errors, or utilize an experimental design for survey questionnaires. Moreover, a choice set can be pre-specified and can include non-existing alternatives. Thus, the stated preference approach is useful for designing a new service, since a research

---


can define choice contexts quantitatively, can include non-existing services, and can experiment by offering different choice contexts to different subjects.

On the other hand, several conditions should be met in order to get reasonable estimates of demand elasticities with a stated preference data. First, responses are highly subject to context effects (e.g. survey phrase) and reference effects (if there is a reference alternative). Since respondents make hypothetical decisions, they may employ simple rules in answering a survey such as choosing always the cheapest mode (simplification effects) or the most familiar mode (prominence effects). Other biases inherent in self-stated responses are social norm effects, policy bias, justification bias, and respondent heterogeneity. Responses may also differ from actual choices due to omitted situational constraints or to lack of external validity.

Second, due to the above problems, researchers should have a reliable survey or laboratory method for eliciting the stated preference of consumers. McFadden and Leonard proposed psychometric robustness, statistical reliability and economical soundness as three criteria for the validity of the stated preference approach. Yet, there seems to be no general agreement about what survey forms provide the most realistic hypothetical choice situations. Recent efforts try to design a multimedia survey rather than a mail survey in order to make questionnaires more realistic. A multimedia survey uses interactive interviews over computer terminals that provide audio-visual descriptions.

Third, sampled respondents should be representative of target markets and should be motivated to answer survey questionnaires. Since most logistics managers are heavily surveyed, researchers have to develop a creative survey form that can facilitate answers or that can prevent respondents from employing a simple rule when answering. Meanwhile, non-response bias is also becoming a big issue, and a follow-up study for a subset sample of non-respondents becomes necessary.

Fourth, there are merits to combine several conjoint experiments in order to take advantage of strengths of different survey forms. Yet, the way of handling the combined estimation is dubious at best. For example, statistical properties of the bridging process that are used in adaptive and hybrid conjoints were not well known. In hierarchical conjoints, scale differences among subexperiments are not explicitly considered. Nor joint estimation of both subexperiments and overall experiment is tried. A research on bundling process has just started. Rigorous econometric analysis that will ensure the same coefficients in the utility function with scale adjustment needs to be performed.

---

15 Section 5.3 includes some discussions on this issue.
Reference

Albrecht, K., At America's Service, *Dow Jones-Irwin, Homewood, 1988*


Peters, Tom (1992). The Union Pacific Railroad: Decimate the Middle Ranks, Liberate the Conductors, and Launch a Counter-attack against the Truckers, Liberation Management, Alfred A. Knopf, Inc.


