

78-2

**DEMAND MODELS FOR U.S.  
DOMESTIC AIR PASSENGER MARKETS**

Steven E. Eriksen

**Massachusetts Institute of Technology  
Flight Transportation Laboratory**

**Report FTL-R78-2  
June 1978**

Singson  
**MIT**

**FTL COPY, DON'T REMOVE  
33-412, MIT .... 02139**

**DEPARTMENT  
OF  
AERONAUTICS  
&  
ASTRONAUTICS**

**FLIGHT TRANSPORTATION  
LABORATORY  
Cambridge, Mass. 02139**

Massachusetts Institute of Technology  
Flight Transportation Laboratory

Report FTL-R78-2

DEMAND MODELS FOR U.S. DOMESTIC AIR PASSENGER MARKETS

Steven E. Eriksen

June 1978

## ABSTRACT

The airline industry in recent years has suffered from the adverse effects of top level planning decisions based upon inaccurate demand forecasts. The air carriers have recognized the immediate need to develop their forecasting abilities and have applied considerable talent to this area. However, their forecasting methodologies still are far below the level of sophistication of their other planning tools. The purpose of this thesis is to develop a set of demand models which are sufficiently sensitive to measure the effects upon demand of policy decisions with respect to such variables as fare and technological and quality of service factors.

A brief overview of transportation demand theory and a survey of recently published research in air passenger demand modeling are presented. Following these is a discussion of the economic nature of domestic air transportation passenger service indicating the demand and service attributes and how they interact in equilibrium. Based upon this background information a multi-equation econometric model is developed. The model is calibrated over subsets of a base of historical data from 180 markets over a six year time frame. The subsets are cross classifications of markets with respect to length of haul and market size. Recently developed techniques in model sensitivity analysis are applied to ensure statistical robustness, and principal components regression is employed to combat the problem of multicollinearity. Numerical examples of applications of the model are provided.

The results indicate that the model performs very well in the

analysis of long and medium haul markets. It is particularly effective in the higher density markets. The model is not equipped to account for the impacts upon air transportation passenger demand of competing modes, and therefore does not perform well in the analysis of short haul (less than 400 miles) markets.

Thesis Supervisor: Nawal K. Taneja

Title: Associate Professor of Aeronautics and Astronautics

## ACKNOWLEDGMENTS

My sincere appreciation is extended to Mr. Louis Williams and Mr. Mark Waters, both of the NASA Ames Research Center, who continuously have shown enthusiasm for this research since its inception three years ago. This study has been sponsored under NASA Grant No. NSG-2129.

I am very grateful for the technical support provided by my committee, whose diverse professional backgrounds and interests reflect the interdisciplinary nature of this thesis. Professor Simpson's profound knowledge of operations research techniques as applied to airline problems coupled with his common sense approach to teaching originally attracted me to the study of this fascinating industry. Professor Welsch's influence upon my work has convinced me that data analysis is the least understood, most often neglected, and most interesting aspect of statistical/economic modeling. Professor Pindyck's expertise in economic model building has had a very positive impact upon the theoretical foundation of my work. Special thanks go to Dr. James Kneafsey for his many comments and suggestions and, most importantly, for his encouragement.

The greatest motivating force behind the technical aspects of this thesis has been provided by Professor Taneja, who first introduced me to the problem of modeling the demand side of the airline industry. He has continually been quick to compliment and quick to criticize my work, and has been a great friend throughout the course of this research.

Many typists have had a hand in preparing the drafts of the various sections of this thesis. My gratitude is extended to all of them. I wish to particularly praise the work of Abby McLaughlin, whose editorial assistance with the final draft was at least as valuable as her outstanding typing skills, and Martha Nikas and Michelle Lang, who manage the front office of the Flight Transportation Laboratory and are major contributors to making this lab such a pleasant place to work.

The people most deserving of my appreciation are my parents and my wife, Peggy, and son, Kris. I hope their sacrifices will be rewarded.

## List of Figures

<u>Figure Number</u>		<u>Page</u>
2.1	Gronau's Model Choice Representation for a Given City Pair	34
2.2	Representation of Development of Models Surveyed in Chapter II	56
3.1	The New York-Miami Market	61
3.2	Typical Itinerary of a Transcontinental Flight	62
3.3	Conceivable Shape of Typical Demand Vs. Fare Curve	89
3.4	Typical Demand Vs. Frequency Curve	91
3.5	ISOMRKT Supply INLP Algorithm	101
3.6	Output of ISOMRKT Program	102
4.1	Flow Diagram of Interaction of the Variables	123
4.2	Empirical Time of Day Demand Distribution for Eastern Airlines' Boston/New York Air Shuttle	136
4.3	Theoretical Time of Day Demand Distribution for Chicago to Los Angeles	141
4.4	Theoretical Time of Day Demand Distribution for Los Angeles to Chicago	142
4.5	Theoretical Time of Day Demand Distribution for Boston to San Francisco	143
4.6	Theoretical Time of Day Demand Distribution for San Francisco to Boston	144
4.7a	Flight Schedule for Boston to Washington	146
4.7b	Level of Service Computations for Boston to Washington	148
4.8a	Flight Schedule for Chicago to Philadelphia	151
4.8b	Level of Service Computations for Chicago to Philadelphia	153

## List of Figures (continued)

<u>Figure Number</u>		<u>Page</u>
4.9a	Flight Schedule for San Francisco to Omaha	155
4.9b	Level of Service Computations for San Francisco to Omaha	156
4.10a	Flight Schedule for San Antonio to Tucson	158
4.10b	Level of Service Computations for San Antonio to Texas	160
4.11	Estimates of Market Shares for Boston to Washington	167
4.12	Estimates of Market Shares for Chicago to Philadelphia	167
4.13	Estimates of Market Share for San Francisco to Omaha	169
4.14	Estimates of Market Shares for San Antonio to Tucson	169
4.15	List of Regions	175
5.1	Estimates of Demand Equation Parameters for All, Long Haul, Medium Haul, and Short Haul Markets	192
5.2	Chow Test for Pooling Markets by Length of Haul	194
5.3	Estimates of Demand Equation Parameters for Long, Long/Large, Long/Medium, and Long/Small Markets	197
5.4	Chow Test for Pooling Long Haul Markets by Level of Socio-Economic Activity	199
5.5	Preliminary Estimates of Service and Demand Equation Parameters for Large Long Haul Markets Before Principal Component Deletions	201
5.6	Plot of Residuals of Preliminary Estimation of the Demand Equation for Large Long Haul Markets	203
5.7	Preliminary Estimates of the Demand Equation Parameters for Large Long Haul Markets After Principal Component Deletions	207



## List of Figures (continued)

<u>Figure Number</u>		<u>Page</u>
5.8	Phase I Estimates of the Parameters of the Service and Demand Equations for Large Long Markets before Principal Component Deletions	214
5.9	Plot of Residuals of Phase I Estimation of the Demand Equation for Large Long Markets	216
5.10	Phase I Estimates of the Demand Equation Parameters for Large Long Haul Markets After Principal Component Deletions	217
5.11	Phase II Estimates of Service and Demand Equation Parameters for Large Long Haul Markets	221
5.12	Plot of Residuals of Phase II Estimation of the Demand Equation for Large Long Haul Markets	223
5.13	Phase II Estimates of the Demand Equation Parameters for Large Long Haul Markets After Principal Component Deletions	224
5.14	Chow Test for Pooling Medium Haul Markets by Level of Socio-Economic Activity	231
5.15	Chow Test for Pooling Short Haul Markets by Level of Socio-Economic Activity	237
5.16	Estimated Demand Equations for Short Haul Markets	239
6.1	Demand Vs. Frequency Curves, Boston-San Francisco and Chicago-New York	248
B.1	Scatter Plot of Hypothetical Data	275
B.2	Regression Function for Hypothetical Data	276

## List of Tables

<u>Table Number</u>		<u>Page</u>
3.1	Characteristics of Fleet	99
3.2	Eight Best Schedule Options	108
4.1	Four Levels of Equivalent Air Service	128
4.2	Empirical Time of Day Demand Distribution for Eastern Airlines Boston/New York Air Shuttle	137
5.1	Principal Component Analysis, Long Large Markets	206
5.2	Diagonal Elements of the Hat Matrix, Demand Equation for Large Long Markets, Phase I of Sensitivity Analysis	210
5.3	Studentized Residuals, Demand Equation for Large Long Markets, Phase II of Sensitivity Analysis	211
5.4	Diagonal Elements of the Hat Matrix, Demand Equation for Large Long Haul Markets, Phase II of Sensitivity Analysis	218
5.5	Studentized Residuals, Demand Equation for Large Long Haul Markets, Phase II of Sensitivity Analysis	219
6.1	Demand vs. Frequency Relationship for Boston to San Francisco	245
6.2	Demand vs. Frequency Relationship for Chicago to New York	246
6.3	Effect upon Demand in Chicago-New York Market of Fuel Efficient Aircraft, Assuming a Resulting 5%-30% Decrease in Fare (Constant Dollars)	252
B.1	Hypothetical Data	273
B.2	Diagnostic Statistics for Hypothetical Data	274
C.1	Perfectly Collinear Data	282
C.2	Singular Values and Condition Indices for the Collinear Data	286

## Table of Contents

	<u>Page</u>
Abstract	2
Acknowledgments	4
List of Figures	6
List of Tables	9
I. Introduction	15
1.1 The Role of Demand Forecasting in the Air Transportation Industry	16
1.1.1 Demand Forecasting and the Airlines	17
1.1.2 Demand Forecasting and the Airport Authorities	18
1.1.3 Demand Forecasting and the Airport Manufacturers	19
1.2 The Need for Policy Sensitive Models	20
1.3 Purpose and Outline of This Thesis	22
II. Overview of Transportation Demand Theory and Domestic Air Transportation Economic Models	25
2.1 Total Disaggregation	26
2.2 Aggregation by Destinations	31
2.3 Aggregation by Incomes	38
2.4 Aggregation by Modes and Destinations	43
2.5 Aggregation by Incomes, Destinations, and Modes	46
2.6 Total Systemwide Aggregation	49
2.7 Summary	55

Table of Contents (continued)	<u>Page</u>
III. The Economic Nature of Domestic Air Transportation	
Passenger Service	58
3.1 Fundamental Economic Aspects of Air Passenger Service	59
3.1.1 The Air Transportation Market	59
3.1.2 The Airline Product -- Not a Good But a Service	63
3.1.3 Attributes of the Quality of Air Passenger Service	67
3.1.3.1 Trip Time	67
3.1.3.2 Comfort and Onboard Amenities	69
3.1.3.3 Safety	70
3.1.3.4 Reliability	71
3.1.3.5 Convenience	73
3.1.3.6 Summary of the Service Attributes and Their Impact Upon Demand	74
3.1.4 The Price of Air Transportation Services	75
3.1.5 Some Aspects of Marketing, Advertising and Competition	79
3.1.6 Summary	82
3.2 A Hypothetical Isolated Market	82
3.2.1 Analysis of Demand	83
3.2.2 Analysis of Supply	90
3.2.3 The Analysis of Equilibrium -- Some Numerical Examples	96
3.2.3.1 Equilibrium Conditions at Optimal Level of Output	100
3.2.3.2 Effect of an Increase in Total Demand	104
3.2.3.3 Effect of an Increase in Fare	105

Table of Contents (continued)		<u>Page</u>
	3.2.3.4 The Decision of Whether or Not to Offer Wide Bodied Service	106
3.3	The Market as a Subset of an Integrated Network	107
	3.3.1 Analysis of Demand	109
	3.3.2 Analysis of Supply and Equilibrium	111
IV.	Description of the Functional Forms, Market Variables, and Experimental Design	118
4.1	Functional Form of the Demand and Service Functions	119
4.2	Description of the Market Variables	122
	4.2.1 Demand ( $Q_D$ )	122
	4.2.2 Level of Service (LOS)	124
	4.2.2.1 Behavioral Assumptions	126
	4.2.2.2 Development of the Index	129
	4.2.2.3 Determination of Nonstop Jet Time	132
	4.2.2.4 Determination of Time of Day Demand Function	134
	4.2.2.5 Examples of Level of Service Calculators	140
	4.2.2.6 Combination of Directional Levels of Service	159
	4.2.3 Fare (F)	162
	4.2.4 Socio-Economic Activity (SE)	163

Table of Contents (continued)		<u>Page</u>
4.2.5	Competition (COMP)	165
4.2.5.1	Development of the Competition Variable	165
4.2.5.2	Examples of the Computation of the COMP Variable	166
4.2.6	Location within Route Structure (RS)	168
4.3	Sample Design and Market Selection	171
4.3.1	The Concept of Region Pairs	171
4.3.2	Sampling Design and Procedure	174
V.	Empirical Analysis of U.S. Domestic Air Passenger Markets	177
5.1	Pooling of Entire Data Set	190
5.2	Analysis of Long Haul Markets	195
5.2.1	Large Long Haul Markets	200
5.2.2.1	Preliminary Analysis	200
5.2.2.2	Sensitivity Analysis: Phase I	208
5.2.2.3	Sensitivity Analysis: Phase II	215
5.2.2	Medium Size Long Haul Markets	225
5.2.3	Small Long Haul Markets	229
5.3	Analysis of Medium Haul Markets	230
5.3.1	Large Medium Haul Markets	232
5.3.2	Medium Size Medium Haul Markets	233
5.3.3	Small Medium Haul Markets	235
5.4	Analysis of Short Haul Markets	236

Table of Contents (continued)	<u>Page</u>
VI. Applications of the Models	240
6.1 Derivation of Demand Vs. Frequency Relationships	241
6.2 The Impact Upon Demand of New Technologically Advanced Aircraft	247
6.2.1 The Introduction of a Supersonic Transport on Long Haul Domestic Routes	249
6.2.2 The Introduction of a Fuel Efficient Subsonic Aircraft on Medium Haul Routes	250
6.3 Aggregate Forecasting	253
VII. Conclusions and Recommendations for Future Research	255
Appendix A. List of Region Pairs by Demographic Stratifications	260
Appendix B. The Hat Matrix in Linear Models	268
Appendix C. Multicollinearity and Related Topics	278
C.1 Singular Value Decomposition	281
C.2 Principal Components	285
Bibliography	291
Biographical Sketch	296

## I. Introduction

The process of forecasting the demand for air transportation services has in recent years become an extremely complex operation. Since the late sixties, when the United States domestic air carriers suffered through a grave financial crisis, industry analysts have recognized the sensitivity of the fiscal strength of the airlines and aircraft manufacturers to their planning process, which is based upon travel demand forecasts. Furthermore, the analysts have realized that traditional forecasting methods such as trend extrapolation are inadequate due to the impact upon travel demand of recent changes within the economic and operating environment such as high inflation rates, escalating fuel and labor costs, and uncertainties with regard to future technology and regulatory conditions.

The past performance record of air traffic demand forecasters has been unimpressive. Forecasting models based upon methodologies with a low level of sophistication have not, as mentioned above, captured the impacts of important demand and/or supply determinants. Models based upon more sophisticated methodologies, such as advanced econometric techniques, have generally been limited by either insufficient understanding of the total air transportation system or by lack of relevant data.

In his introductory remarks in a talk entitled, "Air Transportation -- Directions for Future Research" presented at a workshop on air transportation demand and systems analysis in June 1975, George Avram of Pratt and Whitney Aircraft cited the findings of a survey recently conducted by Brushkin Associates. The findings indicated that people gave a very



favorable rating to the forecasting accuracy of sportswriters, sports announcers, and weathermen. At the opposite end of the spectrum, those forecasters who received the least favorable ratings included stock brokers, astrologers, and economists. Mr. Avram's reaction to this was that "since economics is the cornerstone of our understanding of the present and the future...we (sic) got a lot of work to do!"<sup>1</sup>

Although this dubious honor was directed at economists in general, Mr. Avram was singling out forecasters of the levels of future activity in the air transportation industry. His remarks are indicative of a general dissatisfaction of airline companies, aircraft manufacturers, airport authorities, and regulatory agencies with the accuracy of the forecasters' predictions.

### 1.1 The Role of Demand Forecasting in the Air Transportation Industry

Two fundamental questions arise when considering the role of demand forecasting in the air transportation industry:

- (1) Is it necessary to forecast air traffic demand, and
- (2) What is the importance and role of the forecasting process?

This section will address these questions from the perspective of three different components of the commercial aviation system: the airlines, the airport authorities, and the equipment manufacturers.

---

<sup>1</sup> "Proceedings of the Workshop: Air Transportation Demand and Systems Analysis" (Cambridge: M.I.T. Flight Transportation Laboratory Report R75-8, August, 1975), pp. 452-456.

### 1.1.1 Demand Forecasting and the Airlines

The basic function of demand forecasts to the management of the airline companies is to provide input to the planning processes. The adverse effects of planning decisions made by airline companies based on inaccurate forecasts can be clearly illustrated by examining the plight of the industry during the late 1960's and the early part of the 1970's. Throughout the early and middle sixties, aggregate domestic traffic had been growing exponentially, at a fairly constant rate of between ten and fifteen percent per year. As the sixties progressed, it became necessary for the airlines to continually increase frequency of service to satisfy the growing demand and to maintain reasonable load factors. This expansion of service began to cause serious delays at major airports, resulting in considerable inconvenience to the passengers and expense to the carriers.

Assuming that the levels of passenger demand would continue to grow at the 1960's rate, the airlines decided to introduce wide-body aircraft with two to three times the seating capacity of the narrow-body aircraft in use at that time. The carriers could then greatly reduce frequency while maintaining reasonable load factors, thereby satisfying consumer demand, relieving congestion, and reducing per-passenger cost.<sup>2</sup>

These planning decisions would have provided fruitful results, had the demand forecasts been accurate. What in fact did occur was a sharp decline in passenger growth in the late sixties and beyond, and the airlines were producing a lower level of service (to be formally defined

---

<sup>2</sup> A more detailed discussion of supply factors can be found in Chapter III.

in Chapter IV), load factors declined, and costs soared. This alarming situation caused the carriers to recognize the importance of accurate forecasting and the need for improved forecasting methodologies. This condition is aptly characterized by Harry G. Lehr, Director of Regulatory Affairs of United Airlines:

"We know from hard experience the difficulty airlines have had with [forecasting] and considerable talent has been applied to this problem. Yet, I would characterize the development of our forecasting ability as having reached a level that can only be referred to as Organized Soothsaying. The current economic conditions of the industry and the perishability of our product...dictate a need for a forecasting methodology that is substantially closer to the level of development of our other planning tools." <sup>3</sup>

#### 1.1.2 Demand Forecasting and the Airport Authorities

The need for improved forecasting of air transportation is equally as great for airport authorities as for the carriers. In recent years sophisticated models, based upon operations research and simulation techniques, have been developed to measure airport runway capacity and hence the airport's ability to control delays and delay costs<sup>4</sup> and to measure passenger flow through terminals.<sup>5</sup> However, the ability of

---

<sup>3</sup> "Proceedings of the Workshop", p. 160.

<sup>4</sup> For example, see Gerd Hengsbach and Amedeo R. Odoni, "Time Dependent Estimates of Delays and Delay Costs at Major Airports (Cambridge:M.I.T. Flight Transportation Laboratory Report R75-4, January, 1975).

<sup>5</sup> For example, see Terry P. Blumer, Robert W. Simpson, and John Wiley, "A Computer Simulation of Tampa International Airport's Landside Terminal Shuttles" (Cambridge:M.I.T. FTL Report R-76-5, April, 1976).

these models to evaluate future airport design is only as good as the input data, such as expected number of aircraft movements and passenger enplanements, provided by forecasting models.

George P. Howard, Chief of Aviation Economics of the Port Authority of New York and New Jersey, emphasizes this need for good input data:

"If one considers that Port Authority investment in the three metropolitan airports to date is some one billion dollars and that airlines and other tenants have invested very significant amounts of their own money in these facilities, there is no need to question the desirability, indeed the necessity, of market forecasts which project, as accurately as possible, future trends in the air transportation market in the New York/ New Jersey area." <sup>6</sup>

### 1.1.3 Demand Forecasting and the Equipment Manufacturers

Forecasts of future air traffic is also essential to the manufacturers of aircraft and aircraft engines. The need is particular great in this case not only for forecasts of aggregate traffic (in system revenue passenger miles, for instance), but for forecasts by density and length of haul of the various markets, since design parameters of an individual aircraft type (and its engines) are based upon its desired range and seating capacity. John D. Karraker of General Electric elaborates:

---

<sup>6</sup> George P. Howard and Johannes G. Augustinus, "Market Research and Forecasting for the Airport Market", in Airport Economic Planning, ed. George P. Howard (Cambridge: M.I.T. Press, 1974), p. 109.

"The nature of our product, the jet engine, demands long range planning. It costs in the neighborhood of half a billion dollars and takes from 5 to 10 years to develop a new jet engine from the concept stage to commercial service. As you can well appreciate, an error in determining the potential market for a given engine can have very grave consequences in our business. Obviously, the minimizing or avoidance of such costly errors is a necessity. For this reason, forecasting is recognized in General Electric as one of the major elements of our business." <sup>7</sup>

## 1.2 The Need for Policy Sensitive Forecasting Models

The above arguments all indicate a strong need for models that forecast levels of traffic, whether it be over the entire system or subsets of the system, as in the case of the airlines; or into, out of, and through, particular cities, as in the case of the airport authorities; or over sets of market types, as in the case of the manufacturers. Some of the more recently developed models are quite capable of predicting reasonable estimates of traffic.<sup>8</sup> However, as will be argued in Chapter II, none of the current models is sufficiently policy-sensitive to determine within a given market the impact of such economic alterations as route awards, fare changes, modifications in quality of service, and acquisition of new equipment. So, while traffic can be forecasted based upon exogenous factors assuming that the variables under the control of the carriers and regulators remain basically stable, no conditions have been introduced into existing models to demonstrate how the carriers and regulators can themselves influence traffic levels within the individual markets.

<sup>7</sup> "Proceedings of the Workshop", p. 445.

<sup>8</sup> For example, see "Aviation Forecasts, Fiscal Years 1976-1987" by the Federal Aviation Administration, summarized in Chapter II.

An example of where a policy-sensitive market demand model would be useful is in conjunction with fleet assignment optimization models. Several such models have been developed by equipment manufacturers and by academic institutions. One example is FA-4, a model developed in the Flight Transportation Laboratory at M.I.T.<sup>9</sup> FA-4 is a very sophisticated linear programming model which allocates a mixed fleet of aircraft over a route network so as to maximize the difference between total revenue and the sum of direct and indirect operating costs. The optimization is constrained by a number of economic factors including, among others, prescribed load factor conditions, fleet availability, minimum number of departures in the various markets, and maximum number of departures from the various stations.

While FA-4 is very realistic in its modeling of the attributes of an airline's fleet allocation problem, it is limited by the fact that it requires a set of rather ad hoc piecewise linear demand vs. frequency curves as input. If a demand model existed which adequately estimated the response in passenger traffic as a function of level of service ( as well as other factors such as fare and regional demographics), more confidence could be placed in the output of this otherwise-acceptable fleet assignment model.

---

<sup>9</sup> Swan, William, "The Complete FA-4 Memo" (Cambridge: M.I.T. FTL Technical Memorandum 72-10, August, 1972).

### 1.3 Purpose and Outline of This Thesis

The purpose of this thesis is to develop a set of demand models whose parameters are a function of length of haul and demographic size of the markets. These models will be sufficiently policy sensitive so as to forecast the impacts upon market demand of changes in such factors as quality of service, competition, fares, equipment, and regional demography.

There are basically four reasons why these models will constitute an improvement over existing models. They are as follows:

(1) The models will be multi-equation in structure. This advanced specification was chosen to eliminate the coefficient bias present in single-equation models, caused by simultaneity between supply and demand variables.

(2) The variables will be more complete by definition. Many of the proxies used to calibrate existing models have been too simplistic in definition and do not fully measure the levels of important demand and service-related attributes.

(3) The experimental design will be more complete. This will assure not only a larger sample size, but also a more representative subset of U.S. domestic markets with respect to length of haul and to socio-economic factors.

(4) Advanced techniques will be employed to assure statistical robustness and high precision of coefficient estimates.

The various problems with existing models that were mentioned in this

introductory chapter will be elaborated upon in Chapter II, which reviews some published studies in air transportation demand analysis. Preceding this review will be an overview of general transportation demand theory. The discussion in Chapter II will focus on how the air transportation models have been developed as various aggregations of the general theory and how the models which will be developed later in this thesis fit into the aggregation process.

Based upon the general economic theory of transportation demand, and upon the issues raised in the development of the models surveyed in Chapter II, an economic theory of domestic air transportation service will be presented in Chapter III. This discussion is essential for providing a detailed explanation of errors of omission, for the models surveyed in Chapter II, of important supply and demand related attributes. Furthermore, the contents of Chapter III provide the necessary coverage of issues that will, and will not, be important in the development of the resulting models in this research. The end product of Chapter III is a general specification of the models to be calibrated.

Chapter IV takes the general model developed in Chapter III and more precisely identifies the specification. The complete definitions of the variables are presented in Chapter IV, and numerical examples of the computation of the more complex proxies are provided. The sampling procedure is also detailed in Chapter IV.

Chapter V contains the results of the empirical calibration of the models specified in Chapter IV. Special attention is given to the assurance of proper pooling and to statistical robustness. The end



product of Chapter V is a structural demand analysis equation for each stratification of markets, separate demand equations are produced for forecasting and analytical (policy sensitivity) purposes.

Chapter VI applies the models calibrated in Chapter V to measure the impact upon demand within given markets due to changes in future technologies and fare structures. Chapter VI also outlines how the models may be applied for aggregate demand forecasting. Chapter VII contains the conclusions of this research and recommendations for future research.

## II. Overview of Transportation Demand Theory and Domestic Air Transportation Economic Models

The introductory chapter described the motivation for economic models that forecast the impact upon air passenger demand within domestic markets of policy changes within the industry. Furthermore, it was mentioned therein that the performance of existing models, particularly in the sense of policy analysis, has generally been quite disappointing. In this chapter, an overview of the economic theory of travel demand, particularly with respect to the air mode, will be presented. The development of five existing models based upon this theory will be studied, and their strengths and weaknesses with respect to their reliability for forecasting and analysis will be identified.

The main variant in the structure of transportation demand models is their level of aggregation. A totally disaggregate model would specify the consumption optimization problem for each consumer (or group of equivalent consumers) in the population. The decision of how many air trips from each consumer's origin to a specific destination in a given time period would be a function of not only the characteristics and prices of these trips, but also of the characteristics and prices of all other transportation services available to this consumer. These other services include trips to all other destinations and to the same destination by alternative modes. Summing over all consumers would yield estimates of total demand in all markets and by all modes.

Since the volume of data required to calibrate such a model renders

total disaggregation intractable, researchers are forced to combine some or all factors. Some models of air transportation demand do not consider the effect of competing modes and/or destinations. Other models do not recognize distinct groups of consumers, such as by income levels, and instead focus on aggregate regional or national economic variables. Another level of aggregation is by markets whereby a model would not measure demand from point to point, but over a large mass of origin-destination pairs such as the Northeast Corridor. Total aggregation combines all of these factors and generally forecasts a single variable such as revenue passenger miles for the entire domestic system.

## 2.1 Total Disaggregation

Lancaster, in his recent theory of consumer demand,<sup>1</sup> asserts that individual consumers select goods and services based upon the "characteristics" of these products, where the characteristics are "those objective properties of things that are relevant to choice by people."<sup>2</sup> For example, the scheduled flight time of an aircraft departure is certainly a characteristic of this product, whereas the color of the aircraft is a property that is not likely to be a characteristic.

The primary difference between Lancaster's demand theory and traditional consumer demand theory is that Lancaster's theory states that

---

<sup>1</sup> Kelvin Lancaster, Consumer Demand: A New Approach (New York: Columbia University Press, 1971).

<sup>2</sup> Ibid., p. 6.

a consumer will maximize a utility function which is dependent upon characteristics, rather than upon the goods themselves. Assuming that the characteristics of various products are indeed objective and quantifiable and that their levels are linearly related to the levels of the product, the following relationship holds:

$$z = Bx \quad (2.1)$$

where

$z$  = an  $m$  by  $1$  vector of characteristics

$x$  = an  $n$  by  $1$  vector of products

and  $B$  = an  $m$  by  $n$  matrix of coefficients relating the products to the characteristics called the "consumption technology matrix".

An important point is that the consumption technology matrix,  $B$ , is invariant among consumers. A given product is viewed by each consumer as bearing the same levels of characteristics; however, consumers will react differently to these characteristics. Two reasonable people would assumedly not argue over size, ride quality, handling, performance, etc., of different types of automobiles. However, they may purchase different autos because of their relative desires and preferences for these various characteristics.

Given that each consumer is constrained in consumption by his budget, the following mathematical program is formulated:

$$\begin{aligned} \text{Max } U &= U(z) \\ \text{subject to } z &= Bx \end{aligned} \quad (2.2)$$

$$p^T x \leq k$$

$$x \geq 0$$

where

$p$  = an  $n$  by  $1$  vector of prices of products

and  $k$  = the consumer's income.

Since the consumption technology matrix,  $B$ , and the price vector,  $p$ , are assumed identical for all consumers, and the utility function,  $U(z)$ , and the budget,  $k$ , will vary among consumers, levels of demand for various products can change due to two distinct classes of effects. The "efficiency effect" is the result of a change in product technology,  $b_j$  (the column of  $B$  referring to project  $j$ ), or price,  $p_j$ , and is perceived by all consumers. The "demand effect" is the result of changes in an individual's needs and preferences for various characteristics, manifested in  $U(z)$ , or changes in an individual's income,  $k$ . The demand effect is obviously an individual consumer phenomenon.

Quandt points out that "...travel is viewed as the result of individuals' rational decision making in an economic context."<sup>3</sup> Since this point is very difficult to argue, it appears as though the analysis of travel demand by the consideration of the maximization of individuals' utility functions is a sensible approach. However, if one were to exploit the full power of the above theory, the researcher is forced to develop a set of demand models that are stratified by many factors. The parameters of the model would vary by income group, as it has been shown that the

---

<sup>3</sup> Richard E. Quandt, The Demand for Travel: Theory and Measurement (Lexington, Mass.: Heath Lexington Books, 1970), p. 1.

amount of travel and the modal split are related to personal income.<sup>4</sup> Furthermore, the utility functions will vary by demographic groups. For example, an individual's occupation may affect his utility for travel related characteristics, distinct from the income effect.<sup>5</sup>

Lancaster develops the notion of "natural" or "intrinsic" groups of products, based upon the structure of the consumption technology matrix.<sup>6</sup> An intrinsic group is a subset of the available products that relate to a subset of characteristics, such that no products outside of this subset possess positive levels of any of the related characteristics. Furthermore, no product in the subset possesses positive levels of any characteristics not in the related characteristic subset. If indeed an intrinsic group does exist, the consumption technology matrix can be written as follows:

$$B = \begin{bmatrix} B_1 & 0 \\ 0 & B_2 \end{bmatrix} \quad (2.3)$$

where  $B_1$  is the subtechnology of the intrinsic group and  $B_2$  is the subtechnology of the remainder of the products in the market.

All of the transportation demand models that were surveyed prior to this research assume that transportation services comprise an intrinsic

---

<sup>4</sup> Philip K. Verleger, Jr., "A Point-to-Point Model of the Demand for Air Transportation" (Ph.D. dissertation, M.I.T., 1971), p. 79.

<sup>5</sup> Quandt, op. cit., p. 12.

<sup>6</sup> Lancaster, op. cit., pp. 126-128.

group. Since efficiency substitution effects are the attainment of a new characteristics vector due to a change in the mix of products purchased caused by a shift in the consumption technology matrix or price vector, this assumption implies that efficiency substitutions cannot occur between transportation services and other products. Consequently, prices and characteristics of non-transportation products are generally not included in transportation demand models.

The individual's optimization problem for transportation services can now be formulated as a subprogram of (2.2) as follows:

$$\begin{aligned}
 \text{Max } U_t &= U_t(z_t) \\
 \text{subject to } z_t &= B_t x_t \\
 p_t^T x_t &\leq k_t \\
 x_t &\geq 0
 \end{aligned}
 \tag{2.4}$$

where

$U_t$  = the utility function of characteristics related to transportation service characteristics

$z_t$  = the set of transportation service characteristics

$B_t$  = the transportation technology matrix

$x_t$  = the set of available transportation services

$p_t$  = the price vector of transportation services

$k_t$  = the amount of the individual's income that can be afforded for transportation services

A totally disaggregated demand model must then consider the following elements:

- (1) sensitivity to different consumer groups, particularly with respect to income levels;
- (2) sensitivity to price and characteristics of competing modes;
- (3) sensitivity to price and characteristics of competing destinations.

Since the inclusion of all of the above elements is obviously intractable, researchers are forced to aggregate some or all of them. The remainder of this chapter will provide a survey of models indicating how these aggregations have been performed.

## 2.2 Aggregation by Destinations

Reuben Gronau, in his Ph.D. dissertation at Columbia University which was later published as a book,<sup>7</sup> developed a model sensitive to income groups and modal split, but did not explicitly account for consumer choice of alternate destinations. Because of data restrictions, this model was calibrated for trips in and out of New York City over the 38 most heavily traveled New York markets. Since this is a cluster sample, there would likely be considerable bias if one were to attempt to apply these results to the entire U.S. domestic system.

---

<sup>7</sup> Reuben Gronau, The Value of Time in Passenger Transportation: The Demand for Air Travel, National Bureau of Economic Research Occasional Paper 109 (New York: Columbia University Press, 1970).



Gronau altered Lancaster's theoretical model (2.4) by defining the utility function over an "activity" space, rather than over characteristic space. The activities,  $z_i$ , are packaged combinations of market products,  $x_i$ , and time,  $T_i$ .

$$z_i = f_i(x_i, T_i) \quad (2.5)$$

One type of activity, he states as an example, is a "visit" which is a combination of transportation, hotel and restaurant services, travel time, and time at the destination.<sup>8</sup>

A second modification to Lancaster's theory is that Gronau considers time as a constraint analogous to the income constraint in (2.4). One of the objectives of this research was to evaluate a monetary value of time for various income groups, and, as will be shown, the inclusion of the set of time constraints facilitates this in theory.

The consumer's optimization problem for travel activities is then as follows:

$$\begin{aligned} \text{Max } U &= U(z_1, z_2, \dots, z_n) \\ \text{subject to } \sum_{i=1}^n P_i x_i &= Y \\ \sum_{i=1}^n T_i &= T_0 \end{aligned} \quad (2.6)$$

where  $p_i$  = price of market product  $x_i$   
 $Y$  = the consumer's monetary travel budget  
 $T_i$  = time investment for product  $x_i$

---

<sup>8</sup> Ibid., p. 7.

$T_0$  = the consumer's travel time limit<sup>9</sup>

The Lagrangian of (2.6) can be written as follows:

$$L = U(z_1, z_2, \dots, z_n) + \lambda(Y - \sum P_i x_i) + \mu(T_0 - \sum T_i) \quad (2.7)$$

The first order conditions for optimization of (2.7) are:

$$\frac{\partial L}{\partial z_i} = \frac{\partial U}{\partial z_i} - \lambda P_i \frac{\partial x_i}{\partial z_i} - \mu \frac{\partial T_i}{\partial z_i} \quad (2.8)$$

Gronau then defines  $K = \mu/\lambda$  as the shadow price of time.

Assuming that there are four modes of travel, A, B, C, and D, between two cities, that the only two attributes of modal choice for a consumer deciding to travel between the cities are price and travel time, and that none of these modes is dominated by another in terms of these attributes, the consumer will choose the appropriate mode based upon his value of  $K$ . This is depicted in Figure 2.1, where points A, B, C, and D are the location of the modal attributes in price/travel time space. The line ABCD is the convex envelope of these four points. A generic consumer will select that mode represented by the extreme point of the convex envelope

---

<sup>9</sup> Since, as previously noted, transportation services are generally considered as constituting an "intrinsic group", all products, characteristics, and activities will hereafter be assumed as being related only to transportation services. For simplicity the "t" subscripts included in (2.4) will be suppressed.

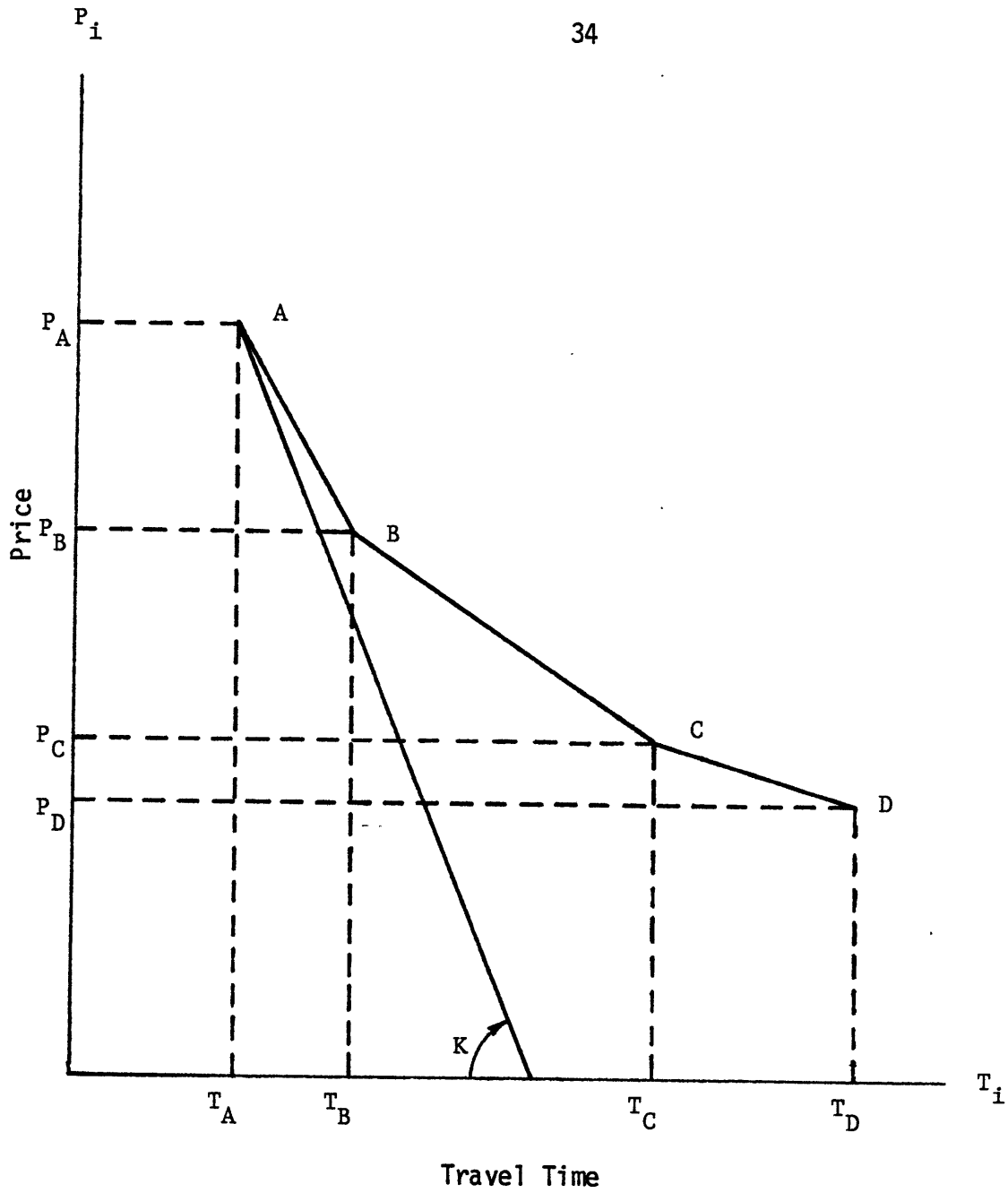


Figure 2.1 Gronau's Modal Choice Representation for a Given City Pair<sup>10</sup>

<sup>10</sup> Ibid., p. 17.

which is tangential to a line with slope  $-K$ . In the example shown in Figure 2.1, this would be mode A.

More formally, if one assumes that the traveling process has no utility in itself, and hence the various modes may be considered as different combinations of the price and time attributes, then the "marginal rate of substitution",  $K^*_{ij}$ , between modes  $i$  and  $j$  can be defined as the absolute value of the slope of the line segment between  $i$  and  $j$ .

$$K^*_{ij} = \frac{P_i - P_j}{T_j - T_i} \quad (2.9)$$

Mode  $i$  will be preferred to mode  $j$  if  $K^*_{ij} > K$ .

Since prices and travel times between cities for any given mode are generally linear functions of intercity distance, Gronau concludes that the general modal choice decision is a function of  $K$  and of distance. For example, based upon numerical estimates of trip times and fares, he concludes that for a 150 mile trip, "a passenger prefers to use air rather than rail transportation....only if his price of time exceeds \$11.80 per hour, he prefers air to bus if his price of time exceeds \$7.10 per hour, and he prefers rail to bus if his price of time exceeds \$5.30 per hour. Bus transportation is, therefore, used for 150 mile trips only for individuals whose price of time is less than \$5.30 per hour."<sup>11</sup>

The specifications of Gronau's demand model are as follows:

---

<sup>11</sup> Ibid., p. 36.

$$x_{ij} = \pi_{ij}^{\beta_{1j}} Y_i^{\beta_{2j}} e^{u_{ij}} \quad (2.10)$$

where

$x_{ij}$  = number of trips to destination  $j$  per family of income group  $i$

$\pi_{ij}$  = generalized trip cost =  $P_j + K_i T_j$

$Y_i$  = average income for income group  $i$

$u_{ij}$  = disturbance term

$B_j$ ,  $\beta_{1j}$ , and  $\beta_{2j}$  are regression coefficients.

Assuming that the price of time is proportional to hourly wages,  $K_i = kW_i$ , and taking logs (2.10) becomes:

$$\log x_{ij} = \log B_j + \beta_{1j} \log (P_j + kW_i T_j) + \beta_{2j} \log Y_i + u_{ij} \quad (2.11)$$

Finally, the assumption is made that  $\beta_{1j}$  and  $\beta_{2j}$  are independent of the destinations and adding an "attractiveness" factor for each destination  $j$ ,  $G_j$ , the model (2.11) becomes

$$\log x_{ij} = \beta_0 + \beta_1 \log (P_j + kW_i T_j) + \beta_2 \log Y_i + \beta_e \log G_j + u_{ij} \quad (2.12)$$

The income and traffic data used to calibrate equation (2.12) was extracted from onboard passenger surveys conducted by the New York Port Authority. These surveys provide demographic and travel frequency data on a sample of local New York passengers (connecting passengers were excluded) conducted from April 1963 through May 1964. Travel time was taken to be the fastest scheduled flight on each route plus average driving time from the center of New York to the airports. The trip price

was standard coach fare plus limousine fare from downtown to the airport.

The level of attractiveness variable,  $G$ , is a function of the population of the destination city's SMSA and the number of phone calls made between the two cities. The latter factor presumably was selected as a result of previous research by Brown and Watkins.<sup>12</sup> While the number of phone calls may be an adequate attractiveness variable for the purpose of Gronau's research, to measure monetary value of time, it is doubtful that this variable would be useful in a forecasting model. The number of phone calls are probably no easier to forecast than the response variable, air passenger trips.

The estimation procedure that Gronau used was a nonlinear search technique in which the variable  $k$  was varied from zero to two -- by increments of 0.25. While the coefficients of all the variables had the correct sign and were highly significant, and the  $R^2$  values were reasonable (all generally between 0.8 and 0.9), the effect upon the fit of varying  $k$  was nearly negligible, usually only changing the third decimal place of  $R^2$ .

A major flaw in Gronau's specification is the definition of travel time (fastest scheduled flight plus driving time from center city to airport). One limiting assumption is that a person desiring to fly will be able to board the "fastest" flight at his convenience; this completely

---

<sup>12</sup> Samuel L. Brown and Wayne S. Watkins, "The Demand for Air Travel: A Regression Study of Time Series and Cross Sectional Data in the U.S. Domestic Market," in Airport Economic Planning, ed. George P. Howard (Cambridge: M.I.T. Press, 1974), p. 99.

ignores the existence of schedule delays. Secondly, the inclusion of driving time from the center city to the airport is appropriate only if all travelers originate in the center city. Finally, the travel time is likely to be a function of route density. This creates a simultaneity problem which requires a multiple equation system to assure unbiasedness in the estimates.

### 2.3 Aggregation by Incomes

Terry Blumer created a city pair air transportation demand model for short haul (less than 400 miles) markets in his master's thesis from M.I.T.<sup>13</sup> His "base model", a common gravity model which aggregates by income groups and does not consider the effect of competing modes and/or destinations, is shown below:

$$T_{ij} = b_0 \frac{M_i M_j}{D_{ij}^{b_1}} I_{ija}^{b_2} \quad (2.13)$$

where

- $T_{ij}$  = air traffic between cities  $i$  and  $j$
- $M_i$  = effective buying income of city  $i$
- $I_{ija}$  = disutility of air travel, "air impedance", between  $i$  and  $j$
- $D_{ij}$  = distance from city  $i$  to city  $j$

The term "gravity model" stems from the comparability between quantity of travel between two cities and gravitational attraction between two

---

<sup>13</sup> Terry P. Blumer, "A Short Haul Passenger Demand Model for Air Transportation" (S.M. Thesis, M.I.T., 1976).

physical bodies. Both are positively related to the mass of units and negatively related to the distance between them.

A major improvement of this model over conventional gravity models is the explicit inclusion of a level of service variable,  $I_{ija}$ . This variable is a prototype of LOS, the level of service variable developed by Eriksen for the research of this dissertation, and explained in detail in Section 4.2. The impedance variable is a generalized trip time which considers not only the number of daily available flights, but also time of day of departure, number of intermediate stops and/or connections, and speed of aircraft.

Blumer then expanded his base model by developing a "mode sensitive" model consistent with the concept developed in Section 2.1 that air travel will be sensitive to the characteristics of competing modes. He defines  $\bar{T}_{ij}$ , the "total transportation impedance", in a given city pair, as follows:

$$\frac{1}{\bar{T}_{ij}^2} = \frac{1}{I_{ija}^2} + \frac{1}{I_{iju}^2} + \frac{1}{I_{ijr}^2} \quad (2.14)$$

where  $I_{iju}$  and  $I_{ijr}$  are impedances for auto and rail, respectively,

The mode sensitive model is then defined as:

$$T_{ij} = b_0 \left( \frac{M_i M_j}{D_{ij}} \right)^{b_1} \bar{T}_{ij}^{b_2} f_{ij}^{b_3} \quad (2.15)$$

where  $f_{ij}$  = share of travelers using the air mode.



The  $f_{ij}$  variable was described several different ways each of which was some function of the relative impedances of air and surface modes.

The next step was the development of a "destination sensitive" model based upon the assertion that "...[the presence of alternate destinations in close proximity to market  $i$  to  $j$ ] has the effect of decreasing the number of trips made to  $j$  and may possibly increase the total number of trips generated by city  $i$  due to [the alternate destination's] own unique attractions."<sup>14</sup> This again is consistent with the theory developed in Section 2.1, which implies that air travel will be sensitive to characteristics of trips to competing destinations.

The destination sensitive model is defined as follows:

$$T_{ij} = b_0 H_{ij} (M_i, M_j, S_i, S_j)^{b_1} I_{ija}^{b_2} F_{ij}^{b_3} \quad (2.16)$$

where

$H_{ij}$  = air travelers generated by the cities  $i$  and  $j$  (to all destinations)

$S_i$  = total attraction of alternate destinations for travelers from city  $i$

and  $F_{ij}$  = fraction of total travelers generated by regions  $i$  and  $j$  that move from  $i$  to  $j$  or from  $j$  to  $i$ .

Finally, Blumer combined the models shown in (2.13), (2.15), and (2.16) to obtain a set of six "mode and destination sensitive" models.

---

<sup>14</sup> Ibid., p. 47.

After calibrating these using various data sets he concluded, on the basis of highest  $R^2$ , that the best model was as follows:

$$T_{ij} = b_0(M_i^{\theta_1} + M_j^{\theta_2})^{b_1} I_{ija}^{b_2} D_{ij}^{b_3} F_{ij}^{b_4} \quad (2.17)$$

where  $\theta_1$  and  $\theta_2$  are constants determined by nonlinear search.

The data used in the calibration of the "best" model, (2.17), was a pooling of two years of cross sectional data. All coefficients were highly significant and had correct signs. The  $R^2$  value was 0.90, and an extensive set of statistical tests indicated no violation of the normal least squares regression assumptions (consistent pooling, homoscedasticity, normality of residuals, etc.).

Blumer addresses the fact that aggregating over income distribution may be a somewhat limiting assumption.

"It might be said that neither population nor income truly measures the ability of a city to generate travelers. Population is poor because a large city would still not produce many travelers if all residents were poor. Income also has shortcomings in that, if used alone, it assumes every dollar has the same propensity to generate a trip -- a [person] who is twice as rich will take twice as many flights. Surveys taken indicate the flight generation is not linear with income, but increases at an increasing rate (for the income levels sampled).<sup>15</sup>

He suggests, as a future research topic, the development of a more general

---

<sup>15</sup> Ibid., p. 78.

mass variable which assumes that the function relating household income levels to number of trips varies by mode:

$$M_i = \sum_{r=1}^R N_i(r) \sum_{m=1}^K H_{im}(r) \quad (2.18)$$

where

$r$  = index of income levels

$R$  = number of predefined income levels

$N_i(r)$  = number of households in city  $i$  in income level  $r$

$m$  = index of modes

$K$  = number of modes

$H_{im}(r)$  = average number of trips by mode  $m$  per household of income level  $r$

An obvious shortcoming in what otherwise is an exceptionally good model is the simultaneity of the response variable,  $T_{ij}$ , and the impedance term,  $I_{ija}$ . Blumer realizes and addresses this weakness again within the rubric of future research.

"Another area to explore is a second [equation in the ] model representing service. Obviously the number of flights depends upon the demand and the demand depends upon the number of flights. This two way causality suggests building a second [equation] representing air service to be solved simultaneously with the first [equation]. This would also eliminate the requirement to forecast air service, since it would become an endogenous variable." 16

---

16 Ibid., p. 112.

## 2.4 Aggregation by Modes and Destinations

In his doctoral dissertation at M.I.T., Philip Verleger constructed a "point-to-point" model of air transportation demand.<sup>17</sup> The purpose of this research was not to construct a forecasting model, but rather a model to test if demand relations vary across markets. This, in turn, would thereby test the validity of cross-sectional and aggregate models. The models proposed and tested by Verleger were similar in the respect that, while they were city pair oriented, they did not consider the effects upon air traffic demand due to the existence of competing modes and destinations.

The structure of the model that was calibrated over all routes in his sample follows:

$$T_{ij} = aP_{ij}^{\lambda}(M_i M_j)^{\gamma} (\text{SEA } 1)^{\alpha_1} (\text{SEA } 2)^{\alpha_2} (\text{SEA } 3)^{\alpha_3} \epsilon \quad (2.19)$$

where

$T_{ij}$  = air travel between cities  $i$  and  $j$

$P_{ij}$  = air fare between cities  $i$  and  $j$

$M_i$  = mass of city  $i$

and SEA 1, SEA 2, SEA 3 = seasonal dummy variables.

The price variable,  $P_{ij}$ , is an index formed by weighting the major airline fares: first class, coach, family, and discounts. For routes in which the fare usage data were not available, they were estimated by taking those values for "similar" routes. "Thus the weights for Cleveland

---

<sup>17</sup> Philip K. Verleger, "A Point-to-Point Model of the Demand for Air Transportation" (Ph.D. dissertation, M.I.T., 1971).

to New York were used for Cincinnati to New York."<sup>18</sup>

A major contribution of this research was Verleger's handling of the income distribution effects. He argues that population alone as a mass variable (a common occurrence in gravity models) assumes that either income distributions are constant among cities or that the proclivity of people to travel is independent of their income. Since the former point is known to be false and the latter has been refuted by many surveys, Verleger chose to disaggregate the traveling populations by income.

The definition of his mass variable is as follows:

$$M_i = \sum_{\ell=1}^N x_i^{\ell} e^{\beta_i Y_i^{\ell}} \quad (2.20)$$

where

- $\ell$  = index of income groups in city  $i$
- $N$  = number of predetermined income groups
- $x_i^{\ell}$  = weighting of group
- $\beta_i$  = regression coefficient
- $Y_i^{\ell}$  = average income of group

The estimation of the parameters of the model (2.19) required a nonlinear regression technique, and a special algorithm was written. An iterative procedure was devised by searching for the value of  $\beta$  ( $\beta_i$  and  $\beta_j$  were assumed equal in a given city pair  $ij$ ) that minimized the variance. In nearly all markets (108 of 115) a positive estimate for  $\beta$  resulted, and in most of these markets it was statistically significant (79 of the

---

<sup>18</sup> Ibid., p. 112.

108 were significant at the 95% level). In addition to this, the overall income coefficient  $\gamma$  was positive in all but one market (Boston-New York) and significant in 105 markets.

The general conclusion arrived at by the author is that air travel is "very income elastic"<sup>19</sup> and only weakly responsive to price changes, as the price elasticity showed no regular pattern and was negative and significant in only 20% of the markets. The mean price elasticity was -0.12 with a variance of 0.45. The author concluded that it is inadvisable to interpret aggregate measures of price elasticity with any confidence.<sup>20</sup>

The author proceeds to analyze the fare effect by density. The results show that in the more heavily travelled markets, the fare elasticities have a tendency to be more uniformly significant, while in the low density markets few are significant. The author suggests that the variance of the price coefficient decreases as traffic increases, and that in aggregate analyses a weighted least squares estimation procedure should be used to eliminate this bias. However, it is possible that the implied weak impact of fares is due to the omission of a level of service variable. In dense markets the service will be greater than in the sparse markets, and inclusion of a level of service variable may therefore strengthen the significance of the price variable by reducing the variance of its coefficient. In other words, the level of service variable would provide

---

<sup>19</sup> Ibid., p. 186.

<sup>20</sup> Ibid., p. 189.

an economically justifiable set of weights and then ordinary or preferably two-stage least squares could be used to estimate the coefficients.

### 2.5 Aggregation by Incomes, Destinations, and Modes

A two-equation city pair economic model of air transportation service was developed by Pat Marfisi in his Ph.D. dissertation at Brown University.<sup>21</sup> The purpose of this research was to model the capacity decision process of an airline firm within a given market when faced with uncertainty of future demand. This study was an innovative contribution to the research in air transportation economics, in that an attempt was made to build a comprehensive model of the supply side that would be compatible with the demand so that both could be solved simultaneously.

Aside from the statistical problems inherent with the simultaneity situation with single equation models, Marfisi states the following:

"Economists argue on the basis of theory that an increase in fare levels will increase market equilibrium capacity. It is argued here that this is not an unambiguous implication of economic theory but rather an empirical question regarding the relative magnitude of certain demand elasticities; a judgement which has never been tested by a statistically consistent procedure. Concurrent with the discussion suggesting a theoretical link between fare levels and capacity, other economists are estimating demand functions for scheduled air transportation based on econometric models that implicitly assume the absence of any linkage between the demand and supply side of the

---

<sup>21</sup> E. Pat Marfisi, "Theory and Evidence on the Behavior of Airline Firms Facing Uncertain Demand" (Ph.D. dissertation, Brown University, 1976).

market (i.e., demand does not constrain or alter supply behavior and vice versa." 22

Marfisi's discussion is segmented into two somewhat disjoint sections. The first section is a quite comprehensive economic theory of the supply side of the industry. The second half is an empirical study of his two-equation model. The major conclusion of the theory is that the use of a single-equation model to estimate price elasticity implicitly assumes that  $\frac{dQ}{dP} = \frac{\partial Q}{\partial P}$ , where  $Q$  is demand and  $P$  is price. Marfisi's theory arrives at the relationship where, if demand is differentiated with respect to price, the result is  $\frac{dQ}{dP} = \frac{\partial Q}{\partial P} + \frac{\partial Q}{\partial C} \cdot \frac{dC}{dP}$ , where  $C$  is capacity. Since it is unclear as to what is the sign of the derivative of capacity with respect to price, then  $\frac{dQ}{dP}$  may well be positive. The result is that positive price coefficients in reduced form equations are indeed consistent with economic theory and explain the somewhat embarrassing results of some economists (e.g., Verleger<sup>23</sup>), who misspecify a reduced form as a structural form system and obtain positive price elasticity estimates.

The structural form of the model used for Marfisi's empirical study is as follows:

$$Q = C^{B_1} P^{\gamma_1} Y^{\gamma_2} \rho^{\gamma_3} \quad (2.20)$$

$$C = Q^{B_2} P^{\gamma_4} W^{\gamma_5} \Sigma^{\gamma_6} N^{\gamma_7}$$

---

<sup>22</sup> Ibid., p. 2-17.

<sup>23</sup> Verleger, op cit., p. 188.



where

$Q$  = origin to destination passenger demand

$C$  = flight frequency

$P$  = fare

$Y$  = per capita income

$\rho$  = population

$W$  = average cost per flight

$N$  = number of competing firms

$\Sigma$  = measure of demand dispersion

The capacity variable is more formally defined as the number of nonstop and one-stop scheduled flights. While the inclusion of a level of service variable is praiseworthy, this definition is restricting. Such important aspects as the relative number of nonstops to one-stops, whether the one-stops are direct or connecting, the speed of the aircraft, and the time of day scheduling, are not reflected in this measure.

The competition variable, number of competing firms (airlines), is also a very important addition in this model, but its definition is also somewhat crude. The variable was assumed to be the number of airlines in a market that provide at least five percent of the capacity. By this specification, a highly competitive market such as Chicago/Los Angeles, which is served by four strong carriers, would be considered the equivalent in competitive structure of a market in which one major carrier provides eighty percent of the capacity while three other minor competitors evenly provide the remainder.

Some very important results regarding pooling of data have emanated

from this study. Marfisi stratified the markets by purpose of trip and length of haul and conducted a series of Chow tests to investigate the appropriateness of aggregating markets across these strata. The conclusions indicate that the underlying structural relationships between market variables are substantially different. This is to be expected, as it is generally believed that business travel is considerably less price elastic than non-business travel. Furthermore, the structural relationship for markets of less than 200 miles are significantly different than that of markets greater than 200 miles. This effect is due presumably to the presence of competing modes in the short haul markets. The analysis showed, however, that no significant differences in the structural relationships exist between medium and long haul markets.

## 2.6 Total Systemwide Aggregation

The most common approach to econometric modeling of air transportation demand is the total systemwide aggregate model. These macroeconomic models predict a single figure for the industry, revenue passenger miles (RPM). While these models are generally very accurate in terms of forecasting, they are restricting in two respects. The first limitation is that they are inappropriate for measuring the impacts of policy decisions (e.g., changes in fare or other regulatory factors, implementation of new technology). The second limitation is that knowledge of the aggregate variable RPM is of little value for the industry planner.

Revenue passenger miles is a term used to measure overall growth of the industry. An individual carrier in its planning process is more concerned with expected growth of its own route structure which is a function of, among other things, the cities it serves, the types of markets it serves (business, vacation, etc.) and length of haul. Aircraft manufacturers are concerned with industry growth, but primarily as a function of length of haul and of market density. The selections of range and seating capacity for aircraft design are based upon this information. Airport operators are concerned with the growth of traffic through their facilities, which are functions of cities' abilities to generate traffic, attractiveness to lure traffic, and their positions within the total route structure.

Verleger creates an analogy between RPM and aggregate commodity output.

"[RPM] has been useful for accounting purposes. It is not, however, comparable to aggregate measures of a commodity such as tons of wheat. In fact, an aggregate measure of travel demand is more like a measure of gross output of a group of commodities, such as the gross output of all agricultural production measured in tons. Such an aggregate would, of course, be rejected by most agricultural economists as a meaningless measure." <sup>24</sup>

While there are far too many of these models to exhaustively summarize them in this chapter, the reader is referred to several listed in the bibliography and a summary paper coauthored by this writer.<sup>25</sup> One example

---

<sup>24</sup>Ibid., p. 49.

<sup>25</sup>Steven E. Eriksen and Nawal K. Taneja, "Directions for Improving Air Transportation Demand Models", (paper presented at the Joint National Meeting of the Operations Research Society of America and the Institute of Management Sciences, Miami, November 1976).

is, however, summarized below.

Jonathan C. Tom developed a macroeconomic forecasting model for the Federal Aviation Authority (FAA)<sup>26</sup> which consists of three equations, two of which are separately calibrated using ordinary least squares, and the third of which is an identity. The structural form of the model is as follows:

$$\begin{aligned}
 \text{RPM} &= f(\text{SRVC}, \text{APSU}, \text{PAT}, \text{REL}, \text{STR}) \\
 \text{ENP} &= g(\text{CMP}, \text{APSU}, \text{PAT}, \text{REL}, \text{STR}) \\
 \text{OPS} &= h(\text{RPM}, \text{LOAD}, \text{SEATS}, \text{STAGE})
 \end{aligned}
 \tag{2.21}$$

where

$$\begin{aligned}
 \text{RPM} &= \text{revenue passenger miles} \\
 \text{ENP} &= \text{revenue passenger enplanements} \\
 \text{OPS} &= \text{air carrier itinerant operations (aircraft departures)} \\
 \text{SRVC} &= \text{consumer expenditures on services} \\
 \text{CMP} &= \text{number of civilians employed in the labor force} \\
 \text{APSU} &= \text{purchases of automobiles} \\
 \text{PAT} &= \text{plant, equipment, and other investment in the air} \\
 &\quad \text{transport industry} \\
 \text{REL} &= \text{relative price of air transportation to other modes} \\
 \text{STR} &= \begin{cases} 1 & \text{if a major airline strike is in progress} \\ 0 & \text{otherwise} \end{cases}
 \end{aligned}$$

---

<sup>26</sup> Federal Aviation Administration, "Aviation Forecasts, Fiscal Years 1976-1987" (Washington: FAA Technical Report FAA-AVP-75-7, September 1975).

LOAD = average load factor

SEAT = average number of seats per aircraft

STAGE = average stage length

The service consumption variable was selected since air transportation is in itself a service, and therefore it is hypothesized that this will represent the income effect upon demand. Although in theory this is a sound hypothesis, it is possible (since this is a totally aggregate system and therefore not sensitive to income distribution variation between markets, such as Blumer's and Verleger's models) that income alone would have been sufficient. Total income is a direct causal factor related to consumer expenditures on services which correlates with air travel demand. Thus, little if any explanatory power would have been lost. The advantage of having used income is that it is probably easier to predict than service consumption, and the small loss in the explanatory power is more than offset by the increased forecasting accuracy.

The number of civilians employed supposedly reflects that portion of the population which would use air services. The rationale behind this variable is that the number of air passengers is a function of the employment figure, while the distance flown, as measured by RPM in the first equation, is a function of total income (as measured by SVRC).

The purchase of automobiles was included to capture the impact of competition by other modes. However, since the decision to purchase an automobile is probably unrelated to air transportation, but more likely to serve purposes that are not competitive with air transportation (i.e., journey to work, shopping for necessities, visiting nearby friends

and relatives), the usefulness of this proxy is somewhat questionable.

The investment variable ostensibly is a measure of level of service. The theory behind its inclusion is that new plant and equipment are purchased to ultimately improve the quality of the operation. If this is indeed the case, then it should probably be lagged.

The relative price of air transportation with respect to other modes is defined as the ratio between a price index for standard coach air fare and the cost of private transportation. This measures a combination of air fare elasticity and auto price cross-elasticity, but does not separate the two impacts. One limitation of this observation is that, if all modes of transportation increased their prices by a fixed percentage (due, for instance, to an overall fuel increase), the demand for travel on each mode, including air, would decline. This impact, however, would not be captured by the price variable in this model.

The strike variable accounts for the decrease in travel due to the service impedance due to a major airline strike. The load factor, available seats, and stage length variables are scaling factors to convert revenue passenger miles into number of operations.

The first two equations are specified in linear form and calibrated using quarterly data, starting with the first quarter of 1964 and continuing through the third quarter of 1974. The results are as follows:

$$\begin{aligned} \text{RPM} = & 12.32 + 0.32 \text{ SRVC} - 0.06 \text{ APSU} + 0.56 \text{ PAT} - 0.17 \text{ REL} - 4.18 \text{ STR} \\ & (17.35) \quad (-3.94) \quad (1.17) \quad (-2.18) \quad (-2.63) \\ R^2 = & 0.955 \quad \quad \quad D - W = 1.791 \end{aligned}$$

$$\text{ENP} = -75.01 + 1.64 \text{ CMP} - 0.04 \text{ APSU} + 1.98 \text{ PAT} - 0.17 \text{ REL} - 5.79 \text{ STR}$$

$$\qquad\qquad (15.75) \quad (-1.76) \quad (3.04) \quad (-1.62) \quad (-2.52)$$

$$R^2 = 0.944 \qquad\qquad\qquad \text{D - W} = 1.487$$

(figures in parentheses are "t" statistics)

In spite of the somewhat questionable theoretical merit of the definition of some of the explanatory variables, the statistics indicate a reasonably good fit. While the estimates of the coefficients are perhaps difficult to interpret, the model appears to be an adequate forecasting tool, assuming that the functional relationships are invariant over time.

The addition of the strike variable improves the fit of the model, but its coefficient is not particularly useful for forecasting purposes. The only observation during the time period over which this model was calibrated that included a major strike was the third quarter of 1966. The coefficients estimated in this calibration indicate that the 1966 strike resulted in 4.18 billion fewer revenue passenger miles and 5.79 million enplanements than would have been expected otherwise. This is not necessarily representative of the impact upon travel of any future strikes which may be longer or shorter in duration and/or occur at other times of year, when general demand patterns are quite different. In fact, it can be shown that the definition of a dummy variable which assumes a positive value at only one data point is statistically equivalent to eliminating that particular data point. A dummy variable may, however, have been useful to account for the redefinition in 1969

of systemwide RMP's to include trips between the continental United States and Alaska or Hawaii.

## 2.7 Summary

Figure 2.2 is a schematic representation of the development of the models surveyed in this chapter. A totally disaggregated model of the demand for transportation service is depicted at the top of the figure. Such a model would, for each market, consider the response of each income level group of consumers, to not only changes in characteristics and price of air service in that market, but also to changes in characteristics and price of services in competing markets and by competing modes. Since this analysis is obviously intractable, the researcher is forced to aggregate some or all of these factors.

The extreme case of aggregation is the macroeconomic model at the bottom of Figure 2.2, which is detailed in Section 2.6. In this setting a single figure, revenue passenger miles (RPM), is generated for the entire industry. Since this figure is of little value as a planning tool, it is clear that forecasting models must necessarily be geared to a more microeconomic level.

In the following chapters a theory and set of models that focus on individual markets will be developed. Due to data limitations, aggregation by modes, destinations, and income groups will be necessary; therefore, this set of models fits next to Marfisi's model in Figure 2.2. The models developed herein will be more complete than that of Marfisi,



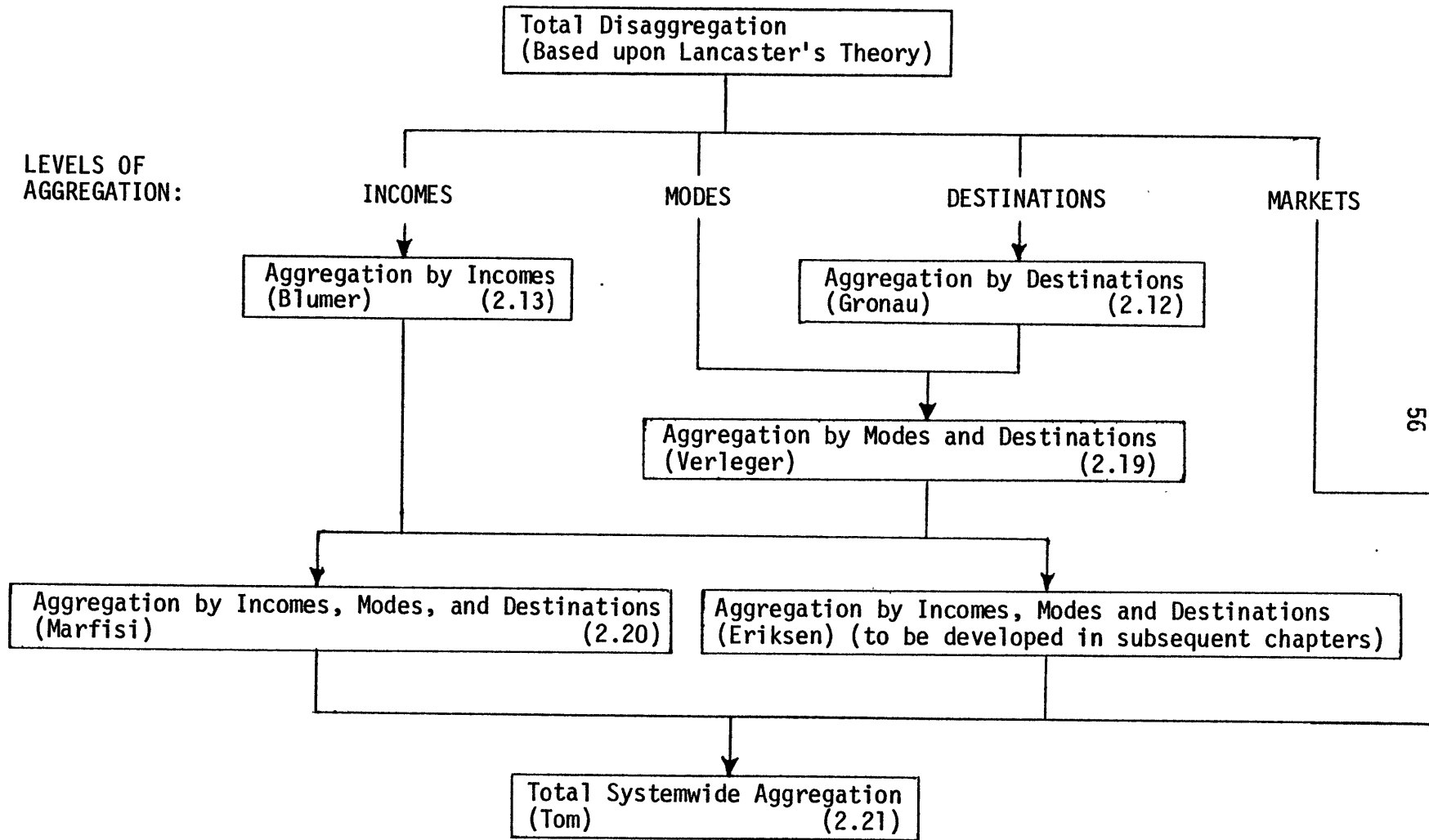


Figure 2,2 Representation of Development of Models Surveyed in Chapter II

due to a more realistically defined set of variables and a more representative sampling procedure. Furthermore, due to detailed utilization of available data and a substantial sample size, the level of sophistication will be adequate to meet the requirements set forth in Chapter I.

### III. The Economic Nature of Domestic Air Transportation Passenger Service

Chapter II contained a summary description of a general theory of demand and how it relates to transportation services. Also included in Chapter II was an overview of how this theory has spawned several models of air transportation demand. In this current chapter, a brief summary of major economic issues pertaining to the nature of air passenger activity within a given market is presented. A more specific theory of demand than that discussed in the previous chapter is developed, the result of which is a two-equation model of demand and service within an air transportation market.

The chapter is divided into three sections, the first of which is a general description of the product, its market place, and its consumption. The second section is an investigation of a hypothetical "isolated market". The third section views the market as an interacting subset of a large and complex network structure. In progressing from the second to the third section, the discussion will indicate why, due to peculiarities of the U.S. domestic air passenger system, the theories drawn from classical economic theories are inappropriate for analysis of this system.

### 3.1 Fundamental Economic Aspects of Air Passenger Service

#### 3.1.1 The Air Transportation Market

The basic market for air transportation is a "region pair" and the level of consumption in a region pair market in a given time period is the number of "origin to destination passengers" that purchase this service. The term "region pair" is in essence identical to the term "city pair" commonly used in transportation systems analysis. However, the former term is used in this study because it emphasizes the fact that a major airport serves a surrounding area which is generally larger than the city itself. Therefore, in the consideration of the potential density of a market, the analyst must consider regional demographics rather than merely those of the central urban area. An inherent difficulty confronting the analyst is the definition of region boundaries. These boundaries may vary for any given airport as a function of the length of haul of the market. For example, the catchment area around the Boston airport certainly includes Providence, Rhode Island when one considers the Boston to Los Angeles market. However, it is less clear that the air service from Boston to New York attracts a significant number of travelers from Providence.

To be consistent with the above definition, the "New York-Miami market" refers to those potential passengers originating from a point in an area surrounding New York who would desire to travel to a destination point in an area surrounding Miami, and vice versa. This is schematically represented, including depiction of the access from the origin to the New

York airport, the line haul portion, and the egress from the Miami airport to the destination, in Figure 3.1. Such references as "the Northeast to Florida market" or the "North Atlantic market" are inconsistent with the definition of a market, since they are actually aggregates of many markets.<sup>1</sup>

The origin to destination passengers in a region pair market are defined as those travelers that originate in one of the regions and terminate in the other region. There are a number of other travelers as well who are not consumers of services in this market, but who appear on board flights that do service the market. Consider the flight whose itinerary, New York to Chicago to San Francisco, is shown in Figure 3.2, and particularly that segment of the flight from Chicago to San Francisco. If one were to survey the passengers on board during that particular segment, the following eight types of passengers may be observed:

- a. those who originated in Chicago (region) and are terminating in San Francisco (region), i.e. local market consumers
- b. those who originated in New York and are terminating in San Francisco
- c. those who connected through Chicago and are terminating in San Francisco
- d. those who connected through New York and are terminating in San Francisco
- e. those who originated in Chicago and are connecting through San Francisco

---

<sup>1</sup> Robert W. Simpson, "A Theory for Domestic Airline Economics", unpublished lecture notes for the M.I.T. course 16.74 -- Air Transportation Economics, Fall 1974, p. 5.

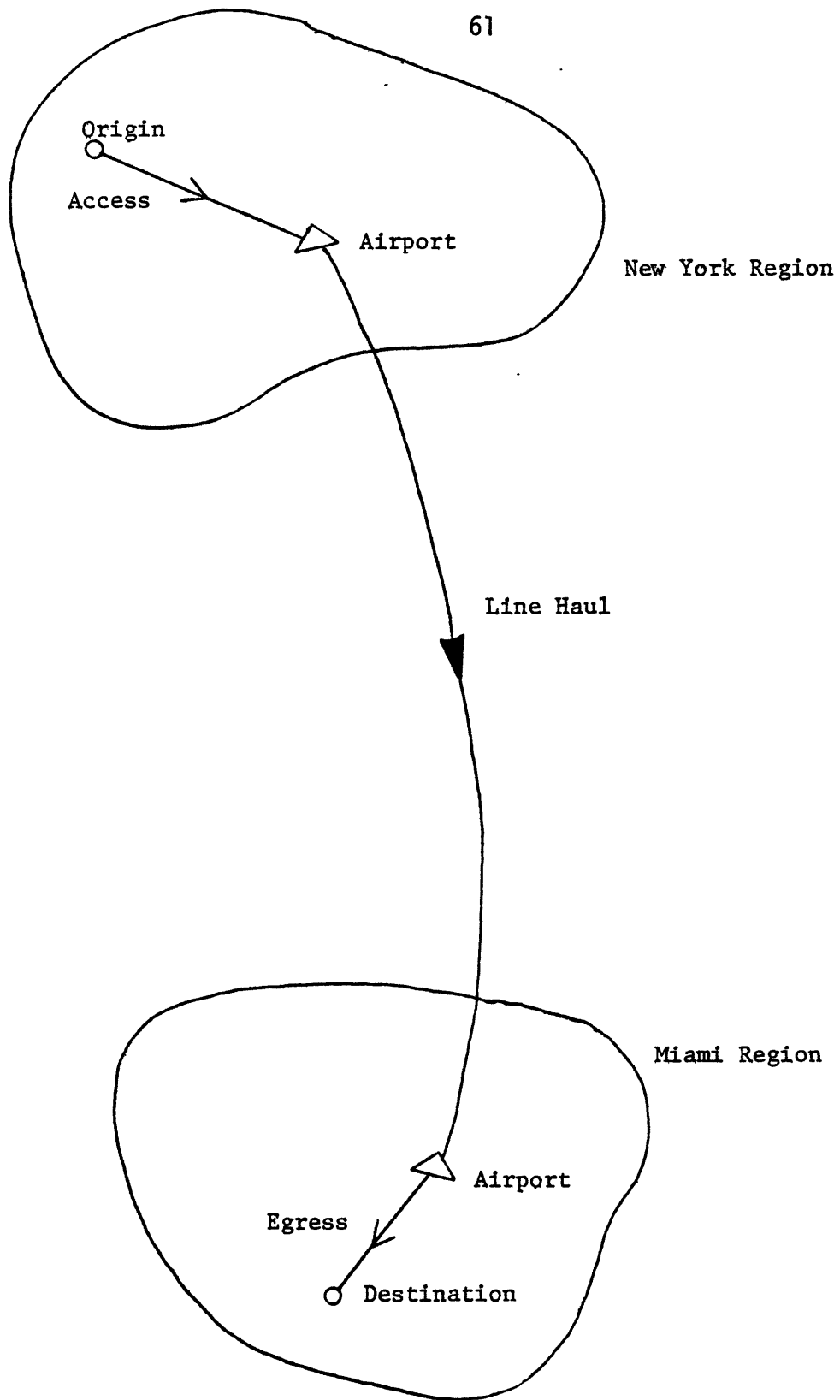


Figure 3.1 The New York-Miami Market

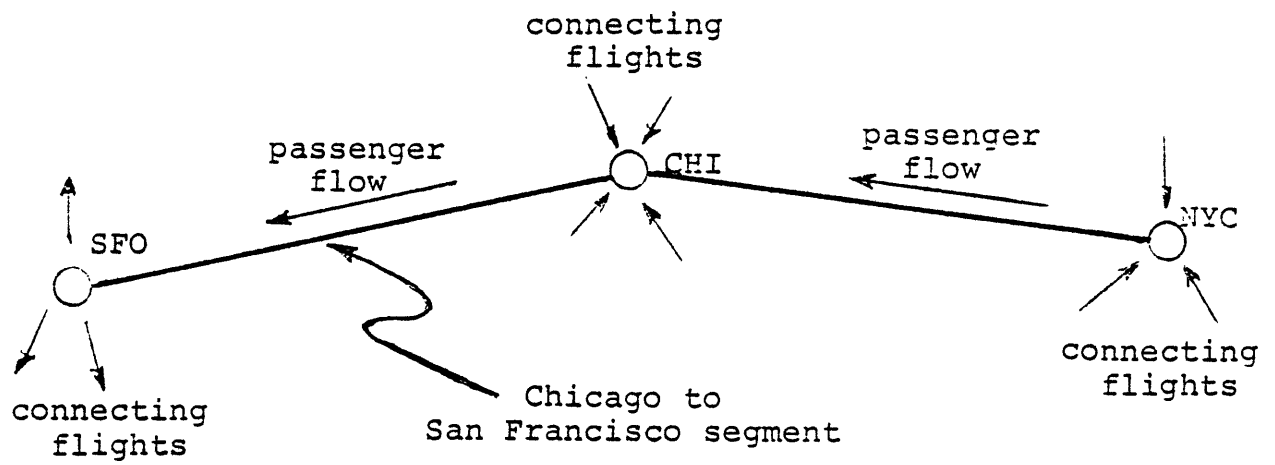


Figure 3.2 Typical Itinerary of a Transcontinental Flight

- f. those who originated in New York and are connecting through San Francisco
- g. those who connected through Chicago and are connecting through San Francisco
- h. those who connected through New York and are connecting through San Francisco

All of the above classes of passengers are considered as "segment flow" traffic, since they are on board during this segment. However, only those passengers that are in the type (a) class are consumers in the local market (origin to destination passengers).<sup>2</sup>

### 3.1.2 The Airline Product -- Not a Good but a Service

The service provided by the air passenger transportation industry is a perishable product similar, in this respect, to a newspaper, a Christmas tree, and a unit of capacity at a performance of an opera. If service

---

<sup>2</sup> The number of these classifications of passengers on a particular segment grows quite rapidly with the number of segments included in a particular itinerary. For example, on the middle leg of a three-segment flight, one could enumerate sixteen such types. In general, it can be shown that the number of such classes on the  $m$ th segment of an  $n$  segment flight is  $4m(n-m+1)$ . This is a major complication that arises when one attempts to analyze the supply of, and demand for, service in a given market. The level of service offered in a segment may reflect not only the origin to destination demand, but also the location of the market with respect to the complex network of the U.S. domestic and international route structures. This issue will be elaborated upon in Section 3.3 and in Chapter IV.



is produced and not sold by a particular point in time (in this case, the scheduled departure time), the cost of producing this service is essentially the same as if it had been purchased, but its salvage value to the would-be seller (in this case, the airline) is zero. While this fact is universally accepted, the definition of the product and whether it is a good or a service are widely debated.

The fact that perishable products such as those mentioned above are generally referred to as "perishable goods" and the fact that the airlines as well as other suppliers of transportation services are frequently considered to be in the business of selling seats (seats being physical objects) may erroneously lead one to believe that the airlines' product is a good. The airlines are not, however, selling the consumers any physical object which becomes the property of the passengers (aside from incidental amenities). The primary product of an airline is the service of the transportation of a passenger from some point A to another point B with a series of attributes to be discussed in Section 3.1.3. This service will hereafter be referred to as a "trip".

Even though a passenger purchasing a trip is provided a seat on the aircraft for his comfort, and the capacity of an aircraft departure equals the number of such seats, this does not imply that the airlines are selling these goods. An analogy can be made with public surface transportation. The reason that a passenger purchases the right to board a bus, subway, or commuter train at some location is to satisfy his desire to be transported to another location. Whether a seat is available or he is forced to stand, he will still be provided the same product -- the service of being transported. While there exists a perceived product

differentiation between a passenger fortunate enough to find an available seat and one who was forced to stand, the basic service purchased, the transportation from A to B, is identical.

The demand for air transportation services is a "derived demand". A consumer generally does not purchase an airline ticket because he or she has an absolute need to sit on an airplane as it travels to its destination. This need is derived from the fact that the passenger has an absolute need to be at his or her destination for business, pleasure, and/or personal reasons, or is returning home after fulfilling such a need. Another example of a product with a derived demand is steel, which is purchased by automobile manufacturers in response to consumer demand for motor vehicles. The marketability of any product with a derived demand is obviously responsive to the absolute demand of the products to which it is an input. In fact, this relationship can be extremely sensitive due to the "acceleration principle" of economics.<sup>3</sup> Just how much the acceleration principle applies to the airline industry is uncertain, but it will suffice to say that the industry is clearly sensitive to business cycles and to fluctuations in the tourism industry.

Most analysts agree that the air transportation product is highly differentiated. However, there is some debate as to the factors by which the product is differentiated. The major differentiation in terms of consumer response (generating origin to destination passengers) is trip time. A trip on a piston aircraft, a trip on a subsonic jet, and a trip

---

<sup>3</sup> Eugene F. Brigham and James L. Pappas, Managerial Economics (Hinsdale, Illinois: The Dryden Press, 1976), pp. 118-121.

on a supersonic transport over the same route, are three clearly different deliveries of the same service. Another trip time consideration is the number of enroute stops and/or connections. Other differentiations which have a negligible effect upon passenger generation but a significant effect upon market shares in competitive markets are distribution channels (making reservations), ground services (ticketing and baggage handling), and cabin services (food and liquor service, entertainment, and attitude and appearance of the cabin crew).

"Those persons who believe that the carriers scarcely dwell on such differences [in terms of marketing cabin services] are reminded of Delta's slogan of professionalism, PSA's stewardesses' sex appeal, TWA's "lasagne over Los Angeles", Western's free champagne, and American's piano bars." <sup>4</sup>

One debatable issue is whether differences in time of day of departure constitute product differentiation. Some analysts would argue that *ceteris paribus* a trip at 5:00 p.m. and one at 2:30 a.m. are differentiated services. Since these two trips offer the same aircraft type, trip time, onboard amenities, etc., they are actually identical services offered at different times. The number of passengers on the two flights may be quite different, due to the time of day dependence of the utility functions of potential purchasers of air transportation services, which will be elaborated upon in Section 3.2.

---

<sup>4</sup> Nawal K. Taneja, The Commercial Airline Industry (Lexington, MA: Lexington Books, 1976), p. 131.

### 3.1.3 Attributes of the Quality of Air Passenger Service

When a consumer purchases an airline ticket, he is buying the right to be delivered the service of a "trip" from an airport in one region to an airport in another region.<sup>5</sup> The quality of this service is comprised of many different attributes, some of which are major factors in generating passenger demand, and others of which have a very small effect. These attributes are herein categorized into five classes: trip time, comfort and onboard amenities, safety, reliability, and convenience. While the set of attributes contained in these classes may not be exhaustive, they include the majority of items considered by the individual in determining whether and when to fly.

#### 3.1.3.1 Trip Time

The distinct service advantage that air transportation offers over other modes, particularly in medium and long haul markets, is faster line haul travel time. For this reason, air is the only sensible mode for any traveler who has a high utility of time and desires to make a trip of any appreciable distance. However, in the process of deciding if, when, and by which mode, a prospective traveler has several trip time components to consider. These are access and egress times, preflight and postflight

---

<sup>5</sup> For the purposes of this analysis, only interregional air transportation will be considered. Very short haul transportation between an airport in one city to an airport in another city within the same region and, in the extreme situation, the ultra short haul traffic between two airports within the same city are separate cases not considered in this research.

processing times, and displacement time, as well as the line haul or flight time.

Access time is the time required to travel by some other mode from the passenger's origin to the curb of the airport from which he or she will depart. Egress time is the amount of time required to travel by another mode from the curb of the arrival airport to the passenger's destination. These time components will obviously have a greater adverse effect upon those passengers whose origins and/or destinations are near the periphery of the regions than those whose origins and/or destinations are near the airport. Access and egress times clearly constitute a greater portion of the total trip time for shorter lengths of haul. These factors account in part for the statement in Section 3.1.1 that the catchment area around an airport is smaller for shorter length markets, hence general region boundaries are very difficult to define.

The preflight processing time is the elapsed time from passenger arrival at the curb to aircraft departure from the gate. Activities included in this component typically consist of walking to the check-in counter, ticketing, checking baggage, walking to the gate, receiving a boarding pass, waiting for boarding, boarding, and waiting for departure. Postflight processing includes waiting to deplane once the aircraft arrives at the destination airport, deplanement, walking to the baggage claim, waiting for and claiming baggage, and walking to the curb.<sup>6</sup>

---

<sup>6</sup> As was discussed in Chapter I, a substantial amount of research is being performed to analyze and improve the flow of passengers through airport terminals.

Displacement time is the difference between actual and desired departure times. This is a function of the time dependence in the individual's processes. This component of total trip time is of substantial analytical importance and will be studied in detail in Section 3.2 and in Chapter IV.

The line haul or flight time is the component of trip time inferred from consulting the airline schedule. It is the elapsed time between aircraft departure from the gate of the origin airport and arrival at the gate of the destination airport. The flight time includes the time spent making intermediate stops and/or connections and is a function of the cruise speed of the aircraft, proximity of the gates to the runways, length of haul, atmospheric conditions, air traffic delays and the output of the airlines' scheduling processes.

#### 3.1.3.2 Comfort and Onboard Amenities

This class of service attributes includes all efforts to make the passenger's experience aboard the aircraft enjoyable. These efforts are designed to promote the image of the industry and the carrier in general, and are necessary for competitive reasons within the given market. These attributes include among others the interior layout and decoration of the cabin, the comfort of the seat (including amount of available space), the appearance and attitude of the flight crew, food and beverage service, and entertainment.

Since, as will be discussed in Section 3.1.4, the carriers cannot

compete by varying fares, onboard amenities have been a variable that the airlines use in advertising to boost their shares of profitable markets. One rarely finds a steak or free champagne in the coach cabin of a flight in a monopolistic market. The exploitation of onboard amenities will be discussed in more detail in Section 3.1.5.

The quality of these attributes is substantially improved if the consumer pays a considerable fare premium to sit in the first class rather than the coach section. However, the first class passengers are generally people who are not paying the fare themselves, or the extremely affluent, and the marginal utility of increased quality of comfort and onboard services is probably not substantially great to the vast majority of consumers. Therefore, it is doubtful that fluctuations in the level of this type of service will significantly affect total market demand. Differences between carriers will have an effect upon market shares when similar prices prevail.

#### 3.1.3.3 Safety

The safety attributes refer to the probability of death or injury. It is questionable if prospective consumers possess an accurate perception of this attribute. Minor fluctuations in the safety performance of the airlines relative to the other modes are likely to go unnoticed by the traveling public and therefore have no effect upon demand.

An unfortunate fact is that an airline crash is a very dramatic event and tends to be highly publicized in the media. While a fatal accident

may cause a short-term effect in the demand for the carrier involved, it is doubtful that a significant aggregate decrease would result. A reduction in total demand due to safety reasons would probably occur only in the event of a series of fatal accidents in a fairly short time span, scarring the safety image of the total industry, and even then it is doubtful that this effect would endure for long.

#### 3.1.3.4 Reliability

Measures of reliability include the probabilities of space being available, cancellation, and on-time performance. Space availability is a function of average load factor.<sup>7</sup> As the load factors increase, the probability of a flight being fully booked will obviously increase.

Simpson states:

"Generally, this [probability] measure of space available is kept very high by all airlines in a market, especially since load factors are generally below 60%...Normally then, it does not become an important variable for quality of service." <sup>8</sup>

Since cancellations of scheduled flights occur very infrequently, and then only in the event of mechanical problems or inclement weather,<sup>9</sup> their

---

<sup>7</sup> The load factor of a flight segment is the ratio of passengers on board to the aircraft capacity and will be further discussed in Section 3.2.

<sup>8</sup> Simpson, op. cit., p. 10.

<sup>9</sup> A scheduled departure cannot be cancelled due to a low level of demand. This point is frequently overlooked when comparing the economics of scheduled and charter air services.



effect upon demand is negligible. Even if a single flight in a given market is cancelled, the passenger can usually secure an alternative. This factor becomes a deterrent to demand in the event of extremely inclement weather and in the event of an airline labor strike. Since such occurrences are basically unpredictable, the forecaster generally constructs his estimates assuming (whether he knows it or not) the absence of such conditions. However, in the analysis of historic data, the analyst must be aware of and make adjustments for these events.

The carriers must periodically report to the Civil Aeronautics Board statistics on the percentage of flights which arrive more than fifteen minutes late. Since these data are publicly available and since the airlines are promoting (or damaging) their image by on-time performance, "these measures are generally kept quite high by all airlines."<sup>10</sup>

In the process of selecting a particular flight, the passenger may assess a prior expected delay and allow for this by choosing a flight that is scheduled to leave earlier than is necessary for him. However, the decision of whether or not to purchase an air trip is generally made sufficiently far in advance of the departure to preclude any specific knowledge of whether or not the flight will be delayed. Furthermore, if a passenger arrives at the gate and discovers the departure has been delayed, the likelihood of his cancelling the trip is probably quite small. Therefore, it is doubtful that this attribute has a significant impact upon total market demand. However, one large trunk carrier seems to believe that this attribute will affect market share and is promoting itself as

---

<sup>10</sup> Ibid., p. 11.

the "on time airline".

#### 3.1.3.5 Convenience

The convenience attributes consist of ease of receiving flight information, making reservations, and securing tickets. In addition, the ease of access and egress, preflight and postflight processing, and the number of enroute stops and connections fall under the rubric of convenience.

The evolution of online computer systems has facilitated the processes of information retrieval, booking, and ticketing to the point where one could hardly be dissuaded from making a trip because of problems related to these functions, except in the rare event of a total computer system breakdown.

The ease of access, egress, and processing and the number of enroute stops are accounted for in the total trip time. However, the consequence of a connection is only partially accounted for by the increment in flight time. The additional burden of an online (connecting to a different flight on the same airline) connection over a through stop includes the probability that the connections will be broken by either the first leg aircraft arriving late or the second leg aircraft being delayed or cancelled,<sup>11</sup> the inconvenience of locating another gate and boarding a different aircraft, and the probability of a mistransfer of baggage. An

---

<sup>11</sup> Anyone who has had an experience similar to that of the author, of sitting in the Denver airport for six hours with a cranky one-year-old and a frustrated wife, will appreciate this peril.

interline (connecting to a different flight on a different airline) connection bears all the drawbacks of an online connection, and in addition, the transfer includes a possible journey to another section of the terminal or even a different terminal building, and a greater likelihood of lost baggage. These considerations will be further addressed in Chapter IV.

#### 3.1.3.6 Summary of the Service Attributes and Their Impact Upon Demand

The service attributes have been categorized into five classes: trip time, comfort and onboard amenities, safety, reliability, and convenience. The impact of the perceived levels of the quality of the various attributes upon demand varies substantially depending upon the attribute. Some attributes have an effect both upon the generation of origin to destination traffic and on market share, some affect only market share, and others have no significant impact except in extreme or unusual circumstances upon either.

The major class of attributes in terms of passenger generation is trip time. Since the relative attractiveness of air over over modes, particularly in the long haul markets, is speed, a decrease in trip time (e.g., faster aircraft, fewer intermediate stops) will increase its share of the modal split. Furthermore, this increase in speed will, by the gravity principle discussed in Chapter II, generate new travelers by effectively bringing the regions closer together. Trip time also has an impact upon market shares; the carrier who offers faster service will *ceteris paribus* experience the greater demand.

While comfort and onboard amenities have little effect in generating traffic, they do (at least temporarily) affect market share. Safety and reliability will have virtually no effect on demand, given that the industry's steady performance in these attributes is not radically altered. Convenience attributes affect demand patterns only insofar as they affect travel time and in the case of connecting service.

#### 3.1.4 The Price of Air Transportation Services

The economics of domestic air transportation services in a given market is somewhat unique, in that the producers (airlines) currently have little power, at least in the short run, to vary prices. The fares are set by the Civil Aeronautics Board (hereafter referred to as the CAB or the Board), which prescribes a piecewise linear concave function of intercity distance for the standard coach fare.<sup>12</sup> First class and discount fares are computed solely on the basis of percentages of the standard fare.<sup>13</sup> Thus air fares at any given point in time are a function solely of market distance and are independent of absolute consumer demand in the market and of fluctuations in this demand. Present regulatory policy prohibits personal, temporal, or locational price discrimination.<sup>14</sup> The intercity

---

<sup>12</sup> For example, the fare structure resulting from Phase 9 of the Domestic Passenger Fare Investigation (1974) was:

\$12.56 + 7.07 cents per mile for the first 500 miles  
           + 5.39 cents per mile for miles 501 to 1500  
           + 5.13 cents per mile for miles 1501 and beyond

(from CAB Domestic Passenger Fare Investigation, p. 763)

<sup>13</sup> For example, first class fare is roughly 150% of standard coach and night coach fare is 80% of standard coach.

distance between New York and Chicago is 721 miles and the intercity distance between Bangor, Maine and Akron/Canton, Ohio is 694 miles. The former market experiences a demand of roughly 1.5 million passengers per year, while the latter market attracts fewer than 100 passengers per year. However, the fares in these two markets are virtually identical.

Furthermore, the airlines have no power to seasonally adjust fares in markets that experience seasonal demand patterns. The standard coach fare from New York to Miami, a highly seasonal vacation market, in February is identical to the standard coach fare in August (barring any interim overall fare changes approved by the CAB). This situation is quite unlike the other components of the travel industry, such as hotels and restaurants, who freely exercise the option of seasonally adjusting prices.

Therefore, in the economic analysis of air transportation, the price variable is an exogenous variable. It is set outside of markets and is not affected, at least in the short run, by fluctuations in any variables that are specific to a given market (such as demand, regional economics, etc.).

Based upon findings during the Domestic Passenger Fare Investigation (DPFI), which was initiated by the Board in 1970 and lasted four years, it was concluded that the fare formula should be constructed on an average systemwide cost basis, so that the average rate of return on investment

---

<sup>14</sup> An exception is that the interstate and commuter carriers may, subject to CAB approval, charge up to 130% of standard coach fare.

for the trunk airlines will be 12.00 percent and for the local service carriers 12.35 percent. Since these guidelines are based upon system averages, there appear to be inherent inequities. Since individual carriers are constrained to operate only in those markets for which they are certified by the Board, those airlines with an abundance of profitable routes have a definite financial advantage over those that are required to serve less fruitful markets. This factor has been a common bone of contention in past route award cases, but it is generally very difficult to ascertain if a carrier's current financial status is a result of its route structure or the quality of its managerial functions.

Another topic for debate related to the passenger fare structure is the issue of differentiated fares and services, since there may be several different fares paid among the various passengers with the same origins and destinations. The different prices charged first class and coach class passengers obviously constitutes fare and service differentiation, but also included are the various discount plans such as military, youth, Discover America, and family fares. Airlines benefit by these discounts, because they generate additional passengers who would otherwise not fly. The penalty incurred by the carriers is that reduced fares divert some existing passengers who would normally pay the greater fare. To combat this problem, the carriers usually establish certain service conditions for discount fares. For example, the Bicentennial discount requires a round trip in no less than seven and no more than thirty days. This is an attempt to prevent, for example, the businessman who makes a one to five day trip or students who are returning home for the entire summer, to avail themselves of the price break.

A frequently-cited fare variable that is sensitive to the responses of fare differentiation is "yield". The yield (in the total system or a market) is the ratio of the total ticket revenues to the number of revenue passenger miles. A typical quoted yield, for example, may be "7.3 cents per mile". If consumer response to the introduction of a discount fare plan is insignificant, the carriers will notice a decline in the yield. Or if there is a shift in the ratio of first class to standard coach passengers, either due to a shift in aggregate consumer tastes or reconfiguration of aircraft (as was the trend during the sixties when a greater number of coach and a fewer number of first class seats were appearing), the yield will react accordingly. Note that the yield may vary, although the basic fare structure is held fixed.

A major summary point in the discussion of the price for airline services in a given market is that it is basically fixed. There is no allowable adjustment for marginal costs of various qualities of service. A nonstop flight is priced the same as a multistop or a connection. If an airline decides for competitive reasons to offer steak dinners in the coach section, there is no reflection in the fare (which is why you rarely find a steak in the coach class in a monopolistic market). Simpson summarizes this point:

"...when one carrier institutes an improved level of service... the lack of quality distinction in fares makes it difficult for other carriers to sell the lower level of services, and encourages a competition in quality of service." 15

---

<sup>15</sup> Simpson, op. cit., p. 24.

### 3.1.5 Some Aspects of Marketing, Advertising and Competition

Many aspects of marketing air transportation services are quite unusual. Included in these peculiarities are the following:

(1) As was discussed in the previous section, the price is fixed and therefore the carriers must compete by altering the marketing mix of the other attributes.

(2) The product can only be delivered in batches (flights).

(3) A seat departure in a given batch is generally available to consumers from many different markets, as shown in the example of Figure 3.2.

(4) The set of markets which a given airline serves may contain a wide range of competitive structures, varying from strictly monopolistic to oligopolistic markets.

(5) The demand for air transportation services is a derived demand.

Individual marketing and advertising strategies generally can be placed on a continuum. At one end of the continuum is primary marketing and advertising, and at the other is competitive marketing and advertising. Primary strategies refer to actions that are intended to increase the overall demand in a market by generating new passengers or convincing existing passengers to fly more frequently. Competitive marketing or advertising refers to a carrier's attempts to increase its share of a given market.

An example of an advertisement that is mostly primary in nature is a television commercial in which Eastern shows vacationers enjoying a



variety of frolics in Florida and the Caribbean. While the effect of this advertisement may increase Eastern's north-to-south traffic, it may also draw passengers to the other carriers operating in these markets. Its major purpose is to increase demand in north-to-south markets. An example of purely competitive advertising is North Central's campaign to attract passengers in the Boston to Detroit market, for which they have recently received certification to serve. Without ever extolling the virtues of visiting Detroit, they promote their schedule of four nonstops a day, their steak and eggs breakfast, and other advantages of North Central's service.

Since a carrier may only produce its services in batches, it is in the best interest of the airline to schedule the departures at times that will attract the maximum number of passengers. This can be a difficult problem in that the network structure in which a carrier operates may be very complex. Since the airline is constrained by the fleet size, scheduling decisions in one market interrelate with those of many markets. Furthermore, for competitive reasons, one carrier's optimal scheduling decisions are a function of the outcome of those of the other carriers in the market.

An example of highly competitive or "head to head" scheduling occurred in the late sixties in the New York-Chicago market. American, TWA, and United were locked into a fierce scheduling war in which each carrier felt the need to match every competitive flight with one of their own at the same departure time. At one point there were nearly ninety nonstop flights offered in each direction every day.

In other markets, "collusive" scheduling occurs, where the competitors schedule in different time slots over the day. An example of this is the Boston-San Francisco market, in which TWA and United each offer one nonstop flight per day. United operates the morning westbound flight, while TWA operates the early afternoon flight. On the eastbound link, TWA departs San Francisco in the morning and United in the afternoon.

A substantial amount of past competitive advertising has focused upon comfort and onboard amenities (United's "Friendly Skies", TWA's "Lasagne Over Los Angeles", Delta's "Champagne Coach", etc.). However, the recent trends are toward informing the consumers of scheduling superiority. Many current newspaper and magazine ads actually publish full schedules for given airlines and competitive markets, reaffirming the conviction that the most important determinant of market share is frequency and speed of service. However, in certain markets, airlines still play to consumers' tastes for onboard amenities. An example of this is the New York to Florida markets in which Delta, National and Eastern are supplying their wintertime vacationers with free movies, drinks, and relatively elegant meals. This practice has reached such an extreme that marketing experts are referring to it as a "frills fracas", "revival of free side-shows", and a "big bingo game".<sup>16</sup>

---

<sup>16</sup> "CAB Frill Investigation Expected", Aviation Week and Space Technology, November 15, 1976, p. 26.

### 3.1.6 Summary

Section 3.1 contained a description of five aspects of the domestic air transportation system: the market, the product, the attributes of service quality, the price, and competitive marketing. The concepts presented in this section will be referred to throughout the remainder of Chapter 3 and in Chapter 4 as the economic model is developed. In the remaining sections of Chapter 3, the concepts of the market, price, and quality of service attributes will be particularly important. Section 3.2 contains an analysis of how demand would be related to price and service in a hypothetical market which is isolated from a network. Section 3.3 considers the more realistic case, in which the market is a subset of large network structure, as described in Section 3.1.1. It will become quite evident that the interrelationships of the economic aspects that have been discussed in this current section become very complex when the market is embedded in the superstructure of the domestic passenger system.

### 3.2 A Hypothetical Isolated Market

Consider a simple network structure which consists of two regions, A and B. The transportation link between these two regions is comprised of air service offered by a single carrier who by regulation must charge a fixed fare, but is allowed to vary the level of service by adjusting the number of flights offered per day. Since these regions are not linked to any others, all operations out of A fly nonstop to B, and vice versa, and all passengers on board any flight on this segment are origin to destination passengers in the A-B market.

### 3.2.1 Analysis of Demand

As discussed in Section 3.1.3 and 3.2.4, the major factors that influence the level of origin to destination demand in a given market are trip time and fare. Certainly the greatest volume of traffic that could currently be generated in a given market would be experienced if the fare level were set to zero and if nonstop jets were continually departing during the traveling day. This hypothetical volume of traffic will be noted as  $Q_{D_0}$ .

If such an ideal (from the consumer's point of view) level of service existed in market A-B, the number of passengers in a given day that would fly from A to B<sup>17</sup> is the number of people in region A (whether residents of A or visiting in region A) whose expected cardinal utility of a visit to region B is greater than the sum of the expected utility of not visiting B and the utility of the trip time from A to B. Therefore, those passengers who travel from A to B feel that what they will gain by being in region B for the duration of their stay outweighs the value of remaining in A plus the time investment of the trip.

The term "visit" refers to the purpose of being in region B, whether it is business, personal, pleasure, a combination of these, or it is returning home to B from a visit in region A (the term visit is hence used rather loosely in this last respect). The term "trip" refers to the service that is purchased from the airline.

---

<sup>17</sup> Without loss of generality, the focus of this analysis will be upon those passengers flying from A to B.

Defining:

$K$  = the set of people in region A (residents and non-residents) in a given day

$k$  = a generic person in region A,  $k \in K$

$t_0$  = the nonstop jet line haul trip time between A and B

$T$  = the set of time points in the traveling day (this may be a continuous scale or a discrete approximation,  $T =$

$\{t_j | j=1,2,\dots,n\}$ , and may include the entire 24-hour day or that part of it during which departures are feasible

$U_k^v(t_j)$  = the expected cardinal utility of a visit to region B (relative to not making the visit) given that a flight is boarded in region A at time  $t_j \in T$  for a generic person  $k \in K$

$U_k^t(t)$  = the cardinal utility function of time for individual  $k$  (assumed to be linear)

$T_k^g(t_j)$  = the expected ground time (access, egress, preflight and postflight processing times) required for person  $k$  boarding a flight at time  $t_j$

$N\{S\}$  = the number of elements in any generic set  $S$ , then

$$Q_{D_0} = N \left\{ k \in K \mid U_k^v(t_j) > U_k^t(t_0 + T_k^g(t_j)) \text{ for some } t_j \in T \right\} \quad (3.1)$$

Since  $U_k^t(t)$  is assumed to be a linear function, equation (3.1) can be written as follows:

$$Q_{D_0} = N \left\{ k \in K \mid U_k^v(t_j) > U_k^i(t_0) + U_k^t(T_k^g(t_j)) \text{ for some } t_j \in T \right\} \quad (3.2)$$

In the case where flights are departing at every instant of the day and the line haul trip time,  $t_0$ , is constant, the rational passenger will elect to depart at that time  $t_j \in T$  at which the difference between his expected utility of the visit,  $U_k^v(t_j)$ , and the utility of ground time,  $U_k^t(T_k^g(t_j))$  is maximized.

Defining:

$$\bar{U}_k^v(t_j) = U_k^v(t_j) - U_k^t(T_k^g(t_j)) \quad \text{and}$$

$$\bar{U}_k^v(*) = \max_{t_j \in T} \left[ U_k^v(t_j) - U_k^t(T_k^g(t_j)) \right],$$

equation (3.2) simplifies to

$$Q_{D_0} = N \left\{ k \in K \mid \bar{U}_k^v(*) > U_k^t(t_0) \right\} \quad (3.3)$$

Before proceeding further into this analysis, it is appropriate to elaborate upon two characteristics of the above derivations. The first of these is that the above approach considered the merits of transportation to region B only in terms of the utility of being in the destination region at a particular time. The utility of the "trip" (the transportation service per se) was not considered. This is consistent with the discussion in Section 3.1.2, in which the demand for air transportation services was described as a "derived demand", not purchased as an end to itself, but rather, as a means for achieving a separate end.

A second and very important point is that the utility function  $U_k^V(t_j)$  is time dependent. If this hypothetical market existed in reality and was, due to the economic nature of the regions, highly business oriented, it would be expected that the demand would be intensified in the early morning and late afternoon. This effect is due to a strong desire for business travelers to arrive at their destination in the early and mid-morning hours and to depart for home as soon as possible at the end of the business day. A late departure in the morning or an early return flight in the afternoon implies a sacrifice of time during the business hours in the destination region, hence a lower relative utility of the visit. Likewise, an inconveniently early departure in the morning or late return flight at the end of the day will result in a loss of personal time in the home region. By similar reasoning, demand will also vary by day of the week and month of the year.

Empirical data on the impact of the time of day utility variation upon demand is for most markets difficult to encounter, since actual passenger flow is decided by imperfect scheduling. However, some markets with very frequent and regular service (such as the Boston-New York shuttle) have provided data which reflect the bimodal time of day demand pattern mentioned above. Time of day demand variations will be analyzed in much greater detail in Chapter IV.

The level of demand for air transportation service in a given market will be lower than  $Q_{D_0}$  if a positive fare is charged, or if the level of service provided is something less than the "perfect" service described above.

Defining,

$F$  = fare for service from A to B

$Q_{D_F}$  = passenger demand given fare  $F$  and perfect level of service

$U_k^m$  = cardinal utility of money for individual  $k \in K$ , and

$W_k$  = current wealth of individual  $k$ ,

then

$$Q_{D_F} = N \left\{ k \in K \mid \bar{U}_k^v(*) > U_k^t(t_0) + U_k^m(W_k) - U_k^m(W_k - F) \right\} \quad (3.4)$$

Since  $F$  is positive and  $U_k^m$  is clearly a strictly monotonically increasing function, then  $Q_{D_F} < Q_{D_0}$ .

Now suppose that instead of a nonstop jet departing at every instant of the day, a less than perfect level of service, a finite set of departures were offered. Defining

$I$  = the set of flights

$i$  = a generic flight,  $i \in I$ , and

$D_i$  = local departure time of flight  $i$

then the displacement time, the amount of time which a person must sacrifice if he desires to depart at time  $t_j$ , is  $|t_j - D_i|$ . A generic person  $k$  will then choose to purchase a trip if there exists some time  $t_j$  during the traveling day in which the utility  $\bar{U}_k^v(t_j)$  exceeds the sum of the utility loss of the fare and the utility of the displacement time and the block flight time for some flight  $i$ . The passenger demand,  $Q_D$ , is then defined as follows:



$$Q_D = N \left\{ k \in K \mid \bar{U}_k^V(D_i) > U_k^m(W_k) - U_k^m(W_k - F) + U_k^t(t_0) \text{ for some } i \in I \right\} \quad (3.5)$$

If the regulator were to set the fare level,  $F$ , to zero, the two  $U_k^m$  terms would cancel, yielding a finite level of demand that is no greater, and depending on the schedule, may be substantially less than  $Q_{D_0}$ . As the fare is increased, the difference  $U_k^m(W_k) - U_k^m(W_k - F)$  will increasingly become positive, ultimately reaching a finite point  $F_0$ , where the inequality condition of equation (3.5) will be satisfied by no member of  $K$  and the demand will vanish. Consequently, the demand vs. fare curve for a given market and a given schedule will meet the axes at finite values. However, it is uncertain what the characteristics of this curve are between the extreme points. A reasonable approximation to the shape of the demand vs. fare curve is the common convex function usually assumed in classical microeconomic analyses, shown in Figure 3.3.

It is quite widely accepted that the parameters of the demand vs. fare curves vary significantly between markets. A common theory is that the elasticity of demand with respect to fare<sup>18</sup> is substantially greater in markets that are primarily pleasure oriented (e.g. Los Angeles-Las Vegas) than those that are basically business markets (e.g., Chicago-New York).

If no service were offered between A and B, the variable  $D_i$  in equation

---

<sup>18</sup> The elasticity of demand with respect to fare,  $\epsilon$ , is defined as the ration of the percentage response of demand to the percentage change in fare.

$$\epsilon \equiv \frac{\Delta Q_D / Q_D}{\Delta F / F} \approx \frac{\partial Q_D}{\partial F} \cdot \frac{F}{Q_D}$$

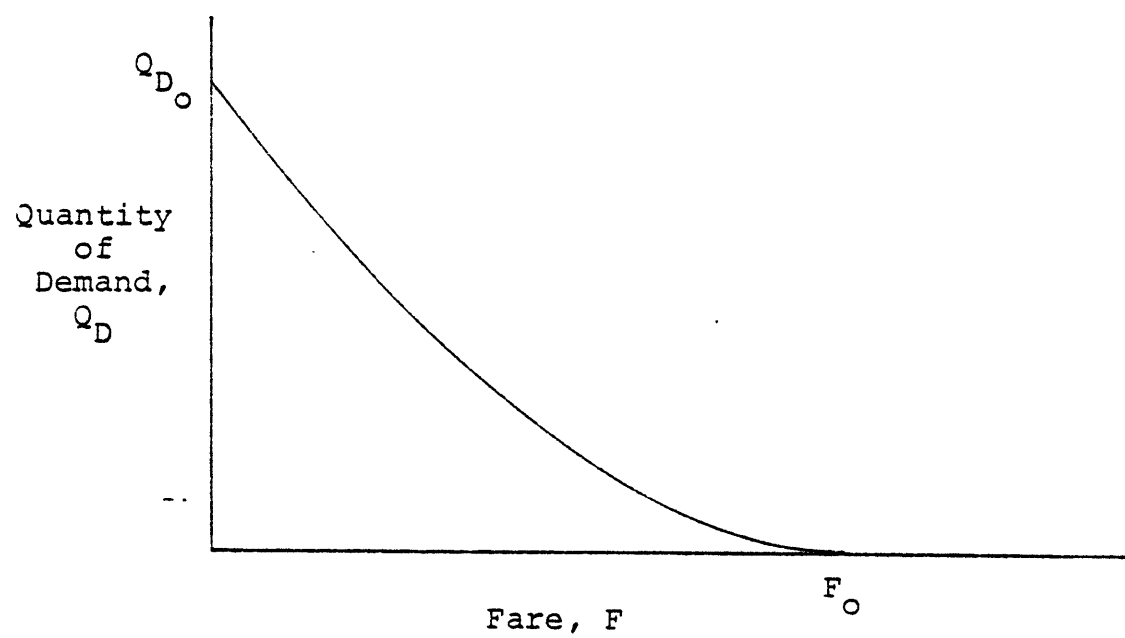


Figure 3.3 Conceivable Shape of Typical Demand Vs. Fare Curve

(3.5) would not exist, and obviously the demand would be zero. Since it is assumed for each member of the population  $K$  that the function  $\bar{U}_k^V(t_j)$  has a maximum value at some  $t_j \in T$ , then

$$\max \bar{U}_k^V(D_i) \leq \bar{U}_k^V(*) \quad \forall k \in K \quad (3.6)$$

Equality in expression (3.6) will in general exist for all members of  $K$  only when the service offered includes flights at every time point  $t_j \in K$ . This corresponds to the condition yielding equation (3.4) in which the demand is  $Q_{D_F}$ . Therefore, as the number of flights,  $N\{I\}$ , increases, the demand asymptotically approaches  $Q_{D_F}$  for some fixed fare  $F$ . This yields a demand vs. number of flights, or demand vs. "frequency", curve, as shown in Figure 3.4.

It can be observed from Figure 3.4 that as  $N\{I\}$  increases, a diminishing return of demand is experienced. This illustrates the fact that eventually as more flights are added, a "saturation frequency" will be reached. This saturation frequency, depicted in Figure 3.4 as  $N\{I\}_{\text{sat}}$ , is defined by Simpson as the number of flights at which the demand reaches 95% of  $Q_{D_F}$ .<sup>19</sup>

### 3.2.2 Analysis of Supply

As discussed in Section 3.1.6, the units of output in air transportation can be produced only in batches. In classical economic analysis of the

---

<sup>19</sup> Simpson, op. cit., p. 17.

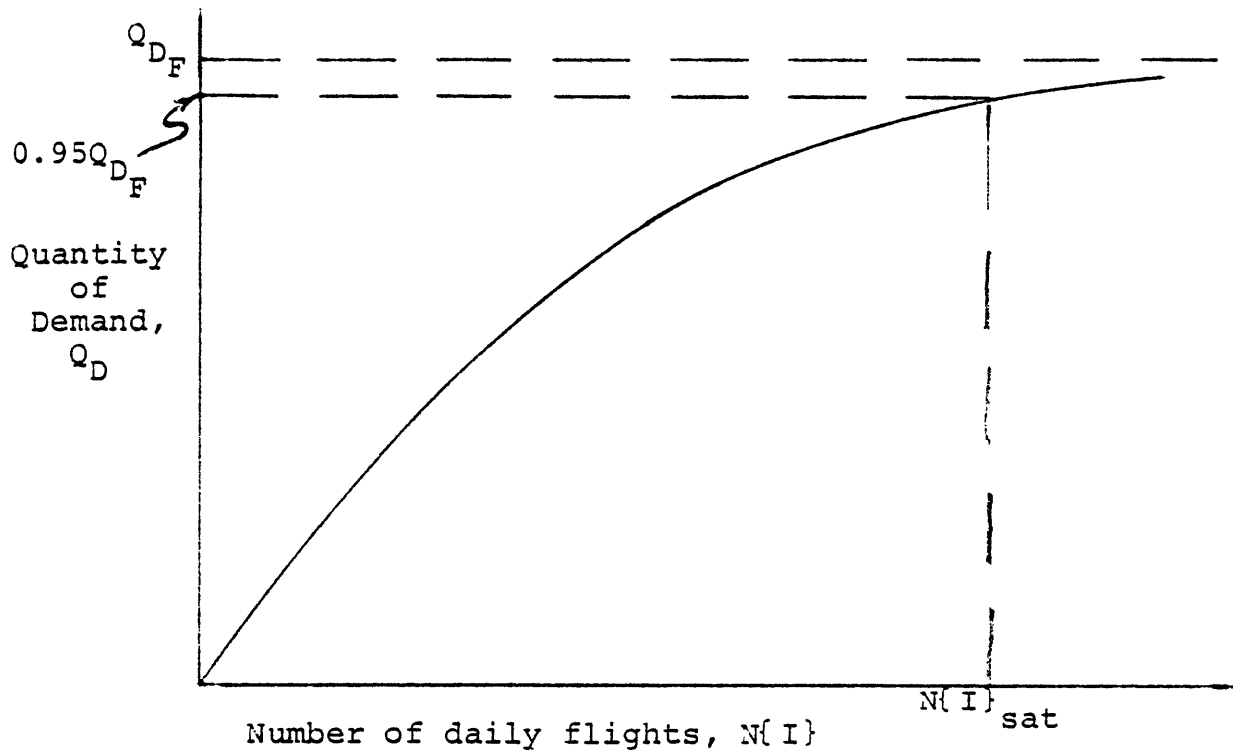


Figure 3.4 Typical Demand Vs. Frequency Curve

production of goods, a manufacturer can increase the production by a desired amount and is afforded substantial flexibility in selecting this amount. Airline trips are produced in batches called "flights" and the number of trips produced in a flight is constrained to the capacity of the aircraft. So the unit decision to increase production is whether to offer an additional flight or to replace an existing departure with a larger aircraft (if one is available in the fleet). In either case, the additional quantity of supply is a discrete and quite inflexible increment.

As mentioned in Section 3.2.1, the airline operating in market A-B will maximize the number of passengers being carried if it offers a nonstop jet departure at every instant of the day. Obviously, since the fleet availability is constrained, the airline must offer something less than perfect service. However, offering the maximum level of service that the fleet size will allow is not necessarily an optimal decision, since there exist diminishing revenue returns, as shown in Figure 3.4, whereas variable costs related to each flight will not diminish. So for a given demand function, cost structure, fare level, and fleet availability, there appear to be an optimal number of flights that the airline can offer in a given day (or other time period) so as to maximize its return. This optimal level of service may or may not be bound by the fleet availability constraint.

Simpson has categorized airline costs into three classes: flight operating costs, ground operating costs, and system operating costs.<sup>20</sup>

---

<sup>20</sup> Ibid., pp. 35-42. Also, by the same author, "An Analysis of Airline Costs", paper given at the M.I.T./NASA Workshop on Airline Systems Analysis, Waterville Valley, July 1972.

Flight operating costs include the costs of flying the aircraft (crew, fuel, maintenance, landing fees, depreciation, etc.). Ground operating costs are the costs per aircraft departure of refueling, dispatching, and aircraft servicing; and the passenger-related ground costs, such as ticketing, reservations, baggage handling, etc. The flight operating costs and the ground operating costs are related to the flight, and are therefore herein considered as variable (by flight) costs. System operating costs are the remaining costs related to the system-wide overhead (executive salaries, ground equipment ownership, etc.). These are considered fixed costs and will not be included in this analysis.

There appears to be an internal economy of scale related to the costs of operating aircraft with different seating capacities. For example, if the flight plus ground operating costs of flying a Boeing 707-320 with 180 seats in a given medium haul market is \$4,000 per flight, the corresponding costs of operating a Boeing 747 with 360 seats is something less than \$8,000 per flight. Thus, if an airline is operating two 707-320s in a given market, it can reduce the cost while providing the same quantity of supply by replacing these flights with a single 747 departure. However, as will be shown, the reduction in revenue caused by one fewer departure may more than offset the cost reduction, causing a net marginal loss to the airline due to this decision. So, for a given market, there are both an optimal number of flights and an optimal seating capacity of the aircraft.

An additional constraint placed upon the airline's scheduling decision problem is that they must maintain an average load factor, the ratio of trips purchased to trips supplied, that is less than some maximum average

load factor (60%, for example). Because of the day-to-day variability of demand, a high average load factor will result in many prospective passengers being unable to purchase service, due to flights being fully booked. For example, if in a given period the demand for a given flight with a capacity of 100 seats were Gaussian distributed with a mean of 85 passengers and a standard deviation of 30 passengers, the flight would be fully booked nearly 20% of the time. If the capacity of the flight were increased to 145 seats, the average load factor would decrease, and the probability of the flight being fully booked would drop to less than 5%.

A similar constraint which in low density markets may affect the airline's scheduling function is a minimum level of service. An example of this would be a requirement that at least two daily flights be offered from A to B, and two from B to A. In this case, a person is afforded the ability to travel from A to B and back (or vice versa) in any given day.

The scheduling problem for the airline operating in an isolated market is to determine the number of flights and capacity of each flight, so as to maximize the difference between revenue and the sum of the flight and ground operating costs. This objective is constrained by the available fleet size, the prescribed average load factor, and the minimum number of daily departures.

Defining:

$A$  = the set of aircraft types available

$a$  = a generic aircraft type,  $a \in A$

$C_a$  = the sum of the flight and ground operating costs of a flight from region A to region B of aircraft type a

$S_a$  = the seating capacity of aircraft type a

$N_a$  = the number of aircraft of type a available for dispatch

$n_a$  = the number of flights offered daily from region A to region B of airline type a (the decision variable)

$Q_S$  = the total number of trips offered daily from region A to region B

LF = average load factor

$LF_{\max}$  = prescribed maximum average load factor

$\{I\}_{\min}$  = prescribed minimum number of daily flights from region A to region B

the scheduling decision problem (or, in economic terms, the airline production function) can be represented by the following integer program:

$$\max_{n_a} \Pi = Q_D \cdot F - \sum_a n_a \cdot C_a \quad (3.7)$$

subject to:

$$LF \leq LF_{\max} \quad (3.8)$$

$$\{I\} \geq \{I\}_{\min} \quad (3.9)$$

$$n_a \leq N_a \quad \forall a \in A \quad (3.10)$$

$$n_a \geq 0, \text{ integer} \quad \forall a \in A \quad (3.11)$$



Since  $Q_D$  is nonlinear with respect to the decision variables (as shown in Figure 3.4), the above integer program is an integer nonlinear program (INLP).

### 3.2.3 The Analysis of Equilibrium -- Some Numerical Examples

In this section, the interactions between the demand variables discussed in Section 3.2.1 and the supply variables discussed in Section 3.2.2 will be illustrated, using some numerical examples. The following demand function is hypothesized:

$$Q_D = Q_{D_0} \cdot F^{\alpha}(1 - e^{\beta F}) \quad (3.12)$$

where  $\alpha$  and  $\beta$  are negative constants.

The demand vs. fare relationship for fixed level of service does not meet the boundary conditions of Figure 3.3, in that equation (3.12) asymptotically approaches the axes as  $F$  approaches zero and infinity (rectangular hyperbola), whereas the curve represented by Figure 3.3 is bounded in both variables. However, since it is reasonable to assume that the equilibrium operating characteristics of the market are not near the extremes of the demand function, this is of no consequence. The advantages of this specification are that in the neighborhood of the equilibrium point the function is convex (since  $\alpha$  is assumed to be negative) as is the function in Figure 3.3, and that  $\alpha$  is the elasticity

of demand with respect to fare.<sup>21</sup> Therefore, if a reasonable intuitive estimate of the elasticity is available, this function can be quite accurately specified.

The level of service factor,  $1 - \exp(-\beta N\{I\})$  was selected since it meets the boundary conditions of the relationship shown in Figure 3.4 and has similar shape. When the number of flights,  $N\{I\}$ , is zero, the demand,  $Q_D$ , will likewise be zero, and the demand will saturate as  $N\{I\}$  gets large.

The following two relationships:

$$Q_S = \sum_a n_a \cdot S_a \quad (3.13)$$

and 
$$Q_D = Q_S \cdot LF \quad (\text{by definition of load factor}) \quad (3.14)$$

constitute the equilibrium conditions. Notice that, unlike the classical economic interpretation of the production and consumption of goods, the quantities of supply and demand are unequal in the airline industry.

The entire system is then as follows:

Demand equation: 
$$Q_D = Q_{D_0} \cdot F^\alpha (1 - e^{-\beta N\{I\}})$$

---

<sup>21</sup> It can be shown that in equation (3.12),

$$\frac{\partial Q_D}{\partial F} \cdot \frac{F}{Q_D} = \alpha$$

$$\text{Supply INLP: } \max_{n_a} \Pi = Q_{D_0} \cdot F^{\alpha+1} \left\{ 1 - \exp \left[ \beta \left( \sum_a n_a \right) \right] \right\} - \sum_a n_a \cdot C_a$$

$$\text{subject to: } LF \leq LF_{\max}$$

$$\sum_a n_a \geq N\{I\}_{\min}$$

$$n_a \leq N_a \quad \forall a \in A$$

$$n_a \geq 0, \text{ integer } \forall a \in A$$

$$\text{Equilibrium equations: } Q_S = \sum_a n_a \cdot S_a$$

$$Q_D = Q_S \cdot LF$$

Suppose that  $\alpha = -1.3$  and  $\beta = -0.5$ , and that the current average traffic is 400 passengers per day with the airline operating five flights from A to B at a fare level of \$50. Then

$$Q_{D_0} = \frac{Q_D}{F^{\alpha} (1 - e^{\beta N\{I\}})} = \frac{400}{50^{-1.3} (1 - e^{-2.5})} = 70,457 \quad (3.15)$$

Suppose that the airline has two types of aircraft,  $a_1$  and  $a_2$ , in their fleet. The capacities of these aircraft, their per flight costs of operation from A to B, and the number of each in the fleet, are shown in Table 3.1.

Table 3.1 Characteristics of Fleet

Aircraft Type <u>A</u>	Passenger Capacity <u>S<sub>a</sub></u>	Per Flight Cost of Operation <u>C<sub>a</sub></u>	Cost Per Seat <u>C<sub>a</sub>/S<sub>a</sub></u>	Number in Fleet <u>N<sub>a</sub></u>
a <sub>1</sub>	120	\$3,000	\$25.00	4
a <sub>2</sub>	180	\$4,000	\$22.22	4

Furthermore, suppose that the airline is required to operate at least two flights per day from A to B, and the average load factor may not exceed 60%.

### 3.2.3.1 Equilibrium Conditions at Optimal Level of Output

Given the numerical values of the parameters as defined above, the supply INLP may be solved to indicate that, for the fleet described in Table 3.1, the current level of frequency of five flights per day is not the optimal level of output. A special purpose algorithm to solve this INLP is described in the flowchart in Figure 3.5. A computer program entitled ISOMRKT has been written in PL/I coding this algorithm. This is a naive algorithm, and while it works very quickly in the solution of a problem this size, it would be very inefficient for a much larger-scale problem. The ISOMRKT program is currently written to handle fleets with up to five aircraft types.

The algorithm looks at each vector  $\underline{n}_a$  which is a set of values of the  $n_a$  variables that is feasible according to constraint sets (3.10) and (3.11), checks for feasibility in constraint (3.9), and if still feasible, evaluates  $Q_D$ ,  $Q_S$ , and LF using equations (3.12), (3.13), and (3.14). It then checks the load factor constraint, (3.8), and if still feasible evaluates the objective function.

Figure 3.6 is the output of ISOMRKT using the numerical values of this example. Note that the optimal level of output is four flights per day, three of the 180 seat aircraft and one of the 120 seat aircraft. The

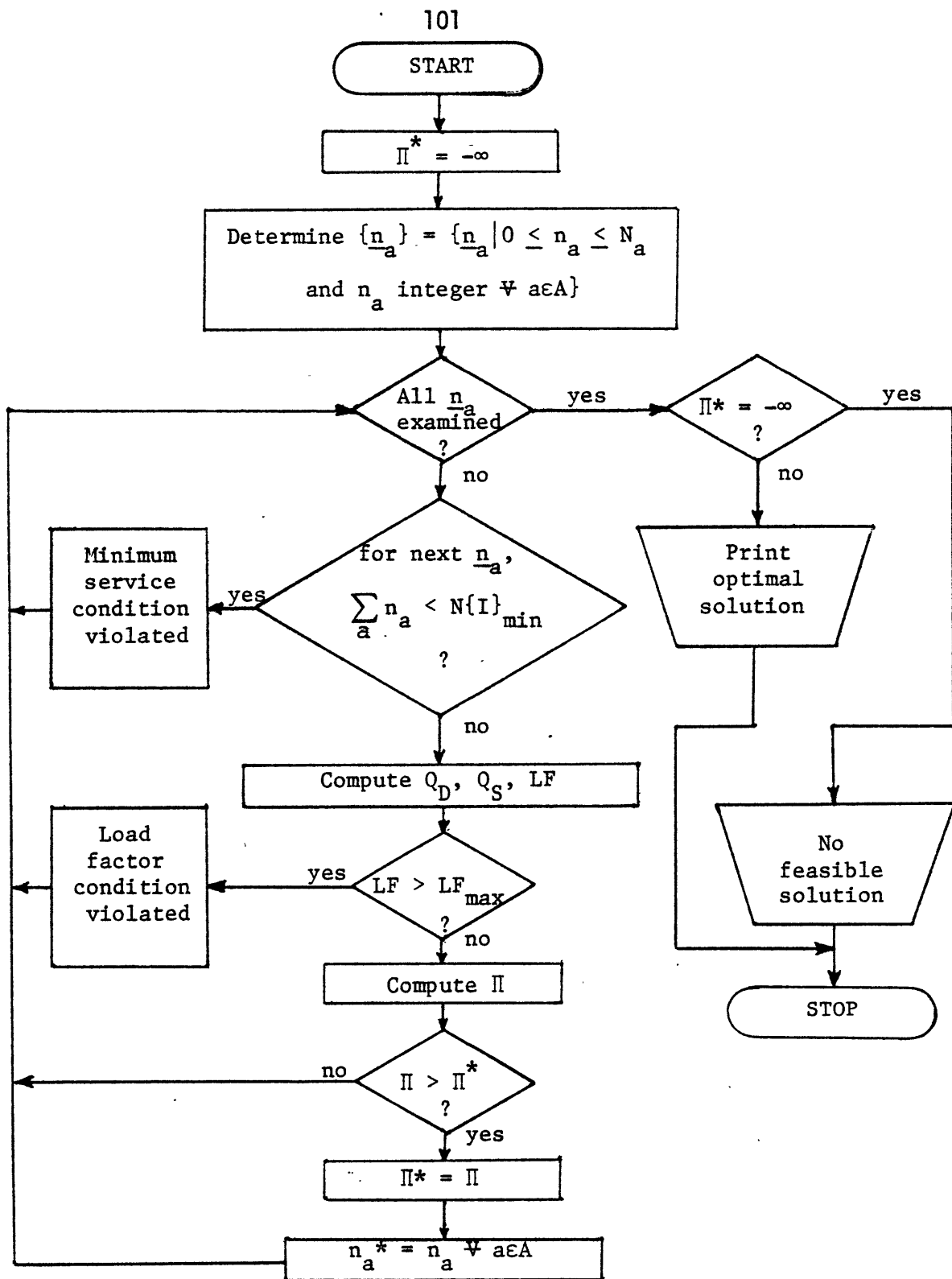


Figure 3.5 ISOMRKT Supply INLP Algorithm

## REGION A TO REGION B

N1	N2	N3	N4	N5	OS	OD	LF	REV	COST	PROFIT
0	0	0	0	0	MINIMUM SERVICE CONDITION VIOLATED					
0	1	0	0	0	MINIMUM SERVICE CONDITION VIOLATED					
0	2	0	0	0	360	275	76.5	LOAD FACTOR CONDITION VIO		
0	3	0	0	0	540	338	62.7	LOAD FACTOR CONDITION VIO		
0	4	0	0	0	720	377	52.3	18800.00	16000.00	2800.00
1	0	0	0	0	MINIMUM SERVICE CONDITION VIOLATED					
1	1	0	0	0	300	275	91.8	LOAD FACTOR CONDITION VIO		
1	2	0	0	0	480	338	70.5	LOAD FACTOR CONDITION VIO		
1	3	0	0	0	660	377	57.1	18800.00	15000.00	3800.00
1	4	0	0	0	840	400	47.6	20000.00	19000.00	1000.00
2	0	0	0	0	240	275	100.0	LOAD FACTOR CONDITION VIO		
2	1	0	0	0	420	338	80.6	LOAD FACTOR CONDITION VIO		
2	2	0	0	0	600	376	62.8	LOAD FACTOR CONDITION VIO		
2	3	0	0	0	780	400	51.3	20000.00	18000.00	2000.00
2	4	0	0	0	960	414	43.1	20700.00	22000.00	-1300.00
3	0	0	0	0	360	338	94.0	LOAD FACTOR CONDITION VIO		
3	1	0	0	0	540	376	69.8	LOAD FACTOR CONDITION VIO		
3	2	0	0	0	720	400	55.6	20000.00	17000.00	3000.00
3	3	0	0	0	900	414	46.0	20700.00	21000.00	-300.00
3	4	0	0	0	1080	423	39.1	21100.00	25000.00	-3900.00

Figure 3.6 Output of ISOMRKT Program

(continued on next page)

4	0	0	0	0	480	376	78.5	LOAD FACTOR CONDITION VIO
4	1	0	0	0	660	400	60.6	LOAD FACTOR CONDITION VIO
4	2	0	0	0	840	414	49.3	20700.00 20000.00 700.00
4	3	0	0	0	1020	423	41.4	21100.00 24000.00 -2900.00
4	4	0	0	0	1200	428	35.6	21350.00 28000.00 -6650.00

REGION A TO REGION B

OPTIMAL SOLUTION

N(1) = 1    N(2) = 3    N(3) = 0    N(4) = 0    N(5) = 0

OPTIMAL PROFIT = 3800.00

Figure 3.6 Output of ISOMRKT Program (continued from previous page)



total supply of seats is 660 and the expected demand is 377 passengers per day for a load factor of 57.1%. The profit<sup>22</sup> to the airline is \$3,850.

### 3.2.3.2 Effect of an Increase in Total Demand

Suppose that the utility functions of the people in region A for visits to region B,  $\bar{U}_v^k(*)$ , were to shift, causing a 20% increase in total demand (increasing  $Q_{D_0}$  to 84,538). Such a shift could be caused by increased income or population in region A, a new attraction (amusement, recreation, or cultural facility perhaps) in region B, a change in season affecting vacation traffic, a promotional or advertising campaign by the airline, or some other factor.

Under the previous solution of four flights per day, the sudden increase in demand will elevate the average load factor to 68.5%, which is in violation of the maximum load factor constraint. Therefore, additional service must be offered to accommodate the 20% increase in demand. Reoptimizing using the supply INLP algorithm results in a level of output of five flights per day, one 120 seat aircraft and four 180 seat aircraft, or an increase of one 180 seat flight. This constitutes an increase in supply to 840 seats per day. The level of demand will increase to 480 passengers per day, which represents a response of 20% (or 75 passengers)

---

<sup>22</sup> Technically, this figure is not the true profit, since system operating costs are not included in the cost base. More appropriately, this is the contribution to system overhead, but will loosely be called profit in this analysis.

due to the utility curve shift, plus an additional 28 passengers due to the increase in frequency. The load factor coincidentally remains at 57.1%, and the airline's profit increases to \$5,000.

### 3.2.3.3 Effect of an Increase in Fare

Consider the case in which the regulatory body has granted the carrier a fare increase of \$10 per trip, raising the price to \$60. Since this market is price elastic ( $\alpha = -1.3$ ), the effect will obviously be a marked decrease in demand. If the airline retains the same schedule as was offered at the last equilibrium point, one small and four large aircraft departures, the demand would drop to 379 passengers per day, load factor would decrease to 45.1%, and profit would fall to \$3,740. This result is consistent with classical microeconomic theory that states that if prices are increased in a price elastic market, then revenue (and hence profit in this case, since cost is held constant) will decline.

However, if the airline reoptimizes its schedule, it will offer four flights per day, two each of the 120 and 180 seat aircraft, resulting in supply and demand levels of 600 and 357 trips per day, a load factor of 59.5%, and a profit of \$7,420. Therefore, the interesting case exists of a fare increase in a price elastic market, resulting, due to a corresponding shift in the production function, in a profit increase. This unusual phenomenon is due primarily to the rigid maximum load factor constraint and the fact that the carrier's product mix options are finite and discrete. Any change in the demand pattern (in this example, it was due

to a fare increase) may allow the airline to offer a different mix that will increase its load factor while not violating the maximum load factor constraint.

#### 3.2.3.4 The Decision of Whether or Not to Offer Wide Bodied Service

Since there appears to be an economy of scale, as was mentioned in Section 3.2.2, of cost per available seat, it may be profitable for the airline, which at the current equilibrium point is offering four narrow body aircraft departures per day, to consolidate some of this service into fewer departures with wide bodied equipment. The advantage of this would be that a similar level of supply (number of trips) would be offered at a lower cost. The disadvantage would be a loss in revenue due to a lower level of demand caused by the reduction in frequency.

Suppose that the airline has an option to introduce into service one or two aircraft of type  $a_3$ , which have a seating capacity of 360 and a per flight cost of \$6,000. The per seat cost is \$16.67, compared to \$25.00 and \$22.22 in the cases of  $a_1$  and  $a_2$ . The question that they face is if the reduced revenue would be offset by the reduction in cost of offering the wide bodied service. The answer to this question, given the market parameters for this example, is no. The current equilibrium level of output, two small and two large narrow body aircraft departures yielding a profit of \$7,420, is still optimal. This is another interesting case in which a low cost technology should not be introduced to the market.

The best solution using some wide bodied service would be to replace the two large narrow body aircraft departures with one flight of the wide body. This would reduce average load factor from 59.5% to 53.4%, resulting in a profit of \$7,260. The extreme case of replacing all four current departures with two wide bodied aircraft departures would reduce load factor to 36.2%, resulting in a profit of \$3,660. Table 3.2 lists the eight best options available to the airline for this example.

### 3.3 The Market as a Subset of an Integrated Network

In Section 3.2, the basic air passenger market was considered to be an isolated entity. All service offered between regions A and B was there to accommodate only the origin to destination traffic in that market, and the carrier's only concern was to maximize its "profit" by optimizing its output in that particular segment. In reality, the domestic air passenger markets are linked by a network of segments known as the airlines' route structure. Any given link serves a plurality of markets, as was indicated in Figure 3.2, and markets are served by a number of different links.

The generalization of a market as being a subset of a complex network, rather than as a single isolation, requires certain minor adjustments to the demand equation developed in the previous section and major restructuring of the supply conditions. Simpson explains:

Table 3.2 Eight Best Schedule Options

<u>Option</u>	$n_{a_1}$ <u>(120 seats)</u>	$n_{a_2}$ <u>(180 seats)</u>	$n_{a_3}$ <u>(360 seats)</u>	$Q_S$	$Q_D$	<u>LF</u>	<u>Profit</u>
1	2	2	0	600	357	59.5%	\$7,420
2 (tie)	0	3	0	540	321	59.4%	\$7,260
2 (tie)	2	0	1	600	321	53.4%	\$7,260
4	4	1	0	660	379	57.4%	\$6,740
5	1	0	1	480	261	54.3%	\$6,660
6 (tie)	1	3	0	660	357	54.1%	\$6,420
6 (tie)	3	0	1	720	357	49.5%	\$6,420
8	1	1	1	660	321	48.6%	\$6,260

"Unlike the demand function which is defined for a [region pair] market, the supply function describes the behavior of suppliers who are operating a scheduled transportation system over a network which simultaneously serves a large set of markets. Thus the various suppliers in a given... market may be serving it jointly with quite different sets of other markets, and so their behavior and other competitive positions may be quite different. Each supplier is optimizing over his network of services, and may not be optimizing in a given market." <sup>23</sup>

The purpose of this section is to indicate the additional complexities of the economic nature of air passenger service in the general case of a market within a complex network structure.

### 3.3.1 Analysis of Demand

In the analysis of the isolated market in Section 3.2, the demand for air transportation services in a given region pair market in a given day,  $Q_D$ , was defined by equation (3.5) as follows:

$$Q_D = N \left\{ k \in K \mid \bar{U}_k^v(D_i) > U_k^m(W_k) - U_k^m(W_k - F) + U_k^t(t) \text{ for some } i \in I \right\} \quad (3.5)$$

This relationship states that the number of people who fly from A to B on a given day is the number of people in region A for whom the utility of a visit to region B via any scheduled departure  $i$  is greater than the marginal utility loss of the fare plus the utility of the flight time.

The only adjustment to be made in the general case of a market within

---

<sup>23</sup> Simpson, op. cit., p. 25.

a network is to account for the fact that the schedule of services offered may no longer ideally consist of a set of  $N\{I\}$  nonstop jet departures. The flights offered in a general market may consist of nonstop, multistop direct, and connecting departures. Furthermore, these departures are not necessarily distributed at convenient times over the traveling day, so displacement time may in general become a major deterrent to demand generation. The trip time component may be further generalized to include the possibility of slower cruise speed aircraft and/or expected delays in line haul travel time due to congestion, so that even in the event of nonstop departures, the block flight time may be something greater than  $t_0$ .

Each of the above considerations will in Chapter IV be lumped into a single proxy called "level of service", LOS. Whereas in the case of the isolated market where the demand relationship is as follows:

$$Q_D = f(Q_{D_0}, F, N\{I\}) \quad (3.16)$$

the general market demand relationship becomes

$$Q_D = f(Q_{D_0}, F, LOS) \quad (3.17)$$

Between markets, the total demand,  $Q_{D_0}$ , will depend upon certain socio-economic characteristics of the regions. Certainly the ability to generate passenger trips from a region will be a function of the population and perhaps of the income distribution of a given region. The ability to attract visits will be a function of a region's industrial, recreational, and cultural activities, among other factors. Combining each of the

above considerations into a socio-economic proxy, SE, as a determinant of total demand, the general functional relationship for the demand for air passenger services in a given market may be expressed as follows:

$$Q_D = f_1 (SE, F, LOS) \quad (3.18)$$

### 3.3.2 Analysis of Supply and Equilibrium

As was previously mentioned, the generalization of the supply side is considerably more complex than that of demand in the analysis of a market embedded in a network, rather than isolated. The fundamental reason for this is, as was explained, that each carrier is attempting to optimize its product mix over that subset of the total route structure to which it is certified to provide service, and not necessarily over a single given market.

One result of this complexity is that service offered in a given segment may be provided for the benefit of a multitude of markets. An extreme example of this is the Birmingham-Atlanta segment, in which there are currently sixteen nonstop daily departures in each direction.<sup>24</sup> Certainly the quantity of origin to destination demand in this market would be adequately accommodated by a substantially lower frequency. This high level of service is provided to feed into the complex at Atlanta, whereby a passenger desiring to travel from Birmingham to virtually anywhere else in the world would fly to Atlanta and connect outward. Therefore, if one were

---

<sup>24</sup> Official Airline Guide, February 15, 1977, p. 87.



to survey the people aboard a flight in that segment, it would probably be discovered that a small proportion of the passengers were Birmingham-Atlanta origin to destination traffic. In fact, during the third quarter of 1975, there were 4082 seats per day flown between Birmingham and Atlanta,<sup>25</sup> while local traffic averaged only 251 passengers per day,<sup>26</sup> roughly 6.2% of the available seats.

A second and quite different result is that, in a given market, a wide variety of service quality may exist. An extreme example of this is that the Official Airline Guide (OAG) currently lists 84 daily departure alternatives from Los Angeles to New York.<sup>27</sup> Thirty of these are direct flights, including thirteen nonstops, five one-stops, six two-stops, five three-stops, and one four-stop. Of the fifty-four connecting flights, forty are online and fourteen are interline. The connection cities number thirteen, including Chicago, Dallas, St. Louis, Minneapolis, Houston, Phoenix, Kansas City, Cleveland, Cincinnati, Pittsburgh, Denver, New Orleans, and Atlanta. Most of these scheduled departures, particularly those of the poorer level of service, are not offered to explicitly serve the Los Angeles to New York demand, but are for intermediate markets. They are listed in the OAG merely because they happen to exist and under unusual circumstances could be selected by passengers in the transcontinental market. This is a major factor that makes  $Q_S$  (available seats) and LF

---

<sup>25</sup> Official Airline Guide, September 1, 1977, p. 91 and p. 148.

<sup>26</sup> Civil Aeronautics Board, Origin-Destination Survey of Airline Passenger Traffic, Table 8, Third Quarter, 1975, p. 77.

<sup>27</sup> Official Airline Guide, February 15, 1977, p. 748.

(load factor) quite meaningless when considering markets in a network, as will be elaborated upon later in this section.

An additional situation that arises due to the network structure is the appearance of wide bodied service in very short haul markets. For example, American Airlines operates a DC-10 daily from Los Angeles to San Diego, a distance of 109 miles.<sup>28</sup> The purpose for this seemingly inefficient utilization of resources is that this flight (flight 11) originates in Boston and the "tag end" link to San Diego is not provided to capture the Los Angeles to San Diego passengers, but rather to attract the high fare paying Boston to San Diego passengers.

The following supply and demand equations, developed in Section 3.2, are not appropriate in the analysis of a market embedded in a network:

$$Q_S = \sum_a n_a \cdot S_a \quad (3.13)$$

and 
$$Q_D = Q_S \cdot LF \quad (3.14)$$

The supply quantity,  $Q_S$ , the number of seat departures offered in a given market, is not an adequate measure of supply. The 84 flight alternatives from Los Angeles to New York comprise something on the order of fifteen to twenty thousand seat departures. However, it is very unclear how many of these are provided for this particular market. Furthermore, if twenty of the connecting flights were deleted from the schedule (a reduction of,

---

<sup>28</sup> Ibid., p. 982.

say, 4,000 seats), the incremental decline in total service in this market would be negligible. However, if ten of the nonstop flights were deleted (a loss of about 2,000 seats), the service reduction would be substantial.

The load factor term, LF, is likewise of no particular value in market analysis, since it is primarily either a segment or a system related figure. For example, in the Birmingham to Atlanta segment the load factors may be quite high. However, the ratio of the Birmingham to Atlanta market demand (number of origin to destination passengers) to the number of available seat departures may be quite small. (As previously mentioned, the latter was 6.2% for the third quarter of 1975.)

Since the production function of an airline is actually a scheduling process over a large network, the INLP developed for the isolated market is far too simplistic to be generally useful. Furthermore, since the purpose of this thesis is to construct a demand model, the suppliers' optimization problem will not be developed herein. This problem has been researched, resulting in a set of "fleet planning" and "fleet assignment" models. An example cited in Chapter I of a fairly comprehensive fleet assignment model is FA-4, developed by the Flight Transportation Laboratory at M.I.T.<sup>29</sup>

One of the generalizations that may be drawn from the analysis of an isolated market is that the amount of service offered is dictated by the

---

<sup>29</sup> Swan, op. cit.

total potential demand. This results in the expectation of a greater level of service in a high density segment (e.g., Chicago-New York) than in a low density segment of equivalent distance (e.g., Bangor-Akron/Canton). This effect is present in time series analysis as well as in cross sectional. The example of Section 3.2.3.2 indicated how an increase in total demand within a market can lead to an increase in level of service. Due to the fact that an increase in total demand may not immediately be perceived as such, and that airline schedules are generally fixed for a period of time, the response in service may tend to lag behind changes in total demand.

A second important generalization from the analysis of an isolated market is that service is responsive to fare changes. The example in Section 3.2.3.3 indicated how a fare increase may affect a single market in time series, due to its impact upon demand and the reoptimization of the carrier's fleet assignment problem. In the network setting, this effect may be very pronounced, due to the variability between markets of the price elasticity. Since the profitability of the more price elastic markets may substantially increase, the reshuffling of the airline's fleet assignment to the new equilibrium point conceivably may result in significant alterations of level of service in many markets.

The effect of lack of competition in airline markets has historically been that the monopolistic or nearly monopolistic carrier has not had to overly concern itself with providing outstanding service to its already captive market. At the other end of this continuum, in potentially profitable competitive segments, the carriers have engaged themselves in

fierce scheduling wars which result in extraordinarily high levels of service to the consumers.

The example of the Birmingham to Atlanta segment containing sixteen nonstop departures per day exemplifies the fact that high levels of service may exist in certain markets for reasons other than to serve the origin to destination passenger demand. Segments such as these usually feed into a large connecting complex (in this case Atlanta). Therefore, the location of a given market with respect to the total U.S. domestic and international route structures will further affect the level of service offered.

By combining the above factors, a quasi-supply function can be stated as follows:

$$\text{LOS} = f_2 (Q_D^*, F, \text{COMP}, \text{RS}) \quad (3.19)$$

where

LOS = proxy for level of service

$Q_D^*$  = lagged origin to destination demand

F = fare

COMP = some measure of the competitive nature of the market

RS = some measure of the market's location with respect to the total U.S. domestic and international route structures

This relationship will hereafter be referred to as the "service function" for a given market. The primary motivation for a service function is the simultaneity between  $Q_D$  and LOS. As will be explained in greater detail in Section IV, this two-way causality renders a single equation

model inappropriate.

Combining equations (3.18) and (3.19) yields the following two equation system which will be specified in more detail in Chapter IV and calibrated for different categories of markets in Chapter V:

$$\text{Demand function: } Q_D = f_1 (SE, F, LOS)$$

$$\text{Service function: } LOS = f_2 (Q_D^*, F, COMP, RS)$$

IV. Description of the Functional Forms, Market Variables,  
and Experimental Design

Chapter III contained an overview of the economic issues that have an impact upon the levels of demand and service in an air transportation market. The end product of Chapter III was the following two-equation model of air transportation activity within a given market:

$$\text{Demand equation} \quad Q_D = f_1(\text{LOS}, F, \text{SE}) \quad (4.1)$$

$$\text{Service equation} \quad \text{LOS} = f_2(Q_D^*, F, \text{COMP}, \text{RS}) \quad (4.2)$$

where

$Q_D$  = origin to destination passenger demand

LOS = level of service

F = fare

SE = socio-economic activity

COMP = level of competition

RS = location within route structure

In this chapter the functional forms of equations (4.1) and (4.2) will be specified. A discussion and a resulting flow diagram will elaborate upon how the individual variables interact within the U.S. domestic air

---

\* The demand variable is lagged in the service equation

transportation system. Each of the variables will then be defined in detail and, where necessary, examples of their calculations will be provided. Finally, a description of the experimental design and sampling procedure is presented.

#### 4.1 Functional Forms of the Demand and Service Equations

The specification of the demand equation is as follows:

$$Q_{D_t} = \beta'_{10} LOS_t^{\beta_{11}} F_t^{\beta_{12}} SE_t^{\beta_{13}} \epsilon'_1 \quad (4.3)$$

where

- $Q_D$  = origin to destination passenger demand
- $LOS$  = level of service
- $F$  = fare
- $SE$  = socio-economic activity

The rationale behind this specification is, as described in Chapter III, that in any time period  $t$  there exists a total potential local demand,  $Q_{D_0}$ , in a region pair market which is determined by socio-economic factors (populations, incomes, amount of recreational facilities, etc.). The flow of this total potential demand is impeded by positive fare levels and less than perfect level of service.

A multiplicative (instead of, say, an additive) form was selected for two reasons. With respect to level of service, this specification satisfies



the necessary boundary conditions in that, if no service were offered (LOS=0) there would be no traffic, and if perfect service were offered (LOS=1) the local demand would be finite ( $Q_D=Q_{D_F}$ , as defined in Section 3.2.1). With respect to fare, the multiplicative, more specifically the log-linear, form was selected to allow for the estimation of the various price elasticities (see footnote 21 in Section 3.2.3). Since it is assumed that  $\beta_{12}$  is negative, if fare values go to zero this specification implies that demand will go to infinity, which is in violation of a boundary condition set forth in Section 3.2.1. However, since the model considers only positive fare values, both in calibration and prediction, this violation is of no consequence.

Since a mutual causality exists between demand and level of service, the demand model as shown cannot appropriately be calibrated using ordinary least squares estimation. In an effort to rectify this problem, a second equation, the "service equation", is developed. The specification of the service equation is as follows:

$$LOS_t = \beta'_{20} (Q_{D_{t-1}} + RS_{t-1})^{\beta_{21}} F_t^{\beta_{22}} COMP_t^{\beta_{23}} \epsilon'_{2} \quad (4.4)$$

The level of service on a given route segment is determined not only by the level of local demand, but also by the level of non-local passenger flow over the segment. The Birmingham to Atlanta route segment example cited in Section 3.3.2 is a classic example of a market in which the amount of service offered is far in excess of what the local demand requires. Therefore, LOS is specified in the service equation as being a function of the total traffic over a route segment.

The total segment traffic ( $Q_D + RS$ ) is lagged in the service equation for two reasons. It is not unreasonable to assume that if traffic (whether local or otherwise) were to increase or decrease in a given route segment, the airlines' response (improving or reducing service) would not be immediate. There would probably be a lag due to the time lapse before the carriers perceive the change in traffic as being significant, and since schedules are normally altered only twice per year, there would certainly be a lag before they could operationalize the schedule change. The second reason for the lag is statistical. The lagged variables are "predetermined", and therefore the simultaneity condition present in the demand equation does not exist in the service equation, and ordinary least squares estimation is appropriate.

The fare variable is included in the service equation to account for the fact that if fares were to increase, it would economically be in the best interests of the suppliers to increase service. This follows from Simpson's<sup>1</sup> and Marfisi's<sup>2</sup> theories and from the theoretical development of Chapter III (specifically, the example in Section 3.2.3.3).

The competition variable has been included in the service equation to account for the commonly held belief that more competition stimulates improved service. This phenomenon was discussed in some detail in

---

<sup>1</sup> Robert W. Simpson, "A Theory for Domestic Airline Economics", unpublished lecture notes for the M.I.T. course 16.74 -- Air Transportation Economics, Fall 1974, p. 68.

<sup>2</sup> E. Pat Marfisi, "Theory and Evidence on the Behavior of Airline Firms Facing Uncertain Demand (Ph.D. dissertation, Brown University, 1976), p. 3-2.

### Section 3.3.2.

Assuming that the specification of the service equation is valid, the predicted values of level of service,  $\hat{L}OS$ , will be highly correlated with the observed values of LOS. Furthermore, since  $\hat{L}OS$  is a (log) linear combination of predetermined variables, it should be uncorrelated with  $\epsilon'_1$  in the demand equation. Therefore,  $\hat{L}OS$  should serve as valid instrumental variable in the demand equation. Substituting  $\hat{L}OS$  for LOS in the demand equation renders ordinary least squares estimation appropriate.

Figure 4.1 is a schematic representation of the interaction of the variables in the demand and service equations.

## 4.2 Description of the Market Variables

### 4.2.1 Demand ( $Q_D$ )

The selected variable for the measure of air passenger traffic activity in a given region pair market is the number of passengers in a given time period that originate in one region and fly to the other region for purposes other than to make a connection to a third region. This variable is declared the true origin to destination passenger traffic, using the passenger intent criterion. The best source for these data is Table 8 of the Civil Aeronautics Board's Origin to Destination Survey.

An unfortunate limitation of employing Table 8 data is that the decision rules selected by the Board for tabulation do not in all cases accurately reflect passenger intent. The net result is that Table 8 data

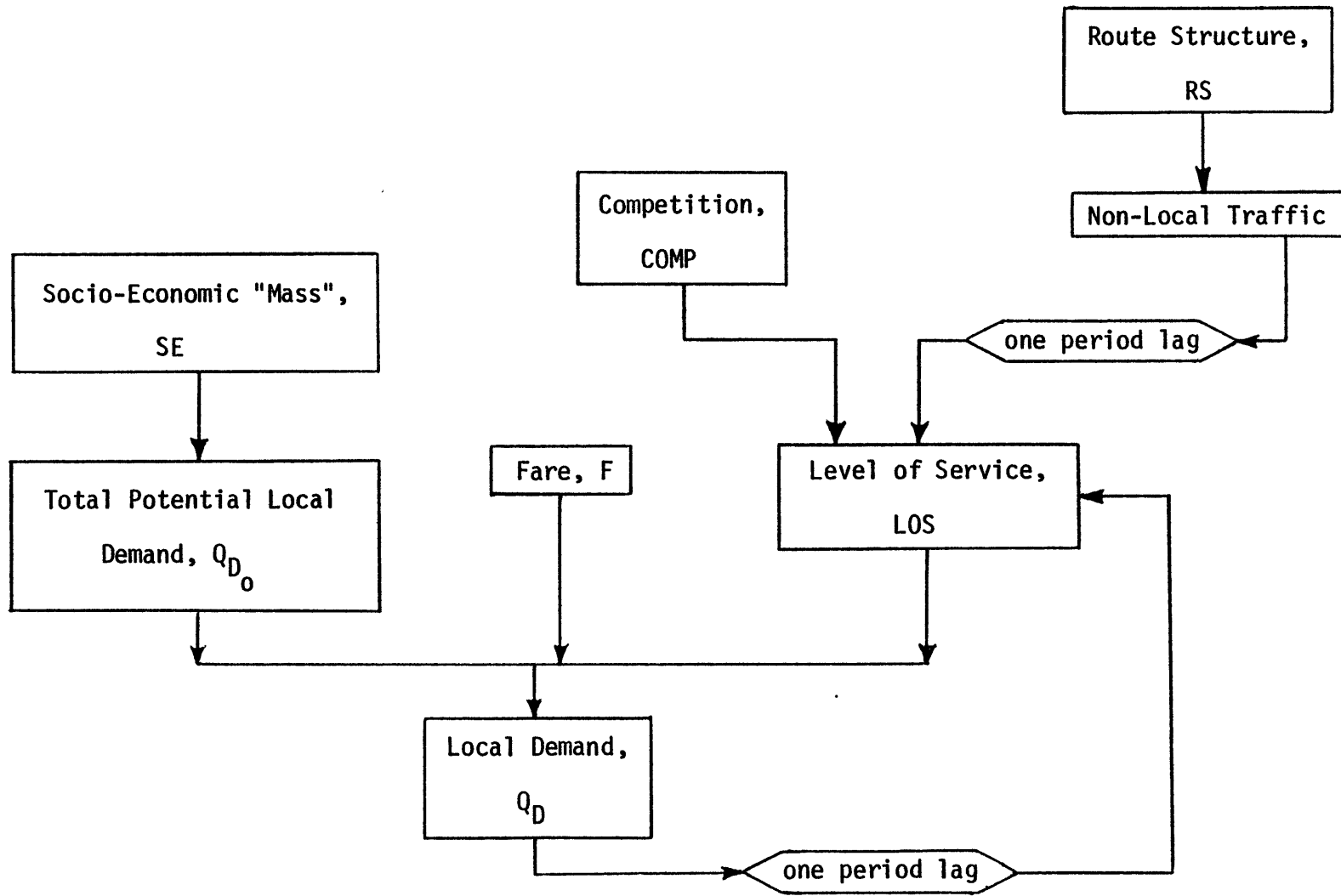


Figure 4.1 Flow Diagram of Interaction of the Variables

have a tendency to be biased by understating true origin to destination traffic flows in long haul markets and overstating them in short haul markets.<sup>3</sup> However, since the bias is slight and unmeasurable, it is assumed to be negligible for the purpose of calibrating this model.

#### 4.2.2 Level of Service (LOS)

An important performance measure to be included in the economic modeling of air transportation services within a given region pair passenger market is the availability of scheduled flights at times when the prospective customers wish to fly. As discussed in Chapter II, many existing models consider only the number of flights offered per day as an indication of availability (e.g., Marfisi<sup>4</sup>). What these models do not consider is the time of day when these departures occur. Time of day not only relates to the needs of the passengers (the consumer value of a departure at 2:30 a.m. may be quite different from that of a departure at 5:30 p.m.), but also to the relationship between number of flights and capacity per flight. (Are three 120-seat aircraft departures at the same time really, in practical terms, three separate consumer alternatives, or the equivalent of one departure of a 360-seat aircraft?)

An additional performance measure frequently overlooked in demand modeling is the type of service offered in a region pair. This quality of service measure, if it is considered at all, is found to be difficult to

---

<sup>3</sup> O'Brien, Frederick J., "C.A.B. Origin-Destination Statistics" (unpublished note, Lockheed-California Company, 1976).

<sup>4</sup> Marfisi, op. cit.

quantify. A Civil Aeronautics Board staff study attempted to address this problem by assigning weights to the different types of service.<sup>5</sup> The study concluded that a two-stop flight is equivalent in consumer value to 0.40 nonstop flights, a one-stop flight is equivalent to 0.55 nonstop flights, etc. This proportionality approach is, however, unrealistic because the weightings are assumed to be independent of stage length. One intermediate stop may be nearly double the block time of a short-haul flight, whereas one stop may increase the block time of a transcontinental flight by merely fifteen to twenty percent. Thus, the proportionality of the penalty paid by intermediate stops decreases as the stage length increases.

The level of service index created herein is developed to systematically account for the above mentioned issues. Basically, the measure is a dimensionless number scaled from zero to one which represents the ratio of the nonstop jet flight time to the average total passenger trip time. The total trip time is the sum of the actual flight time (including stops and connections) and the amount of time the passenger is displaced from when he wishes to fly due to schedule inconveniences.

If "perfect" service were offered in a given region pair (a nonstop jet departing at every instant of the day), there would be no such displacement. The total trip time would be merely the nonstop jet flight time, and the ratio (level of service measure) would be unity. If poor service were

---

<sup>5</sup> C.A.B. Bureau of Operating Rights, "Effect on Total Market Traffic of Changes in Quality of Service (QSI): Explanatory Statement" (Washington: Civil Aeronautics Board Docket 21136).

offered (few flights, multistops, connections, slower aircraft, etc.), not only would block flight time be substantially greater than non-stop jet flight time, but many passengers would be forced to fly at inconvenient times. This inconvenience would be accounted for by the inclusion of significant "displacement" times, and the resulting level of service ratio would be small.

#### 4.2.2.1 Behavioral Assumptions

The basic assumed behavioral pattern in the development of the level of service index is that a generic passenger in accordance with the purpose of his trip will predetermine an optimal or preferred time of departure from the origin airport. Given that he is aware of his preferred departure time and is presented a schedule of available flights, he will then select that flight which minimizes the sum of the "displacement time" and the "adjusted flight time". The displacement time is the absolute value of the difference between the scheduled departure time and the preferred time of departure. The adjusted flight time is defined to be the scheduled flight time (including intermediate stops) for direct flights, the scheduled flight time plus one-half hour for online connections, and the scheduled flight time plus one hour for interline connections.

The motivation for inclusion of the additional time assessment for connecting flights is that the consumer disutility of a connecting flight is greater than merely the increase in flight time. For an online connection, the passenger faces the chance of a broken connection due to a

late arrival of the first leg or cancellation of the second. Also, the passenger is burdened with the inconvenience of having to physically change aircraft. For an interline connection, the passenger faces not only the possibility of a broken connection, but also a greater chance of having his baggage miss the connection. In addition, he faces the burden of not only having to change aircraft but also is frequently forced to walk to a different terminal.

Table 4.1 stratifies four hierarchal levels of types of service, based upon airline scheduling and marketing experience.<sup>6</sup> The table indicates that an online nonstop/nonstop connection, which actually requires only one stop, is equivalent in consumer value to a two-stop direct flight. Hence, the presence of a connection within the same airline is the consumer equivalent of adding an additional intermediate stop. By the same argument it can be inferred that an interline connection bears the equivalent disutility of two additional stops. Assessing an additional one-half hour of effective flight time for each equivalent stop yields the above-mentioned adjustments of one-half hour and one hour for online and interline connections, respectively.

A basic assumption is that the loss function for arrival time displacement is linear and symmetric. In other words, the disutility incurred by being displaced by  $p$  hours is  $p$  times the loss incurred by being displaced one hour. Furthermore, the symmetry of the loss function assumes that the cost of departing late by  $p$  hours is equivalent to the cost of departing early by  $p$  hours.

---

<sup>6</sup> Interview with Frederick J. O'Brien, Lockheed-California Company, Burbank, California, 5 March 1976.



Table 4.1 Four Levels of Equivalent Air Service

<u>Level</u>	<u>Direct</u>	<u>Connecting</u>
1	Nonstop	--
2	One-stop	--
3	Two-stop	Online Nonstop/Nonstop
4	Three-stop	Interline Nonstop/Nonstop

Total trip time, as defined in this analysis, is quite different than what it is commonly referred to in transportation analysis. In addition to merely waiting (or displacement) and line haul travel time, total trip time usually includes access and egress times to and from the line haul terminals. This inclusion is particularly significant in air transportation, in which an airport commonly serves a large geographical region. However, since the purpose of this analysis is to measure the effect of airline scheduling independent of access and egress time, these aspects are not considered.

A further assumption in these models is that of infinite capacity. A passenger who elects (by the governing behavioral assumptions) to board a particular flight may do so without a change of its being overbooked; therefore, load factor is not a consideration in this analysis. This assumption can be justified in light of the fact that generally, if a particular flight is consistently overbooked, the airline(s) serving that market will add capacity or additional service near that particular time of day. Therefore, in most instances, overbooking problems are corrected within a reasonable length of time.

#### 4.2.2.2 Development of the Index

Given the behavioral assumptions described in the preceding section and a published flight schedule for one direction of a particular region pair, the total trip time, defined as the sum of the displacement time plus the adjusted flight time, for a passenger desiring to depart at any

particular time of day can be determined. Then, given a distribution of passenger departure demand over the entire day, the average total trip time, weighted by this distribution, can be generated.

In order to compute the average total trip time, clock time has been divided into a finite number of discrete time points which are separated by equal intervals throughout the traveling day. The time length of these intervals (and hence the number of time points) may be arbitrarily set (perhaps 15, 30, or 60 minutes). The analysis is performed by considering passengers desiring to depart at only these time points rather than continuously. Therefore, the smaller these intervals (greater number of time points) are, the less restricting this approximation will be. However, as the number of time points increases, so does the computation time. Throughout this analysis, the traveling day will usually be separated into 41 time points separated by thirty minutes starting at 4:00 a.m. and ending at midnight. However, for certain markets bearing unusual demand characteristics, this convention may be altered.

The following notation is defined:

- $n$  = number of time points (equally separated) in the traveling day
- $j$  = index of time points  $j = 1$  (start of traveling day),  
2, ...,  $n$  (end of traveling day)
- $t_j$  = time of day at time point  $j$
- $\pi_j$  = proportion of daily passengers preferring to depart at  
time point  $j$
- $m$  = number of daily flights
- $i$  = index of flights  $i = 1, 2, \dots, m$

$$\begin{aligned}
 D_i &= \text{local departure time of flight } i \\
 A_i &= \text{local arrival time of flight } i \\
 Z &= \text{number of time zones crossed (positive if west to east,} \\
 &\quad \text{negative if east to west)} \\
 \gamma_i &= \text{connection adjustment for flight } i = \begin{cases} 0.0 & \text{for direct flights} \\ 0.5 & \text{for online flights} \\ 1.0 & \text{for interline connections} \end{cases}
 \end{aligned}$$

The adjusted flight time for any flight  $i$ ,  $AFT_i$ , is the difference between the arrival and departure times,  $A_i - D_i$ , minus the time zone change,  $Z$ , plus the connection adjustment  $\gamma_i$ .

$$AFT_i = A_i - D_i - Z + \gamma_i \quad (4.5)$$

The displacement time for any passenger preferring to depart at time point  $j$  and selecting flight  $i$ ,  $DT_{ji}$ , is the absolute value of the difference between the departure time of flight  $i$ ,  $D_i$ , and the time of day at time point  $j$ ,  $t_j$ .

$$DT_{ji} = |D_i - t_j| \quad (4.6)$$

By the governing behavioral assumptions described in the preceding section, a passenger preferring to depart at time point  $j$  will select that flight which will minimize the sum of displacement time plus adjusted flight time. This minimized sum is his total trip time  $TT_j$ .

$$TT_j = \min_i (DT_{ji} + AFT_i) = \min_i (|D_i - t_j| + A_i - D_i - Z + \gamma_i) \quad (4.7)$$

The average total trip time,  $\bar{t}$ , is the weighted (by the  $\pi_j$  factors) average of the total trip times of the passengers who prefer to depart at each of the  $n$  time points over the traveling day.

$$\bar{t} = \frac{\sum_{j=1}^n \pi_j TT_j}{\sum_{j=1}^n \pi_j} = \frac{\sum_{j=1}^n \pi_j \min_i (|D_i - t_j| + A_i - D_i - Z + \gamma_i)}{\sum_{j=1}^n \pi_j} \quad (4.8)$$

The level of service index, LOS, is defined as the ratio of the nonstop jet time,  $t_0$ , to the average total trip time,  $\bar{t}$ .

$$LOS = \frac{t_0}{\bar{t}} = t_0 \left[ \frac{\sum_{j=1}^n \pi_j \min_i (|D_i - t_j| + A_i - D_i - Z + \gamma_i)}{\sum_{j=1}^n \pi_j} \right]^{-1} \quad (4.9)$$

#### 4.2.2.3 Determination of Nonstop Jet Time

A necessary component for the computation of the level of service index, LOS, for a given region pair is the nonstop jet time,  $t_0$ . Nonstop jet service is offered in many of the markets selected in the sample for this analysis, and so for these markets this value can be readily determined by examination of the flight schedules. However, for those markets in which nonstop jet service is not offered, a procedure for estimating this value is required.

The following relationship has been hypothesized and calibrated:

$$t_0 = \beta_0 + \beta_1 d + \beta_2 (LO_A - LO_D) + \epsilon \quad (4.10)$$

where

$t_o$  = nonstop jet flight time (hours) between two regions

$d$  = nonstop distance (miles) between the major airports of the two regions

$LO_A$  = decimal equivalent of the longitude of the major airport in the region of arrival

$LO_D$  = decimal equivalent of the longitude of the major airport in the region of departure

The result of the parameter estimation is as follows:

$$t_o = 0.3370 + 0.001976d + 0.009369 (LO_A - LO_D) \quad (4.11)$$

The constant term represents the startup time involved in a flight (taxiing, accelerating to cruise speed, etc.). The second term was included to account for the obvious fact that the trip time is a linear function of distance. The third term measures the effect of the prevailing west to east air flow, resulting in the fact that it requires roughly one additional hour to fly a commercial jet across the country east to west than it takes flying west to east.

The estimation procedure used was least squares multiple regression analysis. The observations consisted of all region pairs selected in the sample for this analysis (see Section 4.3) that offered nonstop jet service in December, 1975. Since in markets where nonstop jet service is offered in one direction it usually is offered in both directions, the correlation between the independent variables was virtually zero. The total sample

size was 213 observations. Some key regression analysis statistics are indicated below:

$$t_o = 0.3370 + 0.001976d + 0.009369 (LO_A - LO_D) \quad (4.12)$$

(111.60)            (13.03)

$$R^2 = 0.984$$

$$n = 213$$

(The figures in parentheses are the corresponding t ratios)

#### 4.2.2.4 Determination of Time of Day Demand Functions

An additional input variable necessary for the computation of the level of service index, LOS, for a given directed region pair is the relative demand for air transportation service as a function of time of day. A uniform time of day distribution is of course rarely, if ever, observed. For example, the daily demand for air transportation in short and medium haul business markets is typically bimodal. There is a peak period between 8:00 and 10:00 a.m. and another between 5:00 and 7:00 p.m. Other markets may observe quite different time of day variations. In transcontinental west to east markets, there actually is a lull in what one would normally expect to be a peak period in the late afternoon. This is caused by the fact that few passengers would choose to arrive at the destination (east coast) at two or three o'clock in the morning. The demand, however, picks up considerably in the late evening for the night flights which arrive on the east coast between eight and ten o'clock the next morning.

The time of day demand distribution unfortunately is, for nearly all markets, virtually impossible to determine. When passengers do fly is, in most cases, a function of the air transportation schedule by which they are constrained. However, data has been provided by Eastern Airlines from their New York/Boston shuttle, a demand-responsive service, which reflects as would be expected the bimodal time of day distribution described above. This distribution is plotted in Figure 4.2 and is used as a basis for deriving theoretical time of day demand distributions for all of the directed region pairs in the sample.

The initial step in this analysis is to discretize time of day into forty-one time points ( $j = 1, 2, \dots, 41$ ) at half hour intervals starting at 4:00 a.m. and ending at 12:00 midnight [ $t(1) = 4.0, t(2) = 4.5, \dots, t(41) = 24.0$ ]. At each time point,  $j$ , a proportion  $p(j)$  of the total number of daily passengers desire to depart from Boston to New York or vice versa, as indicated by the empirical data provided by Eastern and tabulated in Table 4.2.

A basic assumption in these derivations is that the proportion  $p(j)$  of the total daily passengers desire to depart at time  $t(j)$  for one of two reasons:

- (1) The time of day  $t(j)$  is a preferred time to depart
- or (2) The time of day  $t(j + 2)$  is an attractive time to arrive.

The arrival time  $t(j + 2)$  is employed because the time points,  $j$ , are separated by half hour intervals, and one hour is the approximate flight time between Boston and New York.



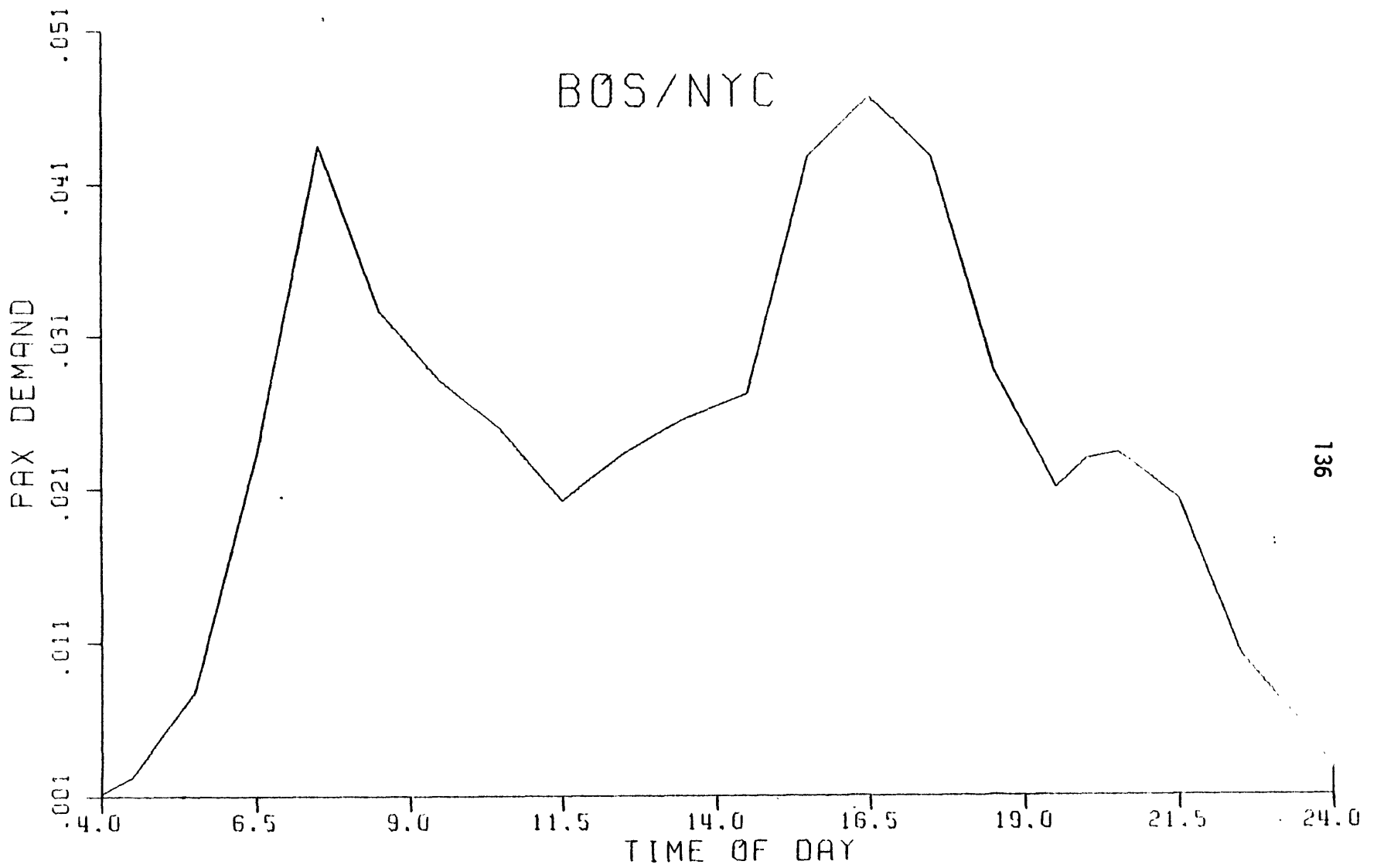


Figure 4.2 Empirical Time of Day Demand Distribution for Eastern Airlines' Boston/New York Air Shuttle

Table 4.2 Empirical Time of Day Demand Distribution for Eastern Airlines  
Boston/New York Air Shuttle

<u>j</u>	<u>t(j)</u>	<u>p(j)</u>	<u>j</u>	<u>t(j)</u>	<u>p(j)</u>	<u>j</u>	<u>t(j)</u>	<u>p(j)</u>
1	4.0	0.001	15	11.0	0.023	29	18.0	0.036
2	4.5	0.002	16	11.5	0.020	30	18.5	0.029
3	5.0	0.005	17	12.0	0.022	31	19.0	0.025
4	5.5	0.008	18	12.5	0.023	32	19.5	0.021
5	6.0	0.016	19	13.0	0.025	33	20.0	0.023
6	6.5	0.023	20	13.5	0.026	34	20.5	0.023
7	7.0	0.033	21	14.0	0.026	35	21.0	0.022
8	7.5	0.044	22	14.5	0.027	36	21.5	0.020
9	8.0	0.038	23	15.0	0.035	37	22.0	0.015
10	8.5	0.033	24	15.5	0.043	38	22.5	0.010
11	9.0	0.030	25	16.0	0.045	39	23.0	0.008
12	9.5	0.028	26	16.5	0.047	40	23.5	0.005
13	10.5	0.026	27	17.0	0.045	41	24.0	0.003
14	10.5	0.025	28	17.5	0.043			

Note:  $\sum_{j=1}^{41} p(j) = 1.00$

In order to project this distribution over all markets, the following two assumptions were made:

(1) The distribution of preferred departure times from any region is  $P_D(j) = p(j)$  for  $j = 1, 2, \dots, 41$

and (2) The distribution of attractive arrival times at any region is:

$P_A(1) = P_A(2) = 0.0$ ,  $P_A = p(j - 2)$  for  $j = 3, 4, \dots, 43$ ,  
where  $t(42) = 12:30$  a.m. and  $t(43) = 1:00$  a.m.

A final assumption in this derivation is that the proportion of daily passengers wishing to depart a given origin for a given destination at time  $t(j)$  is a multiplicative function of the preferability of departure at  $t(j)$ ,  $P_D(j)$ , and the attractiveness,  $P_A(j_{arr})$ , of arriving at the destination at the arrival time,  $t(j_{arr})$ . A multiplicative form was chosen over an additive form after consideration of a typical west to east transcontinental market. Seven o'clock in the evening,  $t(j) = 19.0$ , is, referring to the basic  $t(j)$  distribution, a reasonably preferable time of day for departure. However, a departure from a west coast region for an east coast region at 7:00 p.m. on a nonstop jet would result in an arrival on the east coast at 3:00 a.m. (five hours flying time plus three time zones), which is hardly desirable to anyone. If an additive form were employed, the preferability of departing at 7:00 p.m. would make this flight look desirable, whereas in using the multiplicative form the null attraction of a 3:00 a.m. arrival,  $P_A(j_{arr}) = 0.0$ , will completely eliminate the desirability of this time of departure.

The functional form of  $\pi(j)$  for any given market is as follows:

$$\pi_j = \frac{\sqrt{p(j) \cdot p(j + \partial)}}{\sum_{j=1}^{41} \sqrt{p(j) \cdot p(j + \partial)}} \quad (4.13)$$

where

$$\partial = 2(t_0 + Z) - 2 \text{ (rounded to nearest integer)}$$

$$t_0 = \text{nonstop jet time (hours)}$$

$$Z = \text{number of time zones crossed (positive if west to east, negative if east to west)}$$

The first term in the definition of  $\partial$ ,  $2(t_0 + z)$ , is the local clock time difference, in half hours, between the departure and arrival times of a nonstop jet. The second term,  $-2$ , accounts for the shift in time axis between  $P_D(j)$  and  $P_A(j)$  as mentioned above. The motivation for the radical is that the use of the straight multiplicative form,  $p(j) \cdot p(j + \partial)$ , would not result in the original distribution,  $p(j)$ , for one hour markets such as New York to Boston, where the radical form does. The summation term in the denominator normalizes so that the sum of the  $\pi(j)$  terms over the entire day will equal unity.

Some examples:

$$\begin{aligned} \text{Boston to New York: } t_0 &= 1.0 & Z &= 0 \\ \partial &= 2(t + z) - 2(1.0 + 0) - 2 = 0 \\ \pi_j &= p(j) & j &= 1, 2, \dots, 41 \end{aligned}$$

This results in the original  $p(j)$  distribution, as shown in Figure 4.2.

Chicago to Los Angeles:  $t_0 = 4.0$

$$a = 2(t_0 + Z) - 2(4.0 - 2) - 2 = 2$$

The  $\pi_j$  distribution for Chicago to Los Angeles is shown in Figure 4.3.

Los Angeles to Chicago:  $t_0 = 3.5$                        $Z = 2$

$$a = 2(t_0 + Z) - 2 = 2(3.5 + 2) - 2 = 9$$

The  $\pi_j$  distribution for Los Angeles to Chicago is shown in Figure 4.4.

Boston to San Francisco:  $t_0 = 6.0$                        $Z = -3$

$$a = 2(t_0 + Z) - 2 = 2(6.0 - 3) - 2 = 4$$

The  $\pi_j$  distribution for Boston to San Francisco is shown in Figure 4.5.

San Francisco to Boston:  $t_0 = 5.0$                        $Z = 3$

$$a = 2(t_0 + Z) - 2 + 2(5.0 + 3) - 2 = 14$$

This  $\pi_j$  distribution is shown in Figure 4.6.

#### 4.2.2.5 Examples of Level of Service Calculators

A computer program entitled LOS\_COMP has been written to compute for any given market the values of LOS and the competition variable, COMP. The program is written in PL/I. LOS\_COMP accepts as its input a card containing the region codes, the number of scheduled flights (m), the nonstop jet time

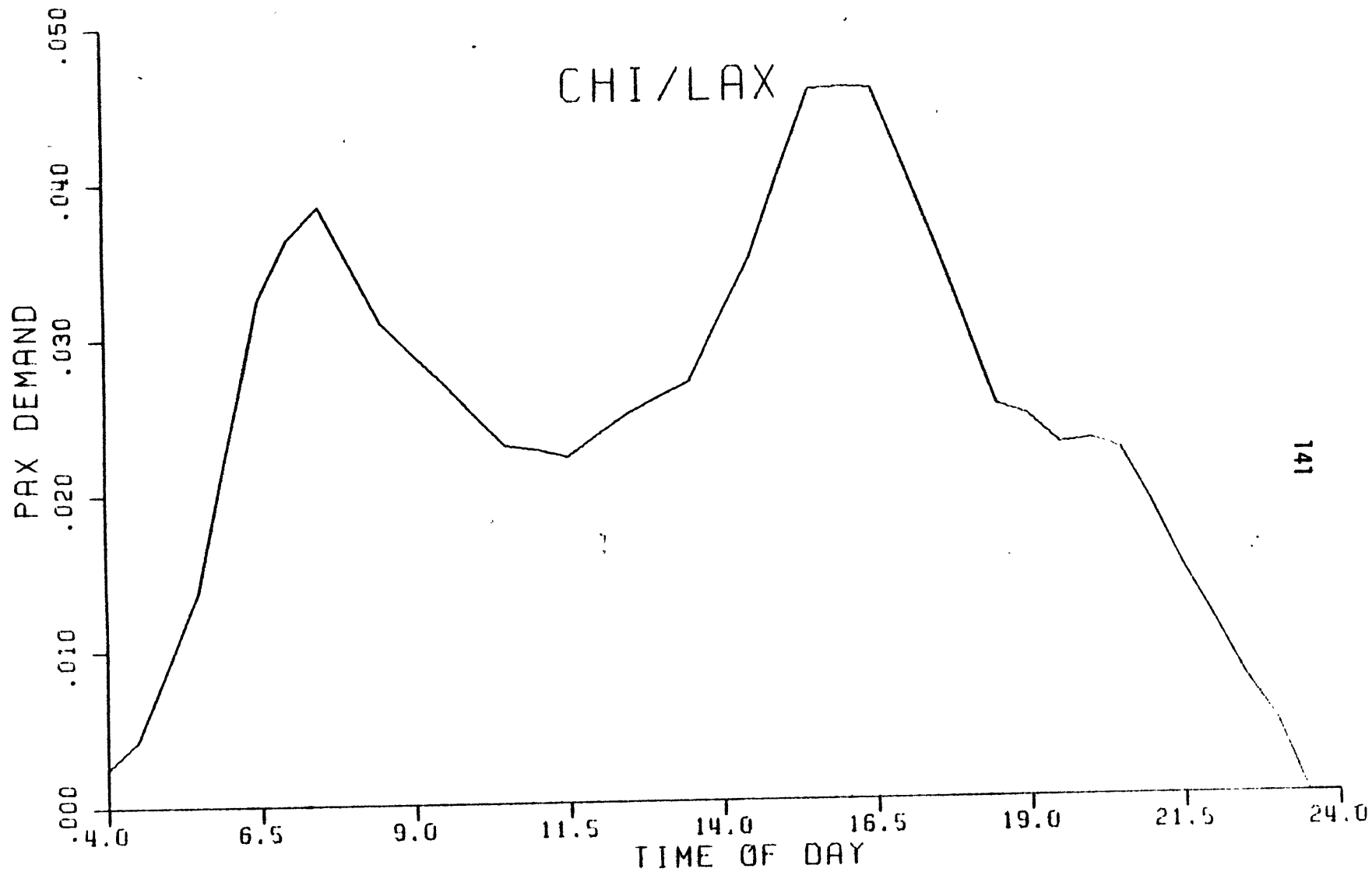


Figure 4.3 Theoretical Time of Day Demand Distribution for Chicago to Los Angeles

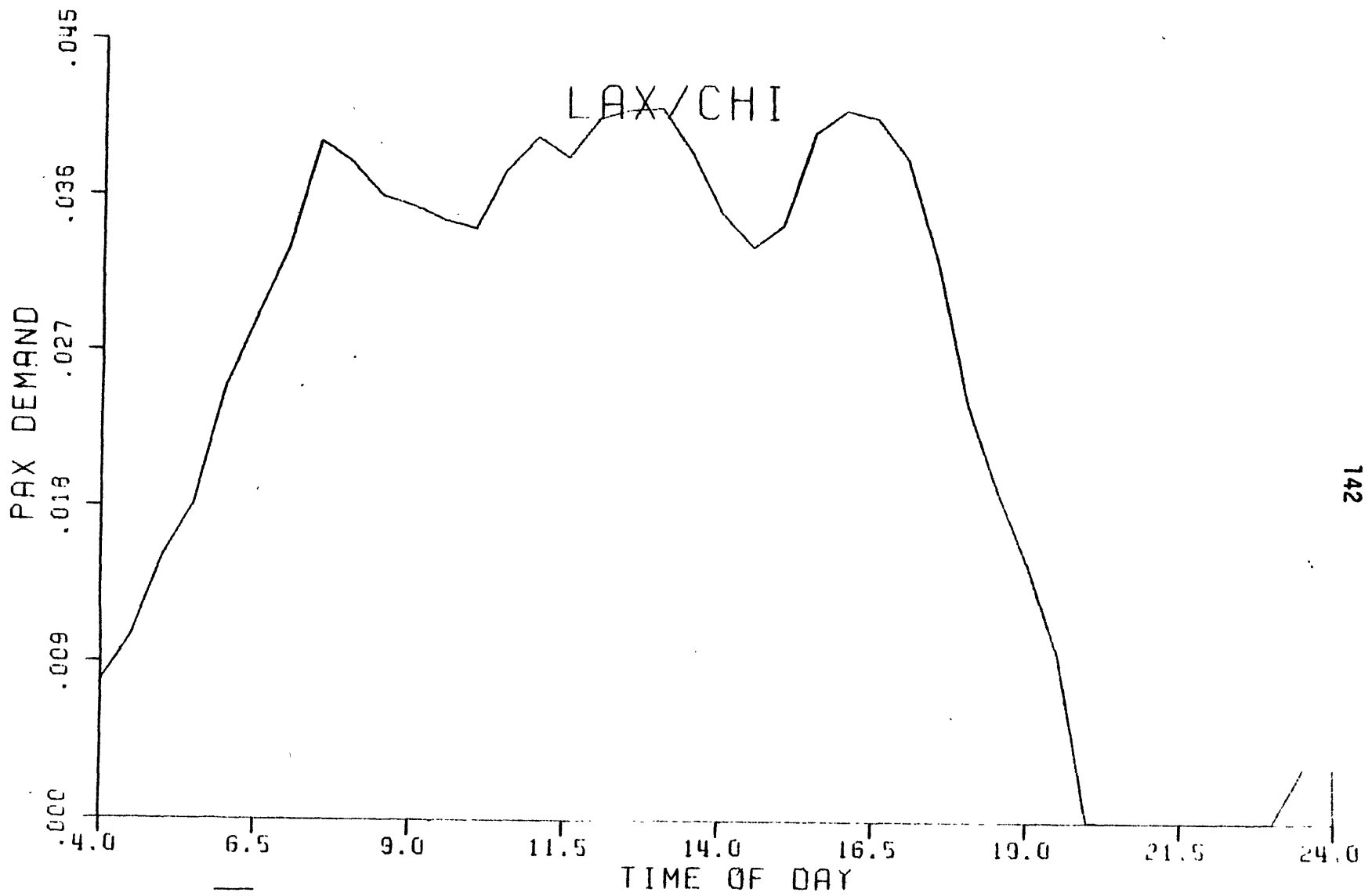


Figure 4.4 Theoretical Time of Day Demand Distribution for Los Angeles to Chicago

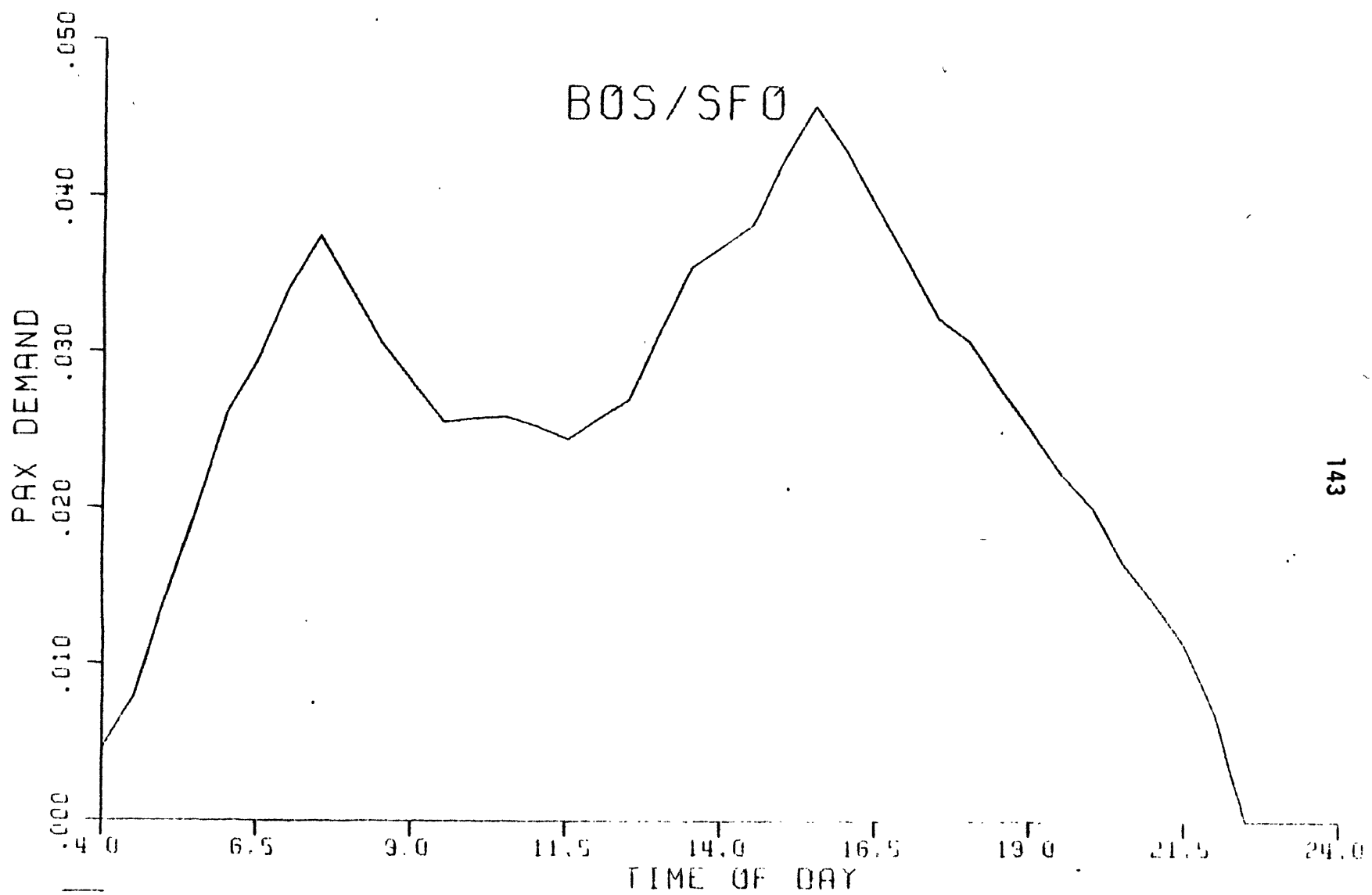


Figure 4.5 Theoretical Time of Day Demand Distribution for Boston to San Francisco



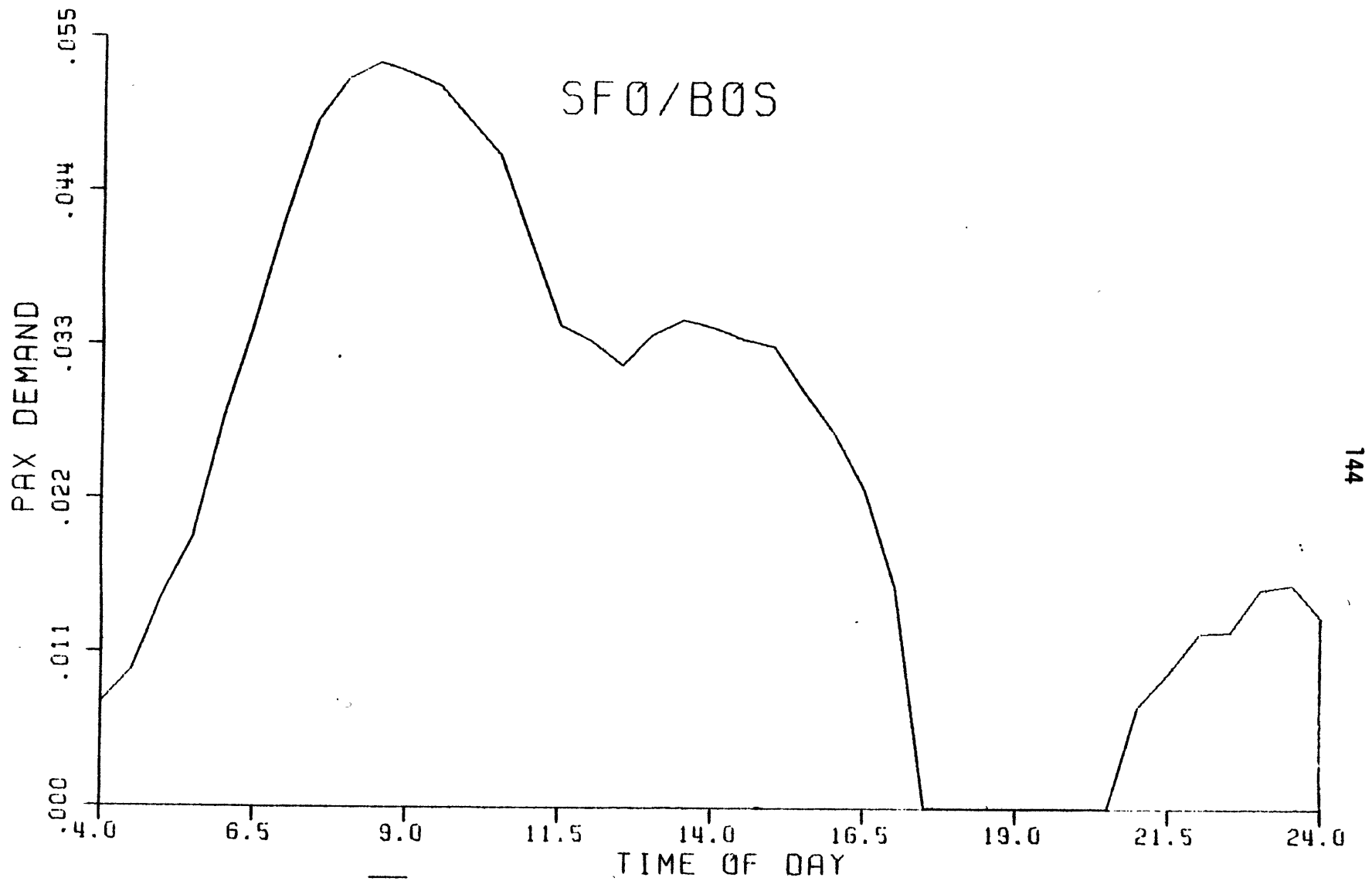


Figure 4.6 Theoretical Time of Day Demand Distribution for San Francisco to Boston

( $t_0$ ), and the time zone difference ( $Z$ ). The cover card is followed by one card for each of the  $m$  scheduled flights which contains the pertinent flight data including local departure and arrival times and the code of the carrier(s).

The program internally determines the time of day demand distribution  $\pi_j$  for  $j = 1, 2, \dots, 41$  and then computes the LOS value using the procedures described in the preceding sections. The output includes a reproduction of the flight schedule and a table of intermediate steps leading to the determination of LOS. The program also prints a table of information on the competitive nature of the market, but discussion of this aspect of LOS\_COMP is deferred to Section 4.2.5.

#### Example 1. Boston to Washington

Boston to Washington is an example of a highly competitive medium haul (406 miles) market, involving two large urban centers which generate a substantial quantity of air passenger demand. Therefore, a high level of service is expected, and referring to Figure 4.7a, the first page of output of the LOS\_COMP program listing the flight schedule,<sup>7</sup> this is indeed the case. Thirty-six flights are offered daily from Boston to Washington; all of these are direct flights, and most are nonstops.

The departure and arrival times are listed in the decimal equivalent of military time. For example, the departure time of the twenty-sixth flight, shown as 16.25, is 4:15 p.m., and the arrival time of the thirty-sixth

---

<sup>7</sup> The flight schedules for each of these examples are extracted from the August 15, 1976 edition of the Official Airline Guide.

Figure 4.7a Flight Schedule for Boston to Washington

FLIGHT SCHEDULE		BOS WAS			
FLIGHT	DEPART	ARRIVE	ADJUSTED FLIGHT TIME	STATUS	CARRIER(S)
1	7.00	8.17	1.17	DIRECT	AA
2	7.00	8.28	1.28	DIRECT	AA
3	7.17	8.35	1.18	DIRECT	DL
4	7.42	8.73	1.32	DIRECT	EA
5	8.00	9.15	1.15	DIRECT	DL
6	8.00	9.20	1.20	DIRECT	DL
7	8.25	9.58	1.33	DIRECT	EA
8	8.75	10.03	1.28	DIRECT	AA
9	9.17	10.80	1.63	DIRECT	AA
10	9.50	10.70	1.20	DIRECT	DL
11	9.92	11.03	1.12	DIRECT	AL
12	10.00	12.67	2.67	DIRECT	AL
13	10.67	11.95	1.28	DIRECT	AA
14	11.58	12.78	1.20	DIRECT	AL
15	12.17	13.45	1.28	DIRECT	EA
16	12.27	13.47	1.20	DIRECT	DL
17	12.30	13.45	1.15	DIRECT	DL
18	12.33	13.62	1.28	DIRECT	AA
19	13.33	16.08	2.75	DIRECT	AL
20	14.17	15.45	1.28	DIRECT	AA
21	14.58	15.78	1.20	DIRECT	DL
22	15.00	16.18	1.18	DIRECT	AL
23	15.62	16.92	1.30	DIRECT	EA
24	16.00	17.28	1.28	DIRECT	AA
25	16.17	17.30	1.13	DIRECT	DL
26	16.25	18.63	2.38	DIRECT	NA
27	16.92	18.27	1.35	DIRECT	EA
28	17.58	18.78	1.20	DIRECT	DL
29	18.17	19.45	1.28	DIRECT	AA
30	18.50	19.70	1.20	DIRECT	AL
31	19.33	21.98	2.65	DIRECT	AL
32	20.00	21.18	1.18	DIRECT	AL
33	20.25	21.48	1.23	DIRECT	AA
34	20.30	21.50	1.20	DIRECT	DL
35	21.00	22.67	1.67	DIRECT	AA
36	22.75	25.50	2.75	DIRECT	NQ

flight, shown as 25.50, is 1:30 a.m. of the following day. The adjusted flight time is merely the scheduled block time; since none of the flights are connections, no adjustments are involved in this particular schedule.

The status of a flight refers to its connection characteristics. Since each of the flights in this schedule is direct, the status is shown as such. Online connections are labeled "ONLINE" and interline connections are labeled "INTLIN".

Figure 4.7b is the second page of output of the LOS\_COMP program, the computation of the level of service related variables. The time of day demand distribution is computed internally and listed in the PI(J) column. For each of the forty-one time points, the computer assigns the passengers preferring to depart at that time to one of the available flights in a manner dictated by the behavioral assumptions discussed in Section 4.2.2.1. For example, those passengers wishing to depart Boston for Washington at 7:00 p.m. (time point 31) are assigned to flight 30 which (referring back to Figure 4.7a) departs at 6:30. Flight 30 is that flight that minimizes the sum of the displacement time (one-half hour) and the flight time. This sum is 1.70 hours as indicated in the TRIP TIME column.

The CONTRIBUTION TO TOTAL TRIP TIME is the product of the PI(J) and TRIP TIME figures, and the sum of this column is the average trip time weighted by the time of day demand distribution. This average, TBAR, is equivalent to the  $\bar{t}$  defined in equation (4.8), and for this example is 1.532 hours.

The level of service index is the ratio of the nonstop jet time,  $t_0$  (listed as "TNJ" in the output), 1.20 hours, to the average total trip

Figure 4.7b Level of Service Computations for Boston to Washington

COMPUTATION OF AVERAGE TOTAL TRIP TIME

J	T (J)	PI (J)	FLIGHT BOARDED	DISPLAC- MENT TIME	ADJUSTED FLIGHT TIME	TRIP TIME	CONTRIBUTION TO TOTAL TRIP TIME
1	4.00	0.001	1	3.00	1.17	4.17	0.005
2	4.50	0.002	1	2.50	1.17	3.67	0.008
3	5.00	0.005	1	2.00	1.17	3.17	0.016
4	5.50	0.008	1	1.50	1.17	2.67	0.021
5	6.00	0.016	1	1.00	1.17	2.17	0.034
6	6.50	0.023	1	0.50	1.17	1.67	0.039
7	7.00	0.033	1	0.00	1.17	1.17	0.039
8	7.50	0.044	4	0.08	1.32	1.40	0.061
9	8.00	0.038	5	0.00	1.15	1.15	0.044
10	8.50	0.033	8	0.25	1.28	1.53	0.050
11	9.00	0.030	8	0.25	1.28	1.53	0.046
12	9.50	0.028	10	0.00	1.20	1.20	0.034
13	10.00	0.026	11	0.08	1.12	1.20	0.032
14	10.50	0.025	13	0.17	1.28	1.45	0.036
15	11.00	0.023	13	0.33	1.28	1.62	0.036
16	11.50	0.020	14	0.08	1.20	1.28	0.026
17	12.00	0.022	15	0.17	1.28	1.45	0.032
18	12.50	0.023	17	0.20	1.15	1.35	0.031
19	13.00	0.025	17	0.70	1.15	1.85	0.045
20	13.50	0.026	20	0.67	1.28	1.95	0.050

Figure 4.7b (continued)

J	T(J)	PI(J)	FLIGHT BOARDED	DISPLACE- MENT TIME	ADJUSTED FLIGHT TIME	TRIP TIME	CONTRIBUTION TO TOTAL TRIP TIME
21	14.00	0.026	20	0.17	1.28	1.45	0.038
22	14.50	0.027	21	0.08	1.20	1.28	0.035
23	15.00	0.035	22	0.00	1.18	1.18	0.041
24	15.50	0.043	23	0.12	1.30	1.42	0.060
25	16.00	0.045	24	0.00	1.28	1.28	0.057
26	16.50	0.047	25	0.33	1.13	1.47	0.068
27	17.00	0.045	27	0.08	1.35	1.43	0.064
28	17.50	0.043	28	0.08	1.20	1.28	0.055
29	18.00	0.036	29	0.17	1.28	1.45	0.052
30	18.50	0.029	30	0.00	1.20	1.20	0.034
31	19.00	0.025	30	0.50	1.20	1.70	0.042
32	19.50	0.021	32	0.50	1.18	1.68	0.035
33	20.00	0.023	32	0.00	1.18	1.18	0.027
34	20.50	0.023	34	0.20	1.20	1.40	0.033
35	21.00	0.022	35	0.00	1.67	1.67	0.036
36	21.50	0.020	35	0.50	1.67	2.17	0.044
37	22.00	0.015	35	1.00	1.67	2.67	0.041
38	22.50	0.010	36	0.25	2.75	3.00	0.030
39	23.00	0.008	36	0.25	2.75	3.00	0.023
40	23.50	0.005	36	0.75	2.75	3.50	0.019
41	24.00	0.003	36	1.25	2.75	4.00	0.011

TBAR = 1.532

$$LOS = TNJ/TBAR = 1.20/1.53 = 0.783$$

time, 1.532 hours, which equals 0.783. The interpretation of this figure is that if "perfect" service, a nonstop jet departing at every instant of the day, were offered, the average total trip time would be 78.3% of its current value.

#### Example 2. Chicago to Philadelphia

Chicago to Philadelphia is similar in certain respects to the Boston to Washington market. It is a medium haul (675 miles) market, involving two major metropolitan regions that generate substantial air passenger traffic. Two major differences between these markets can, however, be noted. The competitive structure of the Boston to Washington markets is comprised of four large carriers (American, Allegheny, Delta, and Eastern), each of which carries a significant passenger load. The Chicago to Philadelphia market is served primarily by two carriers (TWA and United), who schedule head to head, giving each a virtually equal share of the market. The second difference is that the major airport of the origin region in the Chicago to Philadelphia market is notorious for its delays, which are reflected in the schedule by adjustments to block times to compensate for the expected delays.

Figure 4.8a shows the published flight schedule for Chicago to Philadelphia. Of the twenty-three scheduled flights, fifteen are direct and eleven of these are nonstops. These eleven nonstops have scheduled block times of about one hour and fifty minutes (1.83 hours) which, when compared to the block time of about one hour and thirty-five minutes (1.55 hours) from equation (4.11), shows a padding of about fifteen minutes

Figure 4.8a Flight Schedule for Chicago to Philadelphia

FLIGHT SCHEDULE		CHI PHL			
FLIGHT	DEPART	ARRIVE	ADJUSTED FLIGHT TIME	STATUS	CARRIER(S)
1	7.00	9.83	1.83	DIRECT	TW
2	7.00	10.92	3.42	ONLINE	AL/AL
3	7.25	10.08	1.83	DIRECT	UA
4	8.75	12.87	3.12	DIRECT	UA
5	9.75	13.83	3.58	ONLINE	AL/AL
6	10.50	13.37	1.87	DIRECT	TW
7	10.75	15.32	4.07	ONLINE	AL/AL
8	11.67	14.50	1.83	DIRECT	UA
9	12.00	16.93	4.43	ONLINE	AL/AL
10	13.58	17.33	2.75	DIRECT	NW
11	13.75	18.33	4.08	ONLINE	AL/AL
12	14.00	16.83	1.83	DIRECT	TW
13	14.92	17.78	1.87	DIRECT	TW
14	15.17	19.92	4.25	ONLINE	AL/AL
15	15.58	18.47	1.88	DIRECT	UA
16	17.08	20.03	1.95	DIRECT	UA
17	17.75	21.57	3.32	ONLINE	AL/AL
18	17.75	21.78	3.03	DIRECT	NW
19	17.83	20.77	1.93	DIRECT	TW
20	19.50	22.35	1.85	DIRECT	TW
21	20.00	23.98	3.48	ONLINE	AL/AL
22	20.08	22.95	1.87	DIRECT	UA
23	22.07	25.43	2.37	DIRECT	UA



for expected delays out of Chicago.

The level of service measure, LOS, for Chicago to Philadelphia is 0.648 as shown in Figure 4.8b. Note that this figure is about 20% lower than that of Boston to Washington due to fewer flights, a generally inferior scheduling in terms of intermediate stops and connections, and scheduled delays.

Example 3: San Francisco to Omaha

San Francisco to Omaha is a monopolistic (United) long haul (1432 miles) market with few scheduled flights, but a nevertheless reasonably high level of service. Figure 4.9a lists the flight schedule. Eight daily flights are offered, and only three of these are direct. However, the fifth flight listed is United 564 which is a mid-day nonstop 727. Referring to Figure 4.9b, it can be seen that the passengers wishing to depart San Francisco for Omaha between 8:30 a.m. and 1:00 p.m. are assigned to this flight, about 33% of the total daily traffic. The resulting level of service index, LOS, is 0.507.

Example 4: San Antonio to Tucson

The fourth and final example of the computation of level of service is for the San Antonio to Tucson market. This is a low density medium haul (762 miles) market, with a rather low level of service. The only direct flight is the fourth flight listed in Figure 4.10a, Continental 63, a noon departure of a 727-200 which makes one intermediate stop in El Paso. This flight receives the assignment of about 22% of the passengers, as

Figure 4.8b Level of Service Computations for Chicago to Philadelphia

COMPUTATION OF AVERAGE TOTAL TRIP TIME

J	T (J)	PI (J)	FLIGHT BOARDED	DISPLACE- MENT TIME	ADJUSTED FLIGHT TIME	TRIP TIME	CONTRIBUTION TO TOTAL TRIP TIME
1	4.00	0.003	1	3.00	1.83	4.83	0.015
2	4.50	0.006	1	2.50	1.83	4.33	0.027
3	5.00	0.011	1	2.00	1.83	3.83	0.044
4	5.50	0.017	1	1.50	1.83	3.33	0.056
5	6.00	0.027	1	1.00	1.83	2.83	0.077
6	6.50	0.031	1	0.50	1.83	2.33	0.073
7	7.00	0.034	1	0.00	1.83	1.83	0.063
8	7.50	0.038	3	0.25	1.83	2.08	0.079
9	8.00	0.034	3	0.75	1.83	2.58	0.088
10	8.50	0.031	3	1.25	1.83	3.08	0.094
11	9.00	0.029	4	0.25	3.12	3.37	0.097
12	9.50	0.026	6	1.00	1.87	2.87	0.075
13	10.00	0.024	6	0.50	1.87	2.37	0.057
14	10.50	0.024	6	0.00	1.87	1.87	0.045
15	11.00	0.024	6	0.50	1.87	2.37	0.057
16	11.50	0.023	8	0.17	1.83	2.00	0.046
17	12.00	0.025	8	0.33	1.83	2.17	0.053
18	12.50	0.026	8	0.83	1.83	2.67	0.069
19	13.00	0.027	12	1.00	1.83	2.83	0.076
20	13.50	0.031	12	0.50	1.83	2.33	0.073

Figure 4.8b (continued)

J	T(J)	PI(J)	FLIGHT BOARDED	DISPLACE- MENT TIME	ADJUSTED FLIGHT TIME	TRIP TIME	CONTRIBUTION TO TOTAL TRIP TIME
21	14.00	0.035	12	0.00	1.83	1.83	0.064
22	14.50	0.036	13	0.42	1.87	2.28	0.083
23	15.00	0.042	13	0.08	1.87	1.95	0.082
24	15.50	0.046	15	0.08	1.88	1.97	0.090
25	16.00	0.046	15	0.42	1.88	2.30	0.105
26	16.50	0.043	16	0.58	1.95	2.53	0.108
27	17.00	0.037	16	0.08	1.95	2.03	0.076
28	17.50	0.034	19	0.33	1.93	2.27	0.077
29	18.00	0.029	19	0.17	1.93	2.10	0.060
30	18.50	0.027	19	0.67	1.93	2.60	0.070
31	19.00	0.025	20	0.50	1.85	2.35	0.059
32	19.50	0.022	20	0.00	1.85	1.85	0.041
33	20.00	0.022	22	0.08	1.87	1.95	0.044
34	20.50	0.020	22	0.42	1.87	2.28	0.045
35	21.00	0.015	22	0.92	1.87	2.78	0.043
36	21.50	0.013	23	0.57	2.37	2.93	0.038
37	22.00	0.009	23	0.07	2.37	2.43	0.023
38	22.50	0.005	23	0.43	2.37	2.80	0.015
39	23.00	0.000	23	0.93	2.37	3.30	0.000
40	23.50	0.000	23	1.43	2.37	3.80	0.000
41	24.00	0.000	23	1.93	2.37	4.30	0.000

TEAR = 2.390

$$LOS = TNJ/TBAR = 1.55/2.39 = 0.648$$

Figure 4.9a Flight Schedule for San Francisco to Omaha

FLIGHT SCHEDULE SFC CMA					
FLIGHT	DEPART	ARRIVE	ADJUSTED FLIGHT TIME	STATUS	CARRIER(S)
1	8.00	14.60	5.10	ONLINE	UA/UA
2	8.00	14.92	5.42	ONLINE	UA/UA
3	8.75	16.33	5.58	DIRECT	UA
4	10.17	16.97	5.80	INTLIN	WA/AA
5	10.83	15.75	2.92	DIRECT	UA
6	14.00	20.58	5.08	ONLINE	UA/UA
7	14.25	20.58	4.33	DIRECT	UA
8	19.58	29.10	8.52	INTLIN	CC/EA

Figure 4.9b Level of Service Computations for San Francisco to Omaha

COMPUTATION OF AVERAGE TOTAL TRIP TIME

J	T (J)	PI (J)	FLIGHT BOARDED	DISPLACE- MENT TIME	ADJUSTED FLIGHT TIME	TRIP TIME	CONTRIBUTION TO TOTAL TRIP TIME
1	4.00	0.009	1	4.00	5.10	9.10	0.078
2	4.50	0.011	1	3.50	5.10	8.60	0.096
3	5.00	0.015	1	3.00	5.10	8.10	0.124
4	5.50	0.018	1	2.50	5.10	7.60	0.139
5	6.00	0.025	1	2.00	5.10	7.10	0.176
6	6.50	0.029	1	1.50	5.10	6.60	0.194
7	7.00	0.034	1	1.00	5.10	6.10	0.209
8	7.50	0.037	1	0.50	5.10	5.60	0.208
9	8.00	0.033	1	0.00	5.10	5.10	0.168
10	8.50	0.032	5	2.33	2.92	5.25	0.166
11	9.00	0.032	5	1.83	2.92	4.75	0.150
12	9.50	0.031	5	1.33	2.92	4.25	0.132
13	10.00	0.031	5	0.83	2.92	3.75	0.116
14	10.50	0.030	5	0.33	2.92	3.25	0.099
15	11.00	0.029	5	0.17	2.92	3.08	0.091
16	11.50	0.032	5	0.67	2.92	3.58	0.113
17	12.00	0.036	5	1.17	2.92	4.08	0.148
18	12.50	0.038	5	1.67	2.92	4.58	0.176
19	13.00	0.040	5	2.17	2.92	5.08	0.204
20	13.50	0.040	7	0.75	4.33	5.08	0.204

Figure 4.9b (continued)

J	T(J)	PI(J)	FLIGHT BOARDED	DISPLACE- MENT TIME	ADJUSTED FLIGHT TIME	TRIP TIME	CONTRIBUTION TO TOTAL TRIP TIME
21	14.00	0.040	7	0.25	4.33	4.58	0.183
22	14.50	0.037	7	0.25	4.33	4.58	0.170
23	15.00	0.038	7	0.75	4.33	5.08	0.191
24	15.50	0.039	7	1.25	4.33	5.58	0.216
25	16.00	0.036	7	1.75	4.33	6.08	0.221
26	16.50	0.039	7	2.25	4.33	6.58	0.256
27	17.00	0.038	7	2.75	4.33	7.08	0.272
28	17.50	0.036	7	3.25	4.33	7.58	0.275
29	18.00	0.032	7	3.75	4.33	8.08	0.258
30	18.50	0.025	7	4.25	4.33	8.58	0.213
31	19.00	0.019	7	4.75	4.33	9.08	0.171
32	19.50	0.015	8	0.08	8.52	8.60	0.131
33	20.00	0.013	8	0.42	8.52	8.93	0.118
34	20.50	0.009	8	0.92	8.52	9.43	0.089
35	21.00	0.000	8	1.42	8.52	9.93	0.000
36	21.50	0.000	8	1.92	8.52	10.43	0.000
37	22.00	0.000	8	2.42	8.52	10.93	0.000
38	22.50	0.000	8	2.92	8.52	11.43	0.000
39	23.00	0.000	8	3.42	8.52	11.93	0.000
40	23.50	0.000	8	3.92	8.52	12.43	0.000
41	24.00	0.000	8	4.42	8.52	12.93	0.000

TBAR = 5.755

$$LOS = TNJ/TBAR = 2.92/5.75 = 0.507$$

Figure 4.10a Flight Schedule for San Antonio to Tucson

FLIGHT SCHEDULE		SAT TUS				
FLIGHT	DEPART	ARRIVE	ADJUSTED FLIGHT TIME	STATUS	CARRIER(S)	
1	8.83	9.27	2.93	ONLINE	CC/CO	
2	11.00	14.85	6.85	INTLIN	BN/AA	
3	11.33	14.90	6.07	ONLINE	AA/AA	
4	12.00	12.47	2.47	DIRECT	CC	
5	16.17	17.38	4.22	INTLIN	AA/CO	
6	18.00	21.17	6.17	INTLIN	BN/AA	
7	22.33	23.78	3.95	ONLINE	CC/CO	

can be determined from Figure 4.10b. The remaining 78% must use connecting service. The resulting average total trip time, TBAR, is 4.672 hours, which, when compared to a nonstop jet time of 1.96 hours, yields a level of service index of 0.420.

#### 4.2.2.6 Combination of Directional Levels of Service

The values of the level of service variables for a given region pair  $ij$  are computed for the schedules of both directions,  $i$  to  $j$  and  $j$  to  $i$ . For the purposes of model calibration in Chapter V, the two values are combined by taking the geometric mean.

$$LOS_{ij} = \sqrt{LOS_{i \rightarrow j} \times LOS_{j \rightarrow i}} \quad (4.14)$$

This multiplicative form was selected over an additive form to guard against overestimation in markets that may have substantially asymmetric service patterns. If a very high level of service were offered in one direction but little or none in the other direction, it is likely that there would be a very low volume of demand. The multiplicative formulation is sensitive to this, whereas an additive form would not be.



Figure 4.10b Level of Service Computations for San Antonio to Tucson

COMPUTATION OF AVERAGE TOTAL TRIP TIME

J	T (J)	PI (J)	FLIGHT BOARDED	DISPLACE- MENT TIME	ADJUSTED FLIGHT TIME	TRIP TIME	CONTRIBUTION TO TOTAL TRIP TIME
1	4.00	0.000	1	4.83	2.93	7.77	0.000
2	4.50	0.000	1	4.33	2.93	7.27	0.000
3	5.00	0.003	1	3.83	2.93	6.77	0.017
4	5.50	0.004	1	3.33	2.93	6.27	0.027
5	6.00	0.009	1	2.83	2.93	5.77	0.052
6	6.50	0.014	1	2.33	2.93	5.27	0.073
7	7.00	0.023	1	1.83	2.93	4.77	0.111
8	7.50	0.033	1	1.33	2.93	4.27	0.139
9	8.00	0.036	1	0.83	2.93	3.77	0.137
10	8.50	0.038	1	0.33	2.93	3.27	0.126
11	9.00	0.035	1	0.17	2.93	3.10	0.108
12	9.50	0.031	1	0.67	2.93	3.60	0.111
13	10.00	0.029	1	1.17	2.93	4.10	0.119
14	10.50	0.027	4	1.50	2.47	3.97	0.107
15	11.00	0.025	4	1.00	2.47	3.47	0.086
16	11.50	0.023	4	0.50	2.47	2.97	0.068
17	12.00	0.023	4	0.00	2.47	2.47	0.056
18	12.50	0.022	4	0.50	2.47	2.97	0.066
19	13.00	0.024	4	1.00	2.47	3.47	0.082
20	13.50	0.025	4	1.50	2.47	3.97	0.099

Figure 4.10b (continued)

J	T(J)	PI(J)	FLIGHT BOARDED	DISPLACE- MENT TIME	ADJUSTED FLIGHT TIME	TRIP TIME	CONTRIBUTION TO TOTAL TRIP TIME
21	14.00	0.026	4	2.00	2.47	4.47	0.116
22	14.50	0.027	4	2.50	2.47	4.97	0.134
23	15.00	0.031	5	1.17	4.22	5.38	0.167
24	15.50	0.035	5	0.67	4.22	4.88	0.170
25	16.00	0.040	5	0.17	4.22	4.38	0.177
26	16.50	0.046	5	0.33	4.22	4.55	0.207
27	17.00	0.046	5	0.83	4.22	5.05	0.231
28	17.50	0.046	5	1.33	4.22	5.55	0.253
29	18.00	0.041	5	1.83	4.22	6.05	0.247
30	18.50	0.036	5	2.33	4.22	6.55	0.234
31	19.00	0.030	5	2.83	4.22	7.05	0.215
32	19.50	0.025	7	2.83	3.95	6.78	0.170
33	20.00	0.024	7	2.33	3.95	6.28	0.153
34	20.50	0.023	7	1.83	3.95	5.78	0.131
35	21.00	0.023	7	1.33	3.95	5.28	0.121
36	21.50	0.022	7	0.83	3.95	4.78	0.106
37	22.00	0.019	7	0.33	3.95	4.28	0.080
38	22.50	0.015	7	0.17	3.95	4.12	0.060
39	23.00	0.011	7	0.67	3.95	4.62	0.051
40	23.50	0.008	7	1.17	3.95	5.12	0.039
41	24.00	0.005	7	1.67	3.95	5.62	0.026

TBAR = 4.672

$$LOS = TNJ/TBAR = 1.96/4.67 = 0.420$$

#### 4.2.3 Fare (F)

The standard coach (Y) fare has been selected as the price variable. The source of these data is the Official Airline Guide. It can be argued that standard coach fare, while being a common measure of price in air transportation demand modeling, is perhaps improper since it neglects the impact upon demand of discount fare plans. However, the results of a prototype study<sup>8</sup> imply that further sophistication of definition of fare variable produces virtually identical results.

Three cross sectional models, each containing a different fare variable, were calibrated in the prototype study. The models were log-linear specifications with demand as the response variable and level of service, Buying Power Index (a measure of economic activity), and fare as the carriers. The fare proxies in the three models were standard coach fare, an estimated average fare, and the actual average fare (determined by computing an average of the available fares, weighted by the number of passengers who paid each of the fares). The results of the calibration of these models were nearly identical with negligible fluctuations observed in the coefficient estimates, their standard errors, and the standard error of estimate.

In order to ensure against the possible effect of the fare variable measuring a time trend and to capture the true impact of fare levels as perceived by the consumer, it was decided that F should be deflated. Since

---

<sup>8</sup> Steven E. Eriksen, John C. Scalea, and Nawal K. Taneja, "A Methodology for Determining the Relationship between Air Transportation Demand and the Level of Service" (Cambridge: M.I.T. Flight Transportation Laboratory Report R76-3, January 1976), pp. 34-36.

the air transportation product is a service, the selected price deflator is the "implicit price deflator for personal consumption expenditures on services."<sup>9</sup> The deflated fare variable is expressed in terms of constant dollars with 1972 as the base year.

#### 4.2.4 Socio-Economic Activity (SE)

As discussed in Section 3.3.1, it is postulated that the total potential demand for air passenger services in a region pair market is the level of socio-economic activity in the two regions. Two aspects of socio-economic levels are considered in this research. The first of these is the ability of a region to generate air traffic, and is measured by the total personal income of the region.

The second aspect of socio-economic activity is the region's ability to attract air traffic. Quandt points out that

"cities with high concentrations of financial intermediaries, educational and governmental institutions and other service industries give rise to more travel per capita than cities with predominantly manufacturing industries." <sup>10</sup>

Regions such as New York, Las Vegas, and Miami with large, service-oriented economies, tend to draw more traffic relative to aggregate industry than the highly manufacturing regions such as Detroit and Pittsburgh. Therefore,

---

<sup>9</sup> Economic Report of the President (Washington: U.S. Government Printing Office, 1976), p. 174.

<sup>10</sup> Richard E. Quandt, The Demand for Travel: Theory and Measurement (Lexington, Mass.: Heath Lexington Books, 1970), p. 1.

a service industry measure, "total labor and proprietors' income by place of work by industry, service" was selected. These data are published annually<sup>11</sup> by the Bureau of Economic Analysis (BEA) of the Department of Commerce. The data are tabulated by BEA areas which, as will be discussed in Section 4.2, are the defined regions for this study.

The socio-economic attraction from region  $i$  to region  $j$  is then defined as the product of the personal income of region  $i$  and the service income of region  $j$ . The average of the socio-economic attraction in both directions of a given region pair is computed, and the square root of this is taken to convert the units from dollars squared to dollars. The socio-economic variable,  $SE$ , for a generic region pair  $ij$  is then defined as:

$$SE_{ij} = \sqrt{1/2(INC_i \cdot SRVC_j + SRVC_i \cdot INC_j)} \quad (4.15)$$

where

$INC$  = personal income, and

$SRVC$  = total labor and proprietors' income by place of work, by industry, service

The socio-economic variable is deflated by the implicit price deflator for personal consumption expenditures on services with a base year of 1972. This is compatible with the fare variable deflation described in the preceding section.

---

<sup>11</sup> Bureau of Economic Analysis, "Local Area Personal Income 1969-1974, Vol. 1" (Washington: Department of Commerce, 1976).

#### 4.2.5 Competition (COMP)

##### 4.2.5.1 Development of the Competition Variable

A proxy for competition has been developed which assumes the value of 1.0 in strictly monopolistic markets and greater values as the amount of competition increases. The improvement in this variable over what has been used in other studies (e.g., Marfisi<sup>12</sup>), namely "number of carriers in a market", is that it discounts the presence of minor competitors. If a market is served by two equally strong competitors, the COMP variable will have the value of 2.0. If a market is served by one major and one minor competitor, COMP will have a value somewhere between 1.0 and 2.0, depending upon the relative strength of the minor competitor.

The competition variable, COMP, is defined as follows:

$$\text{COMP} = \frac{1}{\sum_{i=1}^m \text{MS}_i^2} \quad (4.16)$$

where

$i$  = index of the carriers serving the market, and

$\text{MS}_i$  = market share of carrier  $i$ .

This variable is the reciprocal of the common Herfindahl index.<sup>13</sup>

---

<sup>12</sup> Marfisi, op. cit.

<sup>13</sup> Frederic M. Scherer, Industrial Market Structure and Economic Performance (Chicago: Rand, McNally, 1971), pp. 51-52.

#### 4.2.5.2 Examples of the Computation of the COMP Variable

The COMP variable is computed in the LOS\_COMP program. As indicated in Section 4.2.2.5, the proportion,  $\Pi_j$ , of the total daily passengers that wish to depart at any time point  $j$  are assigned to a particular flight in accordance with the behavioral assumptions outlined in Section 4.2.2.1. From these assignments the market shares of the various carriers can be estimated, and the value COMP is then calculated from equation (4.16) above.

Consider the flight schedule for Boston to Washington (Figure 4.7a). There are a number of carriers competing for the traffic and providing a high level of service. From the assignments of passengers to flights in Figure 4.7b, the market shares are predicted and printed in the third page of output which is reproduced in Figure 4.1. The major carrier is American, with 38.9% of the market. Delta, Allegheny, and Eastern follow, and a commuter carrier, Cumberland Airlines, picks up a very small percentage. Substituting these market shares into equation (4.16) yields the value for COMP:

$$\begin{aligned}
 \text{COMP} &= \frac{1}{\sum_{i=1}^5 \text{MS}_i^2} \\
 &= \frac{1}{(0.389)^2 + (0.153)^2 + (0.254)^2 + (0.179)^2 + (0.026)^2} \\
 &= \frac{1}{0.2719} = 3.685
 \end{aligned}$$

Figure 4.11 Estimates of Market Shares for Boston to Washington

NUMBER OF COMPETITORS = 5

CARRIER	MARKET SHARE
AA	0.389
EA	0.153
DL	0.254
AL	0.179
NQ	0.026

COMP = 3.685

Figure 4.12 Estimates of Market Shares for Chicago to Philadelphia

NUMBER OF COMPETITORS = 2

CARRIER	MARKET SHARE
TW	0.538
UA	0.462

CCMP = 1.989



Chicago to Philadelphia has only two carriers that are competing for market share. In Figure 4.8a, Northwest and Allegheny have a number of flights scheduled, but both of Northwest's departures involve intermediate stops, and Allegheny's flights are all connecting. TWA and United, however, offer nonstop service on nearly all of their flights and, as indicated in Figure 4.12, are assigned all the traffic. Since their shares of the market are nearly equal, COMP is very close in value to two.

San Francisco to Omaha, as can be inferred from Figure 4.9a, is virtually a monopolistic market. However, the Continental/Eastern connecting flight at the end of the day is assigned a few passengers in Figure 4.9b, so the COMP variable is expected to be slightly above one. Indeed, in Figure 4.13, United receives 96.2% of the market, and COMP is calculated to be 1.079.

Finally, for San Antonio to Tucson, it appears from Figure 4.10 that Continental should capture a large portion of the market. In Figure 4.14, Continental receives 82.5%, and the COMP value is 1.406.

#### 4.2.6 Location within Route Structure (RS)

The route structure variable accounts for the fact that many low density markets have a very high level of service offered to them, because they feed into a complex. An example of this which was mentioned in Section 3.3.2 is Birmingham to Atlanta with sixteen nonstops daily.

The four major connecting hubs in the U.S. domestic system are Atlanta, Chicago, Dallas-Fort Worth, and Denver. Airlines attempt to

Figure 4.13 Estimates of Market Share for San Francisco to Omaha

NUMBER OF COMPETITORS = 3

CARRIER	MARKET SHARE
UA	0.962
CO	0.019
EA	0.019

COMP = 1.079

Figure 4.14 Estimates of Market Shares for San Antonio to Tucson

NUMBER OF COMPETITORS = 2

CARRIER	MARKET SHARE
CO	0.825
AA	0.175

COMP = 1.406

maximize the efficiency of the complexes which they operate at these hubs by scheduling arrivals close together, filling as many of their gates as possible (Delta concurrently operates 33 gates in peak hours at Atlanta), and then closely scheduling departures. The advantage of this mode of operation is that combinatorically there is a very large number of origin-destination pairs that may be served via close connections. The obvious disadvantages are the high variability of demands upon terminal employees and confusion and congestion in the terminals during the peak periods of the day.

A few of the natural characteristics of airports in which complexes are operated are:

- a. A large number of connecting and through passengers are handled.
- b. A large number of aircraft operations are made in a given time period.
- c. A large number of regions are served on a nonstop basis from the airport.
- d. A large percentage of passengers aboard aircraft arriving at or departing from the airport are not local passengers on that particular segment.

In addition to the domestic hubs in which complexes are operated, certain "gateway" airports experience characteristics similar to those listed above, due to a substantial volume of domestic/international connection activity. Most notable of these are New York (Kennedy) and Washington (Dulles) on the east coast and Los Angeles International and San Francisco International on the west coast.

The most meaningful measure of route structure for a given segment is the number of non-local (either continuing or connecting) passengers traveling on that segment in a given time period. However, these data are not readily available. They can only be extracted from the Civil Aeronautics Board's service segment flow data. The acquisition and processing of these data are very expensive, both in terms of cost and time requirements. As a surrogate to the number of non-local passengers, the selected measure of the network effects is the number of connecting passengers. These data are extracted from Table 10 of the CAB Origin to Destination Survey.

### 4.3 Sample Design and Market Selection

#### 4.3.1 The Concept of Region Pairs

As was briefly discussed in Section 3.1.1, an airport generally attracts passengers from a larger area than its respective city or SMSA. Several characteristics of passenger behavior related to this fact are as follows:

(1) Airline passengers may be drawn from cities with air carrier service to more distant airports depending upon the relative levels of service available. For example, consider a passenger desiring to travel from Providence to Cleveland sometime after the only direct flight which leaves at 8:50 a.m. While several connections are available during the rest of the day, a number of nonstops depart from Boston, 60 miles away, and may be as convenient in terms of total trip time. Thus, some of

the Providence-Cleveland demand can be expected to spill over into the Boston-Cleveland statistics solely because of the schedule offered.

(2) Commuter airlines, while becoming a more integral part of the air transportation system since their beginning in the late 1960's, do not report traffic statistics to the C.A.B. in the same detail as do the trunk and local service carriers. While recent C.A.B. actions have attempted to bring the commuters closer to the mainstream of air transportation activity by the introduction of joint fares and airline ticketing, the unregulated commuters began operations in an environment virtually disjoint from the rest of the airline system. Under these conditions, a ticket written from New York to Los Angeles with a connection to Palm Springs on Golden West Airlines would statistically have represented an origin to destination trip in the New York-Los Angeles city pair, while in fact it would be more accurate to consider this the New York-Los Angeles region pair with Palm Springs included within the Los Angeles region.

(3) Due to economic pressures brought before the Board by the airlines, the C.A.B. approved suspensions and deletions of service to a large number of small communities, forcing those passengers formerly served by the suspended flights to use airports further away. If the replacement airport is within the same region as the abandoned one, working with city pairs will show a decline to almost nothing at the abandoned airport and an increase at the replacement airport.

These points support use of regions rather than cities to insure more accurate modeling and analysis of the level of passenger movements. However, this reasoning is highly dependent upon the quality and accuracy

of the delineation of the regions themselves. In 1972, the Bureau of Economic Analysis (BEA) of the Department of Commerce investigated the use of geographical regions delineated by criteria based upon transportation data. By using the journey-to-work data from the 1960 Census of Population, the Bureau divided the country into the 173 self-sufficient regions by minimizing the routine commuting across region boundaries so that labor supply and demand were located in the same region. Region boundaries were restricted to county boundaries and, for the purposes of this work, there is at least one air carrier airport serving each region. Since other geographical delineations considered were not based upon transportation criteria, the BEA regions were adopted for this investigation.

Each region pair is comprised of a set of airport pairs found by enumerating the airports in one region with those in the other. Even if there is more than one airport within a metropolitan area, all airports must be counted and matched with all airports in the other region. The Official Airline Guide aggregates airports within the same city, but for purposes of this research, each airport must be considered separately. The demand in a region pair will be the sum of the demands of the component airport pairs; the supply of service in a region pair will be the aggregate of flights offered in each of the component airport pairs.

#### 4.3.2 Sampling Design and Procedure

The criteria used in the process of selecting a representative sample of region pair markets for this analysis are as follows:

(1) The sample size should be larger than that of most other econometric analyses in this area, preferably on the order of 200 markets.

(2) The stage lengths should be evenly distributed over short (less than 400 miles), medium (400 to 999 miles), and long (1000 miles and longer) haul markets.

(3) To facilitate data collection the total number of distinct regions should be held to a reasonably manageable number (on the order of 50).

(4) A mix of economic levels of the region pairs should be selected within each of the length of haul strata. There should be selections of two large regions, a large and a medium size region, a large and a small region, two medium sized regions, etc.

(5) The markets should be distributed as evenly as possible with respect to geographical location and market type.

The initial step in the sampling procedure was to select fifty distinct regions which were quite evenly distributed across the country. It was hoped that the entire selection of region pairs could be taken using these regions. It turned out to be necessary to add two additional regions, bringing the total number to fifty-two. The regions are listed in alphabetical order in Figure 4.15. For data manipulation reasons, it was necessary to assign each region a two digit code number. These and the standard three letter city codes are included in Figure 4.15.

<u>Region</u>	<u>Codes</u>		175	<u>Codes</u>	
	<u>Number</u>	<u>Letters</u>		<u>Region</u>	<u>Number</u>
Albany	01	ALB	Minneapolis	27	MSP
Atlanta	02	ATL	Minot	28	MOT
Bismarck	03	BIS	Nashville	29	BNA
Boston	04	BOS	New Orleans	30	MSY
Chicago	05	CHI	New York	31	NYC
Cincinnati	06	CVG	Norfolk	32	ORF
Cleveland	07	CLE	Oklahoma City	33	OKC
Dallas	08	DAL	Omaha	34	OMA
Dayton	09	DAY	Philadelphia	35	PHL
Denver	10	DEN	Pittsburgh	36	PIT
Detroit	11	DTT	Portland, Maine	37	PWM
Erie	12	ERI	Portland, Oregon	38	PDX
Fargo	13	FAR	Raleigh	39	RDU
Houston	14	HOU	Reno	40	RNO
Jackson	15	JAN	Richmond	41	RIC
Jacksonville	16	JAX	Rochester	42	ROC
Kansas City	17	MKC	Sacramento	43	SAC
Knoxville	18	TYS	St. Louis	44	STL
Las Vegas	19	LAS	Salt Lake City	45	SLC
Lexington	20	LEX	San Antonio	46	SAT
Lincoln	21	LNK	San Diego	47	SAN
Los Angeles	22	LAX	San Francisco	48	SFO
Lubbock	23	LBB	Seattle	49	SEA
Memphis	24	MEM	Tucson	50	TUS
Miami	25	MIA	Washington	51	WAS
Milwaukee	26	MKE	Wichita	52	ICT

Figure 4.15 List of Regions



The second stage of the sampling procedure was to design a two way stratification using length of haul and economic activity as the stratified variables. The economic variable used was the 1974 Buying Power Index (BPI) for each region. The Buying Power Index is a function of a region's retail sales, population, and total income, and is published annually in the "Survey of Buying Power" edition of Sales Management magazine. The regions were divided into five economic strata, with number 1 being the low BPI regions (e.g., Erie, Reno), and number 5 being the high BPI regions (e.g., New York, San Francisco). This resulted in fifteen economic strata for region pairs:

1-1	2-2	3-4
1-2	2-3	3-5
1-3	2-4	4-4
1-4	2-5	4-5
1-5	3-3	5-5

The lowest economic markets are in the "1-1" category (e.g., Erie-Reno) and highest are the "5-5" markets (e.g., New York-San Francisco).

This two dimensional stratification yields a 3 x 15 matrix with the three rows representing length of haul and the fifteen columns representing economic level. Four regions were selected for each of the forty-five blocks, yielding a total sample of 180 region pairs. The markets were carefully hand-picked in an effort to geographically vary the markets within each block as much as possible.

The markets in the sample are listed by stratification in Appendix A.

## V. Empirical Analysis of U.S. Domestic Air Passenger Markets

The exact specification of the general model developed in Chapter III and defined in detail in Chapter IV is as follows:

$$\text{(Demand equation)} \quad Q_D = \beta'_{10} \text{LOS}^{\beta_{11}} F^{\beta_{12}} \text{SE}^{\beta_{13}} \epsilon'_1 \quad (5.1)$$

$$\text{(Service equation)} \quad \text{LOS} = \beta'_{20} \text{TRAFL}^{\beta_{21}} F^{\beta_{22}} \text{COMP}^{\beta_{23}} \epsilon'_2 \quad (5.2)$$

where

- $Q_D$  = origin to destination passenger demand
- $\text{LOS}$  = level of service
- $F$  = fare
- $\text{SE}$  = socio-economic activity
- $\text{TRAFL}$  = traffic (origin to destination plus connecting passengers)  
lagged one year
- $\text{COMP}$  = competition

Taking the logarithms of equations (5.1) and (5.2) yields:

$$\begin{aligned} \text{(Demand equation)} \quad \log Q_D &= \beta_{10} + \beta_{11} \log \text{LOS} + \beta_{12} \log F \\ &+ \beta_{13} \log \text{SE} + \epsilon_1 \end{aligned} \quad (5.3)$$

$$\begin{aligned}
 \text{(Service equation) } \log \text{ LOS} &= \beta_{20} + \beta_{21} \log \text{ TRAF L} + \beta_{22} \log F \\
 &+ \beta_{23} \log \text{ COMP} + \varepsilon_2
 \end{aligned}
 \tag{5.4}$$

where

$$\beta_{10} = \log \beta'_{10}$$

$$\varepsilon_1 = \log \varepsilon'_1$$

$$\beta_{20} = \log \beta'_{20}$$

$$\varepsilon_2 = \log \varepsilon'_2$$

Both of these equations are of the general linear form:

$$Y = X\beta + \varepsilon \tag{5.5}$$

where

$Y$  is the  $n \times 1$  vector of "response" variable observations

$X$  is the  $n \times p$  matrix of the  $n$  observations of the  $p$  "carrier" variables

$\beta$  is the  $p \times 1$  vector of coefficients

and  $\varepsilon$  is the  $n \times 1$  vector of error terms

The least squares estimators of the  $p \times 1$  vector of  $\beta$  coefficients are as follows:

$$\hat{\beta} = (X^T X)^{-1} X^T Y \tag{5.6}$$

The error terms are random variables assumed to be independent and identically distributed according to a Gaussian distribution with mean of zero and variance  $\sigma^2$ .

$$\epsilon \sim G(0, \sigma^2) \quad (5.7)$$

Therefore the expected values,  $\hat{Y}$ , of the elements of  $Y$  given a set of observations  $X$  are as follows:

$$\hat{Y} = X\hat{\beta} \quad (5.8)$$

The term "residual",  $r_i$ , refers to the difference between the actual value of the  $i^{\text{th}}$  observation of the response variable,  $y_i$ , and its expected or "predicted" value,  $\hat{y}_i$ .

$$r_i = y_i - \hat{y}_i = y_i - x_i\hat{\beta} \quad (5.9)$$

where  $x_i$  is the  $i^{\text{th}}$  row of the carrier matrix,  $X$ .

The purpose of this chapter is to present the process and results of the statistical calibration of the general model over various subsets of the data set. Some of the general conclusions and implications regarding applications based upon the statistical analysis of this chapter will briefly be discussed herein. However, the bulk of the applications and conclusions will be deferred until the two final chapters of this thesis.

The total data set consists of observations from each of the 180

markets in the experimental design for the six years between 1969 and 1974 inclusive. The fare and level of service related data were extracted from the Official Airline Guide (OAG) of September 1 of each of these years.<sup>1</sup> Since many of the markets in the sample are very small, quite often no service was published in the OAG for certain markets in September of one or more of these years. These observations were deleted, and the size of the total sample was consequently reduced to 820 observations.

A series of statistical tests are conducted to determine proper aggregation of subsets of the data. As was expected, the tests affirm that the characteristics of long, medium, and short lengths of haul markets differ sufficiently to warrant separate analyses. Furthermore, it will be shown that within length of haul strata, market characteristics across market size (as defined by demographics) vary significantly. Initially it is unclear why this market size effect is true, but as the analysis proceeds, the reasons become evident. Consequently, nine subsets of the total sample, categorized three ways by length of haul and three ways by market size, are separately analyzed.

Within each of the nine cross-classifications of the data a standard process of parameter estimation will be used throughout this chapter. The first step is to calibrate (estimate the  $\beta_{2j}$  parameters)

---

<sup>1</sup> September was selected because this month least reflects seasonality. Since the demand data is annual, a winter month would overstate nominal level of service for north-south markets, and a summer month would overstate nominal service in east-west markets. September, generally a shoulder period month, seemed to be most representative of nominal level of service scheduling.

the service equation using the ordinary least squares estimation procedure of equation (5.6). The estimated values for the logarithms of the level of service observations are then computed using equation (5.8) and used as a proxy (instrumental variable) for log LOS in the demand equation (5.3) which is then estimated using ordinary least squares. This initial process is referred to as the "preliminary analysis".

Two serious problems in the results of the preliminary analysis of the demand equation will occur throughout. The first of these is a clustering by markets of the residuals. Within each length of haul/size of market stratum it will not be uncommon to observe a number of markets for which the six observations are consistently overestimated and several markets which are consistently underestimated. The second recurrent problem in the preliminary analyses of the demand equations calibrated over the nine subsets is a substantial disparity between the coefficient estimates and judgmental assessments of what are the actual values of these parameters.

The problem of clustering of residuals is hypothesized to be a result of the inclusion of discrepant data points within the various subsamples. Discrepant data points in this pooling of cross-sectional and time series data are of two different varieties. An entire market (six data points) may be improperly inserted into a cross-classification (e.g., a market in the medium haul/large size sample is in reality a member of the medium haul/medium size population), or a single observation within an otherwise properly pooled market may be awry (due to an airline strike, other temporary suspension of service, data error, etc). These two classes of discrepancies are handled separately

(although their ill effects are similar) in Phase I (examination of seemingly discrepant markets) and Phase II (examination of seemingly discrepant single year's observations among otherwise compatible markets) of sensitivity analysis.

A useful instrument for identifying the existence of discrepant data points is the "4x" or "hat" matrix.<sup>2</sup> Recalling that the estimated response variables in the general linear model are

$$\hat{Y} = X\hat{\beta} \quad (5.8)$$

and that

$$\hat{\beta} = (X^T X)^{-1} X^T Y \quad (5.6)$$

then

$$\hat{Y} = H Y \quad (5.10)$$

where

$$H = X(X^T X)^{-1} X^T \quad (5.11)$$

The  $n$  by  $n$  matrix  $H$  is the so-called "hat" matrix. It is a function only of the carrier variables and defines the linear combination of the carriers by which the actual values of the response variable  $Y$  are translated into the estimated values of the response variable  $\hat{Y}$ .

A generic element of the hat matrix,  $h_{ij}$ , when compared to the

---

<sup>2</sup> David C. Hoaglin and Roy E. Welsch, "The Hat Matrix in Regression and ANOVA" (National Bureau of Economic Research Working Paper WP 901-77, January, 1977).

other values in row  $i$  of the hat matrix, measures the relative amount of influence or leverage that the observed value (of the response variable) of data point  $j$  will have upon the estimated value of data point  $i$  (regardless of the value of  $y$ , since  $H$  is a function only of  $X$ ). Of particular interest to the data analyst are the diagonal elements of the hat matrix,  $h_{ii}$ , which for simplicity will hereafter be noted as  $h_i$ . The concept of the hat matrix and its usefulness in sensitivity analysis is elaborated upon in Appendix B.

The situation of data point  $i$  being identified as a high leverage point (large value of  $h_i$ ) is neither a necessary nor a sufficient condition for a point to be labeled as being discrepant. In the sensitivity analysis of the coefficient estimation procedure of a general linear model, it is necessary to evaluate the impact of a high value of  $h_i$  in conjunction with other output statistics.

One such output statistic is the residual,  $r_i = y_i - \hat{y}_i$ , which may be conveniently expressed by dividing by its standard deviation to obtain the standardized residual,  $\tilde{r}_i$ .

$$\tilde{r}_i = \frac{r_i}{\hat{\sigma} \sqrt{1 - h_i}} \quad (5.12)$$

where  $\hat{\sigma}^2$  is the mean square residual.

However, residual analysis must be used with extreme caution. In many circumstances a discrepant data point may have a very small residual, while a data point that is truly representative of the population to which the model is to be fitted may have a very large residual! This



occurs when the discrepant data point has such high leverage that the estimated regression function is forced away from the representative points and through or near the discrepant point. A simple numerical example in Appendix B clearly illustrates this phenomenon.

A more powerful diagnostic tool is the studentized residual,  $r^*_i$ , which is the number of standard errors that the observed value  $y_i$  would lie from the estimated value  $\hat{y}_{(i)}$  if the model were fitted with data point  $i$  deleted from the sample.

$$r^*_i = \frac{y_i - x_i \hat{\beta}_{(i)}}{\hat{\sigma}_{(i)} \sqrt{1 + x_i (X_{(i)}^T X_{(i)})^{-1} x_i^T}} \quad (5.13)$$

where  $\hat{\beta}_{(i)}$  are the estimated coefficients with data point  $i$  deleted  
 $\hat{\sigma}_{(i)}^2$  is the residual mean square with data point  $i$  deleted  
 and  $X_{(i)}$  is the carrier matrix with data point  $i$  deleted

For an example of the usefulness of studentized and standardized residuals, the reader is again referred to the simple numerical example of Appendix B.

Another set of useful diagnostic statistics, if the analyst is concerned with coefficient stability, are the changes in coefficients by removing data point  $i$ ,  $\hat{\beta} - \hat{\beta}_{(i)}$ . If the analyst is concerned with the fit of the equation, he or she may wish to examine the scaled change in fit

due to elimination of data point  $i$ , DFFITS, defined as follows:<sup>3</sup>

$$\text{DFFITS} = \left( \frac{n-p}{p} \frac{h_i}{1-h_i} r_i^* \right)^{1/2} \quad (5.14)$$

Deletion of data points from a sample in order to improve the summary statistics is an extremely risky proposition with regard to maintaining the integrity of the research. A naive analyst may feel smug about removing an observation that has a large residual so as to improve the fit. However, although the summary statistics (standard error of estimate,  $R^2$ , etc.) will imply an improvement, the resulting equation may be further removed from the goal of the research -- an equation or set of equations that is representative of the population over which the model is assumedly being calibrated.

A major objective of this research is to develop a series of statistically robust models. The estimation process should not be sensitive to one or more discrepant data point(s) which will significantly alter the output resulting in a set of equations which is misrepresentative of the underlying population. Equivalently, the sample used in the final calibration of any of the models should not contain a data point which, if removed, would substantially affect the estimates. This is sometimes referred to as the "little a lot" criterion; if a little bit of

---

<sup>3</sup> "TROLL Experimental Programs: Model Sensitivity Analysis" (Cambridge: Computer Research Center, National Bureau of Economic Research, 1976), p. 27. Many other diagnostic statistics are described in this document.

the data is perturbed a lot (or deleted) the estimates should not exhibit substantial variation. Furthermore, the models should satisfy the "lot a little" criterion; if a lot of the data is perturbed a little, the estimates should show similar stability.

With the sometimes-conflicting objectives of research integrity and statistical robustness in mind, the following ground rules for data deletion are laid. A data point, regardless of its apparent ugliness in the sample, must not be deleted unless it satisfies one of the following criteria:

1. The market has obviously been incorrectly grouped (e.g., a market which is more representative of long haul/medium size is in the long haul/large size sample), violating the assumption of homogeneous aggregation.
2. The data point is anomalous, due to an obvious cause not accounted for in the model (e.g., an airline strike).
3. The data point is anomalous for unobvious reasons and its inclusion is causing a major fit problem with many of the other observations (e.g., clustering of residuals).

A data point will not be jettisoned purely because it "improves the fit".

As previously mentioned, the sensitivity analyses were generally conducted in two phases. Phase I searches for, and, if deemed appropriate by the first criterion above, deletes entire markets from the sample. Phase II then involves an investigation of the output statistics from the estimation of the reduced (by Phase I) sample, and searches for individual data points that appear discrepant by criteria 2

and 3 above. The effects of deleting these suspect points are then investigated and a judgmental decision is made as to which points should remain and which should be jettisoned.

This lengthy process is, for illustrative purposes, documented in detail in the calibration of first of the nine stratifications (large long markets). For simplicity, the procedure is merely summarized for the remaining calibrations.

The problem with the disparate values of the coefficient estimates and their judgmental priors was suspected to be an ill effect of multicollinearity. Multicollinearity, a very common problem in regression analysis, is a situation where two or more columns of the X matrix are nearly linearly dependent. For example, if two carrier variables are highly correlated, their respective columns in the X matrix would exhibit this near-dependency condition, and the collinearity problem would exist.<sup>4</sup>

The example above of two variables being highly correlated can easily be diagnosed by observing the off diagonal values of the variance-covariance matrix for X or those of the correlation matrix of X. However, more complex near-dependencies involving more than two columns of X may not be evident by observation of the variance-covariance or correlation matrices. For a simple example of this higher order effect, see Appendix C, which contains a more elaborate discussion of

---

<sup>4</sup> The terms multicollinearity, collinearity, near dependency, and near singularity are used interchangeably throughout the thesis.

multicollinearity.

A more effective diagnostic signal for the presence of collinearity is the  $p$  by  $1$  vector of eigenvalues of the matrix  $X^T X$ . If near singularities do exist in the  $X$  matrix, one or more of the eigenvalues of the cross product matrix will be small relative to the largest eigenvalue.<sup>5</sup> The  $i^{\text{th}}$  condition index of  $X$ ,  $\kappa_i(X)$ , is defined as follows:

$$\kappa_i(X) = \frac{\sigma_1}{\sigma_i} = \frac{\sqrt{\lambda_1}}{\sqrt{\lambda_i}} \quad (5.15)$$

where  $\sigma_i$  = the  $i^{\text{th}}$  largest singular value of  $X$   
and  $\lambda_i$  = the  $i^{\text{th}}$  largest eigenvalue of  $X^T X$ .<sup>6</sup>

The condition number of  $X$ ,  $\kappa(X)$ , is the largest condition index,

$$\kappa(X) = \kappa_p(X) = \frac{\sigma_1}{\sigma_p} \quad (5.16)$$

Generally speaking, the higher the condition number (and condition indices), the greater is the severity of the multicollinearity (see Appendix C).

A major ill effect of multicollinearity is high standard errors of coefficient estimates causing the estimates to be unstable. If the analyst is principally concerned with fit and not so much with individual parameter estimates, multicollinearity may not be too

<sup>5</sup> If one or more exact dependencies exist, the  $X^T X$  matrix will not be of full rank ( $p$ ) and one or more eigenvalue(s) will be equal to zero.

<sup>6</sup> For a more detailed discussion of singular values and eigenvalues, see Appendix C.

consequential. However, for policy analysis it is of utmost importance to produce coefficient estimates that are both intuitively reasonable and are precise (low standard errors). In this sense, the problem of collinearity is indeed a grave concern.

The technique used to correct for the multicollinearity (which was, in fact, detected as the problem) is principal components analysis, and is described in detail in Appendix C. In principal components analysis, the  $p$  dimensional vector space of  $X$  is rotated to the "principal components basis". In the principal components basis, the observations are linear combinations of the original columns of  $X$ , but by definition of the rotation are orthogonal.

As is shown in Appendix C, if small eigenvalues of  $X^T X$  exist, in the principal components basis the components corresponding to these small values are relatively unimportant and may be eliminated. Calibrating the model in the principal components basis using the remaining components and rotating back to the original basis produces a set of estimates  $\hat{\beta}_{pc}$  which may have significantly higher precisions (lower standard errors).

Applying principal components analysis to the ill-conditioned (collinear) data had, in most cases, three impacts:

1. The point estimates of the coefficients became more intuitively reasonable.
2. The precision of the estimates increased substantially (standard errors of coefficient estimates reduced).
3. The fit was poorer (standard error of estimate increased,  $R^2$  decreased, etc.).

Therefore, for each cross-classification of the data, the estimation of the parameters consists of a two-step process. The first step is to calibrate the demand equation using ordinary least squares and then to reduce the sample using sensitivity analysis techniques. The second step is to calibrate the demand equation using principal components on the reduced sample. While the first equation generally produces a better fit, the second provides more reasonable and much more precise estimates of the coefficients. Since the purpose of this research is to generate a set of structural equations for policy analysis, the demand equations estimated using principal components are generally preferred over those obtained by ordinary least squares.

### 5.1 Pooling of Entire Data Set

Throughout the preceding chapters there have been references to the fact that pooling data containing observations over all lengths of haul to calibrate a single equation or set of equations is inappropriate. The inappropriateness is due to the very different market characteristics between length of haul strata (i.e., the model coefficients are functions of length of haul). This effect was a major conclusion of Marfisi's empirical work<sup>7</sup> and is also discussed in detail in Blumer's thesis.<sup>8</sup>

---

<sup>7</sup> E. Pat Marfisi, "Theory and Evidence on the Behavior of Airline Firms Facing Uncertain Demand" (Ph.D. Dissertation, Brown University, 1976). See also the discussion in Section 2.5.

<sup>8</sup> Terry P. Blumer, "A Short Haul Passenger Demand Model for Air Transportation" (S.M. Thesis, M.I.T., 1976).

This fact is verified using the data collected for the present research. If the data could be pooled over lengths of haul, a test of the following null hypothesis would be accepted:

$$H_0 : \beta_{1j}^l = \beta_{1j}^m = \beta_{1j}^s \quad j = 0, 1, 2, 3 \quad (5.17)$$

where the  $\beta_{1j}$  parameters are the coefficients of equation (5.3) and the superscripts refer to long, medium, and short lengths of haul. This hypothesis is testable by Chow's technique.<sup>9</sup>

Four separate regression analyses were conducted: all data, long, medium, and short lengths of haul. The results are shown in Figure 5.1. The test statistic,  $F$ , for the Chow test, as shown in Figure 5.2, is equal to 35.3, which greatly exceeds the critical value, 2.53 at the one percent level of significance.

The conclusion to this is, as expected, that indeed pooling data over length of haul is not appropriate. The implication of this conclusion is that at least one length of haul stratum has different market characteristics (demand equation parameters) than the other two, and it is possible that all three are mutually different. Pairwise Chow tests could be conducted to verify the latter case, but it was assumed a priori that indeed each is different, and so separate models will be

---

<sup>9</sup> Gregory Chow, "Tests of Equality Between Subsets of Coefficients in Two Linear Regressions" (Econometrica, Vol. 28, July, 1960).



Figure 5.1. Estimates of Demand Equation Parameters for All, Long Haul, Medium Haul, and Short Haul Markets

ALL MARKETS			
<u>Carrier</u>	<u>Coefficient</u>	<u>Standard Error</u>	<u>t-ratio</u>
Constant	14.6	0.283	51.7
Level of Service	4.38	0.0651	67.3
Fare	-1.13	0.0217	-52.0
Socio-Economics	0.171	0.0254	6.73

n = 820                       $R^2 = 0.944$                        $F(8/816) = 4610$   
 std. error = 0.339                       $R^2_{adj} = 0.944$                       SSR = 93.8

LONG HAUL MARKETS			
<u>Carrier</u>	<u>Coefficient</u>	<u>Standard Error</u>	<u>t-ratio</u>
Constant	15.2	0.375	40.5
Level of Service	4.15	0.0900	46.2
Fare	-1.41	0.0514	-27.4
Socio-Economics	0.238	0.0372	6.39

n = 232                       $R^2 = 0.980$                        $F(3/228) = 3730$   
 std. error = 0.182                       $R^2_{adj} = 0.980$                       SSR = 7.56

Figure 5.1 (continued)

## MEDIUM HAUL MARKETS

<u>Carrier</u>	<u>Coefficient</u>	<u>Standard Error</u>	<u>t-ratio</u>
Constant	13.9	0.429	32.5
Level of Service	4.07	0.0751	54.2
Fare	-1.07	0.0670	-15.9
Socio-Economics	0.207	0.0311	6.67

$n = 283$                        $R^2 = 0.971$                        $F(3/279) = 3140$   
 $\text{std error} = 0.217$                        $R^2_{\text{adj}} = 0.971$                        $\text{SSR} = 13.2$

## SHORT HAUL MARKETS

<u>Carrier</u>	<u>Coefficient</u>	<u>Standard Error</u>	<u>t-ratio</u>
Constant	10.6	0.606	17.6
Level of Service	4.23	0.115	36.9
Fare	-0.182	0.0973	-1.87
Socio-Economics	0.265	0.0416	6.38

$n = 305$                        $R^2 = 0.942$                        $F(3/301) = 1601$   
 $\text{std error} = 0.402$                        $R^2_{\text{adj}} = 0.041$                        $\text{SSR} = 48.6$

Figure 5.2. Chow Test for Pooling Markets by Length of Haul

Pooled Sample: All markets

- Subsamples:
1. Long haul markets
  2. Medium haul markets
  3. Short haul markets

Total number of observations:  $n = 820$

Number of estimated parameters:  $p = 4$

Number of subsamples:  $k = 3$

$$SSR_{\text{pooled}} = 93.8$$

$$SSR_l = 7.56$$

$$SSR_m = 13.2$$

$$SSR_s = \underline{48.6}$$

$$SSR_{\text{ind}} = 69.4$$

$$F = \frac{\frac{SSR_{\text{pooled}} - SSR_{\text{ind}}}{p(k-1)}}{\frac{SSR_{\text{ind}}}{n - pk}} = \frac{\frac{24.4}{8}}{\frac{69.4}{808}} = 35.5$$

$$F_{\text{crit}} = (8, 808, 0.01) = 2.53$$

calibrated for each length of haul grouping. As will be indicated by the results that follow in this chapter, the coefficient estimates do vary dramatically by length of haul, so this assumption appears to have been very reasonable.

## 5.2 Analysis of Long Haul Markets

The experimental design described in Section 4.2 stratifies markets not only by length of haul, but also by level of socio-economic activity. The first order of business in the analysis of long haul markets was to determine whether pooling over levels of socio-economic activity (within the length of haul grouping) was appropriate. The method of analysis is identical to that of the previous section, where a Chow test was used to ascertain that pooling over length of haul is unacceptable.

The markets were separated into three socio-economic strata -- large, medium, and small. The assignment procedure was somewhat arbitrary. Using the socio-economic labels of Section 4.2, the assignment is as follows:

Large socio-economic: 3-4, 3-5, 4-4, 4-5, 5-5

Medium socio-economic: 2-2, 2-3, 2-4, 2-5, 3-3

Small socio-economic: 1-1, 1-2, 1-3, 1-4, 1-5

It will be discovered later in this section that this arbitrary assignment was, at least for long haul markets, not as representative

as had been hoped, and corrective measures will be applied.

Four separate regression analyses were conducted: all long haul, long/large, long/medium, and long/small, markets.<sup>10</sup> The results of these estimation procedures are shown in Figure 5.3.<sup>11</sup> The Chow test statistic, as shown in Figure 5.4, is equal to 3.29, which exceeds the critical value of F, 2.59, at the one percent level of significance. It is therefore concluded that the characteristics of long haul markets vary by size of market (as measured by socioeconomic levels). The implication of this is that separate long haul models must be estimated for large, medium, and small demographic region pairs.

Unlike the case of pooling markets by length of haul, the economic justification of why market characteristics would vary by market size was at first quite puzzling. Verleger<sup>12</sup> has analyzed the variability of fare elasticities by market density<sup>13</sup> and concluded that the variance of the price coefficient decreases as traffic increases. However, as was

---

<sup>10</sup> "Long/large" means "long haul/large socio-economic", etc.

<sup>11</sup> Comparing the coefficient estimates of the equation for all long haul markets in Figure 5.1 to those of Figure 5.3, an inconsistency is observed. This is due to different level of service estimates. The level of service estimates in the former case were extracted from the service equation calibrated over the entire data set. The level of service estimates in the latter case are from the service equation calibrated over the long haul markets.

<sup>12</sup> Philip K. Verleger, "A Point-to-Point Model of the Demand for Air Transportation" (Ph.D. Dissertation, M.I.T., 1971).

<sup>13</sup> The analysis of Chapter III verifies the strong relationship between socioeconomic activity and (potential) market density.

Figure 5.3 Estimates of Demand Equation Parameters for Long, Long/Large, Long/Medium, and Long/Small Markets

ALL LONG			
<u>Carrier</u>	<u>Coefficient</u>	<u>Standard Error</u>	<u>t-ratio</u>
Constant	24.0	0.567	42.4
Level of Service	7.21	0.165	43.7
Fare	-3.01	0.0615	48.9
Socio-Economics	0.271	0.0385	7.04

n = 232                       $R^2 = 0.978$                        $F(3/228) = 3380$   
 std. error = 0.191               $R^2_{adj} = 0.978$                       SSR = 8.32

LONG/LARGE			
<u>Carrier</u>	<u>Coefficient</u>	<u>Standard Error</u>	<u>t-ratio</u>
Constant	21.7	0.947	23.0
Level of Service	6.78	0.262	25.9
Fare	-2.79	0.0979	-28.5
Socio-Economics	0.391	0.0587	6.66

n = 120                       $R^2 = 0.970$                        $F(3/116) = 1250$   
 std. error = 0.182               $R^2_{adj} = 0.969$                       SSR = 3.86

Fig. 5.3 (continued)

LONG/MEDIUM			
<u>Carrier</u>	<u>Coefficient</u>	<u>Standard Error</u>	<u>t-ratio</u>
Constant	25.5	1.08	23.6
Level of Service	6.92	0.441	16.8
Fare	-3.22	0.0996	-32.3
Socio-Economics	0.190	0.0847	2.25
n = 66	$R^2 = 0.957$	$F(3/62) = 462$	
std. error = 0.154	$R^2_{adj} = 0.955$	$SSR = 1.47$	

LONG/SMALL			
<u>Carrier</u>	<u>Coefficient</u>	<u>Standard Error</u>	<u>t-ratio</u>
Constant	24.6	1.07	23.1
Level of Service	7.09	0.328	21.6
Fare	-3.05	0.206	-14.8
Socio-Economics	0.208	0.102	2.04
n = 46	$R^2 = 0.950$	$F(3/42) = 267$	
std. error = 0.223	$R^2_{adj} = 0.947$	$SSR = 2.10$	

Figure 5.4 Chow Test for Pooling Long Haul Markets by Level of  
Socio-Economic Activity

Pooled Sample: Long haul markets

Subsamples: 1. Long/large

2. Long/medium

3. Long/small

Total number of observations:  $n = 232$

Number of estimated parameters:  $p = 4$

Number of subsamples:  $k = 3$

$$SSR_{\text{pooled}} = 8.32$$

$$SSR_{1l} = 3.86$$

$$SSR_{1m} = 1.47$$

$$SSR_{1s} = \underline{2.10}$$

$$SSR_{\text{ind}} = 7.43$$

$$F = \frac{\frac{SSR_{\text{pooled}} - \sum SSR_i}{p(k-1)}}{\frac{\sum SSR_i}{n-pk}} = \frac{\frac{0.89}{8}}{\frac{7.43}{220}} = 3.29$$

$$F_{\text{crit}}(8, 220, 0.01) = 2.59$$



discussed in Section 2.4, the validity of his analysis was discounted due to specification error (omission of a service variable). The justification of this heterogeneity across market size will, however, become clear after the separate models have been calibrated and analyzed.

### 5.2.1 Large Long Haul Markets

#### 5.2.2.1 Preliminary Analysis

The results of the parameter estimation for the service and demand equations for the long/large markets are provided in Figure 5.5. In the service equation, all coefficients have the expected sign and are statistically significant,<sup>14</sup> with the exception of the competition coefficient which has the wrong sign and is insignificant. This falling out of the competition variable implies that the effect discussed in Section 3.3.2 of greater service in highly competitive routes is not present in this particular sample.

This outcome is not totally surprising, considering that high density long haul markets are generally very profitable. Even if there are as

---

<sup>14</sup> Statistical significance in this case is measured by the coefficient t-ratios. Throughout the remainder of this thesis any allusion to "statistical significance" implies a one percent level of significance, unless otherwise stated.

Figure 5.5 Preliminary Estimates of Service and Demand Equation  
Parameters for Large Long Haul Markets Before Principal  
Component Deletions

SERVICE EQUATION			
<u>Carrier</u>	<u>Coefficient</u>	<u>Standard Error</u>	<u>t-ratio</u>
Constant	-2.82	0.111	-25.5
Lagged Traffic	0.110	0.00574	19.1
Fare	0.297	0.0222	13.4
Competition	-0.0282	0.0171	-1.65

n = 120                       $R^2 = 0.844$                       F(3/116) = 209  
 std. error = 0.0643         $R^2_{adj} = 0.840$                       SSR = 0.480

DEMAND EQUATION			
<u>Carrier</u>	<u>Coefficient</u>	<u>Standard Error</u>	<u>t-ratio</u>
Constant	18.8	0.903	20.9
Level of Service	7.14	0.296	24.1
Fare	-2.22	0.0886	-25.0
Socio-Economics	0.428	0.0615	6.97

n = 120                       $R^2 = 0.966$                       F(3/116) = 1110  
 std. error = 0.194         $R^2_{adj} = 0.965$                       SSR = 4.35

few as two competitors in any given one of these markets, providing a high level of service will likely be in the individual carrier's best interests. There are no monopoly markets in this subset.

The primary measure of goodness of fit of a model to be used for forecasting purposes is the standard error of estimate (abbreviated to "std.error" in Figure 5.5. and all other figures that present model estimations). The standard error for the preliminary estimation of the demand equation for large long haul markets is 0.194. Since this model is to be used for forecasting purposes, the values of the standard error will be closely observed during the fine tuning of the model.

Based upon perusal of the preliminary calibration, two immediate problems with the demand equation have been encountered. The first of these is evident by observation of the residual plot in Figure 5.6. Note that the residuals of many of the markets tend to cluster above (underestimation) or below (overestimation) the axis. Of the twenty markets, five (SAN-SEA, MIA-SEA, HOU-WAS, CHI-LAX, and NYC) are consistently underestimated and four (DAL-PDX, DAL-SEA, MKC-NYC, and STL-SFO) are consistently overestimated.

The residual of a given observation is the difference between the observed (actual) value of the response variable (in this case the logarithm of demand) and the value that the model would predict. As was mentioned in Section 5.1, the residuals are assumed to be independent and identically distributed. The apparent clustering of residuals within markets in Figure 5.6 implies that this assumption has been violated and casts doubts upon the forecasting accuracy of

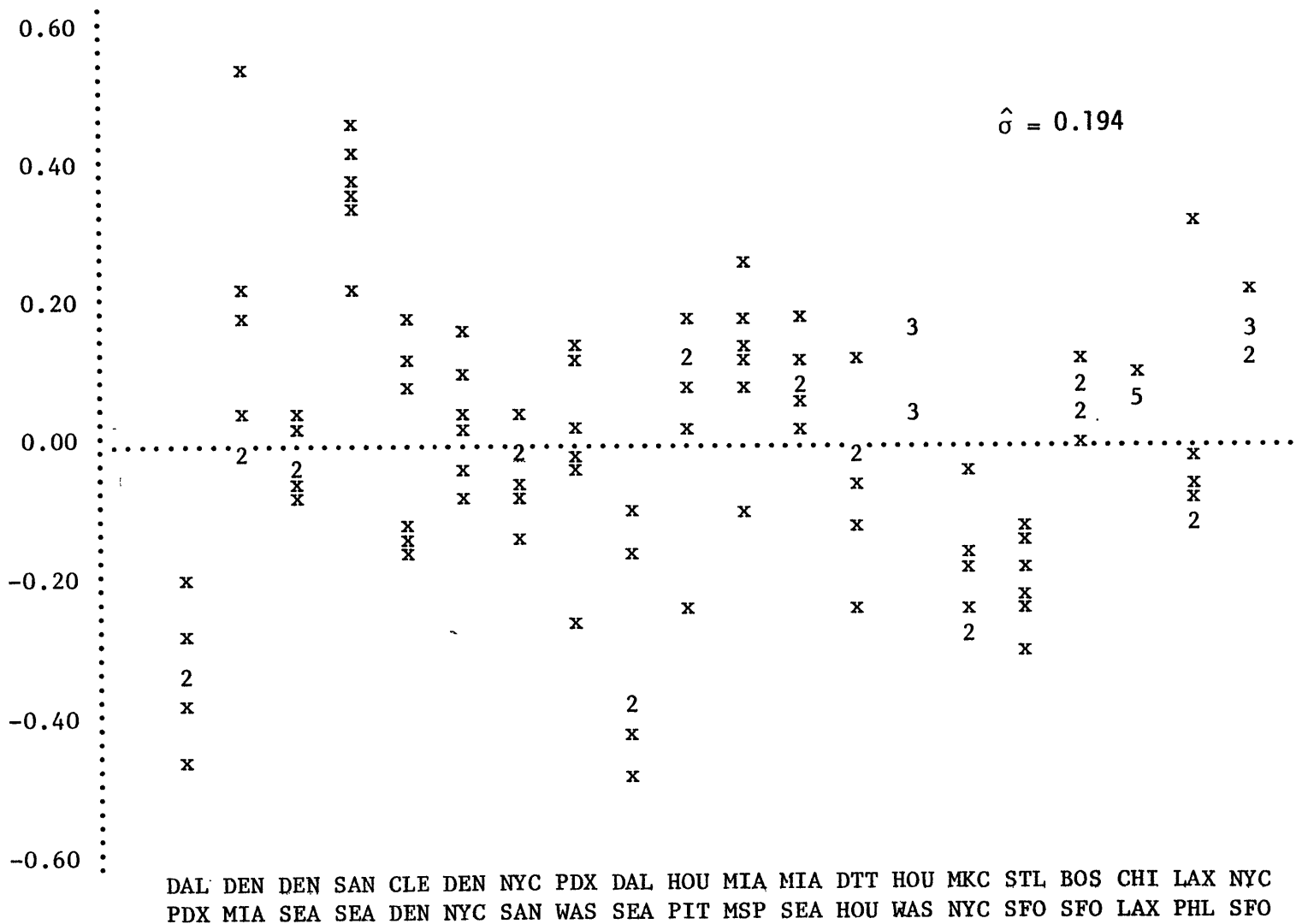


Figure 5.6 Plot of Residuals of Preliminary Estimation of the Demand Equation for Large Long Haul Markets

the model. This serious problem will also be closely scrutinized as the model is fine tuned. Possible causes are misspecification, improper aggregation, and/or the presence of high leverage discrepant observations.

The second immediate problem with the demand model concerns the coefficient estimates themselves. Judgmental estimates of the values of the parameters place them in the following ranges:

$$0.25 < \beta_{11} < 0.60 \quad (\text{service effect})$$

$$-0.90 > \beta_{12} > -1.70 \quad (\text{price elastic})$$

$$1.00 < \beta_{13} < 2.00 \quad (\text{also income elastic})$$

None of the coefficient estimates fall within (or even near) these intervals. This problem does not necessarily endanger the aggregate forecasting integrity of the model (provided that the standard error of estimate is sufficiently low). However, for the model to be used as a policy analysis tool, it is imperative that the coefficient estimates are reasonable and that the standard errors of the coefficient estimates are small.

The suspected culprit behind the coefficient problem is multicollinearity. Multicollinearity is the condition of near-dependencies among columns of  $X$ , the carrier matrix (in this case a column of ones, logs of estimated levels of service, log of fares, and logs of levels

of socio-economic activity). Multicollinearity is detected by observing small eigenvalues or small singular values in the  $X^T X$  matrix.

Principal components analysis (see Appendix C) was employed to detect the presence of multicollinearity and to attempt to correct for it. Summary statistics of the principal components are presented in Table 5.1.

The condition number (ratio of highest to lowest singular value) of 540 indicates a high likelihood that indeed multicollinearity exists (see Appendix C). After deleting the third and fourth principal components, estimating the regression of the log of demand upon the plane defined by the first two principal components, and projecting the results back into the original  $\hat{\beta}$  basis, the results shown in Figure 5.7 are obtained.

A comparison of the results of Figure 5.7 to those of the demand equation estimates presented in Figure 5.5 yields three important results. The coefficient estimates in the post-principal components model make much greater intuitive sense (comparing them to the judgmental intervals). The standard errors of the coefficients in the post-principal components estimation are greatly reduced. The standard error of estimate is increased (more than doubled), and the  $R^2$  value has decreased in the post-principal components equation.

The decision as to which of the two estimated equations constitutes the "better" model involves a tradeoff between fit and parameter reasonableness and precision. The goal of this analysis is to develop a model which will accurately measure the impact upon demand of changes in such factors as level of service and fare. Since little faith can be placed

Table 5.1 Principal Component Analysis, Long Large Markets

<u>Principal Component</u>	<u>Singular Value</u>	<u>Condition Index</u>	<u>t-statistic</u>
1	110.	1.00	-512.
2	3.23	34.1	-30.1
3	1.77	62.3	-7.67
4	0.204	540.	21.2

Figure 5.7 Preliminary Estimates of the Demand Equation Parameters  
for Large Long Haul Markets After Principal Component  
Deletions

<u>Carrier</u>	<u>Coefficient</u>	<u>Standard Error</u>	<u>t-ratio</u>
Constant	-0.124	0.00708	-17.5
Level of Service	0.467	0.0167	27.9
Fare	-1.07	0.0496	-21.5
Socio-Economic	1.64	0.0285	57.7

n = 120

std. error = 0.446

$R^2 = 0.818$

SSR = 23.5



upon the estimates of the coefficients of the model before principal components deletion, the post-principal components equation is clearly the appropriate choice for the resulting model.

The problem of clustered residuals must still be resolved. It is also an intention to reduce the standard error of estimate in the forecasting equation and possibly to reduce the standard errors of coefficient estimates in the analysis equation. These are the objectives of the following two sections.

#### 5.2.2.2 Sensitivity Analysis: Phase I

Two of the possible causes of the residual clustering problem noted in the previous section are improper aggregation of data and discrepant high leverage single observations. These two problems and their ill effects upon estimation are in reality one and the same, but the former terminology will be used to refer to entire markets that are improperly included in a sample, while the latter will refer to a discrepant single year's observation among an otherwise satisfactory market.

As was discussed earlier in the chapter, a tool for investigating the presence of improper pooling or single discrepant points is the hat matrix,  $H$ . The rows of the hat matrix are the weightings by which the individual observations of the response variable are combined to obtain the estimated values. That is,

$$\hat{Y} = HY \quad (5.10)$$

where

$$H = X(X^T X)^{-1} X^T \quad (5.11)$$

A rule of thumb states that, if a diagonal element of the hat matrix,  $h_i$ , is greater than  $2p/n$ , where  $p$  is the number of estimated parameters and  $n$  is the number of observations, then that data point  $i$  is a suspect high leverage point. A more formal rule states that if

$$h_i > \frac{\frac{p-1}{n-p} F(p-1, n-p) + \frac{1}{n}}{1 + \frac{p-1}{n-p} F(p-1, n-p)} \quad (5.18)$$

where  $F(p-1, n-p)$  is the critical  $F$  value with  $p-1$  degrees of freedom in the numerator and  $n-p$  degrees of freedom in the denominator, that point is suspect. The  $2p/n$  rule of thumb will be used as an approximate measure throughout this research, since in all cases studied herein, its value is less than that of the more formal rule (at the 5% level of significance).

Table 5.2 lists the diagonal elements for the hat matrix for  $X$ , the carrier matrix for the demand equation calibrated over large long markets. Two suspicious data points ( $h_i > 2p/n = 0.0667$ ) were discovered in the Denver-Seattle market and one in the Miami-Seattle market. Furthermore, the other  $h_i$  values for these markets are quite high.

Table 5.3 lists the studentized residuals of the preliminary calibration of the demand equation over large long markets. Recall that the studentized residual for data point  $i$ ,  $r^*_i$ , is the

Table 5.2 Diagonal Elements of the Hat Matrix, Demand Equation for  
Large Long Markets, Phase I of Sensitivity Analysis

DAL PDX 69	0.0256	PDX WAS 70	0.0394	MKC NYC 71	0.0420
DAL PDX 70	0.0250	PDX WAS 71	0.0352	MKC NYC 72	0.0523
DAL PDX 71	0.0221	PDX WAS 72	0.0322	MKC NYC 73	0.0624
DAL PDX 72	0.0189	PDX WAS 73	0.0266	MKC NYC 74	0.0561
DAL PDX 73	0.0174	PDX WAS 74	0.0303	STL SFO 69	0.0136
DAL PDX 74	0.0160	DAL SEA 69	0.0299	STL SFO 70	0.0152
DEN MIA 69	0.0329	DAL SEA 70	0.0333	STL SFO 71	0.0143
DEN MIA 70	0.0244	DAL SEA 71	0.0377	STL SFO 72	0.0099
DEN MIA 71	0.0257	DAL SEA 72	0.0257	STL SFO 73	0.0092
DEN MIA 72	0.0177	DAL SEA 73	0.0167	STL SFO 74	0.0090
DEN MIA 73	0.0162	DAL SEA 74	0.0190	BOS SFO 69	0.0297
DEN MIA 74	0.0159*	HOU PIT 69	0.0446	BOS SFO 70	0.0300
DEN SEA 69	0.0668	HOU PIT 70	0.0286	BOS SFO 71	0.0298
DEN SEA 70	0.0664	HOU PIT 71	0.0278	BOS SFO 72	0.0254
DEN SEA 71	0.0681	HOU PIT 72	0.0334	BOS SFO 73	0.0247
DEN SEA 72	0.0618	HOU PIT 73	0.0377	BOS SFO 74	0.0274
DEN SEA 73	0.0586	HOU PIT 74	0.0310	CHI LAX 69	0.0414
DEN SEA 74	0.0532	MIA MSP 69	0.0126	CHI LAX 70	0.0409
SAN SEA 69	0.0460	MIA MSP 70	0.0115	CHI LAX 71	0.0398
SAN SEA 70	0.0480	MIA MSP 71	0.0113	CHI LAX 72	0.0408
SAN SEA 71	0.0411	MIA MSP 72	0.0109	CHI LAX 73	0.0425
SAN SEA 72	0.0382	MIA MSP 73	0.0118	CHI LAX 74	0.0424
SAN SEA 73	0.0397	MIA MSP 74	0.0144	LAX PHL 69	0.0297
SAN SEA 74	0.0410	MIA SEA 69	0.0553	LAX PHL 70	0.0313
CLE DEN 69	0.0272	MIA SEA 70	0.0614	LAX PHL 71	0.0322
CLE DEN 70	0.0165	MIA SEA 71	0.0618	LAX PHL 72	0.0363
CLE DEN 71	0.0230	MIA SEA 72	0.0660	LAX PHL 73	0.0590
CLE DEN 72	0.0195	MIA SEA 73	0.0652	LAX PHL 74	0.0399
CLE DEN 73	0.0181	MIA SEA 74	0.0683*	NYC SFO 69	0.0502
CLE DEN 74	0.0161	DTT HOU 69	0.0591	NYC SFO 70	0.0505
DEN NYC 69	0.0168	DTT HOU 70	0.0403	NYC SFO 71	0.0511
DEN NYC 70	0.0211	DTT HOU 71	0.0335	NYC SFO 72	0.0509
DEN NYC 71	0.0178	DTT HOU 72	0.0447	NYC SFO 73	0.0503
DEN NYC 72	0.0189	DTT HOU 73	0.0444	NYC SFO 74	0.0524
DEN NYC 73	0.0213	DTT HOU 74	0.0518		
DEN NYC 74	0.0213	HOU WAS 69	0.0180		
NYC SAN 69	0.0226	HOU WAS 70	0.0192		
NYC SAN 70	0.0235	HOU WAS 71	0.0159		
NYC SAN 71	0.0234	HOU WAS 72	0.0186		
NYC SAN 72	0.0207	HOU WAS 73	0.0187		
NYC SAN 73	0.0199	HOU WAS 74	0.0169		
NYC SAN 74	0.0220	MKC NYC 69	0.0575		
PDX WAS 69	0.0437	MKC NYC 70	0.0429		

$$* h_i > \frac{2p}{n} = 0.0667$$

Table 5.3 Studentized Residuals, Demand Equation for Large Long Markets,  
Phase II of Sensitivity Analysis

DAL PDX 69	-1.43	PDX WAS 71	-1.37	MKC NYC 73	-1.00
DAL PDX 70	-2.36 *	PDX WAS 72	0.11	MKC NYC 74	-1.47
DAL PDX 71	-1.84	PDX WAS 73	-0.15	STL SFO 69	-0.78
DAL PDX 72	-1.75	PDX WAS 74	0.65	STL SFO 70	-1.14
DAL PDX 73	-1.05	DAL SEA 69	-0.50	STL SFO 71	-1.54
DAL PDX 74	-2.01 *	DAL SEA 70	-0.87	STL SFO 72	-0.91
DEN MIA 69	2.98 *	DAL SEA 71	-2.09*	STL SFO 73	-1.32
DEN MIA 70	1.12	DAL SEA 72	-2.58*	STL SFO 74	-0.59
DEN MIA 71	0.20	DAL SEA 73	-1.94	BOS SFO 69	0.45
DEN MIA 72	0.97	DAL SEA 74	-2.22 *	BOS SFO 70	0.03
DEN MIA 73	-0.09	HOU PIT 69	1.01	BOS SFO 71	0.20
DEN MIA 74	-0.12	HOU PIT 70	-1.25	BOS SFO 72	0.65
DEN SEA 69	0.09	HOU PIT 71	0.36	BOS SFO 73	0.47
DEN SEA 70	0.21	HOU PIT 72	0.15	BOS SFO 74	0.24
DEN SEA 71	-0.21	HOU PIT 73	0.65	CHI LAX 69	0.35
DEN SEA 72	-0.17	HOU PIT 74	0.64	CHI LAX 70	0.34
DEN SEA 73	-0.38	MIA MSP 69	0.70	CHI LAX 71	0.29
DEN SEA 74	-0.34	MIA MPS 70	0.47	CHI LAX 72	0.32
SAN SEA 69	2.26*	MIA MSP 71	0.89	CHI LAX 73	0.34
SAN SEA 70	1.19	MIA MSP 72	0.66	CHI LAX 74	0.54
SAN SEA 71	1.87	MIA MSP 73	-0.51	LAX PHL 69	-0.44
SAN SEA 72	2.49*	MIA MSP 74	1.34	LAX PHL 70	-0.74
SAN SEA 73	1.97	MIA SEA 69	0.96	LAX PHL 71	-0.69
SAN SEA 74	1.92	MIA SEA 70	0.61	LAX PHL 72	-0.08
CLE DEN 69	0.67	MIA SEA 71	0.33	LAX PHL 73	1.65
CLE DEN 70	-0.67	MIA SEA 72	0.47	LAX PHL 74	-0.26
CLE DEN 71	0.88	MIA SEA 73	0.04	NYC SFO 69	0.61
CLE DEN 72	0.39	MIA SEA 74	0.43	NYC SFO 70	0.75
CLE DEN 73	-0.86	DTT HOU 69	0.62	NYC SFO 71	0.62
CLE DEN 74	-0.72	DTT HOU 70	-0.13	NYC SFO 72	0.70
DEN NYC 69	0.83	DTT HOU 71	-0.57	NYC SFO 73	0.77
DEN NYC 70	-0.39	DTT HOU 72	-0.10	NYC SFO 74	1.20
DEN NYC 71	0.27	DTT HOU 73	-1.16		
DEN NYC 72	0.55	DTT HOU 74	-0.34		
DEN NYC 73	0.08	HOU WAS 69	0.20		
DEN NYC 74	-0.19	HOU WAS 70	0.86		
NYC SAN 69	-0.15	HOU WAS 71	0.10		
NYC SAN 70	-0.29	HOU WAS 72	0.76		
NYC SAN 71	-0.57	HOU WAS 73	0.22		
NYC SAN 72	0.21	HOU WAS 74	0.87		
NYC SAN 73	-0.04	MKC NYC 69	-0.16		
NYC SAN 74	-0.19	MKC NYC 70	-0.94		
PDX WAS 69	0.79	MKC NYC 71	-1.24		
PDX WAS 70	-0.08	MKC NYC 72	-1.49		

\*  $r^*_i > 2.0$

(standardized) difference between the observed value of the response variable,  $y_i$ , and the estimated value of the response variable by the equation resulting from calibration over the sample with data point  $i$  deleted.

$$r^*_i = \frac{y_i - x_i \hat{\beta}(i)}{\hat{\sigma}(i) \sqrt{1 + x_i^T (X^T(i)X(i))^{-1} x_i}} \quad (5.13)$$

Since the numerator and denominator of (5.13) are independent,  $r^*_i$  is distributed according to a  $t$  distribution with  $n - p$  degrees of freedom. For the initial perusal of the studentized residuals, those with an absolute value of 2.0 or more are flagged as possible trouble points.

Upon joint investigation of the hat matrix diagonal elements and the studentized residuals, there appears to be a number of questionable points early in the sample. The Dallas-Portland market has high studentized residuals, as do Denver-Miami and San Diego-Seattle. Denver-Seattle exhibits high  $h_i$  values, as previously mentioned. These four markets are the "3-4" markets from the experimental design. These flags imply that perhaps the "3-4" markets, at least in the long haul, may better represent the medium rather than the large size markets. If indeed they are grouped into the wrong size classification, one would expect an improvement in the estimation of the large markets if they are deleted. Furthermore, one would expect little change in the estimation results for the medium size markets if they are added to that sample.<sup>15</sup>

---

<sup>15</sup> This indeed the result of the addition of the four markets to the medium size sample, as will be shown in Section 5.2.2.

The reason that the Miami-Seattle market is so influential may be because of its high value of the fare term. Since this is the longest market in the sample, it is also the highest priced. However, there is no economic reason to believe that this market should be removed from the sample. Furthermore, since neither its standardized nor studentized residuals were large, there were no statistical reasons for deletion of this market.

The implications of the first phase of the analysis were to delete the "3-4" markets from the large size classification and group them into the medium-size sample. One additional market was questionable. The Dallas-Seattle market has highly negative raw residuals and even more negative studentized residuals. While there were no obvious economic reasons for deleting this market, it was suspected that its presence may be contributing to the residual clustering of the other markets. Two regression estimations were conducted, one with only the "3-4" markets deleted and one with the Denver-Seattle market deleted also. While the deletion of the Denver-Seattle market had a negligible effect upon the fit, the residual plots indicated an improvement in the clustering problem. It was then decided to permanently delete this market from the sample.

The ordinary least squares estimates of the parameters of the supply and demand models for long large markets are shown in Figure 5.8. Comparing these results with those in Figure 5.5, it is observed that the fit in both the service and demand equations have both improved substantially with the standard error of estimate in the demand equation being reduced by nearly 40% to 0.119.

Figure 5.8 Phase I Estimates of the Parameters of the Service and Demand Equations for Large Long Markets before Principal Component Deletions

SERVICE EQUATION			
<u>Carrier</u>	<u>Coefficient</u>	<u>Standard Error</u>	<u>t-ratio</u>
Constant	-2.93	0.0990	-29.6
Lagged Traffic	0.113	0.00495	22.8
Fare	0.307	0.0203	15.2
Competition	-0.0107	0.0152	-0.708

n = 90                                       $R^2 = 0.914$                                       F(3/86) = 304  
 std. error = 0.0519                       $R^2_{adj} = 0.911$                                       SSR = 0.231

DEMAND EQUATION			
<u>Carrier</u>	<u>Coefficient</u>	<u>Standard Error</u>	<u>t-ratio</u>
Constant	20.1	1.14	17.7
Level of Service	7.28	0.306	23.8
Fare	-2.25	0.0853	-26.3
Socio-Economic	0.319	0.0752	4.25

n = 90                                       $R^2 = 0.989$                                       F(3/86) = 2470  
 std. error = 0.119                       $R^2_{adj} = 0.988$                                       SSR = 1.22

The residual plot of the Phase I demand equation calibration is shown in Figure 5.9. Comparing this to Figure 5.6, it is apparent that the clustering problem has become much less severe. Only one of the markets, New York-San Francisco, is consistently underestimated, and one, St. Louis-San Francisco, is overestimated. It can also be observed that the spread of the residuals has been substantially reduced.

The demand equation for large markets was calibrated using the reduced data set and principal component analysis. Again the third and fourth principal components were jettisoned, and the results of the estimation in the original  $\beta$  basis are given in Figure 5.10. Comparing these results to the preliminary estimation of the demand equation parameters (Figure 5.7) indicates a decrease in the standard errors of all coefficient estimates. The standard error of estimate of the constant decreased 47.5%, that of level of service coefficient decreased 40.7%, that of the fare elasticity decreased 28.2%, and that of the socio-economic activity decreased 30.2%.

#### 5.2.2.3 Sensitivity Analysis: Phase II

The diagonal elements of the hat matrix for the reduced data set used in Phase I are listed in Table 5.4, and the corresponding studentized residuals are listed in Table 5.5. One data point, Los Angeles-Philadelphia 1973, is flagged in both diagnostic sets. Suspecting a keypunch error, it was surprising to discover that the data point was



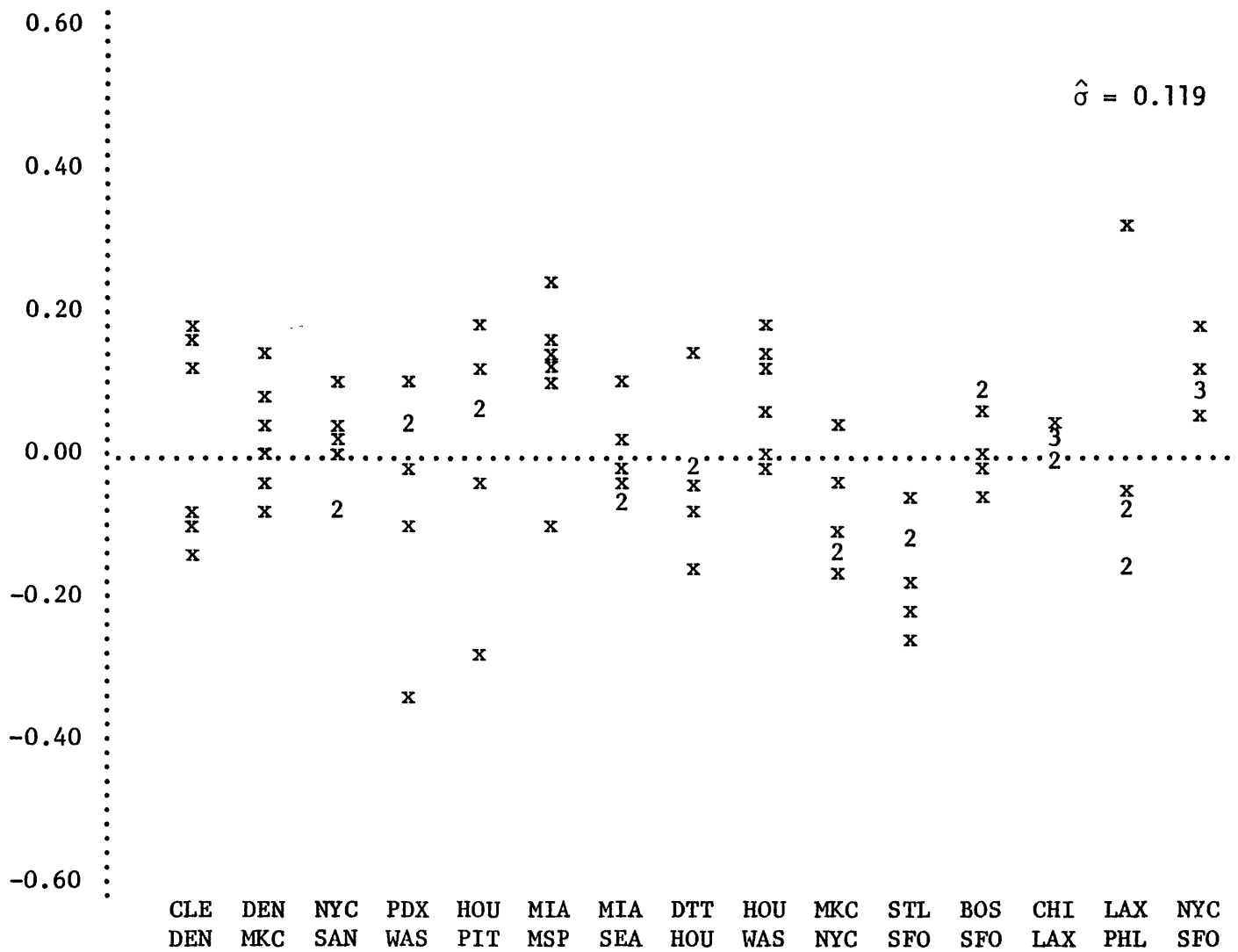


Figure 5.9 Plot of Residuals of Phase I Estimation of the Demand Equation for Large Long Markets

Figure 5.10 Phase I Estimates of the Demand Equation Parameters for  
Large Long Haul Markets After Principal Component Deletions

<u>Carrier</u>	<u>Coefficient</u>	<u>Standard Error</u>	<u>t-ratio</u>
Constant	-0.0881	0.00372	-23.7
Level of Service	0.434	0.00991	43.8
Fare	-1.26	0.0356	-35.4
Socio-Economics	1.73	0.0199	87.0

n = 90

std. error = 0.382

$R^2 = 0.879$

SSR = 12.9

Table 5.4 Diagonal Elements of the Hat Matrix. Demand Equation for Large Long Haul Markets, Phase II of Sensitivity Analysis

CLE DEN 69	0.0417	DTT HOU 70	0.0413	NYC SFO 71	0.0537
CLE DEN 70	0.0439	DTT HOU 71	0.0342	NYC SFO 72	0.0537
CLE DEN 71	0.0318	DTT HOU 72	0.0460	NYC SFO 73	0.0523
CLE DEN 72	0.0340	DTT HOU 73	0.0494	NYC SFO 74	0.0551
CLE DEN 73	0.0360	DTT HOU 74	0.0599		
CLE DEN 74	0.0308	HOU WAS 69	0.0287		
DEN NYC 69	0.0325	HOU WAS 70	0.0247		
DEN NYC 70	0.0424	HOU WAS 71	0.0272		
DEN NYC 71	0.0313	HOU WAS 72	0.0262		
DEN NYC 72	0.0270	HOU WAS 73	0.0283		
DEN NYC 73	0.0294	HOU WAS 74	0.0248		
DEN NYC 74	0.0294	MKC NYC 69	0.0650		
NYC SAN 69	0.0277	MKC NYC 70	0.0467		
NYC SAN 70	0.0273	MKC NYC 71	0.0501		
NYC SAN 71	0.0273	MKC NYC 72	0.0667		
NYC SAN 72	0.0264	MKC NYC 73	0.0883		
NYC SAN 73	0.0234	MKC NYC 74	0.0769		
NYC SAN 74	0.0245	STL SFO 69	0.0433		
PDX WAS 69	0.0548	STL SFO 70	0.0430		
PDX WAS 70	0.0531	STL SFO 71	0.0380		
PDX WAS 71	0.0596	STL SFO 72	0.0219		
PDX WAS 72	0.0427	STL SFO 73	0.0160		
PDX WAS 73	0.0364	STL SFO 74	0.0133		
PDX WAS 74	0.0381	BOS SFO 69	0.0338		
HOU PIT 69	0.0441	BOS SFO 70	0.0347		
HOU PIT 70	0.0317	BOS SFO 71	0.0333		
HOU PIT 71	0.0319	BOS SFO 72	0.0272		
HOU PIT 72	0.0335	BOS SFO 73	0.0259		
HOU PIT 73	0.0375	BOS SFO 74	0.0295		
HOU PIT 74	0.0310	CHI LAX 69	0.0515		
MIA MSP 69	0.0445	CHI LAX 70	0.0518		
MIA MSP 70	0.0407	CHI LAX 71	0.0490		
MIA MSP 71	0.0357	CHI LAX 72	0.0463		
MIA MSP 72	0.0261	CHI LAX 73	0.0467		
MIA MSP 73	0.0189	CHI LAX 74	0.0467		
MIA MSP 74	0.0169	LAX PHL 69	0.0437		
MIA SEA 69	0.0669	LAX PHL 70	0.0465		
MIA SEA 70	0.0735	LAX PHL 71	0.0515		
MIA SEA 71	0.0741	LAX PHL 72	0.0683		
MIA SEA 72	0.0773	LAX PHL 73	0.1578*		
MIA SEA 73	0.0883	LAX PHL 74	0.0805		
MIA SEA 74	0.0965*	NYC SFO 69	0.0525		
DTT HOU 69	0.0697	NYC SFO 70	0.0531		

$$* h_i > \frac{2p}{n} = 0.0889$$

Table 5.5 Studentized Residuals, Demand Equation for Large Long Haul  
Markets, Phase II of Sensitivity Analysis

CLE DEN 69	1.31	DTT HOU 70	-0.22	NYC SFO 71	0.58
CLE DEN 70	-0.86	DTT HOU 71	-0.81	NYC SFO 72	0.72
CLE DEN 71	1.57	DTT HOU 72	-0.15	NYC SFO 73	0.99
CLE DEN 72	0.94	DTT HOU 73	-1.42	NYC SFO 74	1.52
CLE DEN 73	-1.12	DTT HOU 74	-0.39		
CLE DEN 74	-0.70	HOU WAS 69	-0.02		
DEN NYC 69	1.21	HOU WAS 70	1.04		
DEN NYC 70	-0.62	HOU WAS 71	-0.09		
DEN NYC 71	0.29	HOU WAS 72	1.54		
DEN NYC 72	0.81	HOU WAS 73	0.49		
DEN NYC 73	0.03	HOU WAS 74	1.28		
DEN NYC 74	-0.38	MKC NYC 69	0.26		
NYC SAN 69	0.26	MKC NYC 70	-1.26		
NYC SAN 70	-0.54	MKC NYC 71	-1.00		
NYC SAN 71	-0.62	MKC NYC 72	-1.36		
NYC SAN 72	0.78	MKC NYC 73	-0.44		
NYC SAN 73	0.11	MKC NYC 74	-1.27		
NYC SAN 74	0.02	STL SFO 69	-1.01		
PDX WAS 69	0.82	STL SFO 70	-1.58		
PDX WAS 70	-0.78	STL SFO 71	-2.16*		
PDX WAS 71	-3.11*	STL SFO 72	-1.07		
PDX WAS 72	0.19	STL SFO 73	-1.75		
PDX WAS 73	-0.14	STL SFO 74	-0.65		
PDX WAS 74	0.36	BOS SFO 69	0.42		
HOU PIT 69	1.47	BOS SFO 70	-0.45		
HOU PIT 70	-2.33*	BOS SFO 71	-0.01		
HOU PIT 71	0.57	BOS SFO 72	0.74		
HOU PIT 72	-0.38	BOS SFO 73	0.55		
HOU PIT 73	0.95	BOS SFO 74	-0.12		
HOU PIT 74	0.52	CHI LAX 69	0.03		
MIA MSP 69	1.41	CHI LAX 70	-0.16		
MIA MSP 70	1.04	CHI LAX 71	-0.16		
MIA MSP 71	0.94	CHI LAX 72	0.13		
MIA MSP 72	1.14	CHI LAX 73	0.04		
MIA MSP 73	-0.78	CHI LAX 74	0.28		
MIA MSP 74	2.17*	LAX PHL 69	-0.72		
MIA SEA 69	0.82	LAX PHL 70	-1.37		
MIA SEA 70	-0.05	LAX PHL 71	-1.32		
MIA SEA 71	-0.38	LAX PHL 72	-0.50		
MIA SEA 72	-0.49	LAX PHL 73	2.93*		
MIA SEA 73	-0.57	LAX PHL 74	-0.68		
MIA SEA 74	0.15	NYC SFO 69	0.63		
DTT HOU 69	1.14	NYC SFO 70	0.75		

\*  $r^*_i > 2.0$

indeed entered correctly (assuming that the original data sources are accurate). Suspecting that this data point was contributing to the remaining amount of residual clustering, it was deleted and the equation was reestimated. The resulting residual plot indicated that the inclusion of this point was indeed detrimental to many of the remaining market residuals, and so it was deleted. This deletion also slightly improved the overall fit.

The only other point that was flagged in the hat matrix diagonal was Miami-Seattle 1974. Assuming, as in Phase I, that this is due to no other reason than the high fare value, the data point was left alone.

The other points flagged by the vector of studentized residuals were PDX-WAS 69, HOU-PIT 70, MIA-MSP 74, and STL-SFO 71. Individual deletions of each of these showed that with the exception of PDX-WAS 69 none of these was contributing to the residual clustering. The deletion of PDX-WAS 69 did, however, further reduce market residual clustering and was therefore removed. This deletion had a negligible effect upon the fit.

Phase II of the sensitivity analysis of large long haul resulted in the deletion of two data points, LAX-PHL 73 and PDX-WAS 69. The final ordinary least squares estimates of the service and demand model parameters are summarized in Figure 5.11. Note that the fit has improved slightly over the Phase I estimates (Figure 5.8) in both equations, and that the value of the standard error of estimate of the demand equation has been reduced to 0.109. However the values of the estimated parameters remain well outside of the reasonable intervals set in Section 5.2.2.1.

The residual plot of the Phase II ordinary least squares estimation

Figure 5.11 Phase II Estimates of Service and Demand Equation Parameters  
for Large Long Haul Markets

SERVICE EQUATION			
<u>Carrier</u>	<u>Coefficient</u>	<u>Standard Error</u>	<u>t-ratio</u>
Constant	-2.92	0.0958	-30.5
Lagged Traffic	0.112	0.00482	23.2
Fare	0.309	0.0197	15.6
Competition	-.0122	0.0147	-0.829

n = 88                       $R^2 = 0.920$                       F(3/84) = 321  
 std. error = 0.0500       $R^2_{adj} = 0.917$                       SSR = 0.210

DEMAND EQUATION			
<u>Carrier</u>	<u>Coefficient</u>	<u>Standard Error</u>	<u>t-ratio</u>
Constant	21.9	1.13	19.3
Level of Service	7.76	0.304	25.6
Fare	-2.36	0.0835	-28.2
Socio-Economics	0.204	0.0748	2.73

n = 88                       $R^2 = 0.990$                       F(3/85) = 2890  
 std. error = 0.109       $R^2_{adj} = 0.990$                       SSR = 0.997

of the demand equation is shown in Figure 5.12. While New York-San Francisco is still consistently underestimated and St. Louis-San Francisco is still consistently overestimated, most of the other markets' residual sets have become better centered about the zero line. Furthermore, the overall spread has been slightly reduced from the Phase I plot (Figure 5.9).

The results of the principal component regression analysis are presented in Figure 5.13. Comparing these statistics with those of the corresponding Phase I equation in Figure 5.10 indicates a further reduction of the standard errors of the coefficient estimates. It is interesting to notice that the fare and socio-economic elasticity estimates did not change from Phase I to Phase II, and the level of service parameter estimate varied by slightly over one percent. From a robustness standpoint, it is encouraging to see that the deletion of the two data points that appeared to be the most influential had a negligible effect upon the parameter estimates.

The elasticity estimates in the post-principal components demand equation, a fare elasticity of -1.26 and a socio-economic elasticity of 1.73, appear very reasonable. The level of service parameter estimate (0.429) appears also to be within a reasonable range. Although it is difficult to interpret this number by itself, it will be shown in Chapter VI that the demand vs. frequency curve generated using this figure appear to be quite representative of what one might expect in large long haul markets.

The estimated equation of Figure 5.13 is therefore accepted as the

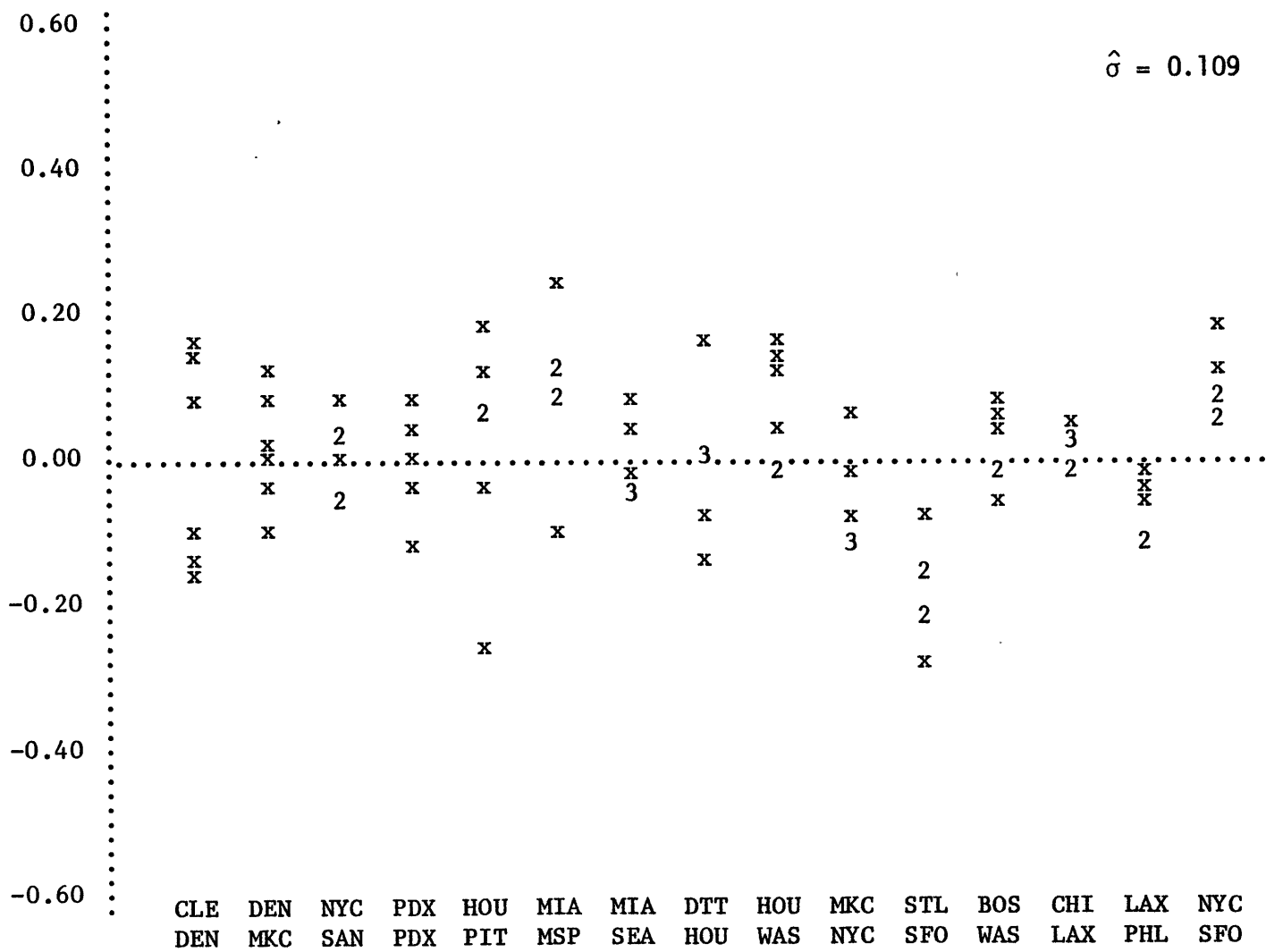


Figure 5.12 Plot of Residuals of Phase II Estimation of the Demand Equation for Large Long Haul Markets



Figure 5.13 Phase II Estimates of the Demand Equation Parameters for  
Large Long Haul Markets After Principal Component Deletions

<u>Carrier</u>	<u>Coefficient</u>	<u>Standard Error</u>	<u>t-ratio</u>
Constant	-0.0859	0.00343	-25.0
Level of Service	0.429	0.00917	46.8
Fare	-1.26	0.0333	-37.9
Socio-Economics	1.73	0.0186	93.1

n = 88	$R^2 = 0.877$
std. error = 0.386	SSR = 12.8

demand model for large long haul markets: The equation obtained using ordinary least squares estimation (Figure 5.11) must, in spite of its better fit, be rejected due to its questionable and relatively imprecise coefficient estimates.

### 5.2.2 Medium Size Long Haul Markets

In Phase I of the sensitivity analysis of long large markets, the smaller region pairs in that cross classification were deleted, suspecting that they may be more representative of medium size markets. The first step of the analysis of the medium size long haul markets was to determine if the inclusion of these rejects from the large markets into the medium sample affected the statistics produced by the regression analysis.

Equation (5.19) is the estimated demand equation using the original long/medium sample, and equation (5.20) is the estimate with the new data added.<sup>16</sup>

$$\begin{aligned} \text{LQD} = & 27.1 + 7.02 \text{ LLOS} - 4.00 \text{ LFARE} + 0.458 \text{ LSE} & (5.19) \\ & (1.81) \quad (0.664) \quad (0.184) \quad (0.109) \end{aligned}$$

$$n = 66 \qquad \text{std. error} = 0.217 \qquad R^2 = 0.915$$

$$\begin{aligned} \text{LQD} = & 22.1 + 5.39 \text{ LLOS} - 3.19 \text{ LFARE} + 0.465 \text{ LSE} & (5.20) \\ & (1.15) \quad (0.433) \quad (0.112) \quad (0.109) \end{aligned}$$

$$n = 90 \qquad \text{std. error} = 0.233 \qquad R^2 = 0.925$$

---

<sup>16</sup> The figures in parentheses are the standard errors of the coefficient estimates. For simplicity the nomenclature of the variables has been modified; LQD is the equivalent of  $\log Q_D$ , etc.

The addition of the "3-4" markets into the medium size long haul sample had a marked effect upon the parameter estimates (although it reduced their standard errors), but only a slight effect upon the fit (7% increase in the standard error).<sup>17</sup> An examination of the residual plot of the original case indicates a certain amount of residual clustering, similar to that of the large long haul market sample. The addition of the new data seemed to partially alleviate this problem. The conclusion drawn from this analysis is that the "3-4" markets seem to be more representative of the medium size classification than of the large, and therefore will remain in the medium size sample (at least for the preliminary analysis).

A model specification problem regarding the service equation was encountered in the analysis of the medium size long haul markets, which casts a doubt on the validity of the resulting estimates. The sample contains a number of region pairs in which the air service is primarily directed through one or more large enroute cities. An example of this type of market is Rochester-San Diego, where there is no nonstop service and most flights, either direct or connecting, are scheduled through either Chicago or Los Angeles. A large portion of the traffic on the segments are not Rochester-San Diego origin to destination passengers and unfortunately are not accounted for by the route structure variable. Since a fair amount of service is offered between Rochester and Chicago

---

<sup>17</sup> As a statistical aside, this is a case where, by adding data points, the standard error increased but (due to a substantial increase in the variance of the response variable) the value of  $R^2$  also increased. This is an example of why, in a forecasting model,  $R^2$  is not a good comparative measure of fit.

and between Chicago and San Diego, many flights through Chicago are published in the OAG. However, the route structure variable does not account for the Rochester to Chicago or Chicago to San Diego local passengers. Therefore, level of service and hence demand is underestimated for these markets.

The service equation estimation consequently produced a poor fit ( $R^2 = 0.487$ ). However, the deletion of the Rochester-San Diego or San Antonio-San Francisco type markets could not be justified, since they are not really misrepresentative of medium size long haul markets. Elimination of these data points would be purely for the purpose of improving the fit which, as was discussed earlier in this chapter, should not be considered appropriate.

The large long haul markets are not subject to this problem, since most of the travellers on these routes fly nonstop, and the majority of the non-local passengers will be accounted for in the route structure variable. One possible exception is the New York-San Francisco market. The OAG has a tendency to publish an enormous number of connecting flights through Chicago and other cities similar to the example of Section 3.3.2, citing the fact that 84 flights (including thirteen nonstops) are scheduled daily from Los Angeles to New York. The service equation will not predict such a high expected level of service, and demand in the New York-San Francisco market was slightly underestimated (Figure 5.12).

The large and medium size medium haul markets and most short haul markets seem to be free of this problem, but the misfits of the service equation, for the same reason, will occur in the

analyses of small long haul and small medium haul markets. This misspecification problem is probably the reason why the Chow test results imply that, within a given length of haul stratum, data may not be pooled over levels of market size.

The only apparent remedy for this problem is to define a more complex route structure variable that accounts for all non-local passengers using any segment of a scheduled flight between two regions. This would require the CAB's service segment flow data, which is very expensive in terms of both cost and time to process. Since a major proportion of U.S. domestic air passenger traffic is in larger markets, for which the route structure variable defined in this research seems appropriate, it is highly questionable whether this effort would be worthwhile.

The conclusion of Phase I of the sensitivity analysis was to delete one market (New York-San Diego). In Phase II, five single observations were removed. As a result, the residual clustering problem was cured and the fit was slightly improved.

The deletion of these eleven data points, however, had a slightly negative effect upon the results of the estimation of the demand equation via principal components with regard to the precision of the coefficient estimates. Therefore, equation (5.21), calibrated over the entire sample, was selected as the demand analysis relationship for medium size long haul markets.

$$\begin{aligned}
 LQD &= -0.0338 + 0.452 LLOS - 2.07 LFARE + 2.20 LSE \quad (5.21) \\
 &\quad (0.00445) \quad (0.0185) \quad (0.0915) \quad (0.0556) \\
 n &= 90 \qquad \qquad \qquad \text{std. error} = 0.551 \qquad R^2 = 0.568
 \end{aligned}$$

The fare and socio-economic elasticities seem to be rather high. This may be due to a greater percentage of vacation markets in this sample than were found in the long haul sample. Also, the standard errors of the coefficient estimates are two to three times as great as those of the large market analysis model. This is consistent, at least in terms of the fare elasticity, with Verleger's findings.<sup>18</sup>

### 5.2.3 Small Long Haul Markets

The route structure problem encountered in the analysis of medium size long haul markets appears to be more severe in the analysis of the small long haul markets. The standard error of the estimate of the preliminary calibration of the service equation for small markets was 0.111 as opposed to 0.086 for medium and 0.064 for large markets.

The residual plot of the demand equation estimated using ordinary least squares indicated considerable clustering, as was the case with the preliminary estimates of the demand equations for large and medium size long haul markets.

The principal component analysis totally destroyed the fit of the demand equation. In an attempt to rectify this situation, the demand

---

<sup>18</sup> Verleger, op. cit.

equation was reestimated by deleting only the fourth principal component. This unfortunately did not appreciably improve the situation, so to be consistent with the analyses of the other cross classifications, equation (5.22) is the estimation result after deleting the third and fourth principal components.

$$\begin{array}{rcccc}
 \text{LQD} & = & -0.105 & + & 0.575 & \text{LLOS} & - & 0.45 & \text{LFARE} & + & 1.27 & \text{LSE} & & (5.22) \\
 & & (0.0194) & & (0.137) & & & (0.150) & & & (0.0864) & & & \\
 n & = & 46 & & & & \text{std. error} & = & 0.834 & & R^2 & = & 0.289 & 
 \end{array}$$

Neither the deletion of the two markets nor that of the two individual observations improved the coefficient estimates of this equation.

### 5.3 Analysis of Medium Haul Markets

The Chow test, performed to determine if pooling markets by size was appropriate in the long haul, was repeated for medium haul markets. The relevant statistics are presented in Figure 5.14. The test statistic,  $F = 7.36$ , exceeds the critical value 2.24 at the 1% level of significance. Therefore, the data may not be pooled by market size. It is suspected that this result is due to the route structure definitional problem discussed in 5.2.2. This problem seems to occur in the small medium haul market sample, but not in the large or medium size medium haul samples. Consequently, separate medium haul models will be developed for the three market size classifications.

Figure 5.14 Chow Test for Pooling Medium Haul Markets by Level of Socio-Economic Activity

Pooled Sample: Medium haul markets

- Subsamples:
1. Medium/large
  2. Medium/medium
  3. Medium/small

Total number of observations:  $n = 283$

Number of estimated parameters:  $p = 4$

Number of subsamples:  $k = 3$

$$SSR_{\text{pooled}} = 13.72$$

$$SSR_l = 4.46$$

$$SSR_m = 2.09$$

$$SSR_s = \underline{4.72}$$

$$\Sigma SSR_{\text{ind}} = 11.27$$

$$F = \frac{\frac{SSR_{\text{pooled}} - \Sigma SSR_i}{p(k-1)}}{\frac{\Sigma SSR_i}{n-pk}} = \frac{\frac{2.45}{8}}{\frac{11.27}{271}} = 7.36$$

$$F_{\text{crit}}(8, 220; 0.01) = 2.24$$



### 5.3.1 Large Medium Haul Markets

The residual plot of the preliminary ordinary least squares calibration of the demand equation for large medium haul markets indicated a rather severe clustering of residuals. Perusal of the hat matrix diagonal elements and the studentized residuals showed that the smaller "3-4" markets generally had high leverage and/or studentized residuals. Therefore, as was the case with the long/large sample, it was decided to delete the "3-4" markets.

This deletion improved the fit, although not as dramatically as in the long haul market analysis (12% reduction in standard error as opposed to 39% in the long haul). The residual clustering problem was reduced somewhat, yet it was still very evident.

A look at the new set of diagnostic elements provided a clue to the existing problem. One very large market, Boston-Washington, and two of the smaller markets in the sample, Albany-Detroit and Detroit-Raleigh, were very highly influential. The studentized residuals of the smaller markets were not very large (only one of the twelve was greater than 2.0), but those of the Boston-Washington market were quite large, ranging from 1.42 to 2.17. By successive deletion of these three markets, it was apparent that each was contributing to the residual clustering problem. With all three deleted, the clustering problem had virtually vanished and the fit had improved.

A common belief in the airline industry is that medium haul traffic is less sensitive to fluctuations in fare and income and more respondent to service than long haul demand. Therefore, one would expect a greater

coefficient estimate for level of service and lower fare and socio-economic elasticities in the medium than in the long haul markets. The judgmental intervals assessed for these parameters for medium haul markets are as follows:

$$\begin{aligned} 0.40 &< \beta_{11} < 0.80 \\ -0.60 &> \beta_{12} > -1.20 \\ 0.80 &< \beta_{13} < 1.60 \end{aligned}$$

Using principal components regression, the estimated demand equation for large medium haul markets is as follows:

$$\begin{aligned} \text{LQD} &= -0.0822 + 0.534 \text{ LLOS} - 0.583 \text{ LFARE} + 1.40 \text{ LSE} \quad (5.23) \\ &\quad (0.00965) \quad (0.0304) \quad (0.0522) \quad (0.0265) \\ n &= 78 \quad \text{std. error} = 0.553 \quad R^2 = 0.699 \end{aligned}$$

The coefficient estimates all seem to be within reason, with the possible exception of the fare elasticity, which appears to be a bit low.

### 5.3.2 Medium Size Medium Haul Markets

In the analysis of the large medium haul region pairs, the "3-4" markets were deleted due to a suspicion that they are more representative of medium than of large size. If this suspicion were indeed true, it would be expected that the diagnostic statistics of the estimation of

the large markets equations would improve by deletion of these observations and that those of the medium size markets would not be significantly altered by the addition of these markets. The former condition was affirmed in the previous section.

The latter condition was investigated by separately obtaining ordinary least squares estimates using the original medium size medium haul sample and then using the same data set plus the "3-4" medium haul markets. The standard error of estimate increased very slightly, and the standard errors of all coefficients decreased. So while the fit of the equation remained initially unchanged, the precision of the coefficient estimates increased. The conclusion is that the region pairs in question are indeed more representative of medium than large size markets.

The residual plots indicated a considerable amount of clustering. The sensitivity analysis diagnostics implied substantial leverage among the smaller of the medium haul markets, similar to the situations encountered in the analyses of large long and large medium haul markets. By deleting the first four "2-2" markets and reestimating, both the fit and the residual clustering condition improved.

Phase II of the sensitivity analysis resulted in the identification of two remaining data points that, because of high  $h_i$  and  $r^*_i$  values, were suspected of being discrepant. Their deletion resulted in a residual plot that appeared virtually free of clustering.

The demand equation for medium size medium haul markets calibrated using principal components regression on the reduced data set is as follows:

$$\begin{array}{ccccccc}
 \text{LQD} = & 0.0144 & + & 0.991 & \text{LLOS} & - & 0.890 & \text{LFARE} & + & 1.56 & \text{LSE} & (5.24) \\
 & (0.00214) & & (0.0265) & & & (0.0320) & & & (0.0185) & & \\
 n = & 100 & & & & & \text{std. error} = & 0.499 & & & & R^2 = 0.636
 \end{array}$$

The coefficient estimates in (5.24) all appear intuitively reasonable, except that the level of service coefficient is a little high.

### 5.3.3 Small Medium Haul Markets

The route structure variable problem that occurred in the analysis of the medium and small long haul markets was expected in the analysis of the small medium haul markets, and it indeed does appear to be present. However, the service equation standard error was 0.139 with  $R^2 = 0.692$ , which is not quite as bad a fit as was observed in the two other cases in which the route structure problem was present.

The residual plot of the ordinary least squares estimate of the demand equation again produced a severe clustering. The hat matrix diagonal elements indicated that three markets, Lincoln-Tucson which had only one observation, Fargo-Milwaukee with two observations, and Boston-Knoxville with two observations, were highly influential. These three markets, a total of only five observations, were deleted.

The residual clustering problem was nearly eliminated, and when the new hat matrix diagonals and studentized residuals indicated no harmful high leverage points, Phase II of the sensitivity analysis was cancelled.

The principal components regression provided the following estimated

demand equation:

$$\begin{aligned} \text{LQD} = & -0.0277 + 0.570 \text{ LLOS} - 0.597 \text{ LFARE} + 1.44 \text{ LSE} \quad (5.25) \\ & (0.0131) \quad (0.0651) \quad (0.0988) \quad (0.0616) \\ n = & 60 \quad \text{std. error} = 0.552 \quad R^2 = 0.572 \end{aligned}$$

The coefficients in equation (5.25) all seem reasonable, except that the fare elasticity again appears to be a little on the low side.

#### 5.4 Analysis of Short Haul Markets

The first step in the analyses of long and medium haul markets was to conduct a Chow test to determine whether the data may be pooled across market size classifications. In both cases, the conclusion was that they may not. In the analysis of short haul markets this process was repeated, and again the test for homogeneity across market sizes failed. The relevant statistics are tabulated in Figure 5.15. The test statistic,  $F = 8.86$ , well exceeds the critical value, 2.23, at the 1% significance level.

Verleger, Marfisi, and Blumer have each cited a problem with calibrating a model that does not account for fare and service characteristics of competing modes over short lengths of haul. The primary advantage of air transportation relative to other modes in medium and long haul trips is its short total trip time. However, as the length of haul decreases to very short distances, the advantage subsides and

Figure 5.15 Chow Test for Pooling Short Haul Markets by Level of  
Socio-Economic Activity

Pooled Sample: Short haul markets

- Subsamples: 1. Short/large  
2. Short/medium  
3. Short/small

Total number of observations:  $n = 305$

Number of estimated parameters:  $p = 4$

Number of subsamples:  $k = 3$

$$SSR_{\text{pooled}} = 8.32$$

$$SSR_l = 8.33$$

$$SSR_m = 5.39$$

$$SSR_s = \underline{25.9}$$

$$\sum SSR_{\text{ind}} = 39.6$$

$$F = \frac{\frac{SSR_{\text{pooled}} - \sum SSR_i}{p(k-1)}}{\frac{\sum SSR_i}{n-pk}} = \frac{\frac{9.2}{8}}{\frac{39.6}{305}} = 8.86$$

$$F_{\text{crit}}(8, 305, 0.01) = 2.23$$

air traffic volumes decrease. In short haul markets there exists a range of distances for which traffic and distance, hence traffic and fares, are positively correlated. This results in fare elasticity estimates that bear incorrect positive signs.

Since the model in this research does not include variables related to other modes, it was believed that positive price elasticity estimates would be obtained in the demand equation. The estimated fare elasticity for large markets was negative but statistically insignificant. The estimated fare elasticities for medium size and small markets were both positive, the former significantly and the latter insignificantly.

The results of the estimation procedure for the demand equations for large, medium size, and small short haul markets appear in Figure 5.16. Again, because of specification error, little faith can be placed in these estimates.

Figure 5.16. Estimated Demand Equations for Short Haul Markets

## LARGE

$$\begin{array}{r}
 \text{LQD} = 0.0348 + 0.0824 \text{ LLOS} - 0.0505 \text{ LFARE} + 1.14 \text{ LSE} \\
 \quad (0.00983) \quad (0.0189) \quad (0.0537) \quad (0.0225) \\
 n = 102 \quad \quad \quad \text{std. error} = 0.603 \quad \quad \quad R^2 = 0.620
 \end{array}$$

## MEDIUM SIZE

$$\begin{array}{r}
 \text{LQD} = -0.150 + 2.81 \text{ LLOS} + 1.80 \text{ LFARE} + 0.605 \text{ LSE} \\
 \quad (0.00702) \quad (0.0792) \quad (0.0393) \quad (0.00701) \\
 n = 88 \quad \quad \quad \text{std. error} = 0.381 \quad \quad \quad R^2 = 0.877
 \end{array}$$

## SMALL

$$\begin{array}{r}
 \text{LQD} = -0.073 + 1.63 \text{ LLOS} + 0.0114 \text{ LFARE} + 1.29 \text{ LSE} \\
 \quad (0.00922) \quad (0.0871) \quad (0.0177) \quad (0.0240) \\
 n = 100 \quad \quad \quad \text{std. error} = 0.986 \quad \quad \quad R^2 = 0.670
 \end{array}$$



## VI. Applications of the Models

The purpose of this research, as was stated in Chapter I, was to develop a set of demand models which are sufficiently sensitive so as to measure the impacts upon market demand of policy decisions. This chapter provides examples of how the models developed in Chapters III and IV and calibrated in Chapter V may be applied to the analysis of demand variations due to changes in quality of service and fare. These changes may be the effects of the introduction of new aircraft technology or of the implementation of managerial strategies within the framework of existing technology. Also included in this chapter is an outline of how the models may be applied for aggregate forecasting purposes.

Chapter VI is segmented into three parts. Section 6.1 is an application of the demand equations to the derivation of demand vs. frequency relationships for large long and medium haul markets. These relationships are very useful to schedule planners for fleet assignment purposes. Section 6.2 estimates the impact upon demand in a large long haul market of the introduction of a supersonic transport aircraft, and the impact upon demand in a large medium haul market of the introduction of a medium range fuel efficient aircraft. Section 6.3 discusses the various necessary tasks that must be performed to apply the models to produce aggregate forecasts of domestic air passenger demand.

## 6.1 Derivation of Demand vs. Frequency Relationships

In Section 3.2.1 the concept of the "demand vs. frequency" relationship was developed. A typical "demand vs. frequency curve" was plotted in Figure 3.4. In Section 1.2, "The Need for Policy Sensitive Forecasting Models", the motivation for the determination of accurate demand vs. frequency relationships was cited in conjunction with fleet assignment models. Many fleet assignment models have been developed in recent years, both within academic institutions and by aircraft manufacturers. One such model is FA-4, developed in the Flight Transportation Laboratory at M.I.T.<sup>1</sup>

FA-4 is a linear programming model which determines the optimal number of daily flights scheduled over each segment of a route structure network. The objective function to be maximized is the difference between total revenue and the sum of direct and indirect operating costs.<sup>2</sup> The optimization process is constrained by a number of economic factors including, among others, prescribed load factor conditions, fleet availability, minimum number of departures in the various markets, and maximum number of departures from the various stations.

Among the necessary input information is a set of demand vs. frequency relationships for the various markets. The frequency variable,  $n$ , in the demand vs. frequency relationship for a given market is the number of

---

<sup>1</sup> Swan, William, "The Complete FA-4 Memo" (Cambridge: M.I.T. FTL Technical Memorandum 72-10, August, 1972).

<sup>2</sup> Direct and indirect operating costs are defined in Section 3.2.2.

daily departures, assuming that each departure is nonstop,<sup>3</sup> that the demand distribution is uniform over time of day, and that the departure scheduling is such that the average displacement time is minimized. It can be shown that for  $n$  daily departures, this optimal scheduling places the departure of each flight  $i$ ,  $D_i$ , at the following times:

$$D_i = \frac{2i - 1}{2n} \quad i = 1, 2, \dots, n \quad (6.1)$$

where the  $[0, 1]$  time scale is defined from the start to the end of the travelling day.

Given the flight schedule implied by equation (6.1), it can be shown that the average displacement time (assuming the passenger behavior patterns cited in Section 4.1.2.1) is as follows:

$$\overline{DT} = \frac{D}{4n} \quad (6.2)$$

where  $D$  = length of the travelling day.

Since the level of service variable LOS is defined as the ratio of nonstop jet block time,  $t_0$ , to the average of the flight and displacement times, then level of service can be defined as a function of  $n$  as follows:

$$LOS = \frac{t_0}{t_0 + \frac{D}{4n}} = \frac{n}{n + \frac{D}{4t_0}} \quad (6.3)$$

---

<sup>3</sup> For simplicity herein, we will assume all departures are nonstop subsonic jets.

The standard value of the length of travelling day used by FA-4 researchers in the development of demand vs. frequency relationships for long and medium haul markets is  $D = 16$  hours.<sup>4</sup> The nonstop jet time for a flight from Boston to San Francisco is roughly  $t_0 = 6.0$  hours. Substituting these values into equation (6.3) yields the relationship between level of service and number of flights (assuming optimal scheduling) for the Boston to San Francisco segment.

$$\text{LOS (BOS-SFO)} = \frac{n}{n + \frac{16}{4(6.0)}} = \frac{n}{n + 0.667} \quad (6.4)$$

Substituting the function (6.4) into the estimated demand equation for large long haul markets (Figure 5.13) yields the demand vs. frequency relationship for Boston to San Francisco,

$$Q_D(\text{BOS-SFO}) = \log^{-1}(-0.0859) \left( \frac{n}{n + 0.0667} \right)^{0.429} \cdot F^{-1.26} \cdot SE^{1.73} \quad (6.5)$$

The volume of passenger demand, given a fixed fare  $F$  and level of socio-economic activity  $SE$ , was defined as  $Q_{DF}$  in Section 3.2.1. By employing this notation equation (6.5) can be non-dimensionalized as follows:

---

<sup>4</sup> Interview with William Swan, M.I.T. Flight Transportation Laboratory, Cambridge, May 19, 1977.

$$\frac{Q_D}{Q_{D_F}} (\text{BOS-SFO}) = \left( \frac{n}{n + 0.667} \right)^{0.429} \quad (6.6)$$

The numerical results of equations (6.4) and (6.6) are presented in Table 6.1. The demand vs. frequency relationship summarized within this table indicates that 80% of the total potential demand will be satisfied with only one daily departure. The 95% saturation frequency discussed in Section 3.2.1 is five daily departures for the Boston-San Francisco market.

Chicago-New York is a large medium haul market with a jet block time of roughly  $t_0 = 2.5$  hours. Substituting this value into equation (6.3) and combining with the results of the estimation of the demand equation for large medium haul markets, equation (5.23), the following results are obtained for the Chicago-New York market;

$$\text{LOS}(\text{CHI-NYC}) = \frac{n}{n + \frac{16}{4(2.5)}} = \frac{n}{n + 1.60} \quad (6.7)$$

$$\frac{Q_D}{Q_{D_F}} (\text{CHI-NYC}) = \left( \frac{n}{n + 1.60} \right)^{0.534} \quad (6.8)$$

The resulting demand vs. frequency relationship for the Chicago-New York market is tabulated in Table 6.2. If a single flight were scheduled, 60% of the potential demand would be satisfied. The 95% saturation frequency for the Chicago-New York market is sixteen flights.

Table 6.1 Demand vs. Frequency Relationship for Boston to San Francisco

<u>Number of Flights n</u>	<u>Level of Service LOS</u>	<u>Percentage of Total Demand <math>Q_D/Q_{DF}</math></u>
0	0,000	0,000
1	0,750	0,803
2	0,818	0,884
3	0,857	0,918
4	0,882	0,936
5	0,900	0,948
6	0,913	0,956
7	0,923	0,962

Table 6.2 Demand vs. Frequency Relationship for Chicago to New York

<u>n</u>	<u>Level of Service LOS</u>	<u>Percentage of Total Demand <math>Q_D/Q_{DF}</math></u>
0	0.000	0.000
1	0.385	0.600
2	0.556	0.731
3	0.652	0.796
4	0.714	0.836
5	0.758	0.862
6	0.789	0.881
7	0.814	0.896
8	0.833	0.907
9	0.849	0.916
10	0.862	0.924
11	0.873	0.930
12	0.882	0.935
13	0.890	0.940
14	0.897	0.944
15	0.904	0.947
16	0.909	0.950
17	0.914	0.953
18	0.918	0.956

The results imply, as expected, that the long haul Boston-San Francisco market will saturate with fewer scheduled departures than will the medium haul Chicago-New York market. The demand vs, frequency curves for these two markets are superimposed in Figure 6.1,

## 6.2 The Impact upon Demand of New Technologically Advanced Aircraft

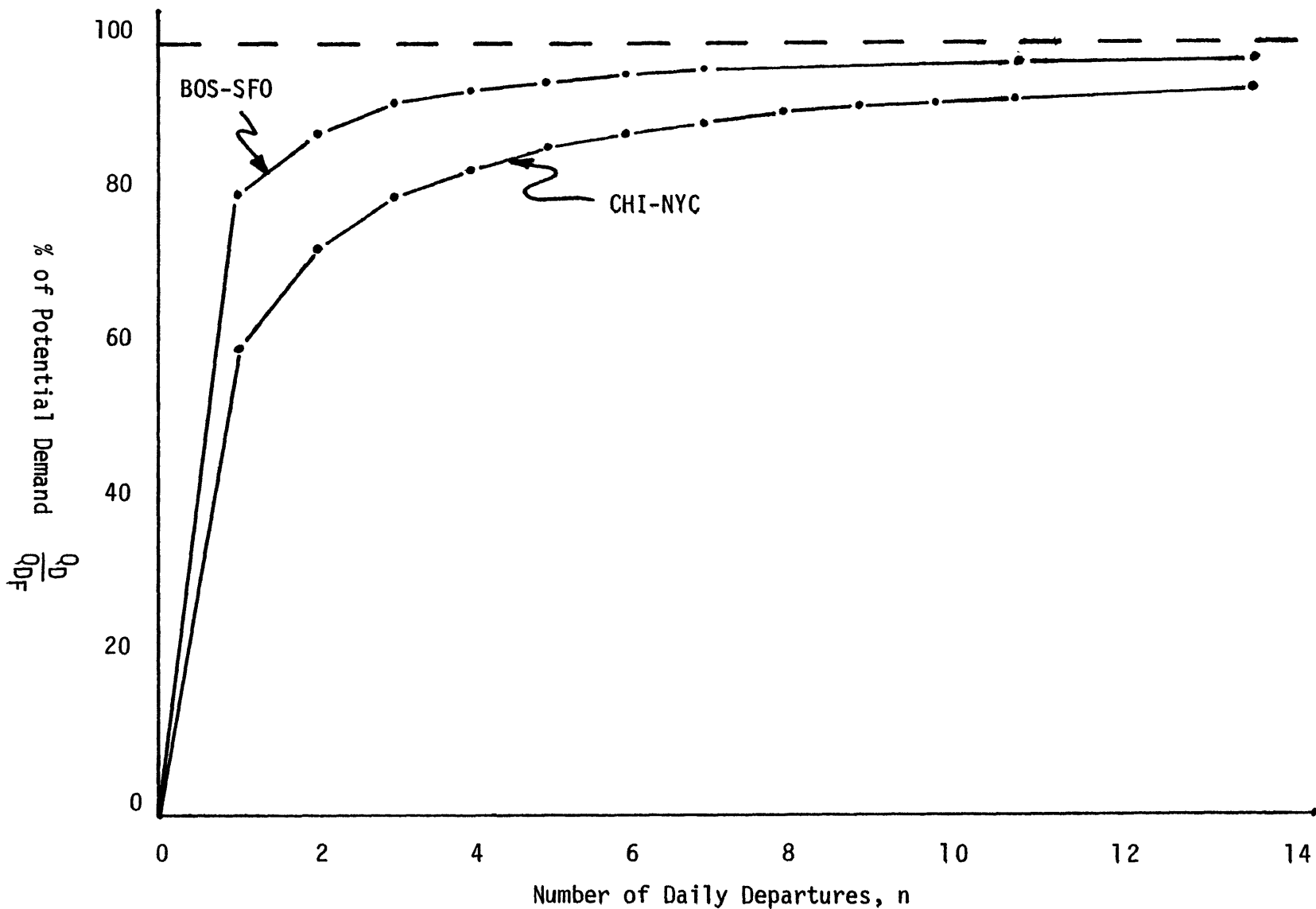
The introduction of a new technologically advanced aircraft will affect the consumers of air passenger transportation in one or two ways. Either the quality of service in a given market will be altered, or the fare structure will change, or both. For example, the introduction of a supersonic transport in long haul markets will improve the level of service by substantially reducing trip time. It may as well result in a price change if a fare premium is charged for the privilege of enjoying this high speed service.<sup>5</sup> If a new fuel-efficient subsonic aircraft were introduced, the savings cost to the airlines would hopefully be passed along to the consumer in the form of either fare reductions or less frequent and/or smaller fare increases. These two hypothetical cases will be investigated in this section.

---

<sup>5</sup> A fare premium was originally charged for jet service when it was first introduced. Also, a substantial surcharge currently exists for transatlantic Concorde flights.



Figure 6.1 Demand vs. Frequency Curves, Boston-San Francisco and Chicago-New York



### 6.2.1 The Introduction of a Supersonic Transport on Long Haul Domestic Routes

In 1974, there were two daily nonstop flights each way between Boston and San Francisco. The value of the level of service variable LOS was 0.792, and approximately 199,000 one-way trips were purchased in this market. In this section, the equipment used for these flights (United's 747 and TWA's L-1011) will be "replaced" by a Boeing SST and the resulting impacts upon demand will be estimated.

Assuming a total of one half hour for taxiway occupancy and acceleration to and deceleration from cruise speed, a cruise speed of 1800 miles per hour, the block time of an SST flight between Boston and San Francisco, approximately 2700 miles, is estimated as

$$t_0 = 0.5 \text{ hours} + \frac{2700 \text{ miles}}{1800 \text{ mph}} = 2.0 \text{ hours} \quad (6.9)$$

This figure is invariant of direction since, at the cruising altitude of the SST, jet stream effects are negligible.

The resulting level of service figures are 1.468 from Boston to San Francisco and 1.321 from San Francisco to Boston. The value of the market level of service is, as defined by equation (4.14), the geometric mean of the two directional values which equals 1.393. This represents a 75.9% increase in LOS.

The coefficient of level of service for large long haul markets is estimated to be 0.429 (Figure 5.13). Assuming no increase in fare, the 75.9% increase in level of service due to the introduction of supersonic

service results in a  $0.429 \times 75.9\%$  or 32.6% increase in traffic, to 264,000 passengers.

The price elasticity for fare on large long haul markets was estimated in Figure 5.13 to be -1.26. Supposing that a 30% surcharge were placed upon SST service, the model implies a  $1.26 \times 30\%$  or 41.1% decrease from the 264,000 passenger figure, to 155,000 passengers. This figure assumes, however, that passengers are offered only the SST as an alternative. If both subsonic and supersonic services were offered (at different prices), the flight selection behavioral process described in Section 4.1.2.1 would involve both trip time and price considerations (as opposed to merely trip time). This is a very complex situation, involving the time value of money, and will be suggested in Chapter VII as a future research consideration.

### 6.2.2 The Introduction of a Fuel Efficient Subsonic Aircraft on Medium Haul Routes

The next generation subsonic aircraft is likely to be a medium-range two or three engine plane with a capacity of about 200 people. It will bridge the gap between the shorter range and smaller capacity narrow-bodies (DC-9, 727, 737) and the longer range and greater capacity wide-bodies (DC-10, L-1011, 747). It will hopefully be substantially cheaper to operate in medium and medium to long haul markets (in terms of direct operating cost per available seat-mile) than the existing four-engine narrow-bodied planes (DC-8, 707).

If the new generation aircraft were introduced, it is reasonable to believe that the cost savings felt by the airlines would be passed on to the consumer over time, in terms of lower fare levels than would be charged if the technology were not introduced. Furthermore, it is possible that level of service could be affected, but this is uncertain and a function of many factors, such as number of planes purchased by the airlines, expected utilization, etc.

It is beyond the scope of this thesis to evaluate the degree to which the introduction of the new equipment will affect fares and quality of service in a given market, particularly since the design parameters of the new aircraft have not as yet been finalized. However, the level of service coefficient and the fare elasticities of the demand analysis equations can provide a clue as to how the service and fare changes caused by the introduction of the new aircraft affect demand.

For example, suppose the new technology aircraft were introduced, resulting in no appreciable change in level of service but, over time, a decrease (in constant dollars) of between 5% and 30% in fares in 700 mile markets, roughly the length of the Chicago-New York market. Since, from equation (5.23), the estimate of price elasticity for large medium haul markets is 0.583, the model would predict the traffic volume increases in that market shown in Table 6.3.

Table 6.3 Effect Upon Demand in Chicago-New York Market of Fuel Efficient Aircraft, Assuming a Resulting 5%-30% Decrease in Fare (Constant Dollars)

<u>Percentage Decrease</u> <u>in Fare</u>	<u>Percentage Increase</u> <u>in Demand</u>
5	2.92
10	5.83
15	8.75
20	11.7
25	14.6
30	17.5

### 6.3 Aggregate Forecasting

The project of applying the models developed in this thesis to aggregate demand forecasting is nearly as complex a task as the development of the models themselves has been. Four major steps are involved in this operation:

Step 1. Determination of Market Sample

Step 2. Gathering of Data

Step 3. Prediction

Step 4. Sensitivity Analysis

The purpose of this section is to outline each of these four steps.

It is not necessary, nor perhaps is it even reasonable, to employ the same sample of markets that was used to calibrate the models for the forecasting process. For the purpose of forecasting aggregate traffic (in, say, RPM's) by length of haul, it is suggested that the samples contain the historically largest (in terms of density) markets in each length of haul grouping, for the following three reasons:

(1) For a fixed sample size, this sampling procedure will provide the maximum ratio of sample RPM's to population RPM's.

(2) The forecasting accuracy of the demand equations (in terms of lower standard errors) appears to be greater for larger markets.

(3) Using a sample for forecasting that is different than the sample used for calibration provides a means for verifying the performance of the model by "forecasting" past aggregate demand and comparing this to actual figures.

The size of the sample is a function of the amount of resources available. The most time and cost sensitive task, with respect to sample

size, will be data gathering.

The necessary socio-economic data are currently being processed by the Bureau of Economic Analysis of the Department of Commerce. The data will include projections of the socio-economic variables (total personal income and income of service industries) through the year 2000.

Scenarios of technological variables can be provided by the aircraft manufacturers and by NASA. Service levels and fares will have to be estimated based upon these technical inputs, by industry predictions of the changes in the various components of direct and indirect operating costs, and by economic forecasts of the appropriate price deflators.

Once the sample has been selected and the data gathered and processed, the estimates of the demand levels for each of the markets for each of the economic and technological scenarios may be obtained by direct substitution into the demand equations. The traffic forecasts may then be summed to obtain aggregate demand forecasts.

Sensitivity analysis is a necessary component to determine how responsive the demand forecasts are to perturbations in each of the factors specified in the technological and economic scenarios. Careful attention must be paid to ensure that the model will not produce bad results if any of the input information is slightly in error.

## VII. Conclusions and Recommendations for Future Research

A series of models have been developed which may be used to forecast future passenger traffic in U.S. domestic air passenger markets. These models are sufficiently policy-sensitive so as to measure the impacts upon market demand due to changes in quality of service, fares, and technological factors. On the surface, the general structure of the models is sufficiently simple so as to be easily communicable to an audience that is unfamiliar with economic theory and econometric modeling. However, the underlying derivations of the components of the model are sufficiently sophisticated so as to capture the important characteristics of this complex industry.

The models are adaptive, in that they may be updated without considerable difficulty as additional data becomes available, although it is not clear that such activity will be necessary. Furthermore, the models are statistically robust in that deletion of any single data points from the samples over which they were calibrated would not substantially alter the estimates. An example of this last fact is the case of the large long haul markets, in which the two data points with the highest apparent leverage of the remaining ninety were removed. Two of three coefficient estimates remained unchanged, while the third changed by one percent.

A common conclusion of other research efforts in this field is that data may not be pooled over lengths of haul to obtain one general demand model. The results of a Chow test in this thesis concurred with this



proposition. Furthermore, the results of Chow tests within length of haul classifications revealed that data may not be pooled by market size (as measured demographically). Therefore, the data were segmented into three lengths of haul and three market size strata. Models were then calibrated over subsets of the data extracted from the markets in each of the nine cross-classifications.

In most of the nine cross-classifications, the equations estimated by ordinary least squares provided a good fit, but did not yield intuitively reasonable estimates of the coefficients. Furthermore, the coefficient estimates were imprecise. The suspected cause of the problem was multicollinearity, and when this suspicion was confirmed, principal components regression was employed to combat the situation. The resulting equations produced reasonable and precise coefficient estimates, but not as good a fit. Since the purpose of this research was to produce a set of models that may be used for policy analysis, it is imperative that the resulting equations bear reasonable and precise coefficient estimates. Consequently, the equations calibrated using ordinary least squares were, in spite of their superior fit, rejected in favor of the equations estimated using principal components deletion.

As was expected, the results of the estimation of the demand equations for short haul markets were unsatisfactory. This is due to the positive relationship between air traffic volume and distance in the short haul because of the supremacy of competing modes for very short distances. Consequently, the fare elasticity was frequently estimated to be a positive

number, as fare is a function only of distance. This reaffirms the need for specialized short haul air traffic demand forecasting models which account for the attributes of surface modes.

For medium and long haul markets, the model seems to perform better for larger markets. This is due to a specification problem regarding the route structure variable. In larger markets a greater percentage of the non-local passengers are accounted for by this variable. Therefore, the service equation estimate produced a poorer fit in the medium size and small long haul markets and the small medium haul markets, than it did in the large long haul and large and medium size medium haul markets. The only apparent remedy for this situation is to define a more complex route structure variable, which would require service segment flow data. However, since these data are very costly to process, and since the majority of the long haul traffic is in large markets and of medium haul traffic is in medium and small size markets (for which the route structure variable as defined herein seems to perform well), it is doubtful whether the benefit of this activity would be worth the resource investment.

Comparing the estimated fare elasticities of long ( $-1.26 \pm 0.067$ ) and medium ( $0.583 \pm 0.104$ ) haul markets, where the error terms are  $\pm$  two standard errors, it appears that air transportation demand is more price elastic in longer haul markets. The results of the generation of demand vs. frequency relationships in Section 6.1 leads to the conclusion that in long haul markets demand will saturate with a fewer number of departures than will demand in medium haul markets. The estimates of the coefficients of the socio-economic variable in all

demand equations for long and medium haul markets imply that air travel demand is very elastic with respect to personal income and the income of service related industries.

The performance of the models in aggregate demand forecasting remains to be seen. The application of this research to medium and long-term forecasting is a process nearly as complex as the development of the models themselves. The accuracy of the forecasts that this research will produce can be only as good as the information received regarding future technological and economic scenarios, and only as good as the methods by which these data are processed to generate predictions of the values of the carrier variables. These applications comprise an obviously ripe area for future research.

The determination of accurate estimates of the relative consequences of displacement time vs. flight time, and of the time/cost tradeoff for air travelers, are other pressing topics of interest related to the research of this thesis. The former can be used to validate the behavioral assumptions adopted herein for the assignment of passengers to flights, and perhaps improve upon the definition of the level of service variable. The latter would provide valuable information for the analysis of markets in which two types of service, one faster and more expensive and one slower and cheaper, exist. This problem was encountered in the analysis of the introduction of domestic supersonic transport service in Section 6.2.

The models developed in this thesis are, as previously mentioned, not effective in the analysis of short haul markets. A complement to this research would be a set of short haul air transportation demand

models that are sensitive to the relative levels of the attributes of competing modes. The thesis by Blumer, surveyed in Chapter II, did an excellent job of laying the groundwork for such models. However, due to specification errors cited in Blumer's conclusions and in Chapter II of this thesis, his work needs modifications.

As a final recommendation, the inclusion of a third stratification, that of market type (business vs. pleasure), would be very insightful, since, as Marfisi indicated in his thesis, the demand equation coefficients are sensitive to the type of traveler predominant in the market. This is a very difficult problem to attack since, while a few markets are obviously highly business-oriented (e.g., Boston-New York, Chicago-Detroit), and some obviously highly pleasure-oriented (e.g., Miami-New York, Las Vegas-Los Angeles), most markets are somewhere on a continuum between the two extremes. Unfortunately, no current data are publicly available that can be used to identify the business/pleasure mix of given markets. The production and dissemination of this data, perhaps by onboard surveys conducted by airlines, would constitute a significant breakthrough for researchers interested in this type of analysis.

## Appendix A. List of Region Pairs by Demographic Stratifications

1-1Short

Bismarck-Minot (106 miles)  
 Knoxville-Lexington (157 miles)  
 Bismarck-Fargo (187 miles)  
 Las Vegas-Reno (345 miles)

Medium

Jackson-Jacksonville (511 miles)  
 Reno-Tucson (709 miles)  
 Las Vegas-Lubbock (775 miles)  
 Lincoln-Tucson (991 miles)

Long

Fargo-Las Vegas (1205 miles)  
 Las Vegas-Lexington (1686 miles)  
 Portland, Maine-Tucson (1825 miles)  
 Erie-Reno (2065 miles)

1-2Short

Lincoln-Omaha (55 miles)  
 Reno-Sacramento (113 miles)  
 Lubbock-Oklahoma City (269 miles)  
 Dayton-Knoxville (282 miles)

Medium

Jacksonville-Norfolk (543 miles)  
 Dayton-Lincoln (665 miles)  
 Minot-Salt Lake City (737 miles)  
 San Antonio-Tucson (762 miles)

Long

Las Vegas-Omaha (1099 miles)  
 Jacksonville-Salt Lake City (1834 miles)  
 Dayton-Reno (1883 miles)  
 Norfolk-Tucson (1999 miles)

\* For an explanation of the demographic stratifications, refer to Section 4.3.2.

1-3Short

Cincinnati-Lexington (70 miles)  
 Jackson-New Orleans (160 miles)  
 Knoxville-Memphis (342 miles)  
 San Diego-Tucson (367 miles)

Medium

Jacksonville-New Orleans (513 miles)  
 Fargo-Milwaukee (516 miles)  
 Denver-Tucson (627 miles)  
 Cincinnati-Portland, Maine (810 miles)

Long

Memphis-Tucson (1224 miles)  
 Las Vegas-New Orleans (1500 miles)  
 Jacksonville-Portland, Oregon (2428 miles)  
 Portland, Maine-San Diego (2623 miles)

1-4Short

Fargo-Minneapolis (223 miles)  
 Lexington-Pittsburgh (289 miles)  
 Dallas-Lubbock (293 miles)  
 Dallas-Jackson (397 miles)

Medium

Minneapolis-Minot (449 miles)  
 Reno-Seattle (566 miles)  
 Dallas-Tucson (839 miles)  
 Atlanta-Lincoln (841 miles)

Long

Bismarck-Seattle (1014 miles)  
 Miami-Portland, Maine (1353 miles)  
 Lubbock-Miami (1400 miles)  
 Atlanta-Las Vegas (1747 miles)

1-5Short

Boston-Portland, Maine (95 miles)  
 Detroit-Erie (155 miles)  
 Las Vegas-Los Angeles (227 miles)  
 Cleveland-Lexington (280 miles)

Medium

Las Vegas-San Francisco (419 miles)  
 Chicago-Lincoln (473 miles)  
 Portland, Maine-Washington 487 miles  
 Boston-Knoxville (830 miles)

Long

Lincoln-Los Angeles (1267 miles)  
 Chicago-Tucson (1441 miles)  
 Jacksonville-San Francisco (2369 miles)  
 New York-Reno (2399 miles)

2-2Short

Norfolk-Richmond (75 miles)  
 Oklahoma City-Wichita (156 miles)  
 Omaha-Wichita (265 miles)  
 Richmond-Rochester (388 miles)

Medium

Norfolk-Rochester (437 miles)  
 Sacramento-Salt Lake City (533 miles)  
 Dayton-Omaha (622 miles)  
 Oklahoma City-Salt Lake City (865 miles)

Long

Dayton-San Antonio (1079 miles)  
 Dayton-Salt Lake City (1461 miles)  
 Sacramento-San Antonio (1463 miles)  
 Norfolk-Salt Lake City (1935 miles)

2-3Short

Raleigh-Richmond (138 miles)  
Cincinnati-Dayton (63 miles)  
Dayton-Milwaukee (285 miles)  
Denver-Salt Lake City (381 miles)

Medium

Denver-Wichita (428 miles)  
Albany-Dayton (576 miles)  
Memphis-San Antonio (626 miles)  
Salt Lake City-San Diego (626 miles)

Long

Portland, Oregon-San Antonio (1714 miles)  
New Orleans-Sacramento (1879 miles)  
Albany-Salt Lake City (1960 miles)  
Rochester-San Diego (2251 miles)

2-4Short

Kansas City-Oklahoma (165 miles)  
Dallas-Oklahoma City (185 miles)  
Dayton-Pittsburgh (215 miles)  
Dayton-St. Louis (339 miles)

Medium

Oklahoma City-St. Louis (462 miles)  
Sacramento-Seattle (608 miles)  
Rochester-St. Louis (729 miles)  
Miami-Richmond (825 miles)

Long

Miami-Rochester (1204 miles)  
Houston-Salt Lake City (1204 miles)  
San Antonio-Seattle (1775 miles)  
Atlanta-Sacramento (2093 miles)



264

2-5

Short

Dayton-Detroit (175 miles)  
Norfolk-Philadelphia (215 miles)  
Boston-Rochester (343 miles)  
Cleveland-Richmond (362 miles)

Medium

Chicago-Omaha (423 miles)  
Chicago-Rochester (522 miles)  
Los Angeles-Salt Lake City (590 miles)  
Detroit-Omaha (660 miles)

Long

Omaha-San Francisco (1432 miles)  
San Antonio-San Francisco (1487 miles)  
Boston-Salt Lake City (2105 miles)  
New York-Sacramento (2510 miles)

3-3

Short

Memphis-Nashville (200 miles)  
Cincinnati-Nashville (230 miles)  
Cincinnati-Milwaukee (318 miles)  
Memphis-New Orleans (349 miles)

Medium

Milwaukee-Nashville (475 miles)  
Albany-Cincinnati (623 miles)  
Denver-San Diego (840 miles)  
Denver-Milwaukee (908 miles)

Long

Denver-New Orleans (1067 miles)  
Albany-Denver (1622 miles)  
Cincinnati-San Diego (1865 miles)  
New Orleans-Portland, Oregon (2050 miles)

265

3-4

Short

Cincinnati-Pittsburgh (256 miles)  
Houston-New Orleans (303 miles)  
Albany-Pittsburgh (367 miles)  
Atlanta-Cincinnati (373 miles)

Medium

Atlanta-New Orleans (425 miles)  
Milwaukee-Pittsburgh (431 miles)  
Miami-Nashville (807 miles)  
Cincinnati-Miami (948 miles)

Long

Denver-Seattle (1020 miles)  
San Diego-Seattle (1052 miles)  
Dallas-Portland, Oregon (1626 miles)  
Denver-Miami (1716 miles)

3-5

Short

Chicago-Milwaukee (74 miles)  
Raleigh-Washington (225 miles)  
Albany-New York (139 miles)  
Albany-Boston (145 miles)

Medium

Albany-Detroit (479 miles)  
Raleigh-Detroit (503 miles)  
Los Angeles-Portland, Oregon (834 miles)  
Denver-San Francisco (956 miles)

Long

Cleveland-Denver (1217 miles)  
Denver-New York (1624 miles)  
Portland, Oregon-Washington (2339 miles)  
New York-San Diego (2435 miles)

4-4Short

Kansas City-St. Louis (229 miles)  
Memphis-St. Louis (255 miles)  
Milwaukee-Minneapolis (297 miles)  
Atlanta-Memphis (332 miles)

Medium

Atlanta-St. Louis (484 miles)  
Houston-Kansas City (643 miles)  
Atlanta-Dallas (721 miles)  
Minneapolis-Pittsburgh (726 miles)

Long

Houston-Pittsburgh ((1124 miles)  
Miami-Minneapolis (1501 miles)  
Dallas-Seattle (1671 miles)  
Miami-Seattle (2725 miles)

4-5Short

Pittsburgh-Washington (193 miles)  
Detroit-Pittsburgh (198 miles)  
Chicago-St. Louis (256 miles)  
New York-Pittsburgh (329 miles)

Medium

Boston-Pittsburgh (496 miles)  
Atlanta-Detroit (602 miles)  
San Francisco-Seattle (671 miles)  
Miami-Washington (920 miles)

Long

Detroit-Houston (1095 miles)  
Kansas City-New York (1098 miles)  
Houston-Washington (1204 miles)  
St. Louis-San Francisco (1736 miles)

5-5

Short

Cleveland-Detroit (94 miles)  
New York-Washington (215 miles)  
Chicago-Detroit (238 miles)  
Boston-Philadelphia (274 miles)

Medium

Boston-Washington (406 miles)  
Boston-Detroit (623 miles)  
Chicago-Philadelphia (675 miles)  
Chicago-New York (721 miles)

Long

Chicago-Los Angeles (1740 miles)  
Los Angeles-Philadelphia (2396 miles)  
New York-San Francisco (2574 miles)  
Boston-San Francisco (2703 miles)

## Appendix B. The Hat Matrix in Linear Models\*

The general linear model is defined as follows:

$$Y = X\beta + \epsilon \quad (\text{B.1})$$

where

$Y$  = an  $n$  by  $1$  vector of response variable observations

$X$  = an  $n$  by  $p$  matrix of the  $n$  observations of the  $p$  carrier variables

$\beta$  = a  $p$  by  $1$  vector of coefficients

$\epsilon$  = an  $n$  by  $1$  vector of errors

Regression analysis is a set of techniques by which the  $\beta$  coefficients are estimated. The most commonly used technique in regression analysis is ordinary least squares where  $\hat{\beta}$ , the estimator of  $\beta$ , is as follows:

$$\hat{\beta} = (X^T X)^{-1} X^T Y \quad (\text{B.2})$$

The elements of the error vector are assumed to have a mean of zero. Therefore, the vector of the expected or "predicted" values of the response variables are as follows:

---

\* A substantial amount of the material in this appendix has been extracted from David C. Hoaglin and Roy E. Welsch, "The Hat Matrix in Regression and ANOVA" (National Bureau of Economic Research Working Paper WP 901-77, January, 1977).

$$\hat{Y} = X\hat{\beta} = X(X^T X)^{-1} X^T Y \quad (\text{B.3})$$

or

$$\hat{Y} = HY \quad (\text{B.4})$$

where

$$H = X(X^T X)^{-1} X^T \quad (\text{B.5})$$

The matrix  $H$  defined in equation (B.5) is called the "hat" matrix since it transforms  $Y$  into  $\hat{Y}$  ("Y-hat").

It is apparent from equations (B.4) and (B.5) that the predicted values of the response variables are linear combinations of the actual values, and that the weights are a function only of the carrier variables.

$$\hat{y}_i = \sum_{j=1}^n h_{ij} y_j \quad (\text{B.6})$$

where

$y_i$  = a generic value of  $Y$  and

$h_{ij}$  = a generic element of  $H$

From equation (B.6) it is observed that the value of  $h_{ij}$  determines the amount of influence or "leverage" that the observed value of the  $j^{\text{th}}$  observation,  $y_j$ , has upon the predicted value of the  $i^{\text{th}}$  observation,  $\hat{y}_i$ . Furthermore, this amount of leverage is irrespective to the value of  $y_j$ , since the hat matrix is a function only of the carrier variables. Therefore, the hat matrix is useful to the data analyst for the identification of sensitive sample data points for which the value of  $y$  will have a substantial impact upon the fit. These points must be closely scrutinized

because a discrepant observed value of  $y$  may have a very detrimental effect upon the results of the estimation.

The amount of leverage that a particular  $y_i$  has upon the overall fit can be most readily ascertained by inspection of the  $i^{\text{th}}$  diagonal element of the hat matrix,  $h_{ii}$ . For simplicity this element will hereafter be referred to as  $h_i$ . The analyst may therefore be concerned with only the  $n$  diagonal elements rather than the total number of elements  $n^2$ . This is certainly much more efficient when analyzing large data sets.

Given a data set with  $n$  observations one may ask how large must a particular  $h_i$  be for it to be considered a high leverage point. Theorem B.2 which follows provides a guideline. The proof of this requires the use of a preliminary theorem from linear algebra.

Theorem B.1: Let  $A$  be an  $m$  by  $n$  matrix and  $B$  be an  $n$  by  $m$  matrix. The trace of  $AB$  is equal to the trace of  $BA$ .

Proof: By definition,

$$\text{trace}(AB) = \sum_{i=1}^m \sum_{j=1}^n a_{ij} b_{ji}$$

$$\text{trace}(BA) = \sum_{j=1}^n \sum_{i=1}^m b_{ij} a_{ji}$$

$$= \sum_{j=1}^n \sum_{i=1}^m b_{ij}^T a_{ji}^T$$

$$= \sum_{j=1}^n \sum_{i=1}^m b_{ji} a_{ij} = \text{trace}(AB)$$

Theorem B.2: The trace of the hat matrix equals  $p$ .

Proof: By definition,

$$\begin{aligned} \text{trace } H &= \text{trace } [X(X^T X)^{-1} X^T] \\ &= \text{trace } [X^T X (X^T X)^{-1}] \quad (\text{by Theorem B.1}) \\ &= \text{trace } I_p = p \end{aligned}$$

where

$I_p$  = a  $p$  dimensional identity matrix

Since by Theorem B.2 the sum of the diagonal elements of the hat matrix equals  $p$ , then the mean value of these elements equals  $p/n$ . A reasonable rule of thumb is that an  $h_i$  value is considered to be "large" if  $h_i > 2p/n$ .

Once a data point  $i$  has been identified as a high leverage point, it is necessary to determine whether its  $y_i$  value is discrepant. An initial clue may be obtained by observing its residual,  $r_i$ , the difference between the observed and predicted values.

$$\begin{aligned} r_i &= y_i - \hat{y}_i \\ &= y_i - X\hat{\beta}_i \end{aligned} \tag{B.7}$$

Removing the scale may add clarity to the residual values. The



standardized residual is defined as follows:

$$\tilde{r}_i = \frac{r_i}{\hat{\sigma} \sqrt{1 - h_i}} \quad (\text{B.8})$$

where

$\hat{\sigma}^2$  = the residual mean square

The residual values may, however, be very misleading. A discrepant point may have a very small residual if its leverage is so great that the estimated regression function is forced away from the representative points and forced through or near the bad point. In such a case, the representative data points will have high residuals.

Consider the example of the data listed in Table .1 and plotted in Figure B.1. It is quite obvious that data point 11 is a discrepant observation. The least squares regression line is as follows:

$$\hat{y} = 6.72 + 0.50x \quad (\text{B.9})$$

This line is plotted in Figure B.2, and the degree of influence of data point 11 is obvious.

The fourth column of Table B.2 lists the diagonal elements of the hat matrix. Since  $p = 2$  and  $n = 11$ , the rule of thumb cutoff point is  $2p/n = 0.364$ . Data point 11 is immediately identified as a high leverage point. It is important to realize that this identification is in no way based upon the discrepant  $y$  value for observation 11. The high value of  $h_{11}$  is due only to its location along the  $x$ -axis which is quite removed from the remainder of the data points.

Table B.1 Hypothetical Data

Observation			
	<u>Value</u>	<u>y</u>	<u>x</u>
1		7.0	8.0
2		9.0	10.0
3		11.0	12.0
4		13.5	14.0
5		15.0	16.0
6		16.5	18.0
7		19.3	20.0
8		21.0	22.0
9		23.0	24.0
10		24.7	26.0
11		16.0	36.0

Table B.2 Diagnostic Statistics for Hypothetical Data

<u>i</u>	<u>y<sub>i</sub></u>	<u>x<sub>i</sub></u>	<u>h<sub>i</sub></u>	<u>r<sub>i</sub></u>	<u><math>\tilde{r}_i</math></u>	<u>r*<sub>i</sub></u>
1	7.0	8.0	0.266	-3.68	-1.02	-1.02
2	9.0	10.0	0.207	-2.67	-0.71	-0.69
3	11.0	12.0	0.160	-1.67	-0.43	-0.41
4	13.5	14.0	0.125	-0.16	-0.04	-0.04
5	15.0	16.0	0.102	0.35	0.09	0.08
6	16.5	18.0	0.092	0.86	0.21	0.20
7	19.3	20.0	0.093	2.67	0.66	0.64
8	21.0	22.0	0.107	3.38	0.84	0.83
9	23.0	24.0	0.133	4.39	1.11	1.13
10	24.7	26.0	0.171	5.10	1.32	1.39
11	16.0	36.0	0.544	-8.56	-2.99	-44.96

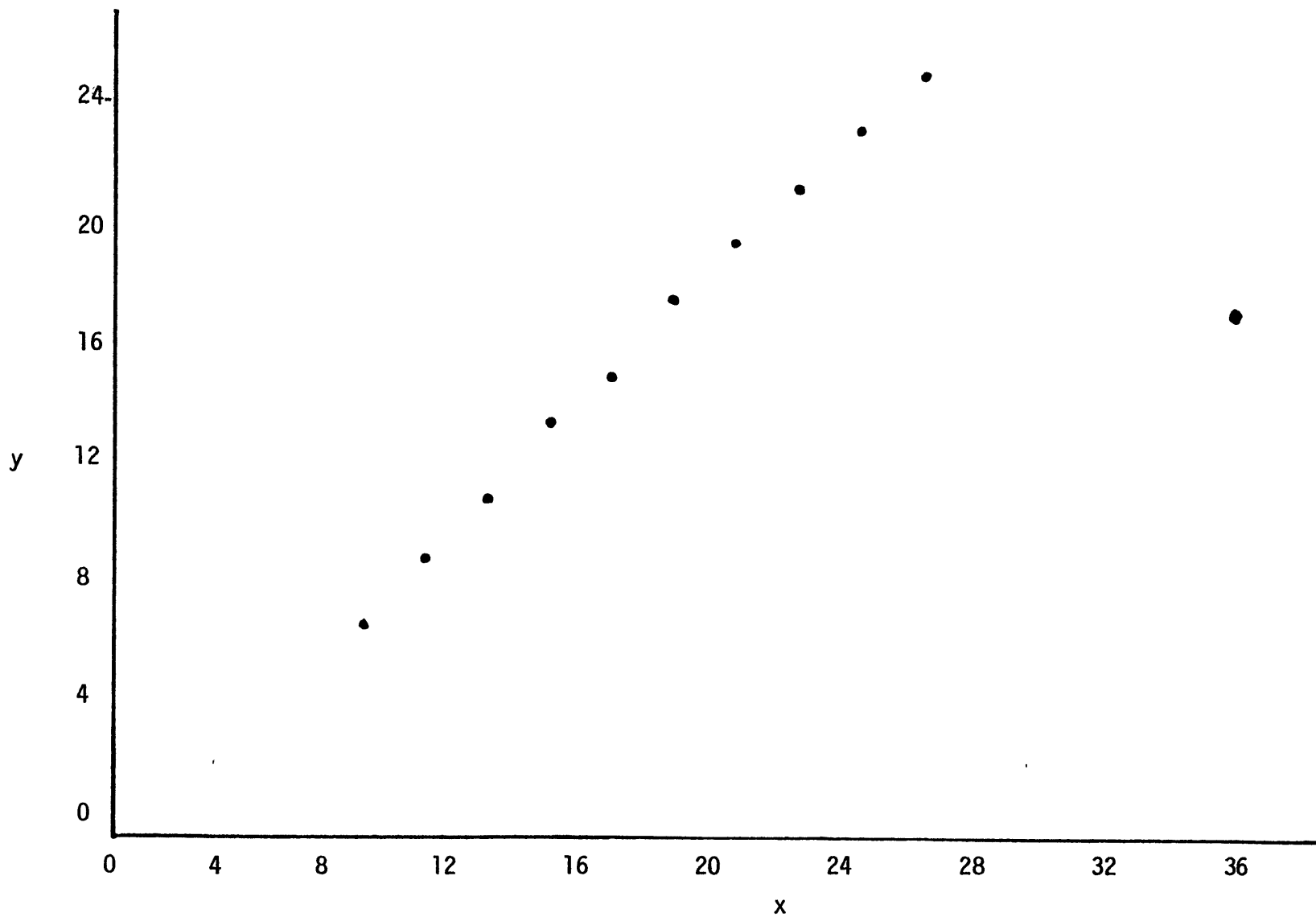


Figure B.1 Scatter Plot of Hypothetical Data

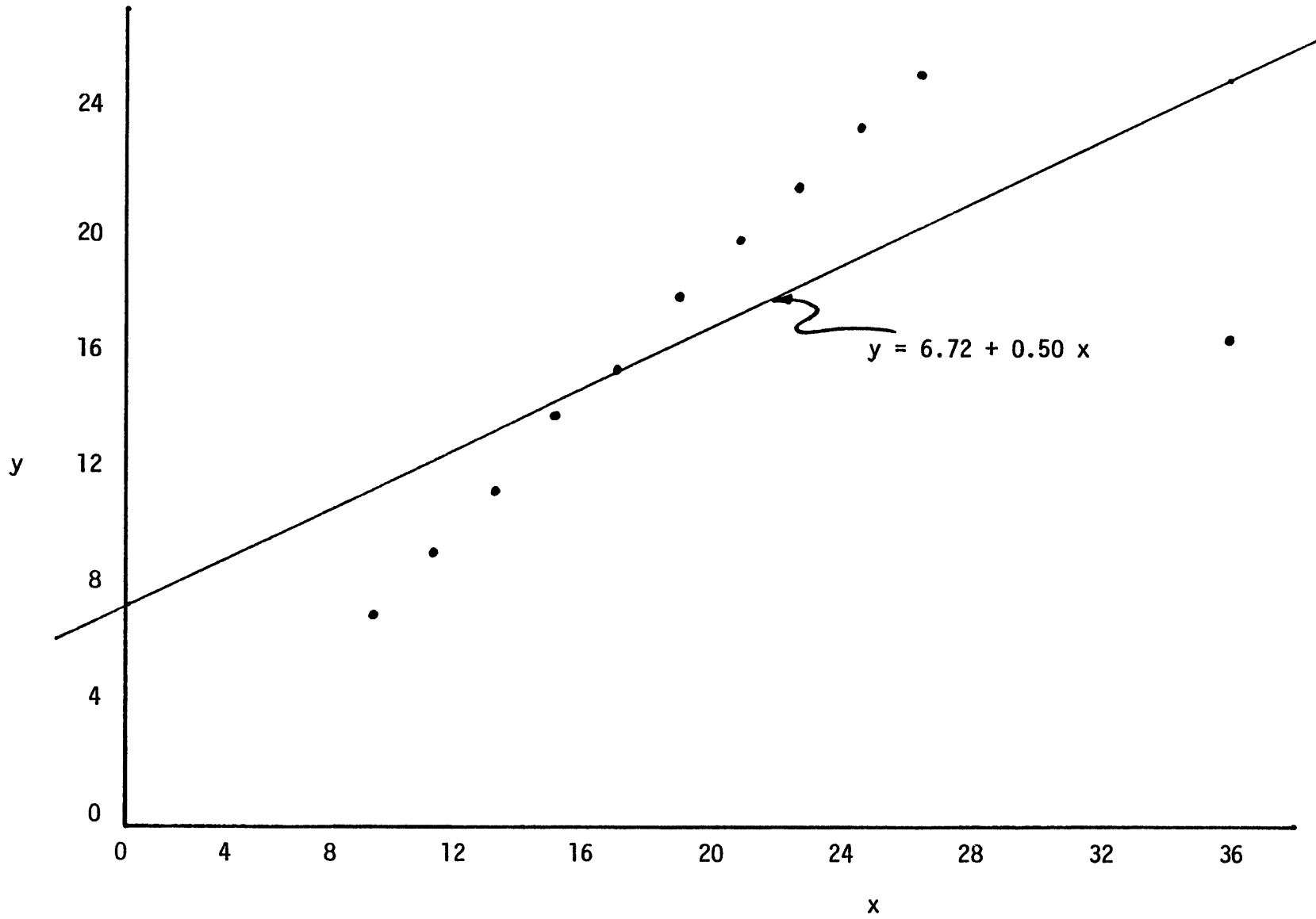


Figure B.2 Regression Function for Hypothetical Data

The fifth and sixth columns of Table B.2 list the residuals and standardized residuals for this example. While data point 11 appears to clearly be an outlier, the relative magnitude of the residuals do not indicate the great degree of influence this observation had upon the overall fit.

A more appropriate diagnostic statistic is the studentized residual,  $r^*_i$ , which is the distance that the observed value of  $y_i$  would lie from the predicted value of  $\hat{y}_{(i)}$  if the model were fitted with data point  $i$  deleted from the sample, divided by its standard error.

$$r^*_i = \frac{y_i - x_i \hat{\beta}_{(i)}}{\hat{\sigma}_{(i)} \sqrt{1 + x_i (X_{(i)}^T X_{(i)})^{-1} x_i^T}} \quad (\text{B.10})$$

where

$\hat{\beta}_{(i)}$  = the estimated coefficients with data point  $i$  deleted

$\hat{\sigma}^2_{(i)}$  = the residual mean square with data point  $i$  deleted

and  $X_{(i)}$  = the matrix of carrier observations with row  $i$  deleted

Since the numerator and denominator of (B.10) are independent,  $r^*_i$  will follow a  $t$  distribution with  $n-p$  degrees of freedom. The studentized residuals of the numerical example are listed in column 7 of Table B.2. The discrepancy of data point 11 is vividly demonstrated by investigation of the studentized residuals.

Appendix C Multicollinearity and Related Topics <sup>1</sup>

Multicollinearity is a problem that frequently arises in practice when the general linear model (C.1) is fitted with empirical data.

$$Y = X\beta + \epsilon \quad (C.1)$$

where  $Y$  = an  $n$  by  $1$  vector of observations of the response variable  
 $X$  = an  $n$  by  $p$  matrix of the  $n$  observations of the  $p$  carrier variables  
 $\beta$  = a  $p$  by  $1$  vector of coefficients  
 $\epsilon$  = an  $n$  by  $1$  vector of errors

Multicollinearity arises when two or more of the carrier variables or linear combinations of them are highly correlated. If this problem is present one may obtain the least squares estimates of the elements of the

---

<sup>1</sup> A substantial amount of the material in this appendix has been extracted from David A. Belsley, "Multicollinearity: Diagnosing Its Presence and Assessing the Potential Damage it Causes Least-Squares Estimation" (Cambridge: National Bureau of Economic Research Working Paper No. 154; October, 1976), and Roy E. Welsch, "Principal Components and the Singular Value Decomposition" (unpublished lecture notes for the M.I.T. Course 15.063/18.457 -- Statistics for Model Building, 1977).

$\beta$  vector,  $\hat{\beta}$ , as defined in equation (C.2), but interpretation of these

$$\hat{\beta} = (X^T X)^{-1} X^T Y \quad (C.2)$$

variables is quite difficult. Suppose, for example, that two carrier variables are highly collinear. Their individual coefficient estimates in the  $\hat{\beta}$  vector are intended to measure the effect upon the response variable due to a change in the carrier variable with all other variables held constant. But since whenever one of these correlated variables changes in the data its related carrier usually changes correspondingly, the data itself provide very little confidence in the intended measure.

The statistical result of multicollinearity is that the standard errors of the coefficients are quite high. If the purpose of the model is to measure the impact upon  $Y$  of changes in the individual  $x$  variables, the analyst desires precise estimates of the  $\beta$  elements, and the presence of multicollinearity can therefore be quite damaging.

One commonly used method for detecting the presence of multicollinearity is observation of the off diagonal elements of the variance-covariance or correlation matrices. However, this procedure may not indicate the true severity of the problem.

Consider the case where the  $X$  matrix is defined as follows:

$$X = [\underline{x}_1, \underline{x}_2, \underline{x}_3] \quad (C.3)$$



where  $\underline{x}_1$  = an n by 1 vector of observations of a random variable with a uniform distribution over  $[-1/2, 1/2]$

$\underline{x}_2$  = an n by 1 vector of observations of a random variable which is independent of  $x_1$  and also uniformly distributed over  $[-1/2, 1/2]$

$$\underline{x}_3 = \underline{x}_1 - \underline{x}_2$$

Since  $x_1$  and  $x_2$  are independent, they are also uncorrelated.

The covariance of  $x_1$  (or  $x_2$ ) and  $x_3$  is as follows:

$$\text{Cov}(x_1, x_3) = E(x_1 x_3) = E(x_1^2 - x_1 x_2) = \text{Var}(x_1) \quad (\text{C.4})$$

The correlation of  $x_1$  (or  $x_2$ ) and  $x_3$  is then:

$$\begin{aligned} P_{x_1, x_3} &= \frac{\text{Cov}(x_1, x_3)}{\sqrt{\text{Var}(x_1) \text{Var}(x_3)}} = \frac{\text{Var}(x_1)}{\sqrt{\text{Var}(x_1) \cdot 2 \text{Var}(x_1)}} \\ &= \frac{1}{\sqrt{2}} = 0.71 \end{aligned} \quad (\text{C.5})$$

The resulting correlation matrix is shown in equation (C.6)

$$P_X = \begin{bmatrix} 1.00 & 0.00 & 0.71 \\ 0.00 & 1.00 & -0.71 \\ 0.71 & -0.71 & 1.00 \end{bmatrix} \quad (\text{C.6})$$

The simple correlations of equation (C.6) may not seem so severe as to imply serious multicollinearity when in fact the linear dependence is perfect. The cross product matrix of equation (C.2),  $X^T X$ , will not in this case invert.

A numerical example of this relationship is the data in Table C.1. The resulting estimate of the correlation matrix is given in equation (C.7).

$$P_X = \begin{bmatrix} 1.00 & 0.08 & 0.73 \\ 0.08 & 1.00 & -0.62 \\ 0.73 & -0.62 & 1.00 \end{bmatrix} \quad (C.7)$$

### C.1 Singular Value Decomposition

The singular value decomposition is a useful technique for computational efficiency and conceptual analysis of least squares estimation. It is included in this appendix primarily for illustrating its application to the diagnosis of multicollinearity.

Any  $n$  by  $p$  matrix,  $X$ , may be decomposed as follows:

$$X = U \Sigma V^T \quad (C.8)$$

where  $U U^T = V V^T = I_p$ , a  $p$  dimensional identity matrix and  $\Sigma$  is a diagonal matrix with non-negative elements. Exactly  $k$  of the elements of  $\Sigma$  will be positive, where  $k$  is the rank of  $p$ . The elements of  $\Sigma$  may be arranged in any order, so without loss of generality it will be assumed

Table C.1. Perfectly Collinear Data

<u><math>x_1</math></u>	<u><math>x_2</math></u>	<u><math>x_3</math></u>
0.070	-0.285	0.355
0.331	0.072	0.259
0.055	-0.190	0.245
-0.030	0.200	-0.230
0.348	0.406	-0.058
-0.420	-0.164	-0.256
-0.135	-0.347	0.212
-0.443	0.141	-0.584
-0.232	0.135	-0.367
-0.074	0.420	-0.494
0.455	0.024	0.431
0.453	0.246	0.207
0.170	-0.458	0.628
-0.325	0.360	-0.685
-0.463	-0.088	-0.375

that they are ordered descendingly,  $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_p$ . The elements of  $\Sigma$  are called the singular values of  $X$ .

The singular value decomposition is similar in concept to an eigen-system. Noting that

$$X^T X = V \Sigma U^T U \Sigma V^T = V \Sigma^2 V^T \quad (\text{C.9})$$

it is clear that the elements of  $\Sigma^2$ , the squares of the singular values of  $X$ , are the eigenvalues of  $X^T X$ , and that the columns of  $V$  are the eigenvectors of  $X^T X$  (and are therefore orthogonal).

The singular value decomposition has useful numerical properties. Rather than inverting the  $X^T X$  matrix to obtain the least squares estimates in equation (C.2), a more efficient procedure is as follows:

$$\begin{aligned} \hat{\beta} &= (X^T X)^{-1} X^T Y \\ &= V \Sigma^{-2} V^T V \Sigma U^T Y \\ &= V \Sigma^{-2} \Sigma U^T Y \\ &= V \Sigma^{-1} U^T Y \end{aligned} \quad (\text{C.10})$$

Furthermore, the hat matrix defined in equation (B.5) may be obtained as follows:

$$\begin{aligned}
H &= X(X^T X)^{-1} X^T \\
&= U \Sigma V^T V \Sigma^{-2} V^T V \Sigma U^T \\
&= U \Sigma \Sigma^{-2} \Sigma U^T \\
&= U U^T
\end{aligned}
\tag{C.11}$$

If multicollinearity is present one or more of the eigenvalues of  $X^T X$ , hence one or more of the singular values of  $X$ , will be very small relative to the largest values. Defining the  $i^{\text{th}}$  condition index of  $X$ ,  $\kappa_i(X)$ , as follows,

$$\kappa_i(X) = \frac{\sigma_1}{\sigma_i} \tag{C.12}$$

and the condition number of  $X$ ,  $\kappa(X)$ , as the largest condition index,

$$\kappa(X) = \kappa_p(X) = \frac{\sigma_1}{\sigma_p} \tag{C.13}$$

results in a more useful measure of multicollinearity than the simple correlation matrix. Generally speaking, the higher the values of the condition number and condition indices are, the more severe is the problem of collinearity. The decision as to how high the condition number must be before one senses danger is somewhat arbitrary and is discussed in detail by Belsley.<sup>2</sup>

---

<sup>2</sup> Belsley, op. cit.



Room 14-0551  
77 Massachusetts Avenue  
Cambridge, MA 02139  
Ph: 617.253.5668 Fax: 617.253.1690  
Email: docs@mit.edu  
<http://libraries.mit.edu/docs>

## **DISCLAIMER OF QUALITY**

Due to the condition of the original material, there are unavoidable flaws in this reproduction. We have made every effort possible to provide you with the best copy available. If you are dissatisfied with this product and find it unusable, please contact Document Services as soon as possible.

Thank you.

**Page numbering error by the author.  
Pages 285-286 do not exist.**

$Z_i$  ( $i=1, 2, \dots, p$ ) where

$$\underline{Z}_i = X \underline{v}_i \quad (\text{C.14})$$

The  $\underline{v}_i$  vectors are the weightings by which each column of  $X$  contributes to the linear combination  $\underline{Z}_i$ . For uniform scaling purposes it will be required that

$$\underline{v}_i^T \underline{v}_i = 1 \quad (\text{C.15})$$

Assuming that the columns of  $X$  have been centered, the sum of the sample variances of the  $p$  columns of  $X$  is

$$\sum_{i=1}^p \hat{\sigma}_{x_i}^2 = \frac{1}{n} \text{trace} (X^T X) \quad (\text{C.16})$$

In order to retain as much variance as possible in the first linear combination,  $\underline{Z}_1$ , it is necessary to maximize  $\hat{\sigma}_{Z_1}^2$  where

$$\hat{\sigma}_{Z_1}^2 = \frac{1}{n} \underline{Z}_1^T \underline{Z}_1 = \frac{1}{n} \underline{v}_1^T X^T X \underline{v}_1 \quad (\text{C.17})$$

The selection of the  $\underline{v}_1$  vector is then the process of maximizing (C.17) subject to the constraint  $\underline{v}_1^T \underline{v}_1 = 1$  (C.15). The Lagrangian formulation is

$$\max L(\underline{v}_1) = \underline{v}_1^T X^T X \underline{v}_1 + \lambda_1 (1 - \underline{v}_1^T \underline{v}_1) \quad (\text{C.18})$$

The first order optimization conditions are

$$\frac{\partial L(\underline{v}_1)}{\partial \underline{v}_1} = (X^T X - \lambda_1 I) \underline{v}_1 = 0 \quad (\text{C.19})$$

A non-trivial solution then requires that

$$\det(X^T X - \lambda_1 I) = 0 \quad (\text{C.20})$$

Therefore,  $\lambda_1$  is an eigenvalue of  $X^T X$  and  $\underline{v}_1$  is its corresponding eigenvector.

To determine which of the  $p$  eigenvectors should be selected for  $\underline{v}_1$ , note that from equation (C.19) it follows that

$$\underline{v}_1^T (X^T X - \lambda_1 I) \underline{v}_1 = 0 \quad (\text{C.21})$$

Therefore,

$$\lambda_1 = \underline{v}_1^T X^T X \underline{v}_1 = n \hat{\sigma}_{z_1}^2 \quad (\text{C.22})$$

Since the objective is to maximize this variance,  $\lambda_1$  must be the greatest eigenvalue of  $X^T X$ . The corresponding eigenvector,  $\underline{v}_1$ , is called the "first principal component."

The second principal component is determined by maximizing  $\underline{v}_2^T X^T X \underline{v}_2$  subject to  $\underline{v}_2^T \underline{v}_2 = 1$ , and so forth. The total variance of the columns of  $X$  is accounted for after all principal components have been



extracted.

$$\lambda_1 + \lambda_2 + \dots + \lambda_p = \text{trace}(X^T X) \quad (\text{C.23})$$

This procedure is an extension of the singular value decomposition (Section C.1) where  $X^T X$  was decomposed to  $V \Sigma^2 V^T$ . In principal components analysis the cross product matrix is "factored" into  $X^T X = L L^T = [V \Sigma][\Sigma V^T]$ . This is one application of a general class of techniques known as "factor analysis".

The regression model in the principal components basis is then

$$Y = Z\alpha + \epsilon = XV\alpha + \epsilon \quad (\text{C.24})$$

The estimates for  $\alpha$ ,  $\hat{\alpha}$ , may be obtained by ordinary least squares:

$$\begin{aligned} \hat{\alpha} &= (Z^T Z)^{-1} Z^T Y \\ &= (V^T X^T X V)^{-1} V^T X^T Y \\ &= V^T (X^T X)^{-1} V V^T X^T Y \\ &= V^T (X^T X)^{-1} X^T Y \\ &= V^T \hat{\beta} \end{aligned} \quad (\text{C.25})$$

If some of the eigenvalues,  $\lambda_i$ , are small it may be decided that the removal of the corresponding principal components will alleviate the collinearity problem (in the original  $\beta$  basis) while retaining the majority of the original variance. Deleting these principal components

is in effect setting their  $\hat{\alpha}_i$  values to zero.

The process of deletion of principal components described above is based solely upon the relationships of the columns of  $X$ . The general rule is that if  $\lambda_i$  is small  $\hat{\alpha}_i$  can be set to zero. However, it is conceivable that the linear combination  $Z_i$  may happen to be highly correlated (relative to the other combinations) with  $Y$ . If this is indeed the case, the deletion,  $\hat{\alpha}_i = 0$ , may destroy the fit. In the principal components basis the  $t$ -statistics are as follows:

$$t_i = \frac{u_i^T Y}{\hat{\sigma}} \quad (C.26)$$

where  $u_i$  is the  $i^{\text{th}}$  row of the  $U$  matrix defined in Section C.1. If both  $\lambda_i$  and the corresponding  $t$ -statistic are low, then setting  $\hat{\alpha}_i$  to zero is a reasonable procedure.

## BIBLIOGRAPHY

## RESEARCH IN AIR PASSENGER DEMAND ANALYSIS

1. Bartlett, H.C. "The Demand for Passenger Air Transportation, 1947-1962". Ph.D. dissertation, University of Michigan, 1965.
2. Blackburn, A.J. "A Nonlinear Model of Passenger Demand." Ph.D. dissertation, Massachusetts Institute of Technology, 1966.
3. Blumer, T.P. "A Short Haul Passenger Demand Model for Air Transportation." S.M. thesis, Massachusetts Institute of Technology, 1976.
4. Brown, S.L., and Watkins, W.S. "The Demand for Air Travel: A Regression Study of Time Series and Cross Sectional Data in the U.S. Domestic Market," in Airport Economic Planning, pp. 82-108. Edited by George P. Howard. Cambridge: M.I.T. Press, 1974.
5. Cochran, D.L. "An Evaluation of the Use of Gravitational Formulae for Estimating Potential Air Passenger Traffic between City-Pairs." Ph.D. dissertation, University of Oregon, 1970.
6. Eriksen, S.E., Scalea, J.C., and Taneja, N.K. "A Methodology for Determining the Relationship Between Air Transportation Demand and the Level of Service." Cambridge: M.I.T. Flight Transportation Laboratory Report R76-3, 1976.
7. Eriksen, S.E., and Taneja, N.K. "Directions for Improving Air Transportation Demand Models." Joint National Meeting of the Operations Research Society of America and The Institute of Management Sciences, Miami, 1976.

8. Federal Aviation Administration. Aviation Forecasts, Fiscal Years 1976-1987. Technical Report FAA-AVP-75-7. Washington: Government Printing Office, 1975.
9. Gronau, R. The Value of Time in Passenger Transportation: The Demand for Air Travel. National Bureau of Economic Research Occasional Paper 109. New York: Columbia University Press, 1970.
10. Howard, G.P., and Augustinus, J.G. "Market Research and Forecasting for the Airport Market," in Airport Economic Planning. pp. 109-127. Edited by George P. Howard. Cambridge: M.I.T. Press, 1974.
11. Joun, Y.P. "The Demand for Air Travel." Ph.D. dissertation, University of Washington, 1966.
12. Keintz, R.M. "The Demand for International Travel to and from the United States." Ph.D. dissertation, Boston College, 1971.
13. Lancaster, K. Consumer Demand: A New Approach. New York: Columbia University Press, 1971.
14. Marfisi, E.P. "Theory and Evidence on the Behavior of Airline Firms Facing Uncertain Demand." Ph.D. dissertation, Brown University, 1976.
15. M.I.T. Flight Transportation Laboratory. "Proceedings of the Workshop: Air Transportation Demand and Systems Analysis." Report R75-8, 1975.
16. Mize, J.L. "An Econometric Analysis of the Demand for Airline Passenger Transportation-Domestic Routes", Ph.D. dissertation, Georgia State College, 1968.

17. Quandt, R.E., ed. The Demand for Travel: Theory and Measurement. Lexington, Mass: Heath Lexington Books, 1970.
18. Verleger, P.K., Jr. "A Point-to-Point Model of the Demand for Air Transportation," Ph.D. dissertation, Massachusetts Institute of Technology, 1971.

#### DATA SOURCES

19. Civil Aeronautics Board. Origin-Destination Survey of Airline Passenger Traffic. Tables 8 and 10.
20. Official Airline Guide. Oak Brook, Ill.: The Reuben H. Donnelley Corporation.
21. U.S. Department of Commerce. Bureau of Economic Analysis. Local Area Personal Income 1969-1974, vol. 1, 1976.
22. U.S. President. Economic Report of the President. Washington: Government Printing Office, 1974.

#### OTHER REFERENCES

23. Belsley, D.A. "Multicollinearity: Diagnosing Its Presence and Assessing the Potential Damage It Causes Least-Squares Estimation." Cambridge: National Bureau of Economic Research Working Paper No. 154, 1976.

24. Blumer, T.P., Simpson, R.W., and Wiley, J. "A Computer Simulation of Tampa International Airport's Landside Terminal Shuttles." Cambridge: M.I.T. Flight Transportation Laboratory Report R76-5, 1976.
25. Brigham, E.F., and Pappas, J.L. Managerial Economics. Hinsdale, Ill.: The Dryden Press, 1976.
26. Civil Aeronautics Board. Bureau of Operating Rights. Effect on Total Market Traffic of Changes in Quality of Service (QSI): Explanatory Statement, Docket 21136, 1972.
27. "C.A.B. Frill Investigation Expected," Aviation Week and Space Technology, 15 November 1976, pp. 26-27.
28. Chow, G. "Tests of Equality Between Subsets of Coefficients in Two Linear Regressions," Econometrica 28 (July, 1960): 591-605.
29. Daniel, C., and Wood, F.S. Fitting Equations to Data. New York: Wiley, 1971.
30. Draper, N., and Smith, H. Applied Regression Analysis. New York: Wiley, 1966.
31. Hengsbach, G., and Odoni, A.R. "Time Dependent Estimates of Delays and Delay Costs at Major Airports," Cambridge: M.I.T. Flight Transportation Laboratory Report R75-4, 1976.
32. Hoaglin, D.C., and Welsch, R.E. "The Hat Matrix in Regression and ANOVA." Cambridge: National Bureau of Economic Research Working Paper WP 901-77, 1977.
33. Mosteller, F., and Tukey, J.W. Data Analysis and Regression. Reading, Mass.: Addison-Wesley, 1977.

34. National Bureau of Economic Research. "TROLL Experimental Programs: Model Sensitivity Analysis." Cambridge: NBER Computer Research Center, 1976.
35. Neter, J., and Wasserman, W. Applied Linear Statistical Models. Homewood, Ill.: Irwin, 1974.
36. O'Brien, F.J. "C.A.B. Origin-Destination Statistics", unpublished note, Lockheed-California Company, 1976.
37. Pindyck, R.S., and Rubinfeld, D.L. Econometric Models and Economic Forecasts. New York: McGraw-Hill, 1976.
38. Scherer, F.M. Industrial Market Structure and Economic Performance. Chicago: Rand McNally, 1971.
39. Simpson, R.W. "An Analysis of Airline Costs." M.I.T./N.A.S.A. Workshop on Airline Systems Analysis, Waterville Valley, 1972.
40. Simpson, R.W. "A Theory for Domestic Airline Economics," unpublished lecture notes for the M.I.T. course 16.74 - Air Transportation Economics, 1974.
41. Swan, W. "The Complete FA-4 Memo." Cambridge: M.I.T. Flight Transportation Laboratory Technical Memorandum 72-10, 1972.
42. Taneja, N.K. The Commercial Airline Industry. Lexington, Mass.: Heath-Lexington Books, 1976.
43. Welsch, R.E. "Principal Components and the Singular Value Decomposition," unpublished lecture notes for the M.I.T. course 15.063/18.457 - Statistics for Model Building, 1977.

## BIOGRAPHIC SKETCH

Steven E. Eriksen was born in Boston, Massachusetts on December 13, 1946. He attended public schools in Needham, Massachusetts, and in 1965 was graduated from Tabor Academy in Marion, Massachusetts.

He did his undergraduate work at M.I.T. where he received a Bachelor of Science degree in 1969. While an undergraduate he worked part-time and during summers as a designer and computer programmer at the M.I.T. Instrumentation Laboratory (now the Charles Stark Draper Laboratory).

After working as an aeronautical engineer at the Aircraft Division of Northrop Corporation in Hawthorne, California, he returned to Massachusetts where he received an M.B.A., with honors, from Boston University in 1972.

While preparing for his Ph.D., Mr. Eriksen has taught courses at the University of Massachusetts, Boston University, Babson College, and Northeastern University, in addition to maintaining numerous teaching and research assistantships at M.I.T. He is currently Assistant Professor of Management at Northeastern University.

Dr. Eriksen is a member of Beta Gamma Sigma, Operations Research Society of America, The Institute for Management Sciences, American Statistical Association, and Air Transportation Research International Forum.