

AIRLINE MARKET SHARE MODELING IN ORIGINATING CITY MARKETS

by

JAMES EDWARD DAVIS

B.S. University of Michigan (1984)

SUBMITTED TO THE DEPARTMENT OF
AERONAUTICS AND ASTRONAUTICS
IN PARTIAL FULFILLMENT OF THE DEGREE OF

MASTER OF SCIENCE

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
August, 1989

© James Edward Davis, 1989. All Rights Reserved

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

Signature of Author _____
Department of Aeronautics and Astronautics
August, 1989

Certified by _____
Dr. Peter P. Belobaba, Thesis Supervisor
Department of Aeronautics and Astronautics

Certified by _____
Prof. Robert W. Simpson, Thesis Supervisor
Department of Aeronautics and Astronautics

Accepted by _____
Prof. Harold Y. Wachman
Chairman, Department Graduate Committee

MASSACHUSETTS INSTITUTE
OF TECHNOLOGY

SEP 29 1989

LIBRARIES

ARCHIVES

AIRLINE MARKET SHARE MODELING IN ORIGINATING CITY MARKETS

by

JAMES EDWARD DAVIS

Submitted to the Department of Aeronautics and Astronautics
on August 9, 1989, in partial fulfillment of the requirements
for the Degree of Master of Science

ABSTRACT

The Airline Deregulation Act of 1978 has not only affected the way airlines compete with each other, it has also changed the distribution channels that were once rigidly controlled by the airlines, travel agents and Civil Aeronautics Board. In recent years, the cost to airlines of having travel agents distribute their tickets has increased significantly. Since a large portion of these costs stems from airlines paying "overrides" to travel agents who exceed a baseline market share, which has been set by airline management, the determination of a carrier's market share in an originating city market can have serious profit implications.

The problem addressed in this thesis is how an airline might predict its share of the passenger market out of an originating city. In this thesis, six different mathematical models, relating a carrier's market share to a set of exogenous variables, are proposed. Using actual airline market share data, each of these models is calibrated and tested in fifteen selected test markets and a statistical determination of each model's accuracy is performed.

Of the six models tested, a non-linear multivariable regression model, relating a carrier's market share to its: seat share; frequency share; proportion of non-stop flights; and proportion of total destinations served, out of an originating city, is preferred. The use of this market share model should help airlines predict their share of the originating city passenger market, and thus, provide them with a means of setting market share quotas based upon a known set of service variables.

Thesis Supervisor: Dr. Peter Belobaba
Title: Research Engineer & Lecturer, Dept. of Aero. & Astro.

Thesis Supervisor: Robert Simpson
Title: Professor of Aeronautics & Astronautics

Contents

	Page
Chapter 1. Introduction	
1.1 Motivation for Thesis	8
1.1.1 Background	8
1.1.2 History - Development of Airline Distribution Channels	9
1.1.3 Current Practices in the Distribution of Airline Tickets	13
1.1.4 Scope of the Problem	15
1.2 Objective of Thesis	17
1.3 Structure of Thesis	19
Chapter 2. Classical Airline Market Share Models	
2.1 Introduction	22
2.2 An Overview of Past Research into Airline Market Share Modeling	24
2.3 Applications of Market Share Models in Originating City Markets	33
Chapter 3. Application of Classical Market Share Models in Originating City Markets	
3.1 Introduction	35
3.2 Description of Data Used in Model Calibrations	38
3.2.1 Market Share Data	39
3.2.2 Frequency Share and Seat Share Data	40
3.2.3 Selected Test Cities	41
3.3 Calibration of MS-FS Model	42
3.4 Calibration of MS-SS Model	57
3.5 A Comparison of the MS-FS and MS-SS Models	70
Chapter 4. Application of Market Share Regression Models in Originating City Markets	
4.1 Introduction	73
4.2 Calibration of the Seat Share Regression Model	77
4.3 Calibration of the Seat Share - Frequency Share Regression Model	92
4.4 Calibration of the Multivariable Multiplicative Regression Model	106
4.5 Calibration of the Multivariable Linear Regression Model	112
4.6 Summary	116
Chapter 5. Summary and Conclusions	
5.1 Overview	119
5.2 Findings and Conclusions	121
5.3 Further Research	126
Selected References	129

List of Figures

	Page
2.1 Classical Market Share vs. Frequency Share Curves	28
3.1 Market Share vs. Frequency Share in Boston	44
3.2 Market Share vs. Frequency Share in Chicago	44
3.3 Market Share vs. Frequency Share in Detroit	45
3.4 Market Share vs. Frequency Share in Huntsville	45
3.5 Market Share vs. Frequency Share in Los Angeles	46
3.6 Market Share vs. Frequency Share in Madsion	46
3.7 Market Share vs. Frequency Share in Memphis	47
3.8 Market Share vs. Frequency Share in Milwaukee	47
3.9 Market Share vs. Frequency Share in Minneapolis	48
3.10 Market Share vs. Frequency Share in Omaha	48
3.11 Market Share vs. Frequency Share in Orlando	49
3.12 Market Share vs. Frequency Share in Rochester	49
3.13 Market Share vs. Frequency Share in Saginaw	50
3.14 Market Share vs. Frequency Share in Seattle	50
3.15 Market Share vs. Frequency Share in St. Louis	51
3.16 Market Share vs. Seat Share in Boston	60
3.17 Market Share vs. Seat Share in Chicago	60
3.18 Market Share vs. Seat Share in Detroit	61
3.19 Market Share vs. Seat Share in Huntsville	61
3.20 Market Share vs. Seat Share in Los Angeles	62
3.21 Market Share vs. Seat Share in Madison	62
3.22 Market Share vs. Seat Share in Memphis	63
3.23 Market Share vs. Seat Share in Milwaukee	63
3.24 Market Share vs. Seat Share in Minneapolis	64
3.25 Market Share vs. Seat Share in Omaha	64
3.26 Market Share vs. Seat Share in Orlando	65
3.27 Market Share vs. Seat Share in Rochester	65
3.28 Market Share vs. Seat Share in Saginaw	66
3.29 Market Share vs. Seat Share in Seattle	66
3.30 Market Share vs. Seat Share in St. Louis	67
4.1 $\ln(\text{Market Share})$ vs. $\ln(\text{Seat Share})$ in Boston	80
4.2 $\ln(\text{Market Share})$ vs. $\ln(\text{Seat Share})$ in Chicago	80
4.3 $\ln(\text{Market Share})$ vs. $\ln(\text{Seat Share})$ in Detroit	81
4.4 $\ln(\text{Market Share})$ vs. $\ln(\text{Seat Share})$ in Huntsville	81
4.5 $\ln(\text{Market Share})$ vs. $\ln(\text{Seat Share})$ in Los Angeles	82
4.6 $\ln(\text{Market Share})$ vs. $\ln(\text{Seat Share})$ in Madison	82
4.7 $\ln(\text{Market Share})$ vs. $\ln(\text{Seat Share})$ in Memphis	83
4.8 $\ln(\text{Market Share})$ vs. $\ln(\text{Seat Share})$ in Milwaukee	83
4.9 $\ln(\text{Market Share})$ vs. $\ln(\text{Seat Share})$ in Minneapolis	84
4.10 $\ln(\text{Market Share})$ vs. $\ln(\text{Seat Share})$ in Omaha	84
4.11 $\ln(\text{Market Share})$ vs. $\ln(\text{Seat Share})$ in Orlando	85
4.12 $\ln(\text{Market Share})$ vs. $\ln(\text{Seat Share})$ in Rochesier	85
4.13 $\ln(\text{Market Share})$ vs. $\ln(\text{Seat Share})$ in Saginaw	86
4.14 $\ln(\text{Market Share})$ vs. $\ln(\text{Seat Share})$ in Seattle	86
4.15 $\ln(\text{Market Share})$ vs. $\ln(\text{Seat Share})$ in St. Louis	87

4.16	ln(Market Share) vs. ln(Frequency Share) in Boston	94
4.17	ln(Market Share) vs. ln(Frequency Share) in Chicago	94
4.18	ln(Market Share) vs. ln(Frequency Share) in Detroit	95
4.19	ln(Market Share) vs. ln(Frequency Share) in Huntsville	95
4.20	ln(Market Share) vs. ln(Frequency Share) in Los Angeles	96
4.21	ln(Market Share) vs. ln(Frequency Share) in Madison	96
4.22	ln(Market Share) vs. ln(Frequency Share) in Memphis	97
4.23	ln(Market Share) vs. ln(Frequency Share) in Milwaukee	97
4.24	ln(Market Share) vs. ln(Frequency Share) in Minneapolis	98
4.25	ln(Market Share) vs. ln(Frequency Share) in Omaha	98
4.26	ln(Market Share) vs. ln(Frequency Share) in Orlando	99
4.27	ln(Market Share) vs. ln(Frequency Share) in Rochester	99
4.28	ln(Market Share) vs. ln(Frequency Share) in Saginaw	100
4.29	ln(Market Share) vs. ln(Frequency Share) in Seattle	100
4.30	ln(Market Share) vs. ln(Frequency Share) in St. Louis	101

List of Tables

	Page
3.1 Selected Test Cities	41
3.2 Results of Frequency Share Analysis	55
3.3 Results of Seat Share Analysis	69
4.1 Results of Seat Share Regression Analysis	90
4.2 Results of SS-FS Regression Analysis	103
4.3 Results of Multivariable Multiplicative Regression Analysis	111
4.4 Results of Multivariable Linear Regression Analysis	115

Acknowledgements

I would like to thank my advisors Dr. Peter Belobaba and Professor Robert Simpson for all the help they have given me, not only during this research but also during my entire academic stay at MIT. My special thanks to Peter, who kept this project moving in the right direction at times when I seriously doubted how it would all come together. His generous contributions of time and assistance, during this research and during my coursework at MIT, have certainly been invaluable.

I would also like to thank The General Electric Company for their support during this endeavor. In particular, their financial support and their willingness to allow me time away from work to pursue my studies is greatly appreciated. I also extend thanks to Northwest Airlines for allowing me to engage in this research and for the materials necessary to complete this project.

Lastly, I extend my heartfelt thanks to my Editor in Chief and best friend, Janet. Her patience, support, and love during this long arduous process, along with her impeccable editorial assistance, have certainly contributed to my academic achievements. Her assistance will always be remembered.

CHAPTER 1

INTRODUCTION

1.1 Motivation for Thesis

1.1.1 Background

The Airline Deregulation Act of 1978 (ADA) has not only affected the way airlines compete with each other, it has also changed the distribution channels that were once rigidly controlled by the airlines, travel agents, and Civil Aeronautics Board (CAB). Deregulation altered the rules of play among these three groups. Prior to deregulation, both the airlines and the CAB were responsible for accrediting travel agents, the only source of airline tickets aside from the airlines. The CAB regulated fares for all interstate airlines and set the commissions that airlines were to pay travel agents for the tickets they issued.

"The role of travel agents in the distribution system was expected to decline after deregulation as airlines experimented and ultimately adopted methods of distribution not sanctioned by regulation."¹ Commissions and

commission rates were expected to decline as airlines decreased fares and entry into the travel agency business became freer. Alternative distribution channels were expected to alleviate the reliance of airlines and passengers on the services of travel agents.

What has transpired since 1978, however, is the complete opposite of what was expected. Deregulation has strengthened the role of the travel agent in the distribution process, largely due to the proliferation of fares and their frequent revisions by airlines. Not only have commission rates increased since deregulation, but the development of "commission override programs" has further increased the cost to airlines of travel agency services.² "Travel agency commissions paid by U.S. major carriers on domestic flights increased from \$527 million in 1977 to slightly more than \$3.1 billion in 1985."³

To understand the current relationship between airlines and travel agents, and why it is costing airlines such a large sum of money to do business with these travel agents, this chapter presents a brief history of how the distribution channels evolved and a description of the current practices that govern the distribution of airline tickets

1.1.2 History - Development of Airline Distribution Channels

In 1945, the Air Transport Association of America, comprised of the

major U.S. airlines, created the Air Traffic Conference (ATC) to deal with the problems and dilemmas of distributing and marketing air transportation services. Under agreements adopted by the member airlines, the ARC established methods by and conditions under which accredited travel agents could sell airline tickets. Regulations guiding the accreditation of travel agents by airlines, including standards for business practices, financial requirements, personnel training, and agency location, were also established by the ATC. A key aspect of these agreements was the exclusiveness provision that restricted the sale of airline tickets to accredited travel agents and to the airlines themselves, agreements that the CAB exempted from U.S. antitrust laws. The CAB also acted as a third party in determining the level of commission rates paid by airlines to travel agents. Since commission rates were a fixed percentage of ticket prices and the CAB regulated prices were a function of length of haul, travel agents would earn more money on tickets sold for long haul flights than on those sold for short haul flights. This rigidly controlled system remained intact for 34 years with few changes made.

In September of 1979, one year after the passage of the Airline Deregulation Act, the CAB initiated an investigation into the competitive marketing of air transportation services to determine if and how competition could be introduced into the system. The investigative order charged that:

"...the means by which, and the locations from which, the retailers may compete for sales with suppliers are limited, and the classes of consumers with which the retailers may do business are circumscribed. The result is little or no competition on price at any level, and carefully controlled competition between suppliers and retailers. Competition

among retailers is largely confined to providing services, which consumers may not want or need."⁴

During the course of this investigation, the CAB voted in 1980 to eliminate fixed industry-wide commissions paid to travel agents. This paved the way for individual negotiations of commissions between carriers and agents. Contrary to the CAB's expectation, however, the deregulation of sales commissions led to an increase in the average commissions paid by the airlines (8.3% in 1977 to 9.6% in 1985).

The rapid rise in commission rates was due to several factors. First, the failure to deregulate distribution networks while allowing carriers to negotiate their own commission levels led not only to an increase in the commission rates, but also to an initiation of commission override programs. Since travel agents were still the airlines' only source for distributing tickets (by 1979 travel agents sold 53% of all airline tickets), providing monetary incentives was a method of influencing their choice, and ultimately the consumer's choice, when making a travel decision. In addition, carriers without their own computer reservation systems (CRS) saw this as an opportunity to alleviate some of the bias that favored the host CRS carrier. Since the non-owning carriers were usually not given the best position in the display, an additional payment to a travel agent might prompt her to search for that carrier's particular flight.

A second reason for the increase in commission rates was the confusion that existed in the marketplace as a result of deregulation. With many new

carriers entering the marketplace and with the development of additional fares and fare classes, the carriers found it necessary to rely on an expert to search for "the best deal." "...[T]he confused marketplace of deregulation allowed for agents to sell information to customers. Deregulation increased the value to consumers of having an expert search for them, since it was more likely that a good agent could find a new airline service, or a fare that the customer could not find herself."⁵ Travel agents' share of the total domestic and international air travel increased from 57 to 74 percent between 1978 and 1984, and in dollar volume, their sales increased from \$10.8 billion to \$24.5 billion.⁶ Clearly, travel agents achieved a dominant role in the distribution of airline services, one that enabled them to negotiate higher commission fees from the airlines.

In 1982, the CAB, in its report entitled "Competitive Marketing Investigation", ended the exclusiveness provision in the marketing of air transportation services. This allowed carriers to develop and utilize other forms of distribution networks during the two year phase out of the exclusiveness provision. The report also provided easier entry into the travel agency industry by requiring the ATC to liberalize the process of accrediting new applicants, a process that had already been under way since 1979. But by 1984, the first year carriers were totally free to distribute tickets in any manner, the more than 25,000 travel agents operating in the U.S. (up from 15,000 in 1978) were selling over 70% of all tickets.⁷ Clearly, travel agents had control of the airline industry distribution network and "... any carrier that upsets its relationships with its travel agents, either by reducing commission

rates or by establishing an alternate distribution channel, is vulnerable to reduced travel agent bookings⁸ and ultimately reduced profits.

1.1.3 Current Practices in the Distribution of Airline Tickets

The changes in the distribution channels that evolved after the passage of the ADA in 1978 and the deregulation of travel agent services in 1982 are, for the most part, still intact today. At this point in time, there has still been no establishment of alternate distribution channels although some airlines (World Airways, in particular) have experimented with selling tickets through Ticketron. Some large corporations and other businesses have established their own business travel departments or hosted in-house travel agencies enabling them to deal directly with the airlines. In this manner, they have been able to negotiate price discounts for advance purchases of block tickets (especially on routes that are frequently traveled by executives). Aside from these small changes, however, travel agents have remained in control of the distribution channels, still selling over 70% of all airline tickets.

In the past few years, the merger of many travel agencies and the buy-out of smaller agencies by larger ones have also served to strengthen the control travel agents have over the distribution channels. Many of these new "mega travel agencies" have been able to exert a significant amount of control over the airlines they serve. Through these mergers, large travel agencies,

with many offices throughout a city, have been able to capture a large percentage of that city's originating air traffic. By obtaining large market shares in originating cities, travel agencies are able to gain negotiating power with the airlines they serve.

As an example of this power, Pan Am, in 1982, was encountering severe financial difficulties and, as a result, made the decision to slash its commissions and commission override programs in an effort to reduce its operating costs. Although successful in reducing its costs, Pan Am also saw a sudden drop in its advance bookings as a result of this action. Management at Pan Am quickly realized the power of their travel agents and publicly stated that "...[W]e have reinstated these bonuses and overrides emphasizing our commitment to travel agents as our airline's primary distribution channel."⁹

United Airlines also proposed a new rate structure between carriers and agents after the CAB deregulated commission payments in 1979. Since United's routes were generally longer than those flown by their competitors, its commissions per ticket tended to be higher than the industry average. United's proposal of a fixed rate of \$8.50 on all tickets was opposed so vigorously by the travel agents that it never found its way into the marketplace.

As previously stated, the means by which most airlines compensate travel agents for the sale of tickets is through both straight commissions and commission override programs (also known as bonus incentives). Straight commissions are simply payments based on a fixed percentage of the ticket

price. For example, if the carrier-agent commission rate is 9 percent, and the travel agent sells a ticket with a \$400 fare, then the airline pays the agent \$36. Commission override programs are generally one of two types. The first type of override program is "...usually in the form of additional payments from an airline to a travel agency of up to 5 percent of the ticket price when agency sales of that airline's tickets exceed a certain volume."¹⁰ The second type of override program involves an additional payment from an airline to a travel agency, at a fixed percent of the ticket price, when the travel agency's market share of that airline's ticket sales exceeds a certain percentage. For example, carrier A may agree to pay Agent X an additional 5 percent of the total ticket price on all tickets sold if 25 percent of Agent X's sales are for carrier A. Thus, if 26 percent of Agent X's sales are for carrier A, then Agent X will receive an additional 5 percent commission on the tickets sold for Carrier A.

1.1.4 Scope of the Problem

In 1985, the Air Traffic Conference estimated that 60 percent of all domestic tickets and 80 percent of all international tickets sold through travel agents benefited from some type of override program. In that year alone, total ticket sales by travel agents amounted to more than \$32 billion.¹¹ At a minimum, if 60 percent of that total revenue received an average override commission of 3 percent (a modest estimate), then airlines would have paid travel agents more than \$576 million in override commissions alone in that

year. This obviously represents a staggering loss in potential profits for the airline.

How airlines structure these override programs can significantly affect their profit potential. Whether the override program is based on a volume of sales by the travel agent or a specific market share achieved by the agent, the setting of a baseline goal that a travel agent attempts to meet is of critical importance.

Determining a baseline level of sales or market share in an originating city market can be a very difficult process in today's deregulated environment. Frequent changes in scheduling, destinations served, and fares can significantly alter an airline's share of the market in an originating city. Seasonality in air transportation demand can also have an effect on an airline's share of the market, especially if the carrier is strictly a North-South or East-West carrier (e.g. Eastern Airlines or United Airlines).

The problem to be addressed in this thesis is how an airline might predict its share of the market in an originating city and how it may determine its relative market dominance in that city. Making a forecast based upon some exogenous variables which are relevant to both the particular airline and market would allow management to set baseline sales goals for travel agencies based upon a known set of variables. Instead of designing incentive programs based simply upon a travel agency's past performance (prior month's sales), incentive targets could be based upon an expected level

of service that an airline has committed to or to an expected future market dominance condition.

1.2 Objective of Thesis

Market share modeling was an important aspect of airline management in the previously regulated environment. With the number of competitors and fare levels fixed by the CAB in each origin-destination (O-D) market, competition between air carriers was most prevalent in the areas of flight scheduling and passenger service. As a way of studying the competitive behavior of airlines, market share models were developed mainly to determine the relationship between flight scheduling and market share. Understanding this relationship and the ability to predict the response of competitors to frequency changes were fundamental to an airline's profitability.

Although these market share models were developed to study and understand the competitive behavior of airlines in O-D markets (see Chapter Two), these same models or similar models might be adapted for the analysis of a carrier's market share or market dominance in an originating city. As previously stated, an airline's ability/inability to predict the share of originating passengers it can attract in a particular city can have a large impact on its profit margins. In addition, information regarding a carrier's

relative strength or dominance in a city can be used to set advertising budgets and to establish travel agency commission levels. If a carrier has achieved a dominant position in a city with few competitors, perhaps there is no need to match the commissions that other carriers are paying.

The objective of this thesis is to apply previously developed market share models and new derivatives of these models to originating city markets. Using actual airline market share data, several of these models will be calibrated and a statistical determination of the accuracy of each model will be made. The results of these modeling efforts will also provide information as to which variables are most important in predicting an airline's share of the originating passengers in a city.

In today's unregulated environment, competition between carriers can vary widely from city to city. Factors such as whether the city is used as a hub by one or more airlines, the number of carriers operating out of the city, and the size of the originating market (in terms of number of departures) can all have an affect on the competition between air carriers. In determining which models are the most accurate for predicting market share, a large cross-section of cities will be analyzed so that the selected model will produce satisfactory results in any type of originating market.

1.3 Structure of Thesis

The remainder of this thesis will be presented in four chapters. The second chapter of this thesis will present an overview of classical airline market share models. This chapter will review the current literature on market share-frequency share models and their relevance in a regulated environment. The author will then show how these models may be applied in a deregulated environment, specifically to an originating city market.

In Chapter Three of this thesis, the author will begin to apply both the market share-frequency share model and the market share-seat share model in several originating city markets. As discussed in Chapter Two, both of these models were very effective in predicting a carrier's market share in an origin-destination (O-D) market. Before fitting these two models to the actual market share-frequency share data, the author will present an overall description of the data used in the analysis including both the limitations and the accuracy of the data being used. Using simple recursive methods, these two models will be applied to the data, and the exponents, which are unique to each city, will be determined. Finally, an error analysis will be presented with a discussion of the results.

In Chapter Four, the author will apply a simple non-linear market share model, first presented by Bond [2], to the originating city markets

analyzed in Chapter Three. Although this model was not extremely effective in O-D markets, its use in an originating city market will prove to be of more importance. A simple logarithmic transformation will be made so that this model can be fit to the data using a linear regression technique. The author will then add to Bond's model additional variables that may be used in predicting a carrier's market share in an originating city market. Finally, an error analysis will be presented and the accuracy of the models used in this chapter will be discussed.

In the concluding chapter of this thesis (Chapter Five), the author will summarize the results of the six models presented in Chapters Three and Four. A comparison of the prediction accuracy of all six models and a determination of which model offers the best results will then be made. The author will also briefly discuss how these results may be used in a practical way by airline management. Finally, a discussion of additional areas of research, including further extensions of the models presented for market share modeling, will be included.

1 Michael E. Levine, "Airline Competition in Deregulated Markets: Theory, Firm Strategy, and Public Policy", Yale Journal of Regulation, Volume 4: 363, 1987, p. 414.

2 Commission override programs usually involve a payment, from an airline to a travel agency, of a certain percentage of the ticket price when an agencies' sales of that airline's tickets exceed a certain volume or market share.

3 John Meyer and Clinton Oster, Deregulation and the Future of Intercity Passenger Demand, (Cambridge, Ma.: The MIT Press, 1978) p. 125.

4 Meyer and Oster, p. 127

5 Levine, p. 127.

6 James O. Lect, "Transportation in the Era of Deregulation," Transportation Quarterly, Volume XXXIX: 2, April, 1985, p. 180.

7 Meyer and Oster, p. 127.

8 Elizabeth E. Baily, David R. Graham, and Daniel P. Kaplan, Deregulating the Airlines (Cambridge, Ma: The MIT Press, 1985), p. 185.

9 Pan Am Corporation Annual Report, 1982, p. 2.

10 Meyer and Oster, p. 128.

11 Meyer and Oster, p. 128.

CHAPTER 2

Classical Airline Market Share Models

2.1 Introduction

Competition between airlines under regulation was most intense in the areas of flight scheduling and passenger service. With both fares and the number of competitors in each market determined by the Civil Aeronautics Board (CAB), and little variation in the equipment operated by the airlines, market share was determined by factors such as frequency of service, departure and arrival times, and the overall image of the carrier. Even service amenities offered by the different airlines were similar. If one carrier upgraded its passenger service in an origin-destination (O-D) market, it was usually matched by the other carriers so that the final product offered was not highly differentiated.¹

Competition was thus left to the area of flight scheduling. It was in this arena "... where the competitive energies of the trunk carriers were unleashed, and where the boom and bust cycles of airline profitability were causally linked to the competitive environment."² Understanding the game of flight scheduling and trying to predict the competitive behavior of competing air carriers is what prompted airline management and academics to develop airline market share models.

"A market share model shows the relationship between the share of passenger traffic for an airline in a given competitive market and the factors that describe the quality of service offered in the market by the carrier."³ Early research in this area indicated that the most significant variable in explaining a carrier's market share was its frequency share. It was a common notion that he who flies the most seats gets the most passengers. It was this notion that led to many of the over-capacity problems that plagued the airline industry under regulation. When carrier A increased frequency in a given market (attempting to increase its market share) carrier B would, if equipment was available, match its competitor's move. If carrier B was able to match, the end result would usually be lower load factors for both carriers and increased costs per passenger mile.

The next section of this chapter will outline some of the previous research that has been conducted in the area of market share modeling. As stated previously, this area of research was particularly useful to airline management in the era of regulation. As such, little, if any, new research has evolved since 1978, the year the airline industry was deregulated. In section three of this chapter, the author will discuss why and how these models may still be used in a deregulated environment, particularly in originating city markets. It is in this area that the author will attempt to provide original ideas, since all previous work has focused on market share modeling in origin-destination (O-D) markets.

2.2 An Overview of Past Research into Airline Market Share Modeling

Prior to any analytical or statistical work, it was commonly perceived by most airline managements that there was a linear relationship between market share (MS) and frequency share (FS). In a market with two carriers, one with a 60 percent frequency share and the other with a 40 percent frequency share, it was generally assumed that their market shares would also be split 60-40. Any deviations from this $MS=FS$ relationship were assumed to be caused by the existence of certain carrier attributes, not easily measured, that made one airline more attractive than another.

Taneja [15] noted that although the dominant variable used to explain market share was frequency share, the relationship was not linear. Using data from the top 50 O-D U.S. domestic markets in 1966, he demonstrated the existence of an S-shaped curve between market share and frequency share. Using the example above, it was found that the carrier (in a two competitor market) that offered 60 percent of the frequency share would obtain a market share greater than 60 percent, and the carrier with a 40 percent frequency share would obtain a market share of less than 40 percent. Thus, a carrier with a greater share of frequency than its competitors should receive a more than proportional share of the traffic.

Having noted this phenomenon, Taneja attempted to model this relationship in various ways. Using the model,

$$MS = A \times FS^3 + B \times FS^2 + C \times FS + D \quad (2.1)$$

where A,B,C, and D are constants determined from a multiple regression, Taneja found that the only significant variables (using the 1966 top U.S. O-D markets) were,

$$MS = C \times FS + D \quad (2.2)$$

However, this did not account for the S-shaped curve.

To account for the S-shaped curve, Taneja proposed additional models which, along with the variable frequency share, also included the number of competitors in the specific O-D markets. One of these models,

$$MS = FS^\alpha \times N_c^\beta \quad (2.3)$$

where N_c is the number of competitors and α and β are estimated parameters, accounted for the S-shaped relation in his selected data; however, the accuracy of the results (predicted market share compared to actual market share) was not overwhelming. Additionally, the coefficients α and β were neither market dependent nor carrier dependent.

Bond's [2] proposed market share-frequency share model,

$$MS = \alpha_i \times FS_i^\beta \quad (2.4)$$

where α_i and β_i are constants for carrier i , postulated that the market frequency relationship is carrier dependent rather than route dependent. These carrier dependent parameters were "... functions of the measurable attributes of the carrier's defined product policy and the attributes of the product as perceived by the air traveler."⁴

Bond calibrated the model for the top U.S. domestic trunk carriers using market share and frequency share data for the first quarter of 1968 in the top 50 O-D markets. Although this model did not fully account for the S-shaped phenomenon noted by Taneja and others, the model did highlight the importance of perceptions that the traveling public held for each air carrier. These perceptions or image factors, Bond hypothesized, contribute to the determination of a carrier's market position.

Research by Simpson [12] also confirmed that market share is an S-shaped function of frequency share with frequency being the dominant explanatory variable. Except for station activity (the number of departures out of the origin city), he found very little correlation with any other variables. He did note, like Taneja, that the number of competitors in a market affects the structure of the S-shaped market share curve.

Simpson proposed that for airline i competing with competitors j, k, \dots

$$\frac{MS_i}{MS_j} = \left(\frac{FS_i}{FS_j} \right)^\alpha \tag{2.5}$$

where MS_i equals the market share of passengers for airline i and FS_i equals the frequency share of departures for airline i , and α is determined by

regressing data points covering a sample of n-competitor markets. For airline i competing with n carriers in a market, Equation 2.5 reduced to

$$MS_i = \frac{FS_i^\alpha}{\sum_{j=1}^n FS_j^\alpha} \quad (2.6)$$

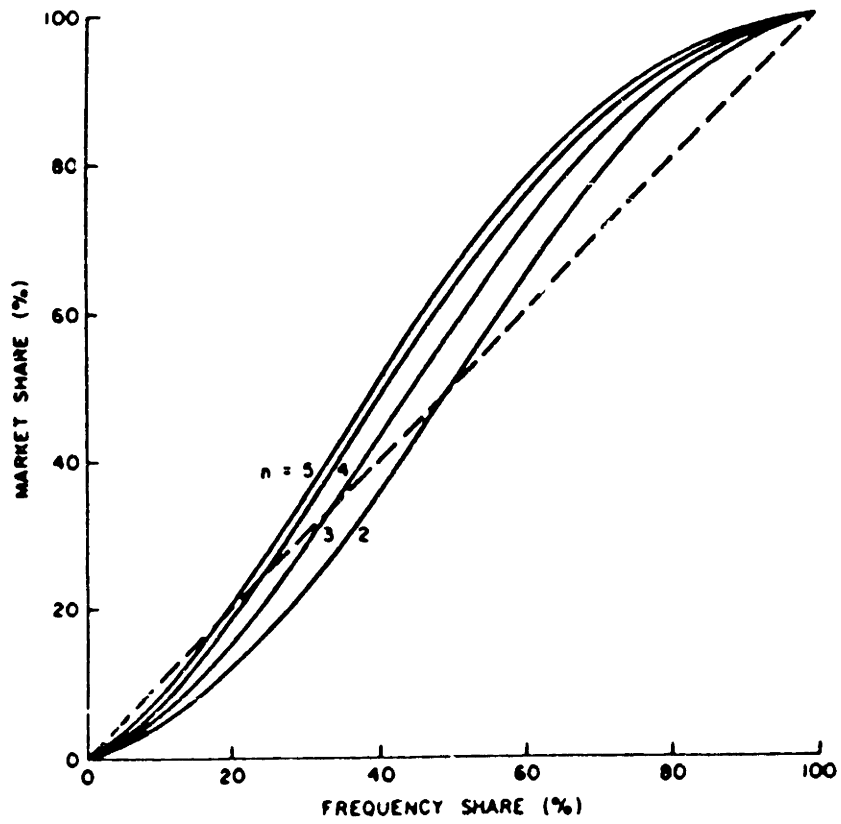
This new model produced a family of curves for n=2,3,4,... competitor markets as shown in Figure 2.1. In developing these curves, it was assumed that if carrier i is faced with three competitors (n=4), for example, each will split the remaining frequency evenly. If, in fact, one of the three competitors had a very low frequency share, the model would reduce to the n-1 curve. Thus the model would produce points in the areas bounded by the curves. Note also that each curve in the family crosses the main diagonal at a point 1/n, which is the case of all competitors having equal frequencies.

Using market share data for 1966 and 1967 in the largest U.S. O-D markets, Simpson estimated the coefficient α in Equation 2.6 and determined that a value of $\alpha=1.45$ produced the best fit. Though the exponent α was neither carrier dependent nor market dependent, Simpson did note that the standard error in the model left room for additional explanatory variables such as a carrier's station activity in the originating market or a carrier's local image and advertising effort.

In his Master's Thesis "Competition in Air Transportation, An Economic Approach", Renard [10] presented five different formulations of Equation 2.6 involving variations of the coefficients according to either the

FIGURE 2.1

Classical Market Share vs. Frequency Share Curves



number of competitors in the market or to the airline's identity or to both simultaneously. Like Simpson, Renard's first formulation with this model assumed that α was simply a constant, neither carrier nor market dependent. The results obtained with this formulation were identical to Simpson's.

Renard's second formulation of the model in Equation 2.6 assumed that the exponent α depended upon the number of competitors in the market. Using linear regression techniques, Renard determined a best α for the markets with two to six competitors. Although there was no correlation between the value of α and the number of competitors, n , the results of his regression analysis did indicate that by breaking up the markets according to the n competitor status, a better model could be fit to the data.

The next formulation of the model in Equation 2.6 that Renard examined assumed that α depended on the specific airline. Thus Equation 2.6 became:

$$MS_i = \frac{FS_i^{\alpha_i}}{\sum_{j=1}^n FS_j^{\alpha_j}} \quad (2.7)$$

where α_j was a different index for each airline. Renard estimated the value of this coefficient for the top U.S. trunk carriers in various O-D markets. Having found values of α ranging from .30 to 1.0, he noted that an α close to 1.0 in this S-shaped model corresponded to an airline's ability to achieve a proportionate share of the market. He also noted that there was not a strong relationship between a carrier's profitability and its competitive position as

determined by α (i.e. a large value of α did not correspond to large profits earned by the carrier.).

Combining methods two and three above, Renard conducted an analysis assuming that the coefficient α varied with both the carrier and the number of competitors. For each of the U.S. trunk airlines, Renard produced a matrix of exponents for each of the carriers in the n-competitor markets (n equals 2 through 6) in which that carrier operated. Although this formulation produced results with the lowest residual error of all methods considered, there was no clear relationship between the carrier's competitive position, α , and the size of the market in which it operated. The results did, however, lead Renard to conclude that a carrier's unmeasurable attributes played a significant role in its ability to capture a proportionate share of the traffic in a market. This conclusion led to his proposal of "a multiple K, single α market share model."

Renard's multiple K single α market share model allowed for differences in attractiveness between airlines by the inclusion of multiplicative coefficients in Equation 2.6. The model then became:

$$MS_i = \frac{K_i \times FS_i^\alpha}{\sum_{j=1}^n K_j \times FS_j^\alpha} \quad (2.8)$$

where the K_j represented the level of attractiveness for each carrier.

Calibrating this model, using the identical data bases analyzed in the previous formulations, Renard determined K factors for each of the trunk carriers and

a single α covering any n-competitor market. The results produced by this formulation led Renard to conclude that this type of model, which includes a multiplicative coefficient representing the attractiveness of each carrier, best represents the share of the market obtained by an airline in a regulated environment.

Further research by Simpson [13] led to a new model that further quantified Renard's idea of carrier attractiveness. In this new model, Simpson introduced the idea of image factor, both on a global basis and on a local market basis. This new model became:

$$MS_i = \frac{I_i \times FS_i^\alpha}{\sum_{j=1}^n I_j \times FS_j^\alpha} + \Delta_i \quad (2.9)$$

where I_j represents the system wide image factor for each airline, and Δ_j represents the local image correction in the market.

The addition of the local image correction factor in the market was a way of further quantifying the residuals that were present in Renard's model. This local image factor was used to account for the differences in, for example, the amount of advertising by a carrier in one market versus another market. Simpson also found that the statistic "seat share," a carrier's share of the total seats offered in a market, was statistically equivalent to frequency share in the domestic markets. This notion of seat share is an important statistic that will be used in market share models for originating city markets, models which will be presented in Chapters Three and Four.

The market share models proposed by Taneja, Bond, Renard, and Simpson summarize the past research conducted in this area. However, since the advent of deregulation, little, if any, new research has appeared in the literature. In a regulated environment, airline market share modeling between an origin and destination city was a much simpler task. With both passenger fares and the number of competitors in a market fixed by the CAB, an airline's market share was dependent on only a few variables, as the above research has indicated. Now, however, the playing field is much more complicated. Airlines are free to determine their own fares and choose the markets in which they will operate, factors that would indeed complicate the modeling process. More importantly, the idea of a market share model in an origin-destination city pair has been complicated by the development of connecting hubs throughout the United States. A carrier's share of the market from Boston to Los Angeles, for example, is not only influenced by the number of competitors operating this route directly, but it is also influenced by the number of carriers offering one-stop or connecting service (through Chicago, Dallas, Denver, or Memphis) or even two-stop service. It is these factors that would make market share modeling in specific O-D markets very difficult today.

2.3 Applications of Market Share Models in Originating City Markets

As stated in Section 2.2, the intent of market share modeling is to describe the factors that determine an airline's ability to attract passengers in a given competitive market. Although the past research into market share modeling has involved only origin-destination (or city pair) markets, there is no reason why these same models, or similar ones, cannot be applied to originating city markets. Since a market is simply an environment in which two or more parties conduct business, the exact nature of that market is of little importance when trying to model that relationship mathematically.

Under regulation, the need for market share modeling in an originating city did not really exist. This is most likely why modeling efforts were not applied to these markets. The liberalization of the airline-travel agency relationship today, however, has resulted in the need for an airline to predict its share of the market in an originating city.

Factors described by Taneja, Bond, Simpson, and Renard will also be applicable in describing a carrier's market share in an originating city market. Obviously, the number of seats offered by a carrier, the frequency of flights, and the number of carriers operating out of a particular city will all influence a carrier's market share in this type of market. Other factors such as the number of non-stop markets served and total number of markets served out of a particular city will influence a carrier's share of the market.

Because many of the factors that influence a carrier's market share in an O-D market also influence a carrier's market share in an originating city, models identified in Section 2.2 may be applicable to this research. As such, the use of these models, along with variations of them, will be explored in Chapters Three and Four of this thesis. Through this analysis, factors that influence a carrier's share of the passengers in an originating city market will be identified.

¹ Origin-Destination (O-D) markets refer to the demand and supply of air transportation between two cities X and Y, where city X represents the origin of that demand and supply and city Y represents the destination of that demand and supply.

² William E. Fruhan, The Fight for Competitive Advantage: A Study of the U.S. Domestic Trunk Carriers. (Boston, Ma: Division of Research, Harvard Business School, 1972), Part III, p. 1.

³ Nawal K. Taneja, The Commercial Airline Industry. (Lexington, Ma. D.C. Heath and Company, 1976), p. 143.

⁴ Richard O. Bond, Frequency Competition on Domestic Air Routes. Master of Science Thesis, MIT Sloan School of Business, 1970, p. 70.

CHAPTER 3

Application of Classical Market Share Models in Originating City Markets

3.1 Introduction

In Chapter Two, several different types of market share models were presented. Broadly speaking, all of these models can be grouped into two categories: those that can be calibrated using linear regression techniques and those that cannot be calibrated using these techniques. Analysis of the equations presented in the previous chapter reveals that Equations 2.1 through 2.4 can be solved by regression techniques, and that Equations 2.5 through 2.9 require some other approach. These latter models (2.5-2.9), which Renard referred to as interactive market share models, will be the focus of this chapter.¹

Although the phrase "originating city market" has been used extensively up to this point, it would be helpful, before proceeding with any model calibrations, to explicitly define its meaning in the context of the airline industry. An originating city market is a market "... which contains all the residences and business locations for consumers"² who have a demand for

scheduled air transportation service. For example, the originating city market in Boston refers to all consumers who choose to depart out of Logan Airport. There might be some consumers in the metropolitan Boston area, however, who choose to depart out of Bradley Field in Hartford. Thus, originating city markets do not necessarily refer to all consumers who reside within a certain geographic area (or live within a fixed radius of the airport), rather, it refers to a group of consumers who choose to depart out of that city's airport. How airlines share this group of consumers is the topic addressed in Chapters Three and Four.

In the previous chapter, several different classical airline market share models were presented. These models were applied strictly to origin-destination markets (O-D) and have never been tested in originating city markets. The purpose of this chapter is to calibrate two of these models in originating city markets. These models, which have essentially the same formulation but contain different independent variables, are:

$$MS_i = \frac{FS_i^\alpha}{\sum_{j=1}^n FS_j^\alpha} \quad (3.1)$$

and

$$MS_i = \frac{SS_i^\alpha}{\sum_{j=1}^n SS_j^\alpha} \quad (3.2)$$

where,

MS_i = market share of carrier i

FS_i = frequency share of carrier i = $\frac{\text{total \# of departures carrier i}}{\text{total \# of departures all carriers}}$

SS_i = seat share of carrier i = $\frac{\text{total \# of seats offered by carrier i}}{\text{total \# of seats offered by all carriers}}$

α = fitted exponent of frequency share or seat share

j = indices representing single airline operating in the market

n = total number of airlines in the market.

Given market share and frequency share data (or seat share data) for carrier i as well as frequency share data for all other carriers in the particular market, the objective of this analysis is to determine a value of α such that the mean residual squared error (defined in Section 3.3) between actual and predicted market share is minimized. A prediction of carrier i's market share will be made by choosing a value of α in Equations 3.1 and 3.2 once the frequency shares and seat shares have been determined for all carriers in the market.

The models presented in Equations 3.1 and 3.2 were developed by Simpson and, as discussed in Chapter Two, explained the variation in market share in O-D markets. Simpson also found that in these O-D markets, seat share was statistically equivalent to frequency share (i.e., the models in Equations 3.1 and 3.2 produced similar results). Since the structure of O-D markets and originating city markets are vastly different, these same models will need to be applied in originating city markets to determine if: (1) this

general model is useful in explaining the variation in market share in originating city markets and (2) the two models are statistically equivalent.

The remainder of this Chapter is divided into four sections. In Section 3.2, a description of the data used in this analysis (and that used in the analysis presented in Chapter Four) will be presented. In Sections 3.3 and 3.4, the results of the model calibrations using Equations 3.1 and 3.2 will be presented. Finally, Section 3.5 will compare the accuracy of the two models and discuss the results presented.

3.2 Description of Data Used in Model Calibrations

To calibrate the models presented in Equations 3.1 and 3.2, three different types of data were required: market share data, frequency share data, and seat share data. For a more meaningful analysis, it was also required that some minimum number of observations be obtained for each variable. Obviously, large numbers of past observations will increase the confidence in the prediction of α , which is important if the results of the calibration are to be used for forecasting future market share. However, observations obtained over too large a period of time could lead to a distorted prediction of α , especially if there are large variations in the data. Since carriers can significantly alter their quality of service in a short period of time, it is

important that the past observations reflect the carriers' current quality of service in that market. This is especially true if these calibrated models are to be used for forecasting market share.

The data used in this analysis was provided by Northwest Airlines. Market share, frequency share, and seat share data was obtained for fifteen different cities in which Northwest operated passenger flights over a twenty-one month sample period. The next three subsections define exactly what the data encompasses as well as the sources of information used by Northwest to collect the observations.

3.2.1 Market Share Data

The market share data compares travel agencies' plated sales for Northwest with the the total plated sales for all carriers operating in the originating city markets, as provided by the Airline Reporting Corporation (ARC).³ The ARC is a clearinghouse used by most airlines to handle the transfer of funds between travel agents and airlines, and between the air carriers themselves. In defining which travel agencies fall into which markets, Northwest groups agencies by zip codes. Agencies who share the first three digits of the five digit zip code are all grouped into one market. In cities such as Boston the market is comprised of over 1000 agencies, while in smaller cities such as Saginaw the market is made up of only 50 agencies.

Although there are various methods for measuring a carrier's market share, this method, based on travel agency sales, was selected for a two reasons. First, it is data readily available from the ARC and represents one of the few methods for measuring a market's total size (using total plated sales). Calculating market share based upon the total number of originating passengers in a market might appear to be a better statistic, however, such a statistic is not universally collected and reported. The second reason for collecting market share in terms of travel agency sales was provided in Chapter One. Since one of the uses of these market share models would be to establish travel agency sales goals, it is only logical to calibrate the model based upon the past performance of travel agencies as measured in terms of ticket sales.

3.2.2 Frequency Share and Seat Share Data

The frequency share and seat share data provided by Northwest were obtained from monthly reports published by the Official Airline Guide (OAG). For each originating market selected, the total number of average daily flights and the number of average daily seats were provided for each carrier in the market. Averages were based on weekdays only (Monday through Friday). Frequency shares and seat shares were easily obtained from:

$$FS_i = \frac{\text{total \# of departures carrier } i}{\text{total \# of departures all carriers}}$$

and,

$$SS_i = \frac{\text{total \# of seats offered by carrier } i}{\text{total \# of seats offered by all carriers}}$$

3.2.3 Selected Test Cities

Fifteen originating city markets were selected as test cases for the calibration of the models in Equations 3.1 and 3.2. These fifteen selected cities represent a cross-section of markets in which Northwest operates domestic passenger flights. Although there are many ways of grouping or classifying the fifteen cities selected, Table 3.1, below, categorizes these cities based upon Northwest's average market share (MS) over the twenty-one months for which data was available.

Category A (MS < 10%)		Category B (10% < MS < 40%)		Category C (MS > 40%)	
City	Avg MS (%)	City	Avg MS (%)	City	Avg MS (%)
Los Angeles	4.23	Huntsville	24.40	Detroit	45.66
Orlando	3.59	Omaha	14.30	Saginaw	48.10
Chicago	5.34	Milwaukee	34.95	Memphis	44.64
Rochester	5.66	Madison	36.04	Minneapolis	36.04
St. Louis	5.68				
Boston	7.53				
Seattle	9.53				

Table 3.1: Selected Test Cities

Notice that category C contains the three connecting hubs that Northwest Airlines currently operates, each of which has more than 200 daily Northwest departures. Smaller cities, such as Saginaw and Rochester, average less than ten departures a day. While the groupings of cities are not of the utmost importance, what should be noticed is the wide variation in Northwest's market share for the selected cities. How the models perform in this diverse cross-section of cities will be the true test of their adequacy.

3.3 Calibration of the MS-FS Model

In Chapter Two it was noted that the relationship between market share and frequency share in most of the origin-destination (O-D) markets analyzed was S-shaped. That is, a carrier with a greater frequency share received a more than proportional share of the market. The theoretical market share model:

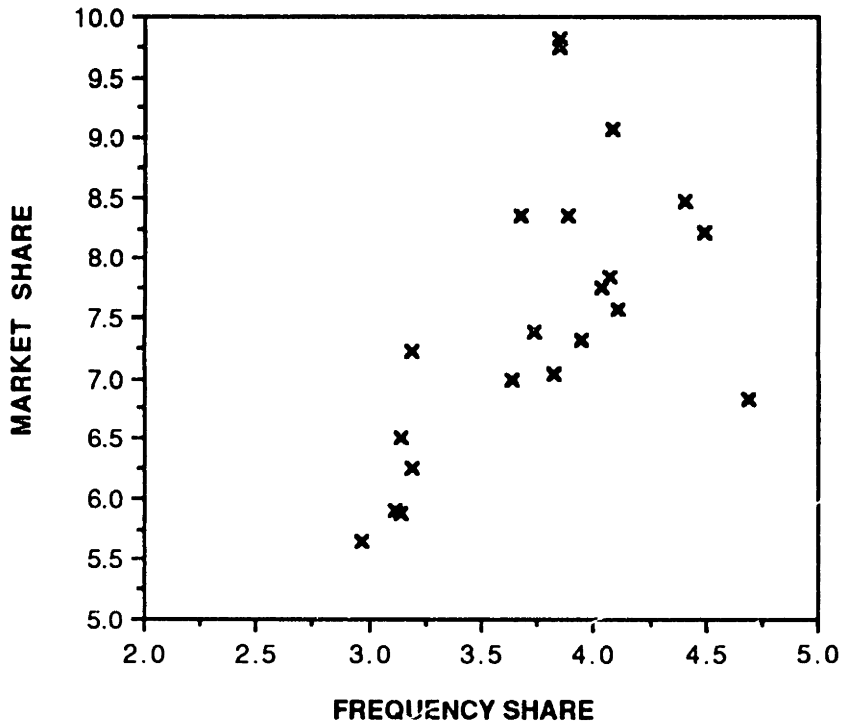
$$MS_i = \frac{FS_i^\alpha}{\sum_{j=1}^n FS_j^\alpha} \quad (3.1)$$

first presented by Simpson, which was used to generate Figure 2.1, accounted for this S-shaped relationship.

It would seem reasonable to check whether this assumption of an S-shaped relationship exists in originating city markets before applying this model to the fifteen selected cities. Clearly, if no such relationship exists then the model given in Equation 3.1 will not be adequate. However, to construct such a figure, the market share and frequency share for all of the n competitors in a given market would have to be known. Unfortunately, market share data like the type obtained from Northwest Airlines (data described in Section 3.2.3) is not made available by the different air carriers.

Some generalizations can be made, however, by examining the relationship between market share and frequency share for carrier i (Northwest). Figures 3.1 through 3.15 show this relationship in each city for the twenty-one months of data available. The figures for the three hub cities; Detroit, Memphis, and Minneapolis, show frequency shares that are 10 to 40% higher than the corresponding market shares. These figures also do not indicate positive linear correlation between the two variables, as one might expect, but such observations can be characteristic of an airline's hub operations. Although a carrier can be expected to dominate operations (frequency) in a hub city, it usually will not be able to match this frequency share with corresponding market share. This is simply due to the fact that many passengers departing out of a hub do not originate in the hub city and are therefore, not part of the measured market share, at least as market share is measured in this analysis.

**FIGURE 3.1
MS VS. FS IN BOSTON**



**FIGURE 3.2
MS VS. FS IN CHICAGO**

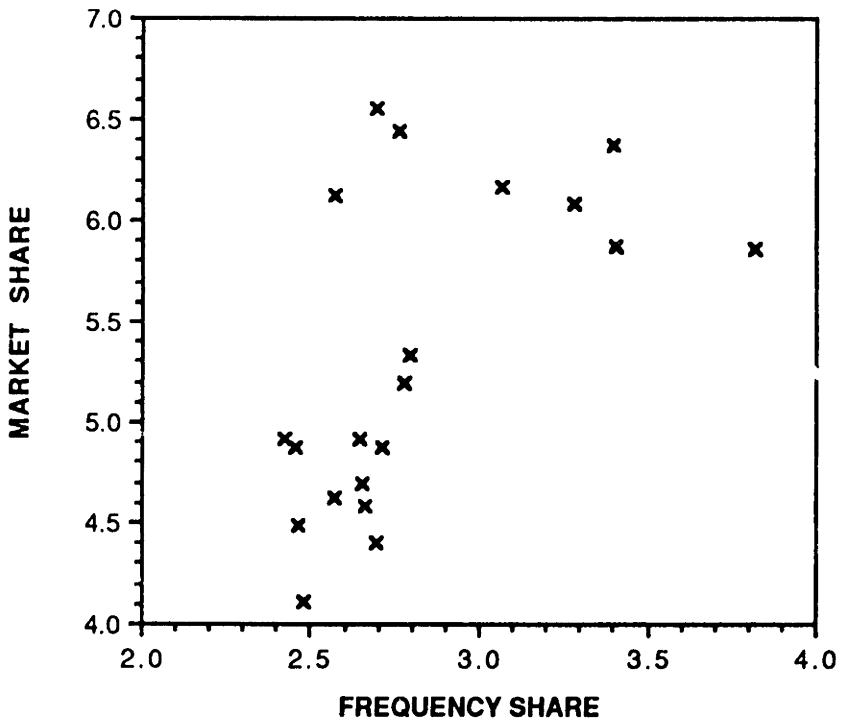


FIGURE 3.3
MS VS. FS IN DETROIT

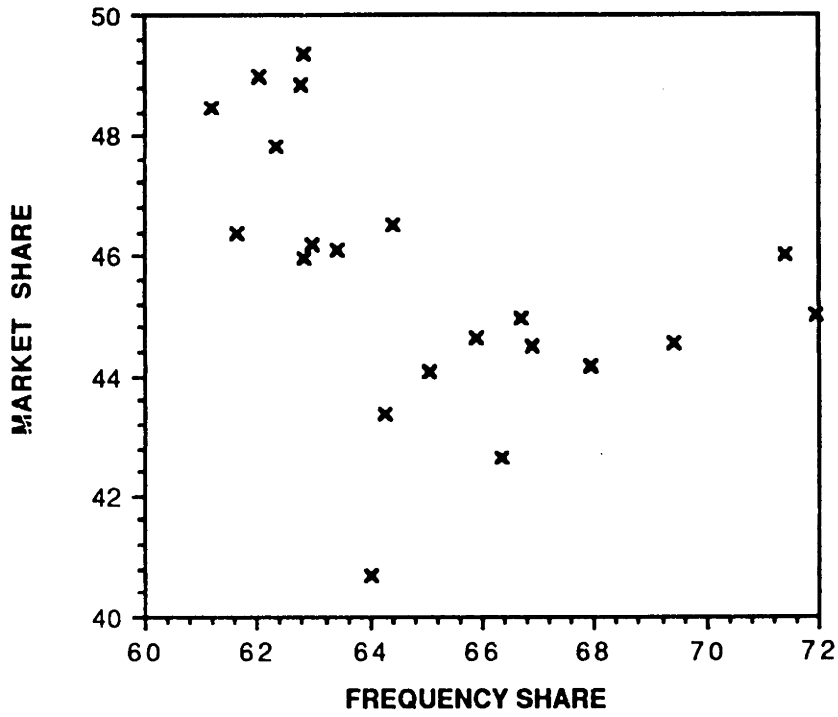
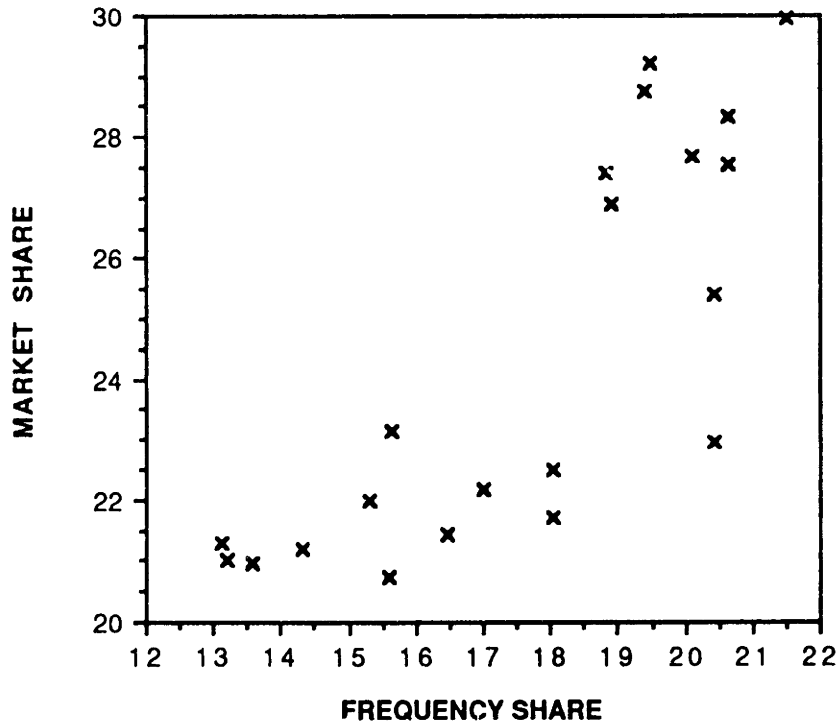
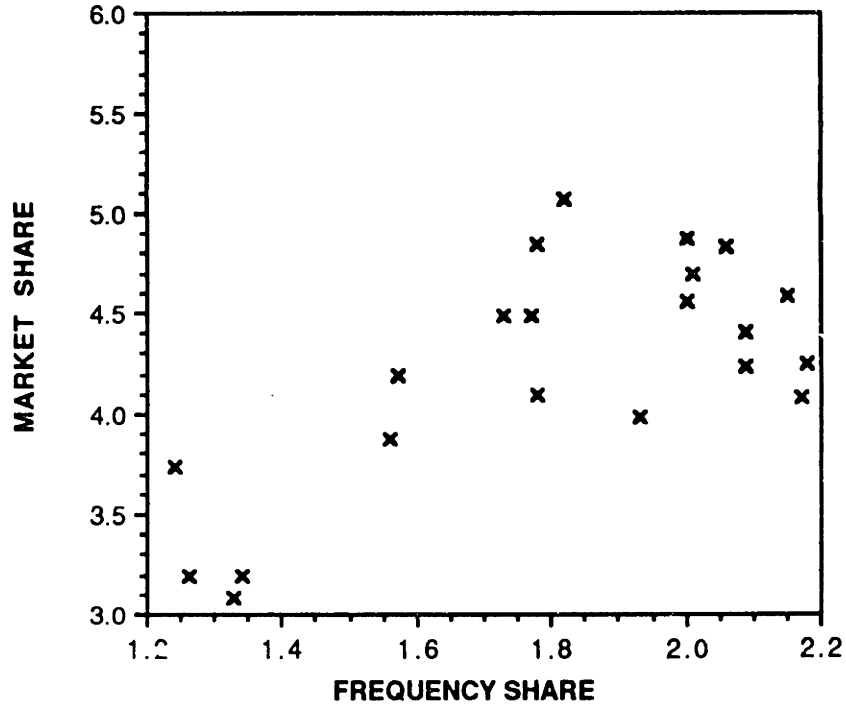


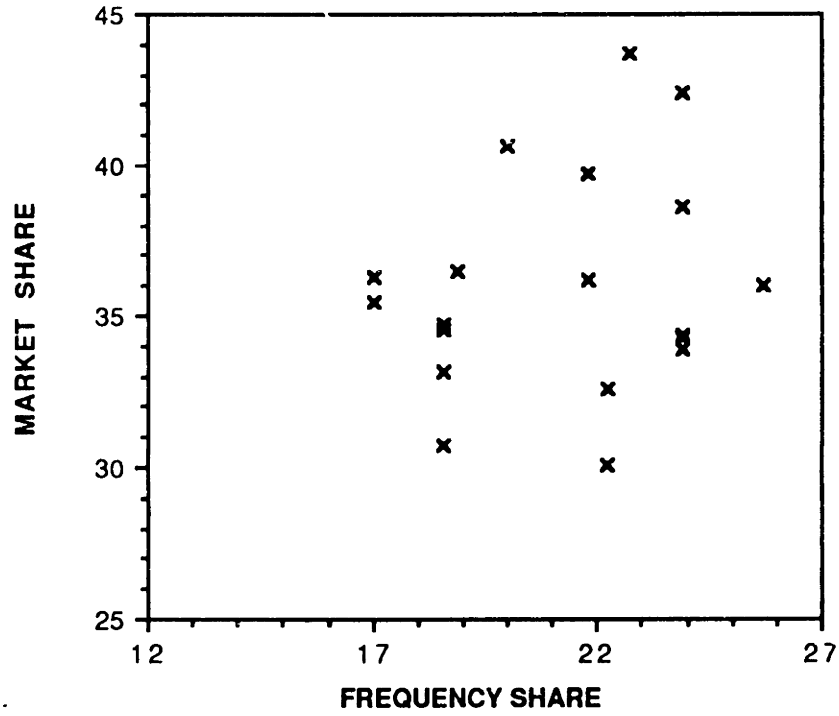
FIGURE 3.4
MS VS. FS IN HUNTSVILLE



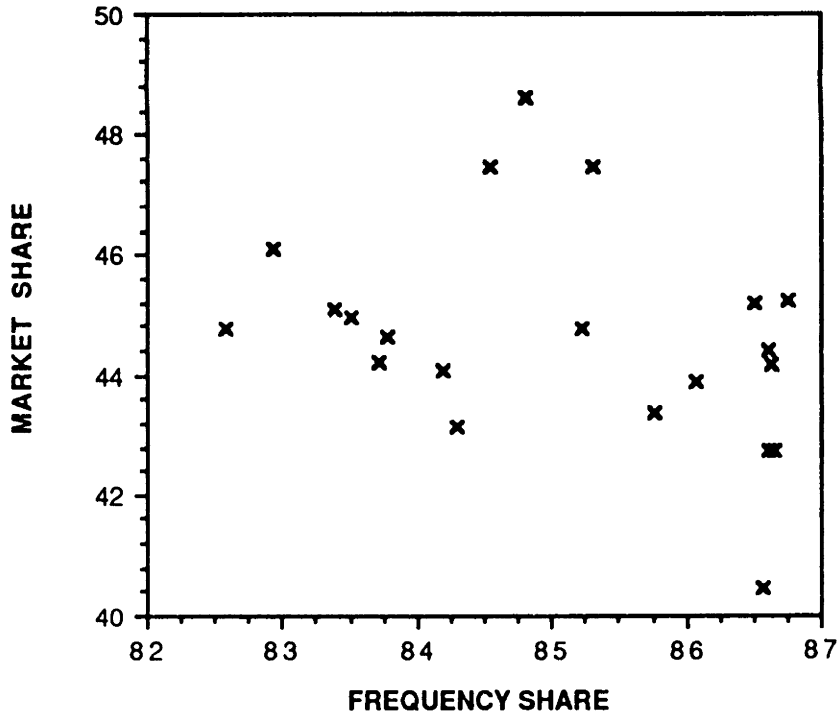
**FIGURE 3.5
MS VS. FS IN LOS ANGELES**



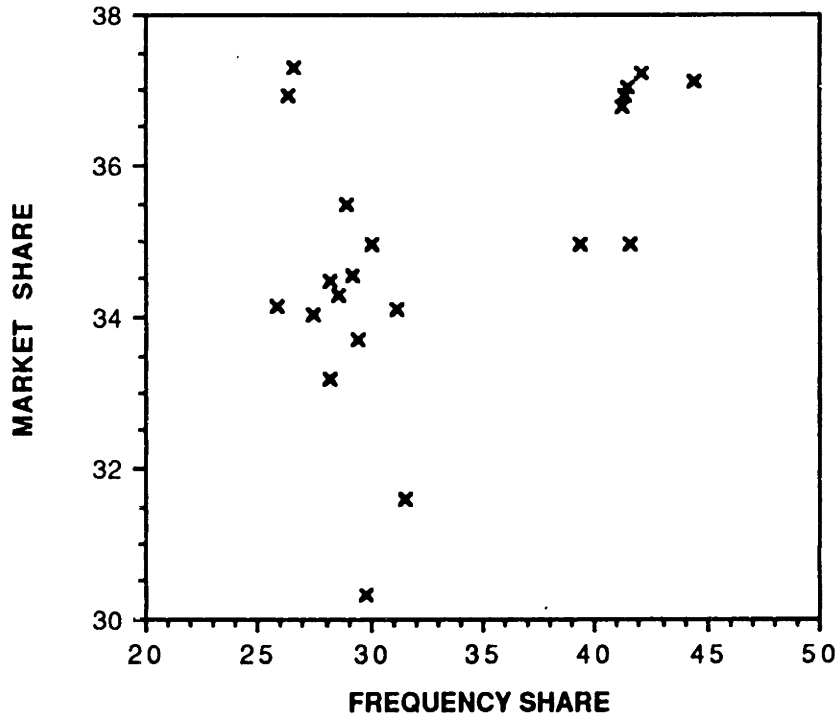
**FIGURE 3.6
MS VS. FS IN MADISON**



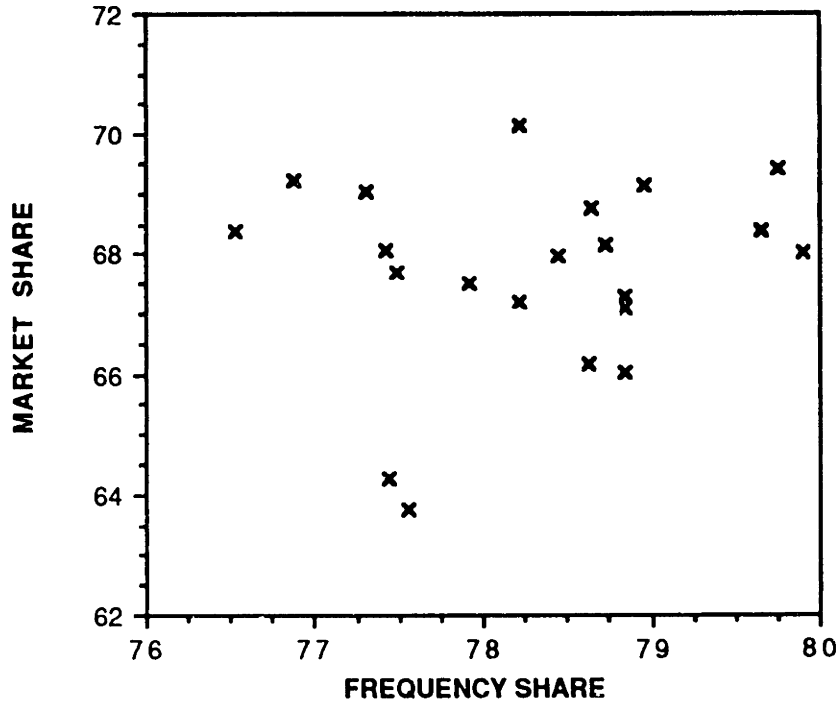
**FIGURE 3.7
MS VS. FS IN MEMPHIS**



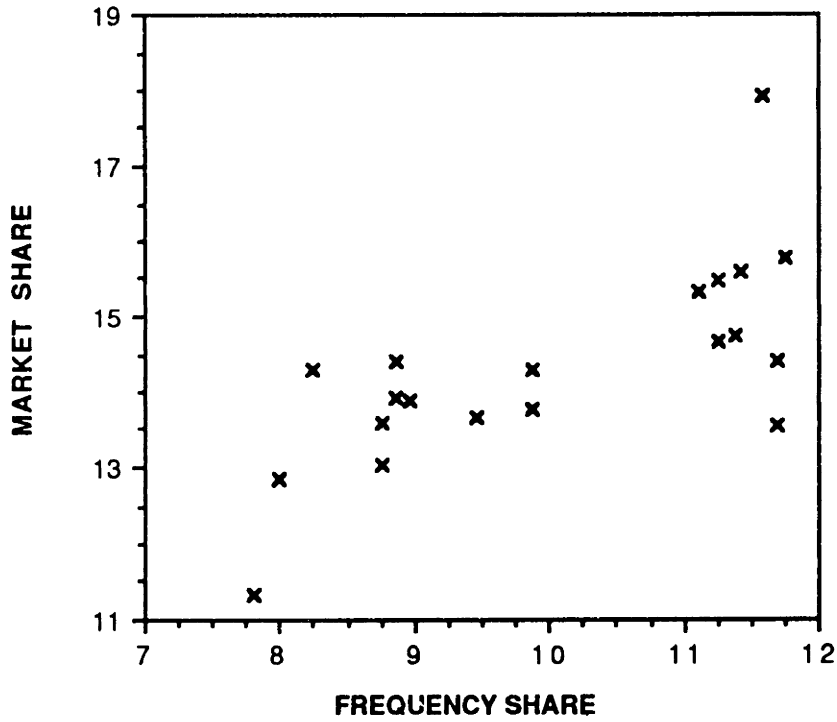
**FIGURE 3.8
MS VS. FS IN MILWAUKEE**



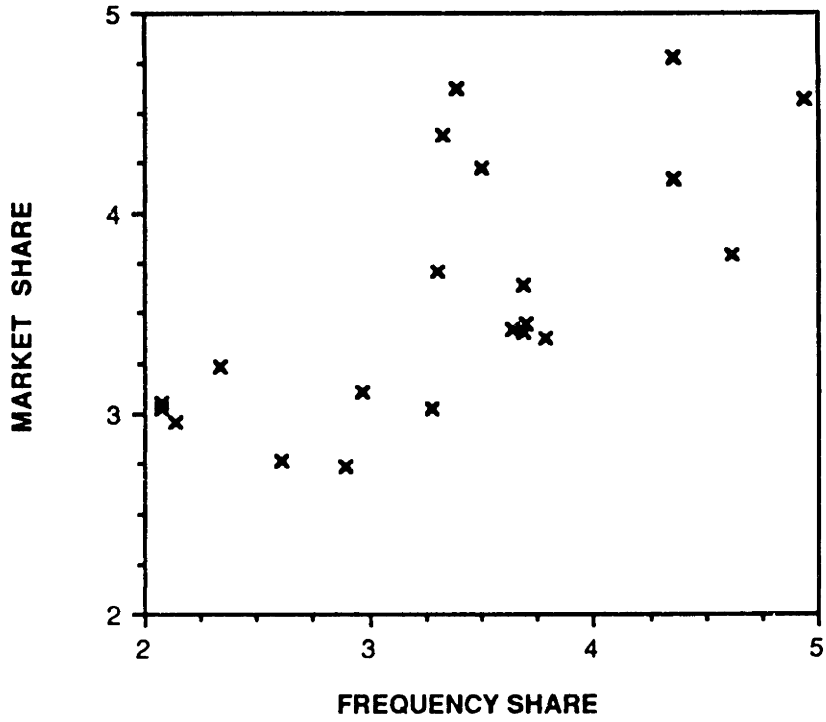
**FIGURE 3.9
MS VS. FS IN MINNEAPOLIS**



**FIGURE 3.10
MS VS. FS IN OMAHA**



**FIGURE 3.11
MS VS. FS IN ORLANDO**



**FIGURE 3.12
MS VS. FS IN ROCHESTER**

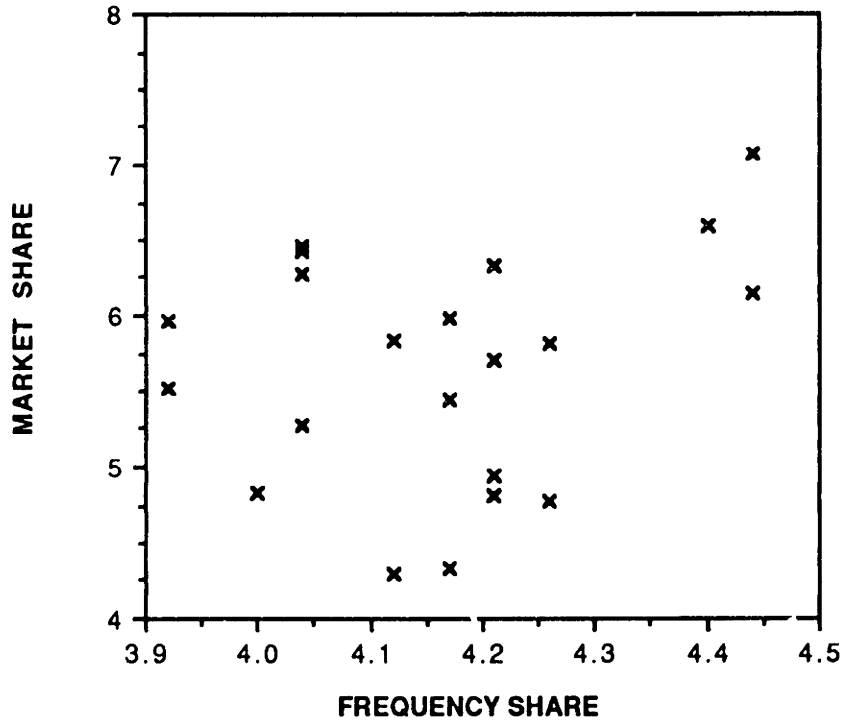


FIGURE 3.13
MS VS. FS IN SAGINAW

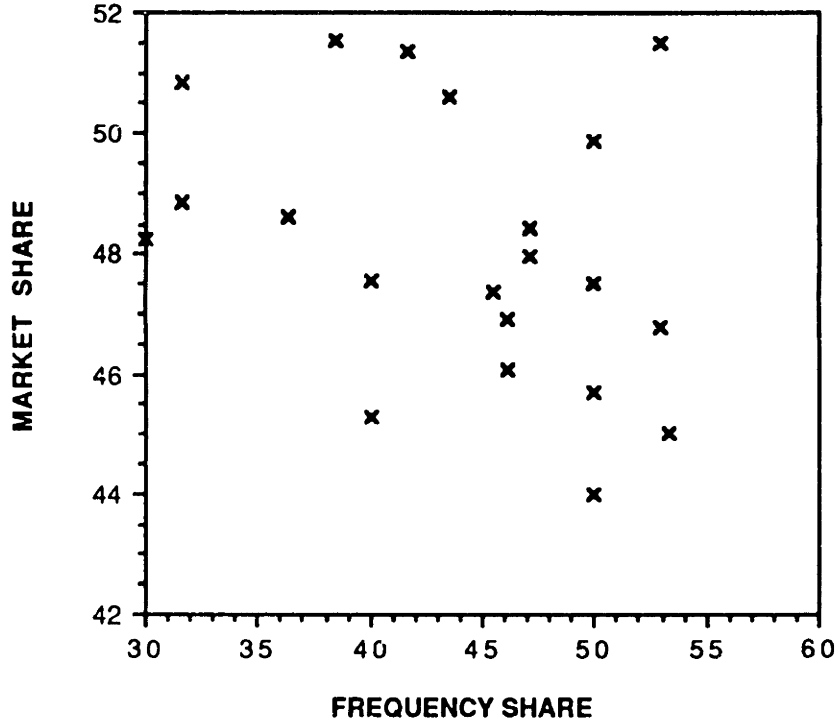


FIGURE 3.14
MS VS. FS IN SEATTLE

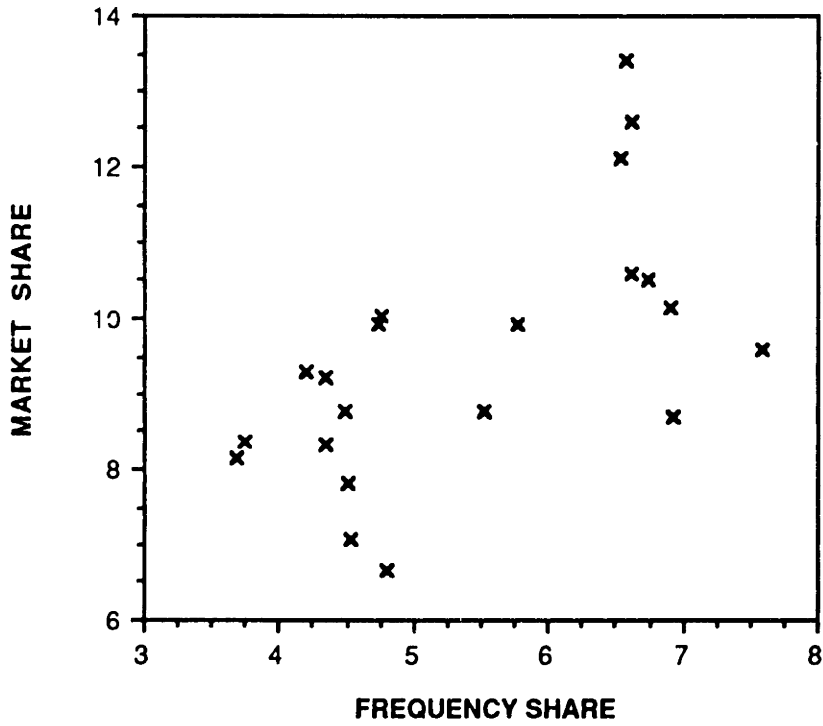
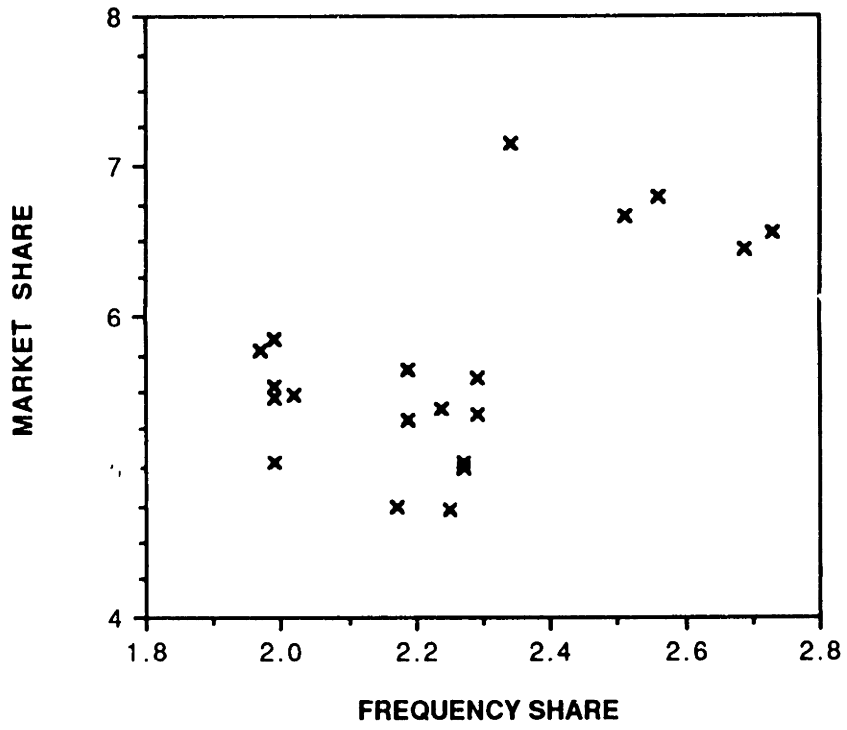


FIGURE 3.15
MS VS. FS IN ST. LOUIS



Different relationships between Northwest's market share (MS) and frequency share (FS) can be observed in the figures for Huntsville, Madison, Milwaukee, and Saginaw. In Huntsville, the market share is nearly double the frequency share, yet Northwest's frequency share is rather low (<20%). The MS-FS relationship shown in Saginaw indicates that higher frequency shares corresponded to lower market shares, which is contrary to what would be expected, and the figures for Madison and Milwaukee show almost no correlation between MS and FS. The model given in Equation 3.1, which assumes a positive correlation between MS and FS (for positive α), may not perform well in these markets.

When estimating the model in Equation 3.1, a unique α can be determined for each of the fifteen selected cities. The coefficient α is selected such that the root mean square error (RMSE) between the actual and predicted market share is minimized over the time for which data exists. The RMSE between the actual and predicted market share is given by

$$RMSE = \frac{\sqrt{\sum_{k=1}^m (MSP_i - MSA_i)_k^2}}{m - 1} \quad (3.3)$$

where MSP_i is the predicted market share for carrier i , MSA_i is the actual market share for carrier i , and m is the number of months over which data is available.

Expanding Equation 3.1 to an n-competitor market,

$$MSP_i = \frac{FS_i^\alpha}{FS_1^\alpha + FS_2^\alpha + \dots + FS_n^\alpha} \quad (3.4)$$

a predicted market share for carrier i, MSP_i , can be calculated by choosing some α given the frequency shares of the n-competitors. The algorithm presented below describes the procedure for selecting α :

1. Calculate the frequency shares for each of the n-competitors in the selected market.
2. Initialize the exponent α ($\alpha > 0$), and calculate MSP_i from Equation 3.3 over each of the $k=1$ to m months.
3. Given the actual market share data for carrier i over each month, MSA_i , calculate the RMSE for the selected α from Equation 3.2.

To determine whether the α selected in Step 2 is the α that minimizes the RMSE, another value of α must be selected and the process repeated. The process (an iterative search process) of selecting an α and calculating a RMSE continues until a minimum RMSE is found. This minimum is found only when perturbing α by $+x\%$, and $-x\%$ does not further decrease the value of the RMSE.

There are several minimization or optimization techniques that are available to carry out the above procedure; however, most of these techniques require the use of a mainframe computer for which the computational costs

can be high. Using a personal computer and a statistical spreadsheet-type software package which allows for large quantities of data to be analyzed, these same techniques can be carried out interactively in a short period of time.⁴ After selecting an initial value of α and then perturbing that value in both the positive and negative directions an optimum value of α can be found very quickly by comparing the RMSEs generated by each α .

Using the algorithm described above, α was determined for thirteen of the fifteen selected cities. Table 3.2 lists the exponent α and the RMSE for each of these cities. Also included in this table are the average predicted market share, the average actual market share, and the standard deviation of the actual market share for Northwest in each of the test markets.

Several conclusions can be drawn from Table 3.2. The first is that this market share model does not perform well in cities where the market share far exceeds the frequency share. Such was the case in Huntsville and Madison where the value of the RMSE was so much larger than the standard deviation of the actual market share, σ_{MSA} , that it was meaningless to report an exponent α . The large values of the RMSE statistic indicate that the model did not explain the variation in market share in these types of markets. In the cities Saginaw, Milwaukee, Omaha, and Seattle, market shares were ~5% higher than the respective frequency shares in each case. Here again, the RMSE statistic was significantly greater than σ_{MSA} , indicating that the model performed poorly in these types of markets.

<u>CITY</u>	<u>PREDICTED EXPONENT</u>	<u>AVG. PREDICTED MARKET SHARE</u>	<u>RMSE</u>	<u>AVG. ACTUAL MARKET SHARE</u>	<u>STD DEV OF ACT. MARKET SHARE</u>
BOSTON	.23	6.34	1.61	7.53	1.19
CHICAGO	.48	5.22	.75	5.34	.73
DETROIT	.73	45.28	4.38	45.66	2.21
HUNTSVILLE	.-	.-	.-	24.40	3.29
LOS ANGELES	.30	4.09	.59	4.23	.57
MADISON	.-	.-	.-	36.04	3.48
MEMPHIS	.43	44.20	3.71	44.64	1.79
MILWAUKEE	1.02	35.57	6.14	34.95	1.89
MINNEAPOLIS	.86	67.81	2.08	67.70	1.59
OMAHA	1.10	10.11	4.44	14.30	1.29
ORLANDO	.95	3.57	.60	3.59	.64
ROCHESTER	.74	5.65	.71	5.66	.77
SAGINAW	1.05	44.98	9.24	48.10	2.24
SEATTLE	.51	6.67	3.28	9.53	1.70
ST. LOUIS	.60	5.56	1.12	5.68	.70

TABLE 3.2: RESULTS OF FREQUENCY SHARE ANALYSIS

High values of the RMSE statistic, compared with σ_{MSA} , in the hub cities (Detroit, Memphis, and Minneapolis) indicate that the model was also inadequate in these types of markets. Only in markets where the frequency share and the market share were relatively close in value to each other, and in markets where these values were relatively small (<10%), did this model perform adequately. Examples of these markets include Rochester, Orlando, Los Angeles, Chicago, St. Louis, and Boston.

The inability of the model to properly explain the variation in market share in many of the originating cities was the result of several factors. Although curves of market share vs. frequency share in each of the originating city markets could not be produced (like Figure 2.1), the data in Table 3.2 clearly indicates that Northwest's MS and FS could not be part of an S-shaped relationship. In the hub cities, where Northwest's frequency share dominates, its market share was proportionately smaller, which is contrary to the S-shaped assumption. The violation of this major assumption accounted for most of the model's poor performance.

A second factor that affected this model's performance was the measurement of frequency share. In cities like Huntsville, Saginaw, and Madison (results were poor in each of these markets), Northwest and its competitors operate a variety of aircraft. For example in Madison, Northwest operated jet aircraft while some of its competitors were operating regional turboprop aircraft with much smaller seating capacities. Although the

carriers may have had equal frequency shares, clearly, the airline operating DC-9's can offer many more seats in the market than the carrier operating SF-340s, for example. Since no consideration was given to the type of aircraft operated, large differences between a carrier's frequency share and market share resulted. These large differences also resulted in the violation of the S-shaped assumption and thus contributed to the model's poor performance in many of these markets.

The results of this analysis indicate that either the general model introduced in Equation 3.1 or the explanatory variable frequency share, or both, are inadequate. In the next section of this chapter, a similar analysis will be conducted using the same model formulation presented in Equation 3.1, except that seat share will be used as the independent variable. The results from this analysis will then be compared to the results presented in this section.

3.4 Calibration of the MS-SS Model

As noted in Chapter Two, Simpson found that the statistic seat share, a carrier's share of the total seats offered in a market, was statistically equivalent to frequency share in the top origin and destination (O-D) markets that he surveyed. The reason for this statistical equivalence was largely due to

the fact that in the top domestic O-D markets, carriers operated very similar equipment (capacity-wise). Also, there were very few, if any, regional carriers operating in the top O-D markets, and as shown in the last section, these can significantly alter the market share - frequency share relationship.

Perhaps, then, seat share would be a better explanatory variable. In small markets such as Huntsville, Madison, and Saginaw, the statistic seat share might account for differences in the equipment operated by the major and regional carriers, and, would be more reflective of a carrier's true market position. In large cities (e.g. Boston, Chicago, and Los Angeles), however, where the major carriers dominate both frequency and seat share, the effect of the regional carriers' operations will not significantly alter a carrier's true market position. Only in the markets where overall frequencies are low do the effects of the regional carriers' operations skew the relative market position of the carriers in the market.

The market share model

$$MS_i = \frac{SS_i^\alpha}{\sum_{j=1}^n SS_j^\alpha} \quad (3.2)$$

is identical to the MS-FS model presented in Equation 3.1, only now the independent variable is each carrier's seat share (SS). Given the similar structure of the two models, the exponent α can be determined in the same manner described in the previous section. The objective here, again, is to

choose an α such that the RMSE,

$$RMSE = \frac{\sqrt{\sum_{k=1}^m (MSP_i - MSA_i)_k^2}}{m-1} \quad (3.4)$$

is minimized.

Before proceeding with the calibration procedure in each of the fifteen selected cities, it is again informative to analyze Northwest's market share vs. seat share over the given time period. Figures 3-16 through 3-30 show this relationship in each of the originating city markets. Similar to the MS-FS curves, the MS-SS curves for the hub cities Detroit, Memphis, and Minneapolis all have seat shares that are 10 to 40% higher than the corresponding market shares. These figures also indicate that very little correlation exists between market share and seat share in these hub cities, which is contrary to the assumption upon which the MS-SS model was constructed.

The MS-SS curves for Saginaw and Huntsville also exhibit some unexpected results. Figure 3.28 (Saginaw) indicates that market share increased as seat share decreased over the twenty-one months surveyed. Obviously, seat share does not explain this variation in market share, a variation that is most likely the result of levels of advertising and/or levels of fares. Figure 3.19 (Huntsville) indicates that Northwest's market share is significantly higher than its seat share (~5%-8%) over the time period

FIGURE 3.16
MS VS. SS IN BOSTON

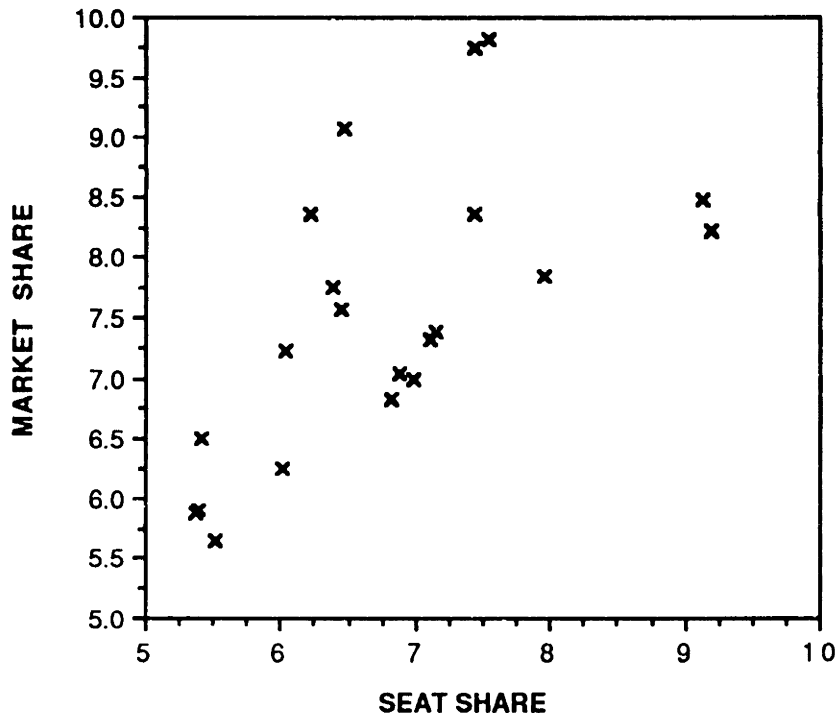


FIGURE 3.17
MS VS. SS IN CHICAGO

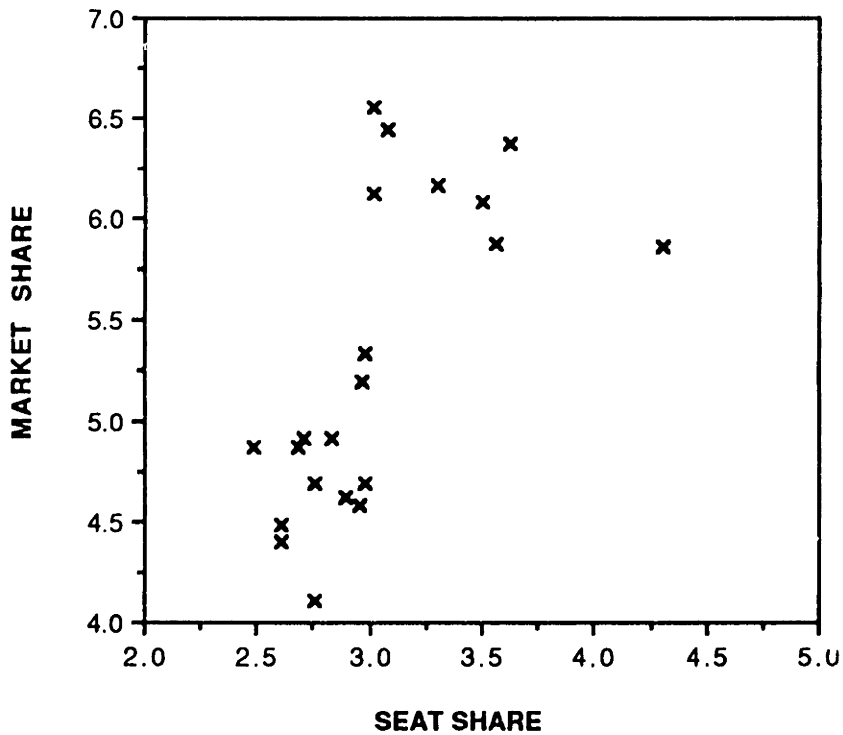


FIGURE 3.18
MS VS. SS IN DETROIT

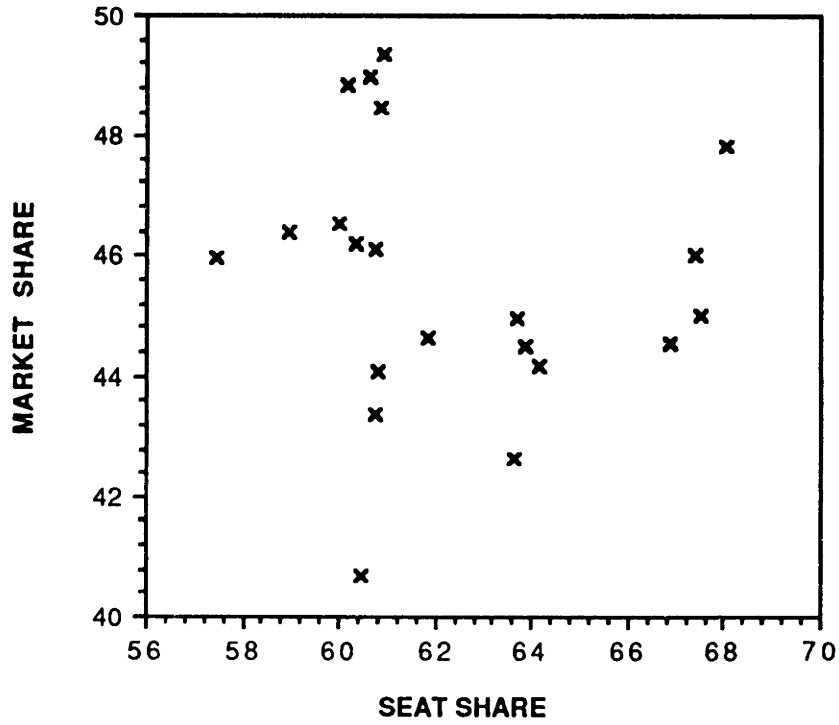
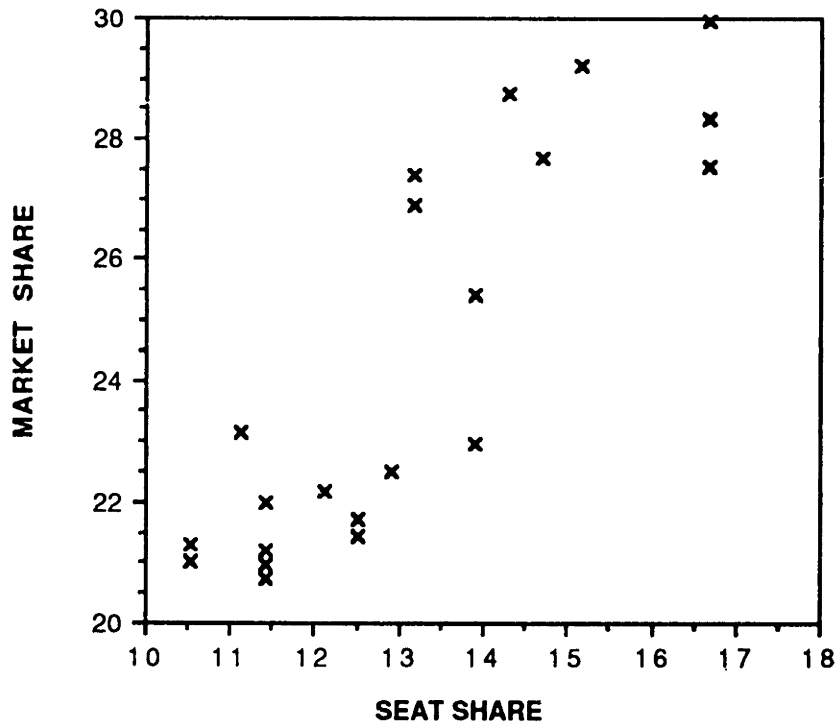
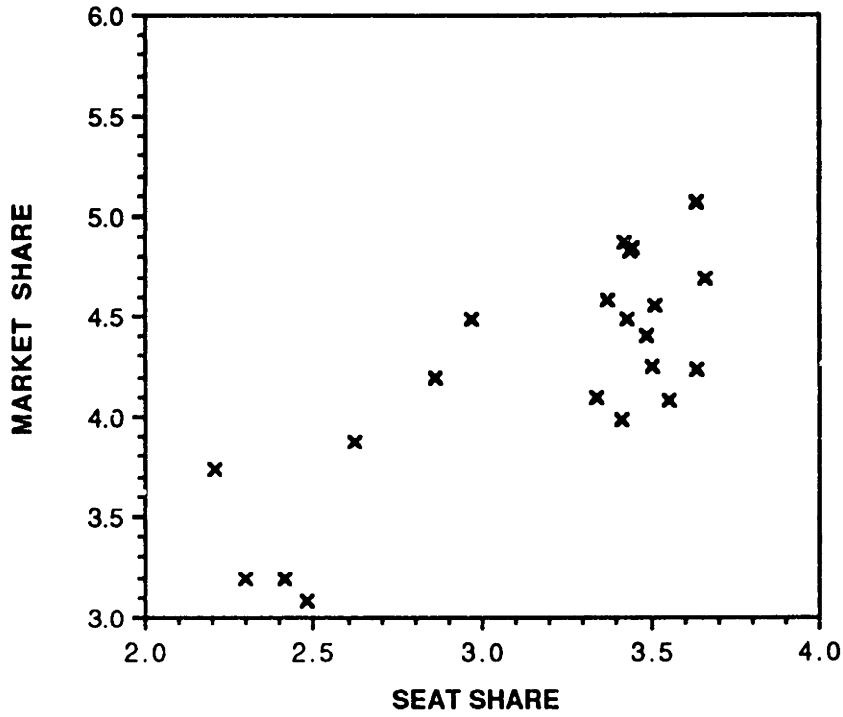


FIGURE 3.19
MS VS. SS IN HUNTSVILLE



**FIGURE 3.20
MS VS. SS IN LOS ANGELES**



**FIGURE 3.21
MS VS. SS IN MADISON**

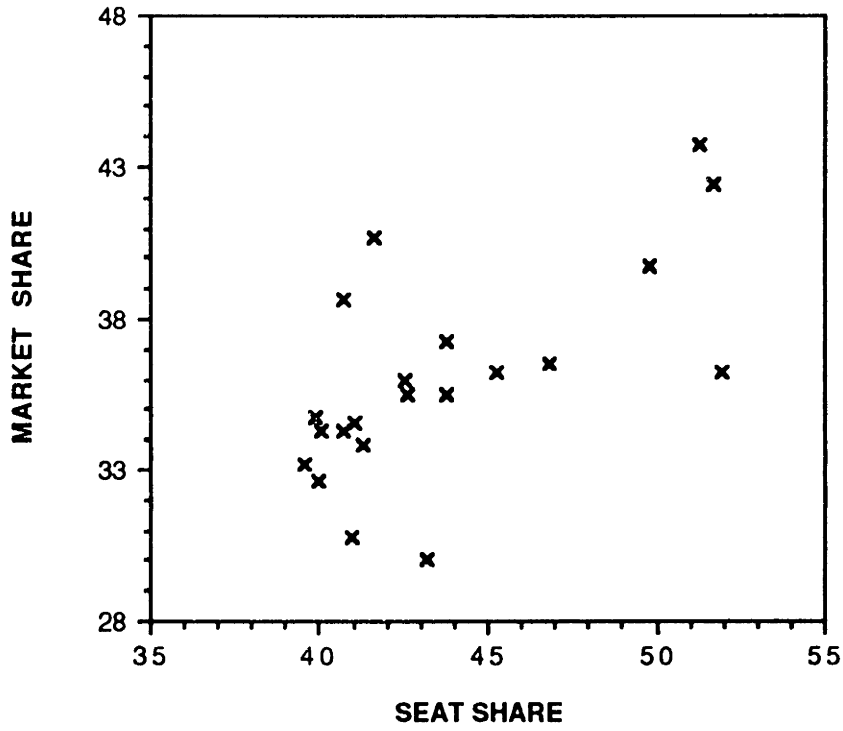


FIGURE 3.22
MS VS. SS IN MEMPHIS

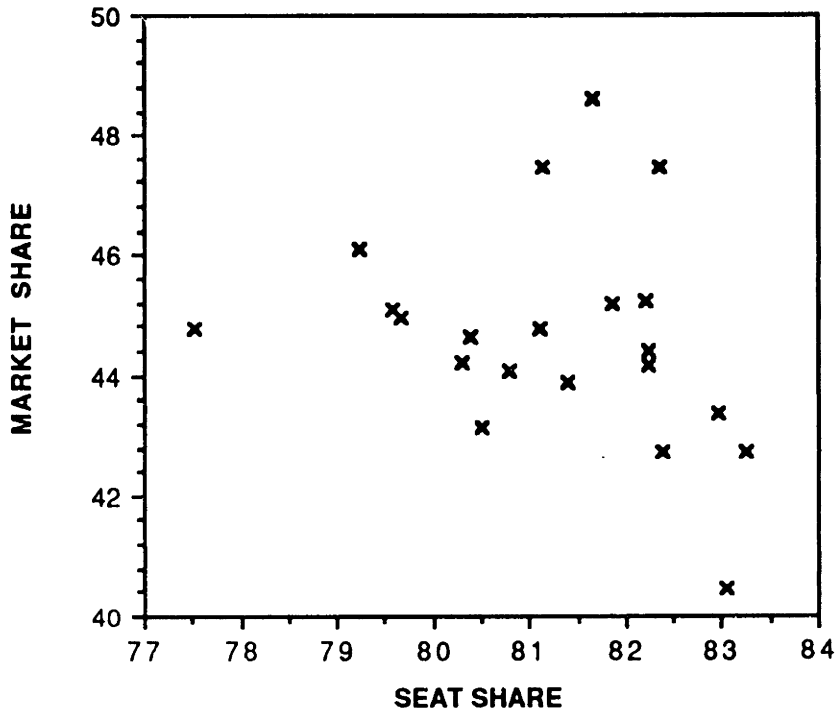


FIGURE 3.23
MS VS. SS IN MILWAUKEE

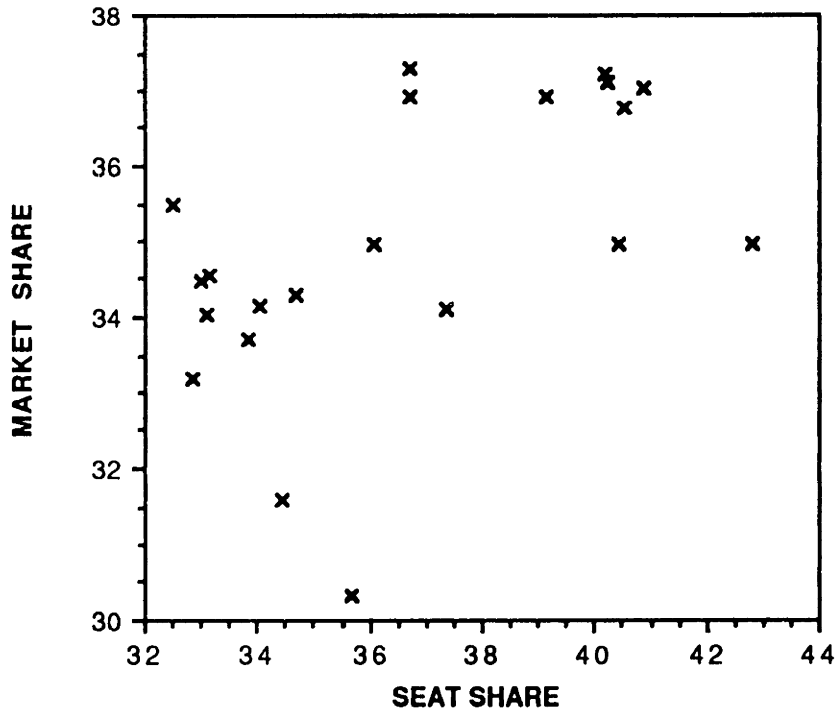


FIGURE 3.24
MS VS. SS IN MINNEAPOLIS

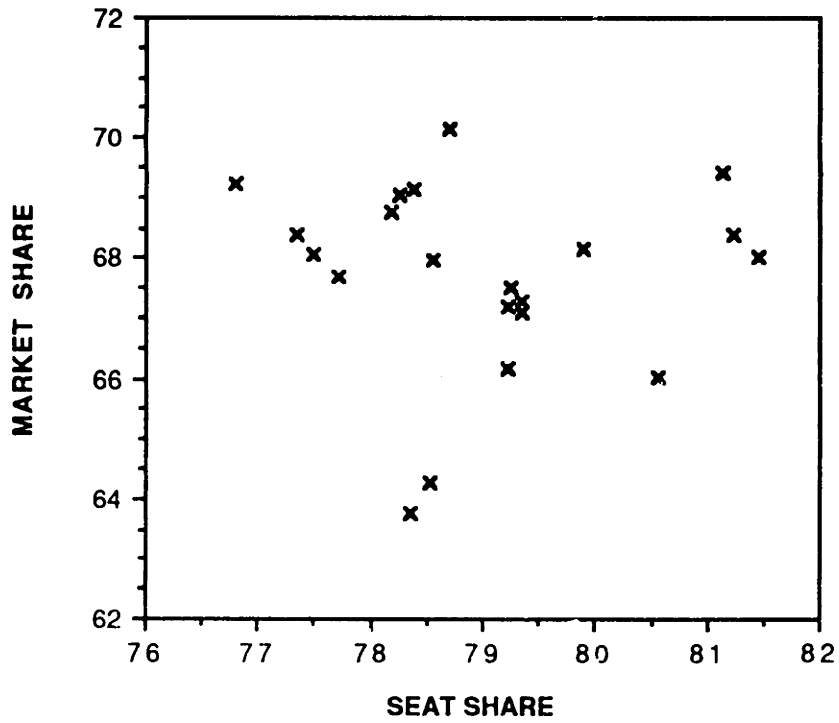
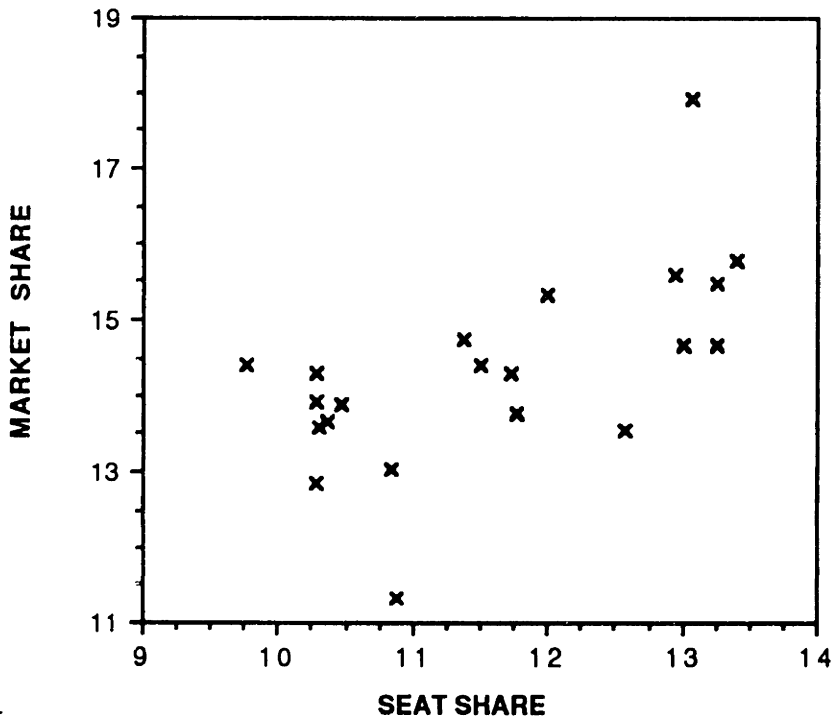
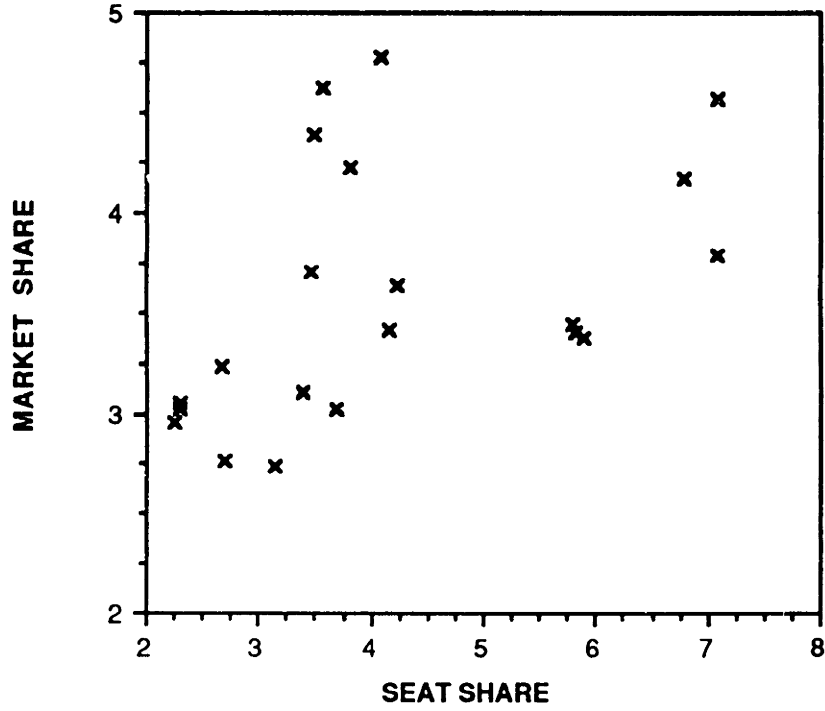


FIGURE 3.25
MS VS. SS IN OMAHA



**FIGURE 3.26
MS VS. SS IN ORLANDO**



**FIGURE 3.27
MS VS. SS IN ROCHESTER**

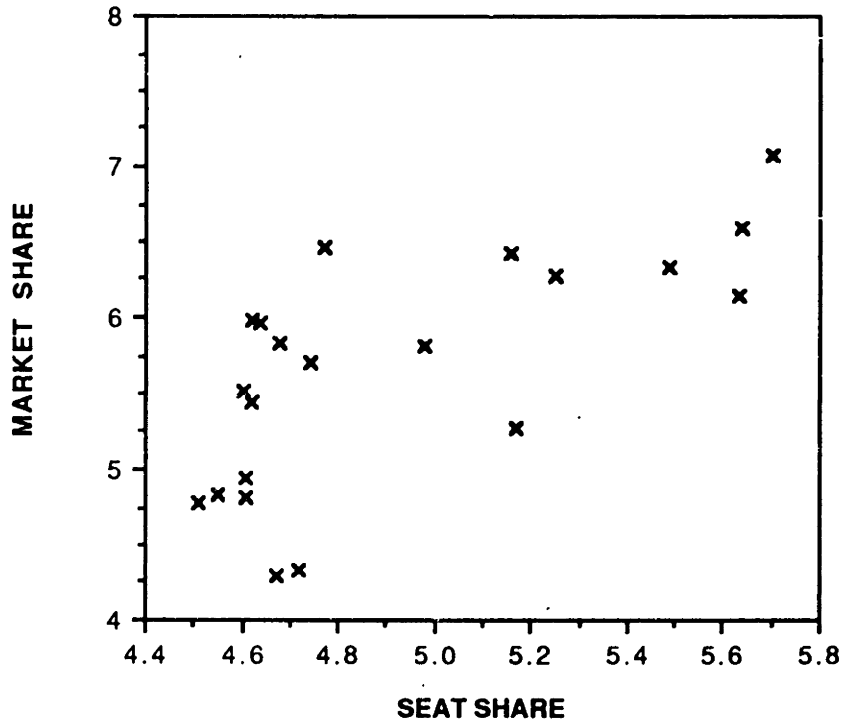


FIGURE 3.28
MS VS. SS IN SAGINAW

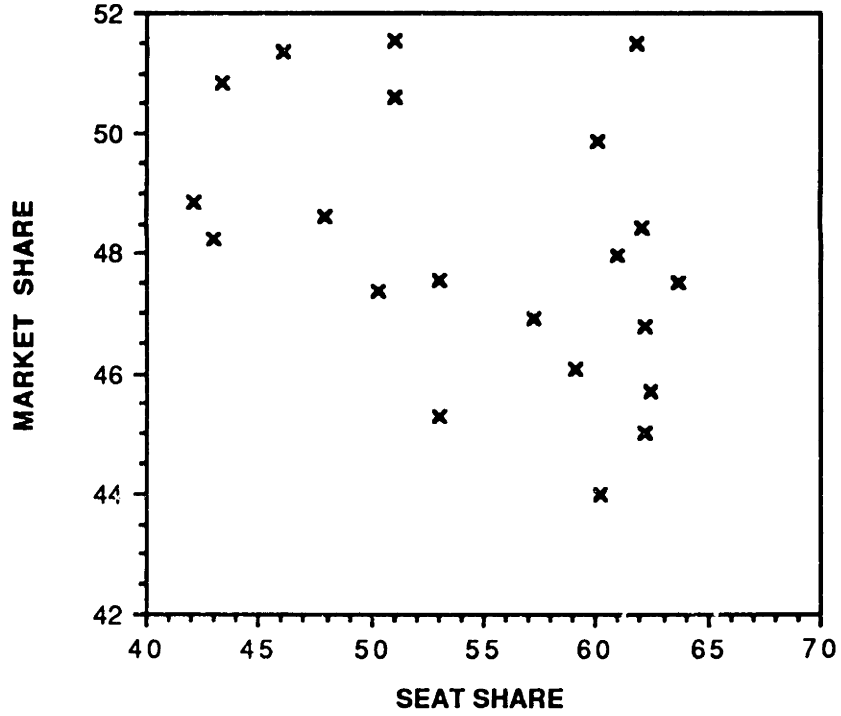


FIGURE 3.29
MS VS. SS IN SEATTLE

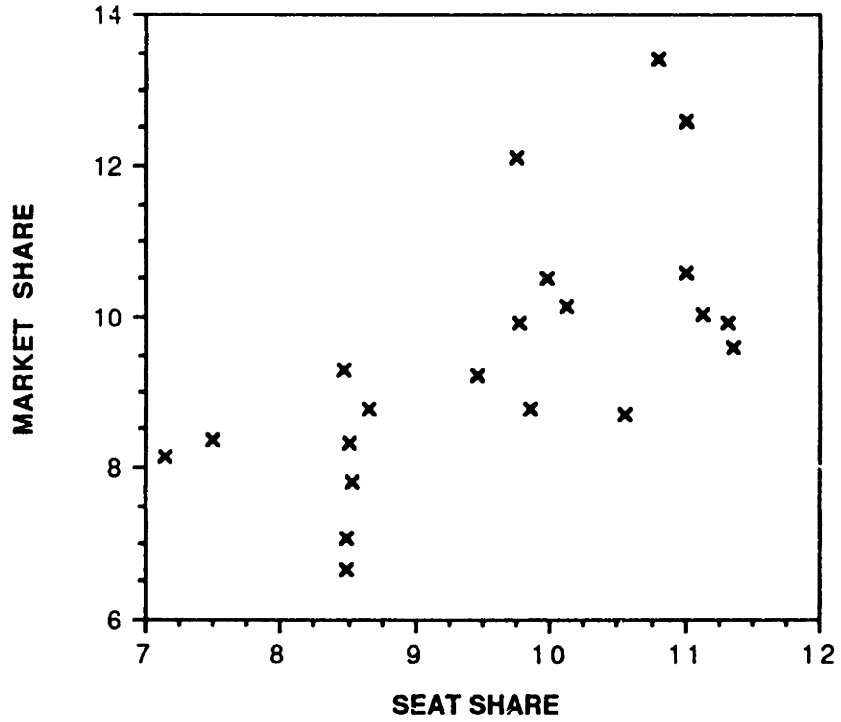
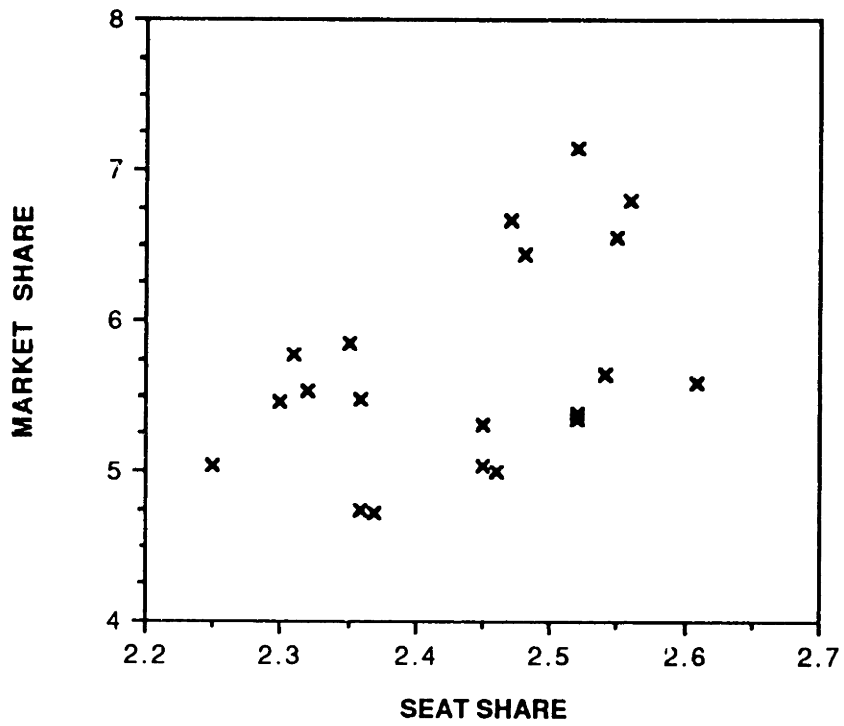


FIGURE 3.30
MS VS. SS IN ST. LOUIS



surveyed. The same trend was also noted in this market's MS-FS relation. Although both of these trends clearly demonstrate Northwest's dominance in Huntsville, they also invalidate the MS-SS assumption. Because of these trends, modeling Northwest's market share in both of these cities will, with the proposed model, present some difficulties.

Using the algorithm outlined in Section 3.3, the exponent α was determined for each of the selected markets. Table 3.3 lists the exponent α and the resultant RMSE for each of these cities. Also included in this table are the average predicted market share, the average actual market share, and the standard deviation of the actual market share, σ_{MSA} , for Northwest in each of the selected cities.

The results shown in Table 3.3 indicate that many of the same conclusions drawn from the MS-FS analysis are also applicable to this analysis. First, comparing the RMSE with σ_{MSA} in each of the hub cities indicates that the statistic seat share does not fully account for the variation in market share in these types of markets (at least as the variable seat share is used in this specific model). As expected, the RMSE of the model calibrations in Huntsville and Saginaw was significantly larger than the σ_{MSA} . As noted previously, this type of model is not sufficient in markets where there are large differences between market share and seat share (i.e. MS>SS). Such was also the case in both the Omaha and Madison markets.

<u>CITY</u>	<u>PREDICTED EXPONENT</u>	<u>AVG. PREDICTED MARKET SHARE</u>	<u>RMSE</u>	<u>AVG. ACTUAL MARKET SHARE</u>	<u>STD DEV OF ACT. MARKET SHARE</u>
BOSTON	.59	7.51	.89	7.53	1.19
CHICAGO	.59	5.29	.64	5.34	.73
DETROIT	.77	45.87	3.54	45.66	2.21
HUNTSVILLE	.56	18.35	6.71	24.40	3.29
LOS ANGELES	.62	4.18	.34	4.23	.57
MADISON	.44	36.08	2.68	36.04	3.48
MEMPHIS	.46	44.26	3.28	44.64	1.79
MILWAUKEE	.95	34.89	2.62	34.95	1.89
MINNEAPOLIS	.84	68.10	1.74	67.70	1.59
OMAHA	.94	11.59	3.00	14.30	1.29
ORLANDO	1.15	3.63	1.47	3.59	.64
ROCHESTER	.88	5.66	.56	5.66	.77
SAGINAW	1.05	44.98	9.24	48.10	2.24
SEATTLE	1.16	9.65	1.36	9.53	1.70
ST. LOUIS	.64	5.58	.74	5.68	.70

TABLE 3.3: RESULTS OF SEAT SHARE ANALYSIS

The MS-SS model did perform adequately, however, in five of the fifteen cities examined. The RMSE that resulted from the choice of α was less than σ_{MSA} in Boston, Los Angeles, Chicago, Rochester, and St. Louis. Similar to the MS-FS model, this model seems to perform satisfactorily in markets where Northwest's market share and frequency share are small (<10%) and also, where the difference between the two variables is small.

3.5 A Comparison of the MS-FS and the MS-SS Models

The results presented in Tables 3.2 and 3.3 clearly indicate that the general market share model proposed in Equation 3.1 (with either frequency or seat share used as the explanatory variable) is not useful in certain types of originating city markets. The inability of this model to correctly explain the variation in a carrier's market share is largely due to the lack of an S-shaped relationship between market share and frequency share, and between market share and seat share. Since this assumption of an S-shaped relationship was the cornerstone upon which this model was built, it is apparent why the model behaves poorly in many of the markets analyzed.

The results also suggest that there are vast differences between originating city markets and origin-destination (O-D) markets. As Figures 3.1 through 3.30 indicate, Northwest's market share in many of the originating cities did not always show positive correlation with frequency share or seat share. These occurrences suggest that other variables, which characterize an originating city market, need to be identified and modeled.

Although neither model (Equations 3.1 and 3.2) was particularly useful in explaining the variation in a carrier's market share, a comparison of the results of the two model calibrations does reveal some interesting conclusions. By comparing the RMSEs in each market in Tables 3.2 and 3.3, it is clear that seat share is a more useful variable in explaining market share (except in Orlando). This result suggests that the statistical equivalence between seat share and frequency share found in O-D markets is not valid in originating city markets.

With the wide variety of equipment operated by airlines today, this result is of no great surprise. Consider the simple case, for example, where carriers A and B are the only two carriers operating in a market. Carrier A operates six DC-9 flights a day and carrier B operates six B747 flights a day. Although their frequency shares are equal, carrier B's seat share is 75% while carrier A's is only 25%. Clearly, carrier B has a competitive advantage, one that should be reflected in its market share. This does not suggest, however, that frequency share is not important. When carriers are operating the same or similar types of equipment (capacity wise), frequency share is of great

importance. But in markets where carriers operate a diverse fleet of aircraft, seat share is a more useful variable in explaining market share. This conclusion will be of use in the next chapter.

¹ In his Master's Thesis, "Competition in Air Transportation, An Economic Approach", Gilles Renard refers to models that can not be solved using linear least squares methods as "interactive models." The non-linear structure of the models presented in Chapter Two, Equations 2.5-2.9 are clearly of this type. See Renard, p.34.

² Robert W. Simpson, Unpublished Course Notes for Air Transportation Economics, Chapter Three, "The Demand for Air Transportation Services", Massachusetts Institute of Technology, Cambridge, Ma.: September, 1982 p.1.

³ A travel agency plated sale simply refers to the sale of an airline ticket by a travel agent to a consumer. In this analysis, Northwest's market share in an originating city market is the ratio of travel agency plated sales on Northwest to travel agency plated sales on all carriers.

⁴ The statistical software package "SOLO" was used in this analysis. SOLO is distributed by BMDP Statistical Software; Los Angeles, California.

CHAPTER 4

Application Of Market Share Regression Models In Originating City Markets

4.1 Introduction

The literature review presented in Chapter Two revealed two broad categories of market share models that were applied to origin-destination (O-D) markets. In Chapter Three, one such category of these models, classical frequency share - seat share models, was applied to originating city markets. Although Simpson and Renard both found that these types of models were useful in explaining the variation in market share in O-D markets, they did not prove to be extremely useful in originating city markets. The results of the model calibrations in Chapter Three suggest that different types of models are necessary to better explain the variation in market share in originating city markets.

The classical frequency share and seat share models calibrated in Chapter Three predicted Northwest's market share as a function of both the airlines's frequency share (or seat share) and its competitors' frequency share (or seat share) in the particular market. The marginal results obtained from these

calibrations suggest, however, that a carrier's market share in an originating city market cannot be fully explained in terms of these statistics alone. Additional independent variables that reflect the number of markets served out of these originating cities and the manner in which they are served (e.g. non-stop, or one-stop) are needed to better explain a carrier's market share. These new variables or a completely different type of model, or both, are necessary to improve upon the results obtained in Chapter Three.

A second type of market share model discussed in Chapter Two is one that can be calibrated using linear regression techniques. Bond's [2] proposed market share model,

$$MS = \alpha \times FS^{\beta} \quad (4.1)$$

where α and β are carrier dependent, is one such model that can be easily calibrated using linear regression methods after the model has undergone a logarithmic transformation. Regression models of this type are attractive for two reasons. First, unlike the models presented in Chapter Three, this type of regression model does not require the input of statistics for every carrier operating in the market. To calibrate Equation 4.1, only market share data and frequency share data pertaining to the particular carrier are needed. Computationally, this model is much simpler than the models presented in Chapter Two. Secondly, regression models can be easily calibrated with very basic statistical software programs. Time consuming iterative techniques like those used in the previous chapter are not necessary to calibrate regression models.

In this chapter, four regression market share models will be calibrated for the selected originating city markets. In the second section of this chapter, using a variation of Bond's market share model,

$$MS = \alpha \times SS^{\beta} \quad (4.2)$$

(where the statistic seat share is used instead of frequency share, based upon the results of the last chapter), the constant α and the exponent β will be determined in each of the originating city markets given Northwest's market share and seat share over the twenty-one month sample period.

In the third section of this chapter, the variable frequency share will be added to the above formulation as a second explanatory variable. In the model

$$MS = \alpha \times SS^{\beta_1} \times FS^{\beta_2} \quad (4.3)$$

the constant α and the exponents β_1 and β_2 , the elasticity of market share with respect to seat share and frequency share, will be determined for Northwest's service in each of the originating city markets. The results of a linear regression analysis will be used to determine whether the variables seat share, frequency share or both are needed to explain the variation in Northwest's market share.

Two new variables, a carrier's proportion of total non-stop markets, NSS, and a carrier's proportion of total possible markets, PMS, served out of an

originating city market will be included in Equation 4.3 in the fourth section of this chapter. Using this new model,

$$MS = \alpha \times SS^{\beta_1} \times FS^{\beta_2} \times NSS^{\beta_3} \times PMS^{\beta_4} \quad (4.4)$$

where,

NES = proportion of total non-stop markets served by all carriers out of an originating city in which the chosen carrier offers non-stop service;

PMS = proportion of the total possible markets served by all carriers out of an originating city in which the chosen carrier offers non-stop, direct, or feasible connecting service.

β_3 = elasticity of market share with respect to NSS; and

β_4 = elasticity of market share with respect to PMS,

a step-wise regression analysis will be conducted to determine which of the above variables are most useful in explaining the variation in a carrier's market share.

The final model presented in this chapter uses the same variables as the model in Equation 4.4; however, the structure of the model is somewhat different. Instead of the non-linear form of the three previous models, the structure of this model,

$$MS = A + B_1SS + B_2FS + B_3NSS + B_4PMS \quad (4.5)$$

is linear with respect to market share, and no transformations are required to perform a regression analysis. The results of this step-wise regression analysis will be presented in the fifth section of this chapter.

Finally, in the last section of this chapter, a comparison of the results of these four model calibrations will be made. By examining the root mean square errors (RMSE) of each of the four models, a determination will be made as to whether these regression type models are more useful than the models calibrated in Chapter Three. Conclusions will also be made as to which type of regression model as well as which of the four independent variables examined are most useful in explaining a carrier's market share in an originating city market.

4.2 Calibration of Seat Share Regression Model

Although neither of the models tested in Chapter Three fully explained the variation in Northwest's market share, it was concluded that of the two explanatory variables, frequency share and seat share, the latter was of greater use. Based on this conclusion, the first regression market share model to be calibrated in this chapter is based solely on the variable seat share. By restricting this first model to one explanatory variable, a comparison can be made between the classical seat share model (Equation 3.2) and the regression seat

share model (Equation 4.2). This comparison will give rise to an assessment of the two different model structures.

It was noted in Chapter Three that the relationship between market share and seat share was not always linear in the fifteen originating city markets that were examined. Figures 3.16 - 3.30 clearly indicated this non-linear relationship. The seat share regression model,

$$MS = \alpha \times SS^\beta \quad (4.2)$$

should be useful in modeling this relationship because of its non-linear structure (unless β is equal to 1.0). In attempting to fit this model to the data shown in Figures 3.16 - 3.30, the parameter β can be varied to account for the different non-linear relationships in each of the originating city markets.

As stated in the Introduction, all of the market share models presented in this chapter can be calibrated through linear regression. Thus, before calibrating the seat share regression model, Equation 4.2 must be linearized. Taking the natural log of both sides of this equation yields,

$$\ln(MS) = \ln(\alpha \times SS^\beta) \quad (4.6)$$

which simplifies to

$$\ln(MS) = \ln(\alpha) + \beta \times \ln(SS) \quad (4.7)$$

After taking the natural log of the market share data and the seat share data in each of the fifteen selected cities, the parameters α and β can be determined through a linear regression analysis.

The linearized model in Equation 4.7 assumes a linear relationship between $\ln(\text{MS})$ and $\ln(\text{SS})$. Of course, if this assumption is false, then the results of the model calibration will be poor. Figures 4.1-4.15 show the relationship between $\ln(\text{MS})$ and $\ln(\text{SS})$ in each of the fifteen selected cities. Examination of these scatter plots indicates that in Detroit (Figure 4.3), Memphis (Figure 4.7), and Minneapolis (Figure 4.8), linear relationships between $\ln(\text{MS})$ and $\ln(\text{SS})$ cannot be detected. Based on these figures, the accuracy of the model calibrations in these cities will most likely be suspect. Also of interest is Figure 4.13, which shows $\ln(\text{MS})$ vs. $\ln(\text{SS})$ in Saginaw. Although there appears to be a linear relationship between these variables, it is, as noted in Chapter Three, contrary to what is expected in the airline industry.

Applying the method of linear least squares to Equation 4.7, the intercept and the slope, α and β respectively, are determined by minimizing

$$S(\alpha, \beta) = \sum_{i=1}^m [\ln(\text{MS}_i) - \ln(\alpha) - \beta \times \ln(\text{SS}_i)]^2 \quad (4.8)$$

where $i=1, m$ is the number of months over which data is available. The minimization of Equation 4.8 is relatively straightforward. First, the partial derivatives of S with respect to α and β are calculated and set equal to zero, i.e.,

FIGURE 4.1
ln(SS) VS. ln(MS) IN BOSTON

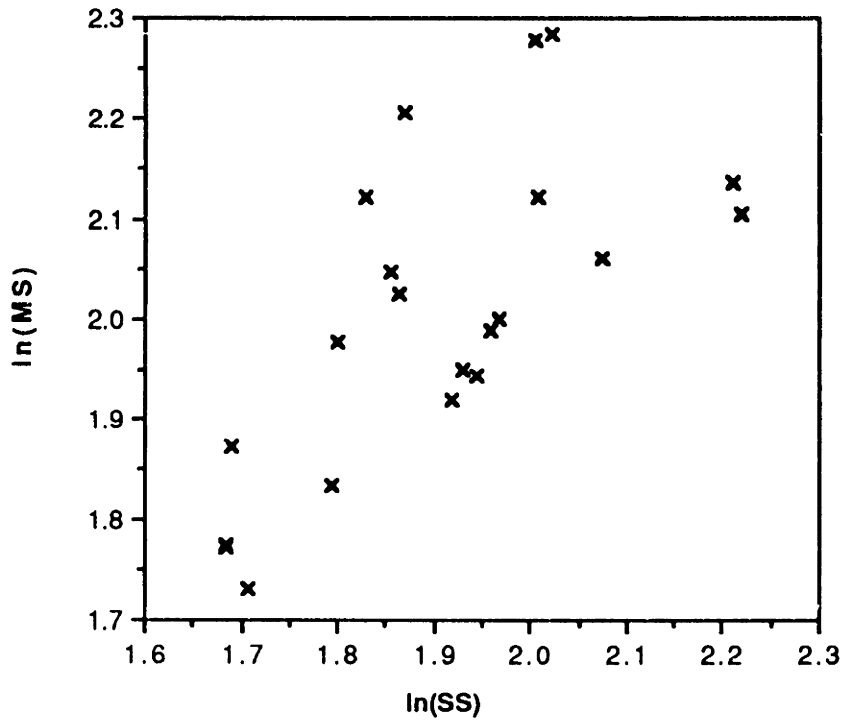


FIGURE 4.2
ln(SS) VS. ln(MS) IN CHICAGO

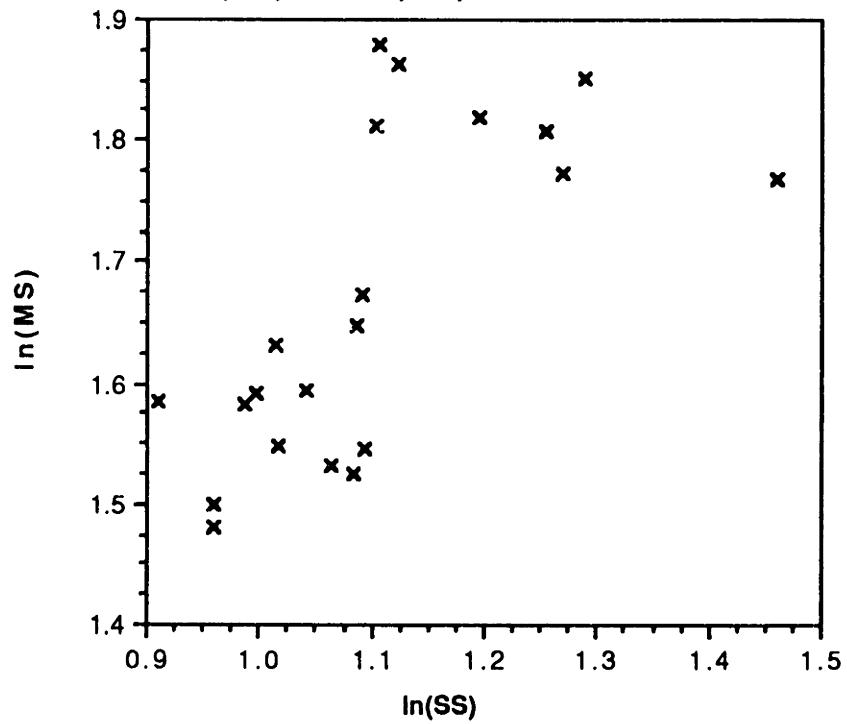


FIGURE 4.3
ln(SS) VS. ln(MS) IN DETROIT

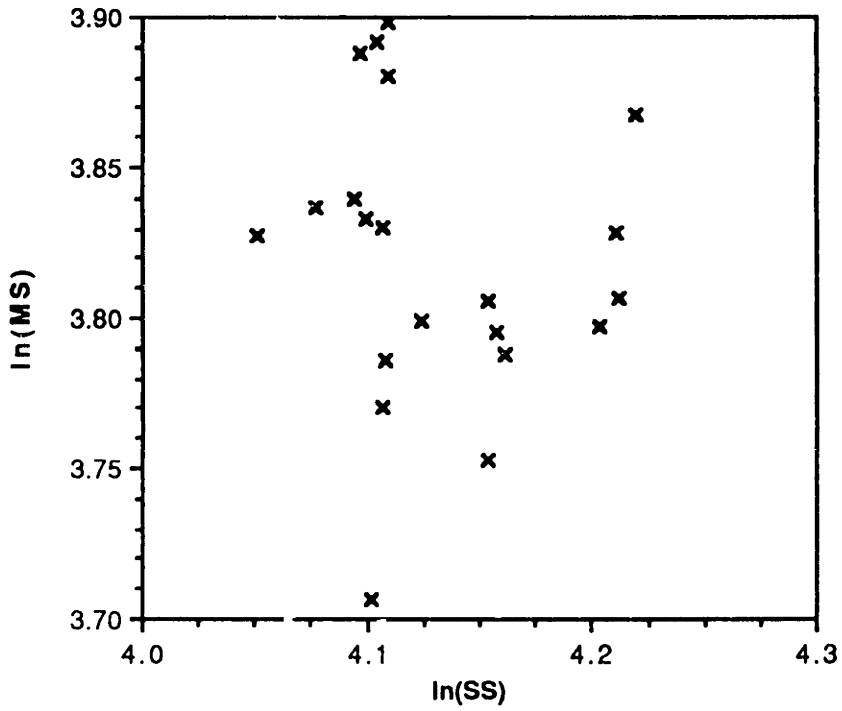


FIGURE 4.4
ln(SS) VS ln(MS) IN HUNTSVILLE

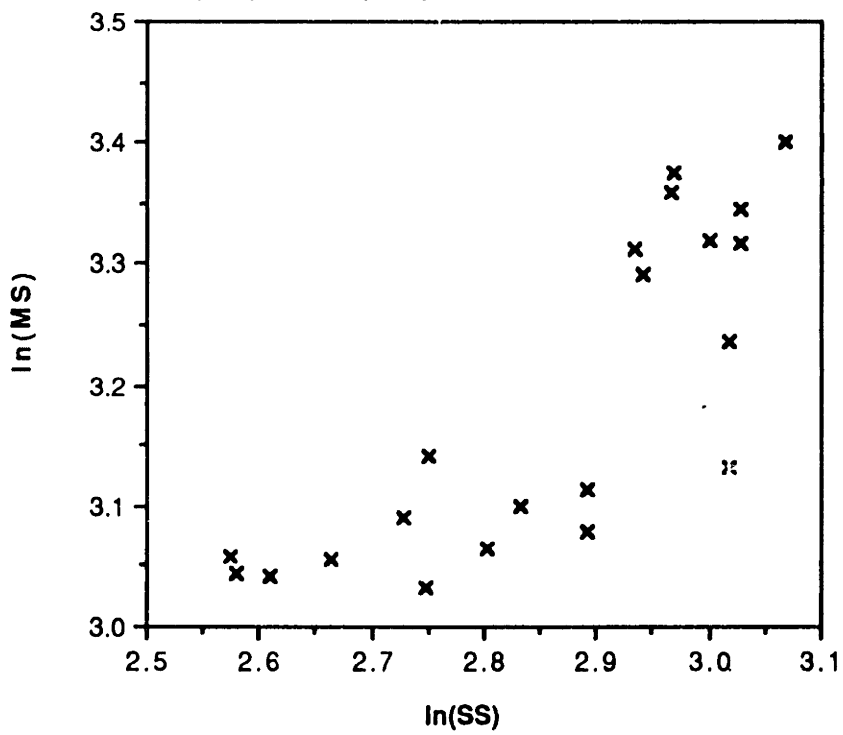


FIGURE 4.5
ln(SS) VS. ln(MS) IN LOS ANGELES

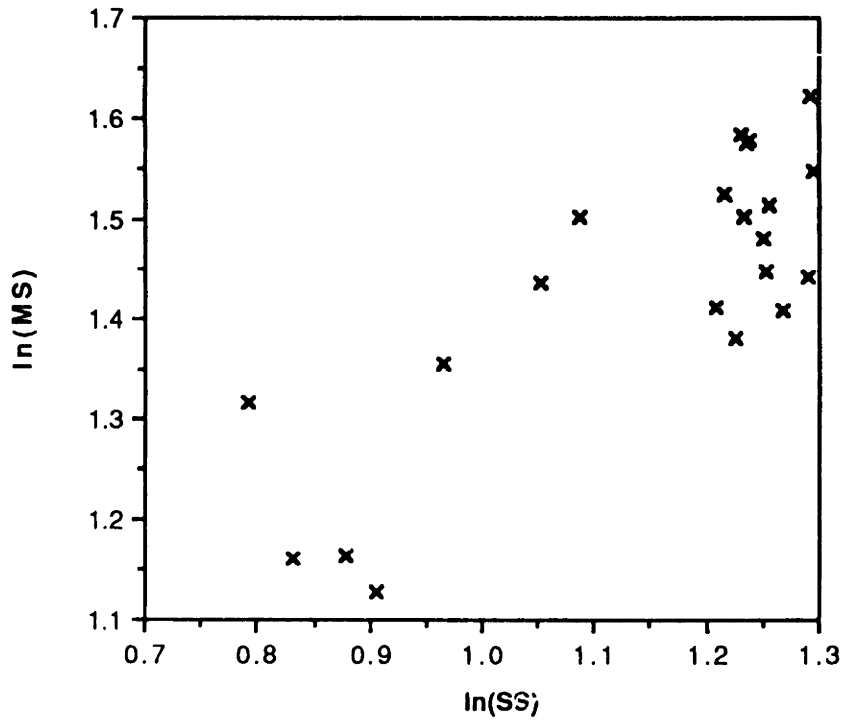


FIGURE 4.6
ln(SS) VS. ln(MS) IN MADISON

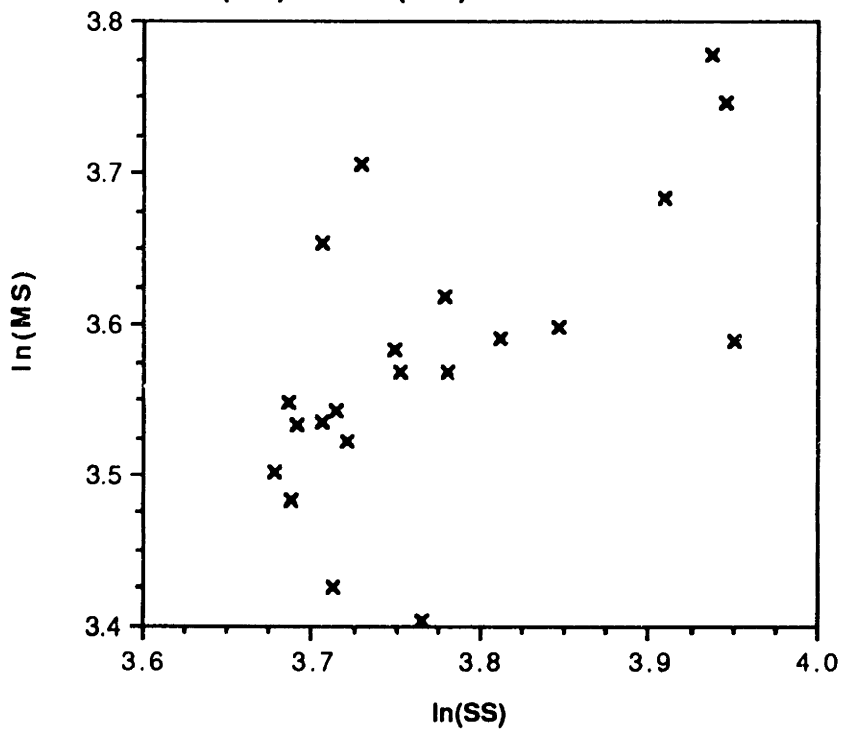


FIGURE 4.7
ln(SS) VS. ln(MS) IN MEMPHIS

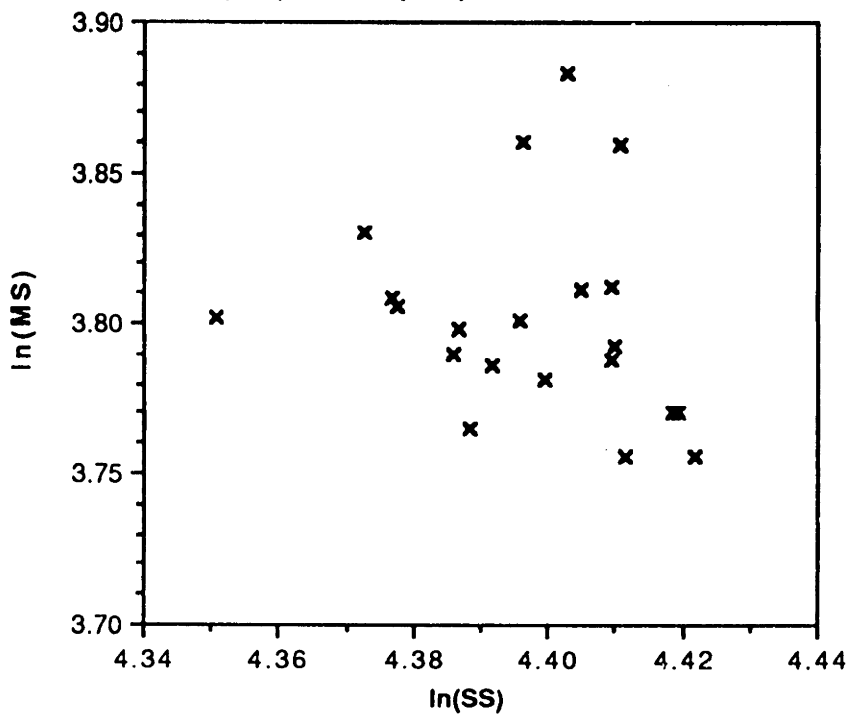


FIGURE 4.3
ln(SS) VS. ln(MS) IN MILWAUKEE

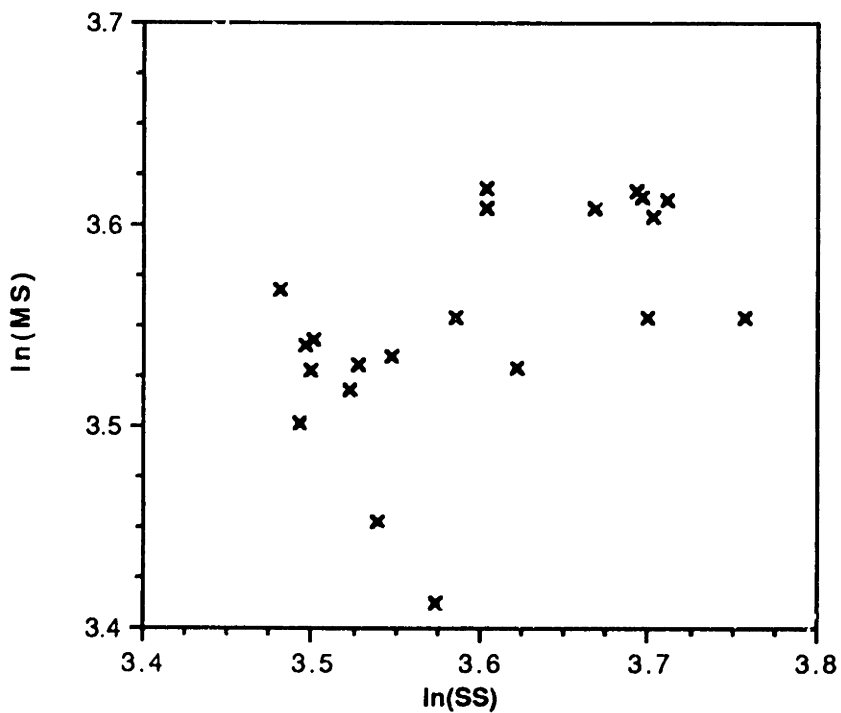


FIGURE 4.9
ln(SS) VS. ln(MS) in MINNEAPOLIS

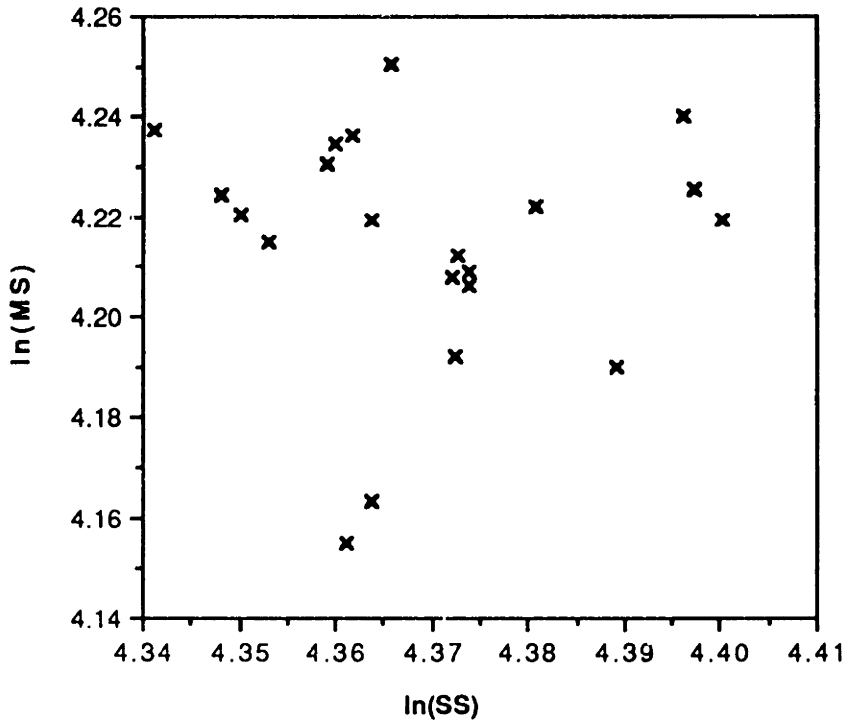


FIGURE 4.10
ln(SS) VS. ln(MS) IN OMAHA

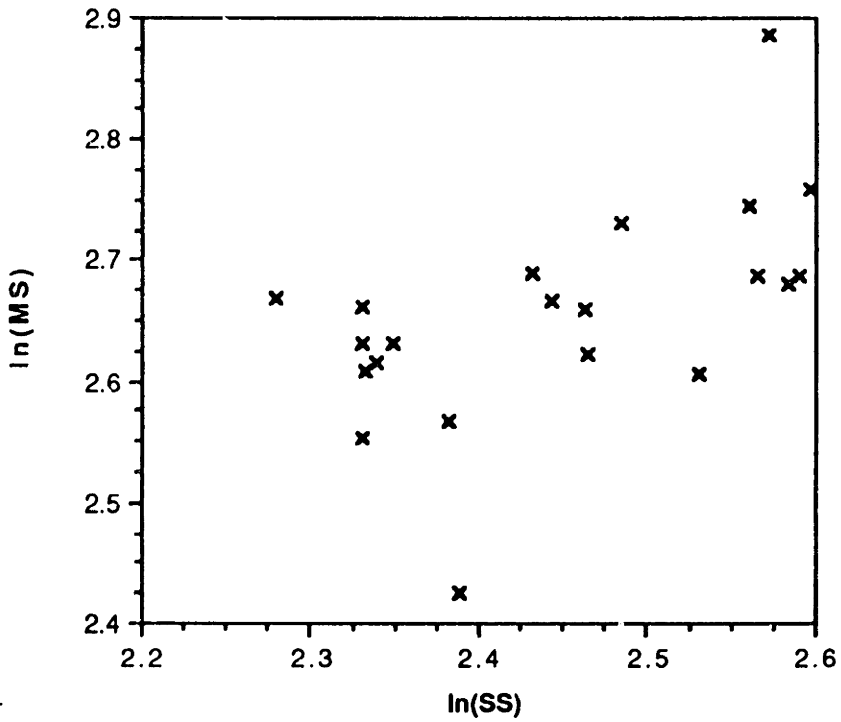


FIGURE 4.11
ln(SS) VS. ln(MS) IN ORLANDO

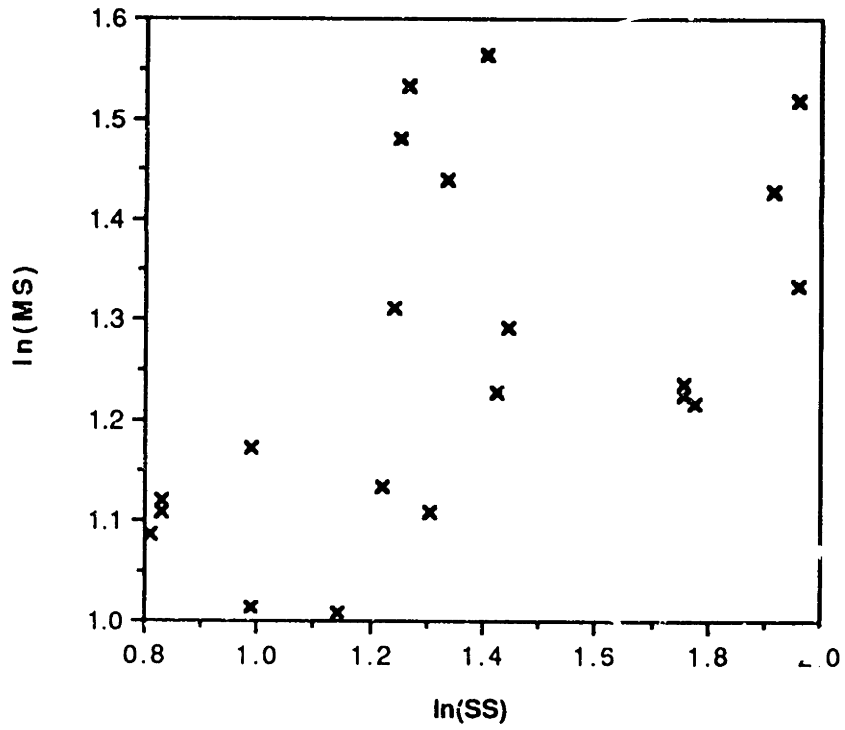


FIGURE 4.12
ln(SS) VS. ln(MS) in ROCHESTER

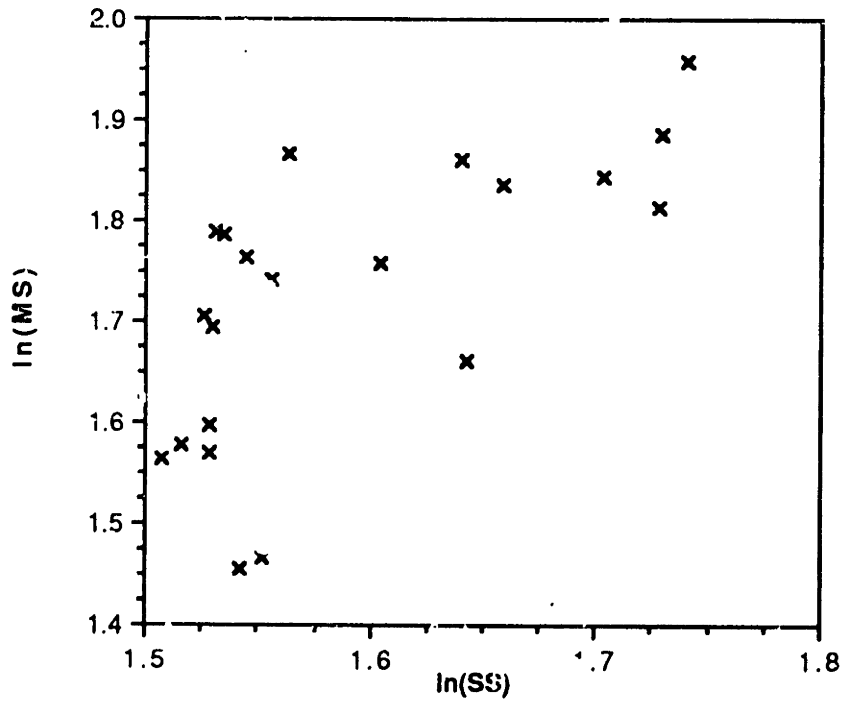


FIGURE 4.13
ln(SS) VS. ln(MS) in SAGINAW

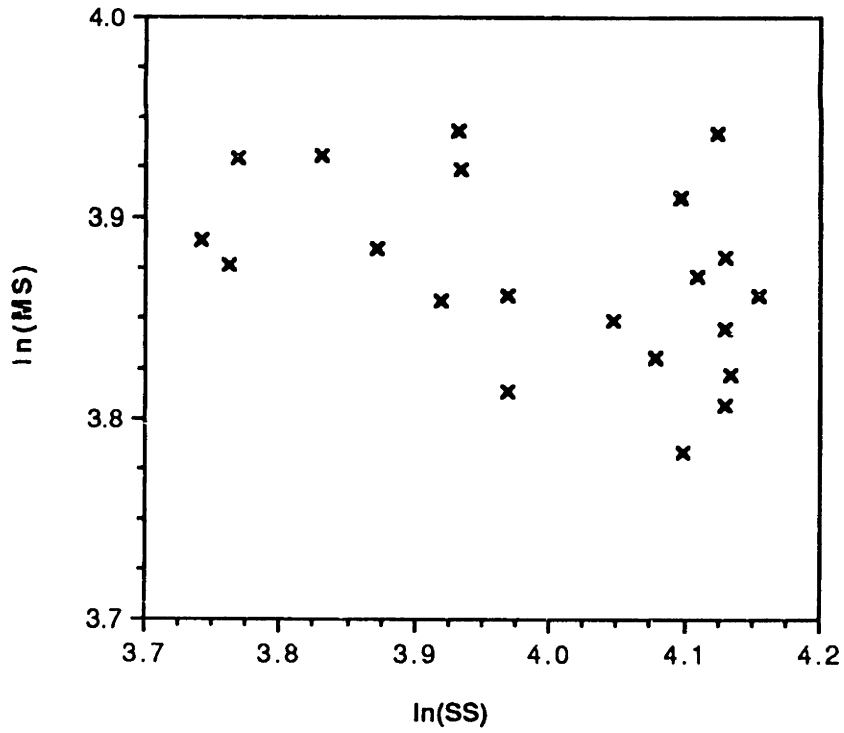


FIGURE 4.14
ln(SS) VS. ln(MS) IN SEATTLE

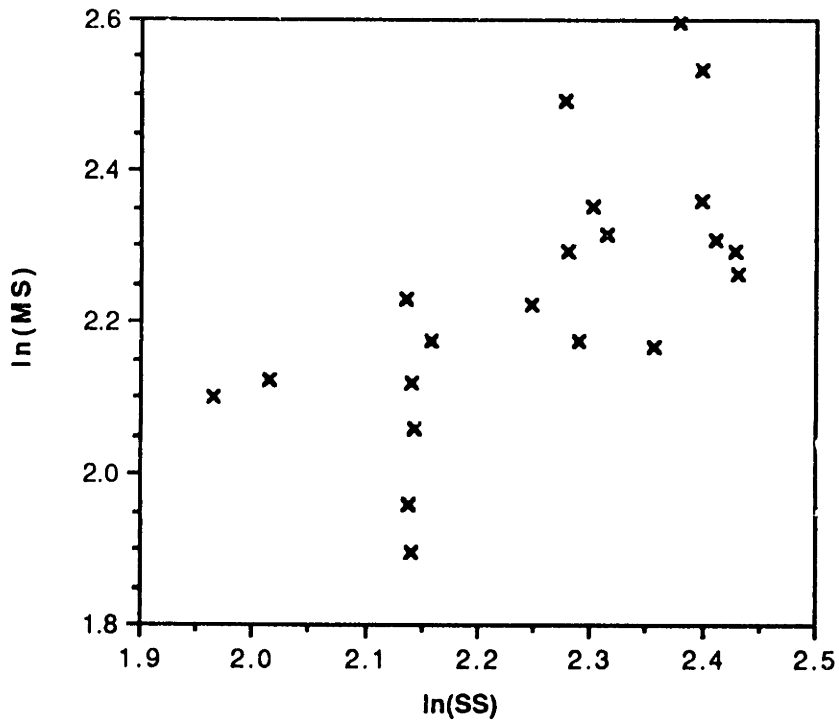
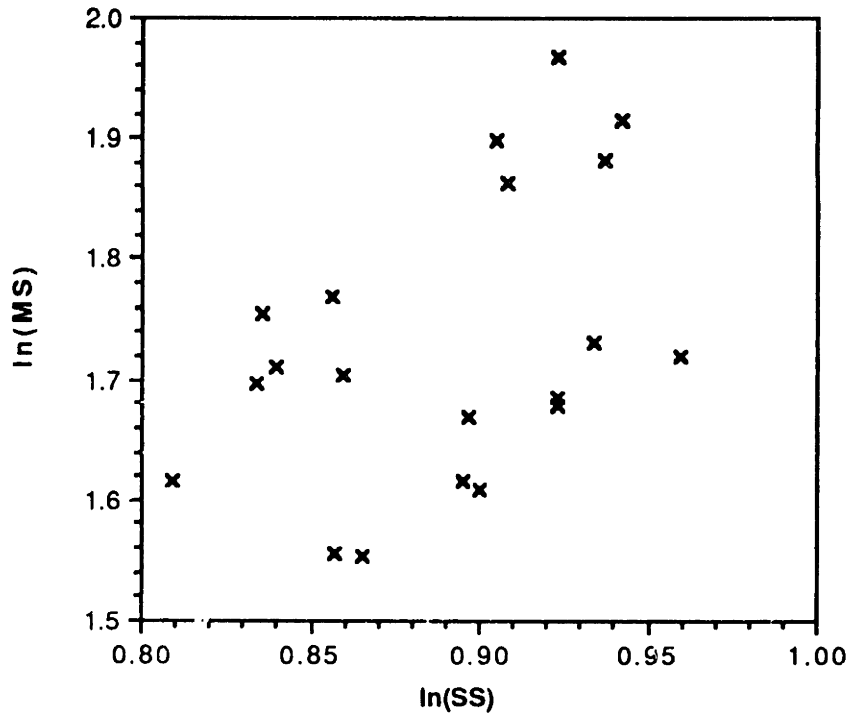


FIGURE 4.15
ln(SS) VS. ln(MS) IN ST. LOUIS



$$\frac{\partial S}{\partial \alpha} = 0 \quad (4.9)$$

and

$$\frac{\partial S}{\partial \beta} = 0 \quad (4.10)$$

This set of linear equations can then be solved simultaneously to find the parameters α and β .

Although this minimization technique can be employed manually without great difficulty, computer algorithms are available which not only carry out this procedure, but also calculate model performance statistics.¹ As in Chapter Three, the main statistic that will be used to compare the performance of these models is the root mean square error (RMSE) which is given by,

$$RMSE = \sqrt{\frac{\sum_{i=1}^m (MS_i - \alpha \times SS_i^\beta)^2}{m - k}} \quad (4.11)$$

where α and β are determined from the least squares regression and k is equal to the number of estimated parameters (in this case $k=2$). This statistic is identical to that used in Chapter Three (Equation 3.3) since MS_i is the actual market share and $\alpha \times SS_i^\beta$ is the predicted market share.

After transforming the market share and seat share data provided by Northwest (described in Chapter Three), a linear regression was performed for

each of the fifteen test markets. Table 4.1 shows the results of this linear regression. For each city, the coefficients α and β are listed, followed by the RMSE; the adjusted coefficient of determination, R_c^2 ; the average predicted market share, \overline{MSP} ; the average actual market share \overline{MSA} ; and the standard deviation of the actual market share, σ_{MSA} , over the sample period. Also presented in Table 4.1 are the t-statistics for the estimated parameter β (in parentheses underneath the parameter). As in Chapter Three, the RMSE statistic should be compared with the standard deviation of the actual market share data. At a minimum, the RMSE should be lower than σ_{MSA} ; the lower it is, the better the fit.

The results in Table 4.1 show that in thirteen of the fifteen cities, the RMSE is lower than σ_{MSA} ; only in Detroit and Minneapolis was the RMSE higher. However, in Memphis and Orlando, the RMSE was equal to and only slightly lower than σ_{MSA} , respectively. Similar to the case in Chapter Three, this model does not explain Northwest's market share in its three major hubs very well. As indicated earlier, this result was not unexpected since the figures showing $\ln(MS)$ vs. $\ln(SS)$ indicated a very weak linear relationship in these three cities.

The t-statistics for each of the estimated β s indicate the significance of the seat share variable in explaining Northwest's market share.² Values of this statistic which are less than -2.0 and greater than 2.0 reflect an acceptably high level of confidence (~95%) that the estimated coefficient (β) is significantly different from zero. Analysis of these statistics in Table 4.1 indicates a low level of confidence in the coefficient β only in the hub cities. This again suggests that the model performs poorly in these types of markets.

CITY	α	β	RMSE	R_c^2	\overline{MSP}	\overline{MSA}	σ_{MSA}
BOSTON	1.98	.6936 (4.09)	.94	.44	7.49	7.53	1.19
CHICAGO	2.48	.6897 (4.03)	.57	.43	5.32	5.34	.73
DETROIT	79.79	-.1333 (-.59)	2.24	.02	45.63	45.66	2.21
HUNTSVILLE	3.56	.6699 (6.02)	1.99	.64	24.32	24.40	3.29
LOS ANGELES	1.90	.6931 (6.10)	.35	.64	4.21	4.23	.57
MADISON	2.99	.6588 (3.54)	2.68	.37	35.94	36.04	3.48
MEMPHIS	712.9	-.6303 (-1.28)	2.73	.03	44.61	44.64	1.79
MILWAUKEE	10.66	.3297 (2.70)	1.63	.24	34.92	34.95	1.89
MINNEAPOLIS	86.64	-.0565 (-.17)	1.63	.0	67.69	67.70	1.59
OMAHA	4.46	.4748 (2.95)	1.08	.28	14.26	14.30	1.29
ORLANDO	2.56	.2380 (2.60)	.64	.22	3.63	3.59	.64
ROCHESTER	.85	1.1851 (3.91)	.58	.42	5.63	5.66	.77
SAGINAV,	86.44	-.1469 (-2.12)	2.07	.15	48.06	48.10	2.24
SEATTLE	1.41	.8411 (3.83)	1.34	.40	9.45	9.53	1.70
ST. LOUIS	1.82	1.2725 (2.15)	.62	.16	5.65	5.68	.70

TABLE 4.1: RESULTS OF SEAT SHARE REGRESSION ANALYSIS

As stated in the Introduction, the exponent β in Equation 4.2 represents the elasticity of market share with respect to seat share. A value that is positive and greater than 1.0 implies a more than proportional increase in market share with an increase in seat share. For the case, $0 < \beta < 1.0$, an increase in seat share implies a less than proportional increase in market share. The implication of a negative β (the only significant $-\beta$ was in Saginaw) is a loss in market share with increasing seat share; this is, of course, counter-intuitive. The overall significance of the exponent β will be discussed in more detail in the final chapter.

Finally, in Table 4.1, the adjusted coefficient of determination, R_C^2 , from the linear regression analysis is listed for each test market. This statistic provides a relative measure of the strength of the relationship that has been fit by least squares.³ The R_C^2 statistic can also be interpreted as the proportion of the variability of the dependent variable (market share) that is explained by the independent variables (in this model, only seat share). As expected, since this model performed poorly in the three hub cities, the R_C^2 values in these markets are all close to zero. Disregarding the three hub cities, R_C^2 varied from .16 in St. Louis to .64 in L. Angeles and Huntsville. This higher value implies that 64% of the variation in Northwest's market share in these two cities is explained by seat share. These lower values suggest that other explanatory variables might be included in Equation 4.2 as a means of further explaining a carrier's market share out of an originating city.

4.3 Calibration of Seat Share-Frequency Share Regression Model

Of the three models tested thus far, clearly, the simple exponential seat share model from the last section has provided the best results. Except for the three hub cities, this model has produced satisfactory results (significant β coefficients and $RMSE < \sigma_{MSA}$) in all of the originating city markets. However, analysis of the R_c^2 statistic in Table 4.1 seems to indicate that other independent variables are necessary to further explain the variation in a carrier's market share in an originating city market.

In Chapter Three, it was concluded that seat share was a more useful variable than frequency share in modeling the variation in a carrier's market share. However, as discussed previously, if two carriers operating in a market have equal seat shares yet different frequency shares, then the carrier offering more departures is likely to have a competitive advantage. Since the results of the last section seem to indicate that further independent variables are needed in the market share model, including frequency share in Equation 4.2 may be of value.

The structure of the market share model in Equation 4.2 is extremely useful because it allows one to include further independent variables in a manner that still permits calibration by linear least squares. The exponential model in Equation 4.2 can be extended by simply multiplying the right hand side of this equation by the new independent variable raised to its elasticity. This new model,

$$MS = \alpha \times SS^{\beta_1} \times FS^{\beta_2} \quad (4.3)$$

where β_2 is the elasticity of market share with respect to frequency share, will be calibrated in the fifteen test cities to determine if frequency share can be used to further explain the variation in a carrier's market share.

The model in Equation 4.3 must first be linearized, however, before the method of linear least squares can be applied. Taking the natural log of both sides of this equation yields

$$\ln(MS) = \ln(\alpha \times SS^{\beta_1} \times FS^{\beta_2}) \quad (4.12)$$

which simplifies to

$$\ln(MS) = \ln(\alpha) + \beta_1 \times \ln(SS) + \beta_2 \times \ln(FS) \quad (4.13)$$

Equation 4.13 not only assumes a linear relationship between $\ln(MS)$ and $\ln(SS)$ as in the previous section, but it also assumes a linear relationship between $\ln(MS)$ and $\ln(FS)$. This further assumption is easily checked by plotting $\ln(MS)$ vs. $\ln(FS)$ in each of the fifteen test markets. Figures 4.16 - 4.30 depict this relationship for the twenty-one months of data described in Chapter Three.

Similar to the plots of $\ln(MS)$ vs. $\ln(SS)$ in the previous section, the scatter plots in the hub cities do not exhibit a strong linear relationship. In addition to

FIGURE 4.16
ln(FS) VS. ln(MS) IN BOSTON

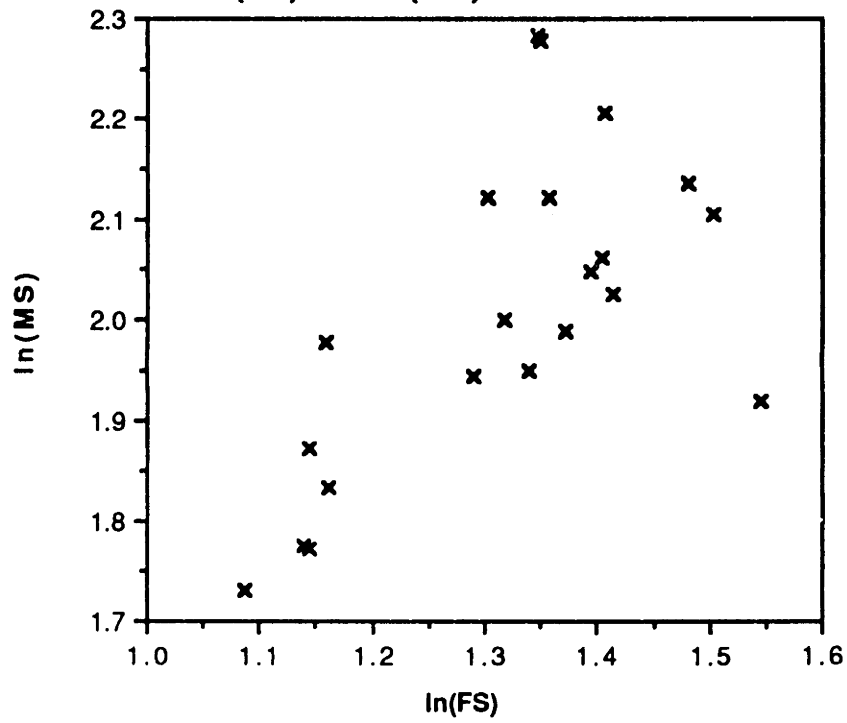


FIGURE 4.17
ln(FS) VS. ln(MS) IN CHICAGO

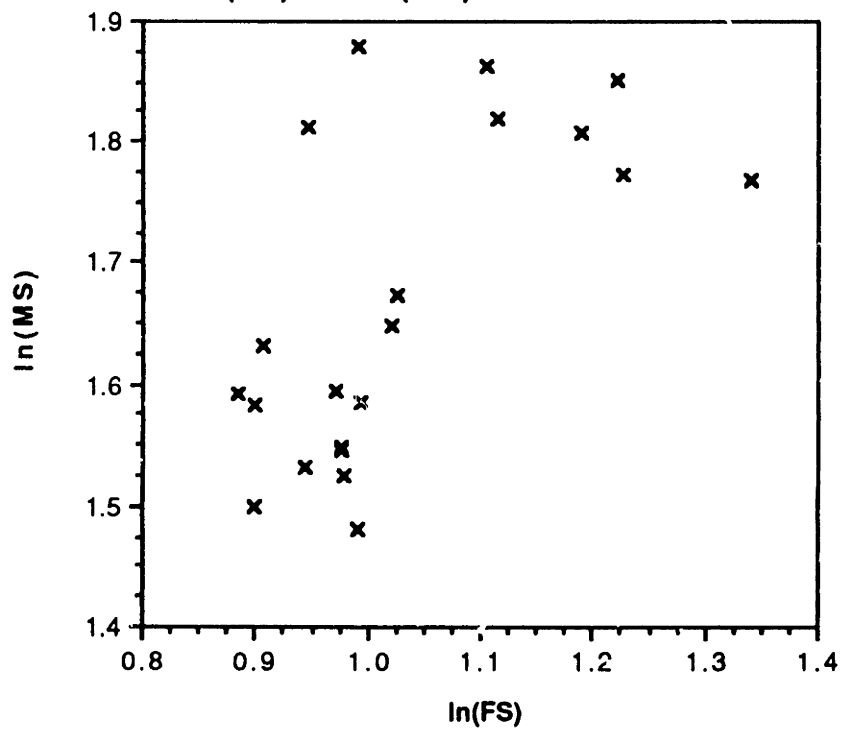


FIGURE 4.18
ln(FS) VS. ln(MS) IN DETROIT

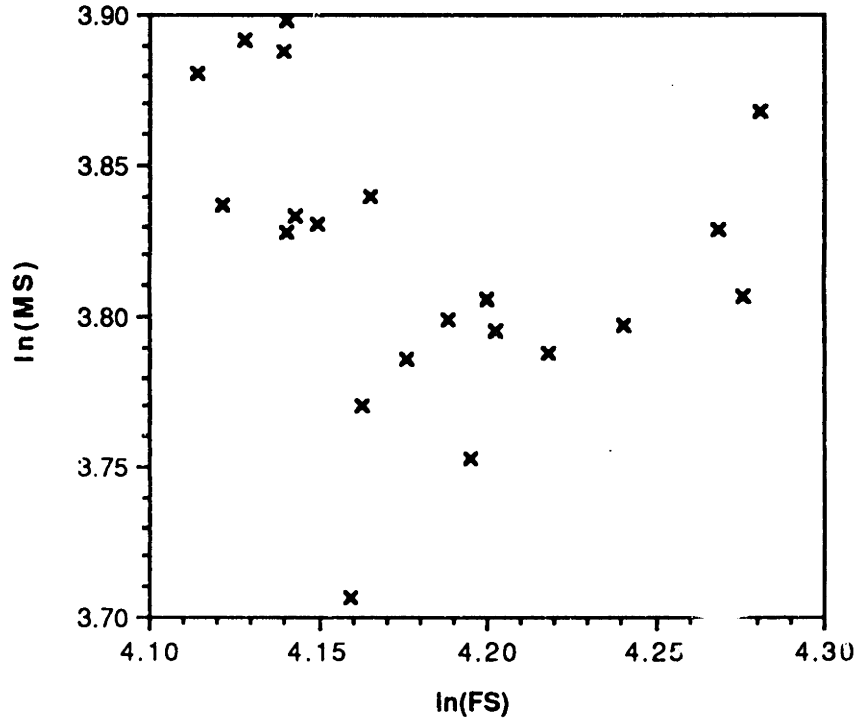


FIGURE 4.19
ln(FS) VS. ln(MS) IN HUNTSVILLE

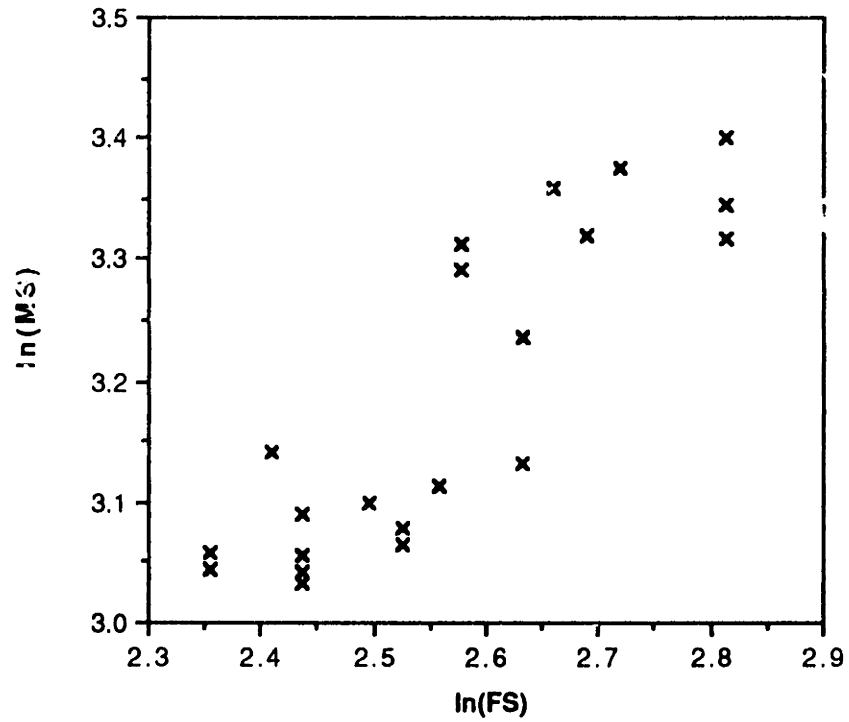


FIGURE 4.20
ln(FS) VS. ln(MS) IN LOS ANGELES

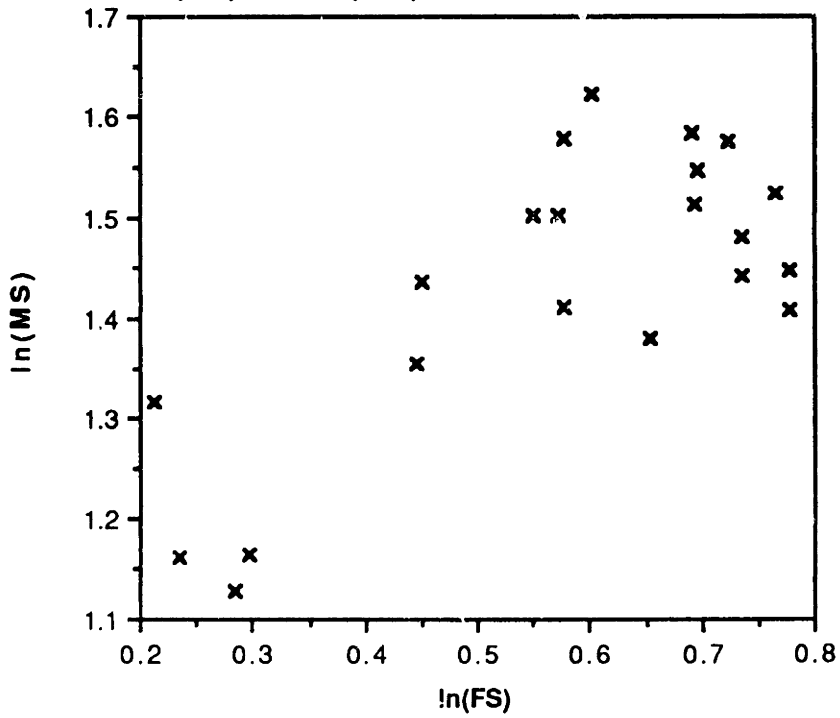


FIGURE 4.21
ln(FS) VS. ln(MS) IN MADISON

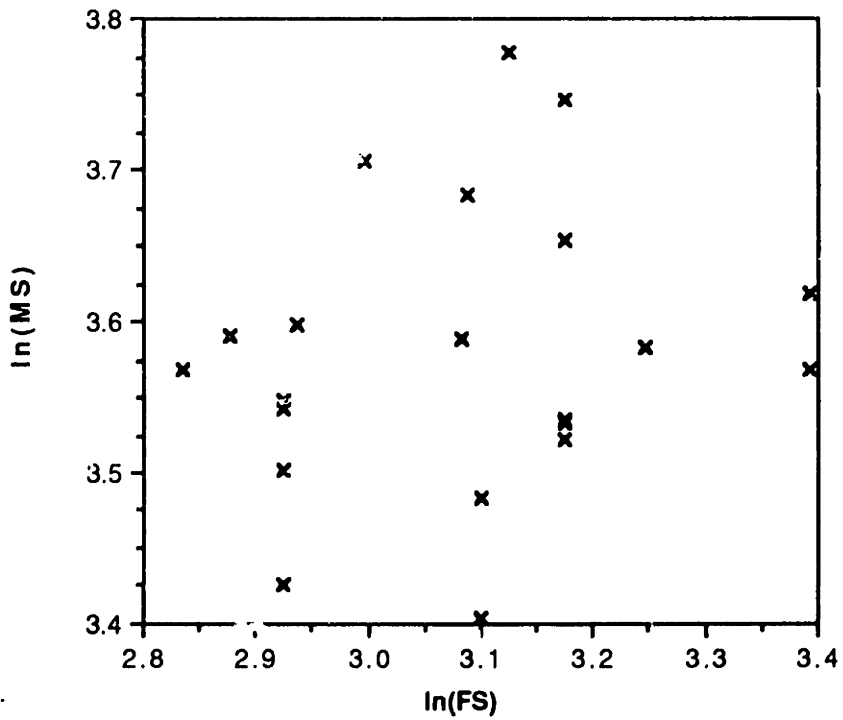


FIGURE 4.22
ln(FS) VS. ln(MS) IN MEMPHIS

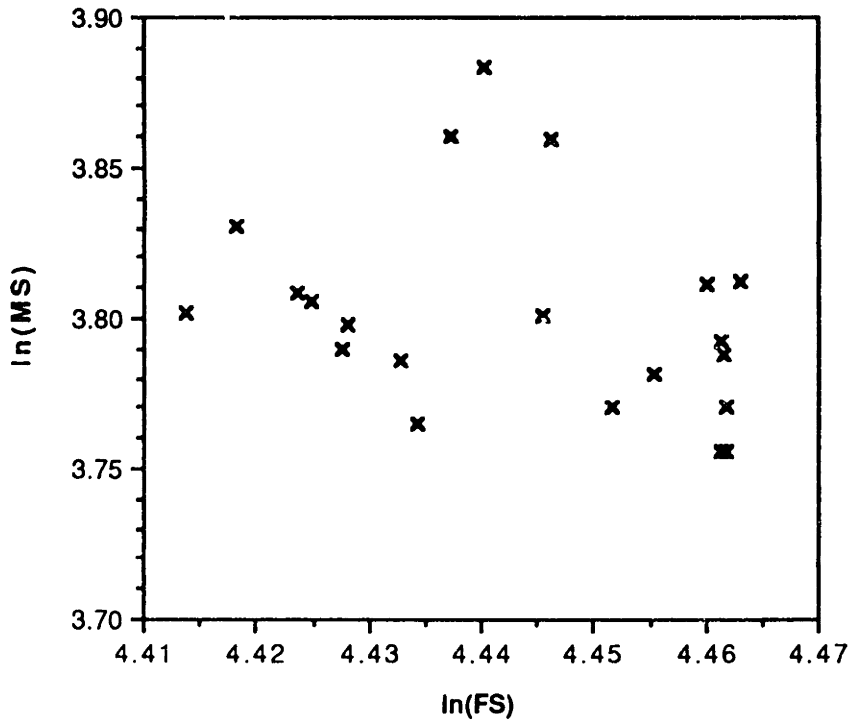


FIGURE 4.23
ln(FS) VS. ln(MS) IN MILWAUKEE

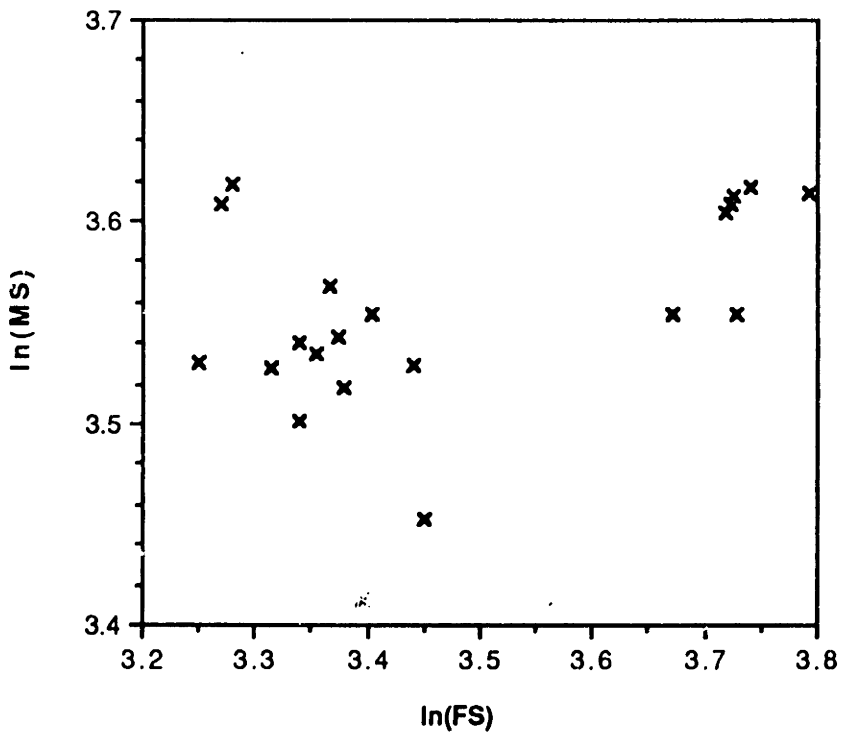


FIGURE 4.24
ln(FS) VS. ln(MS) IN MINNEAPOLIS

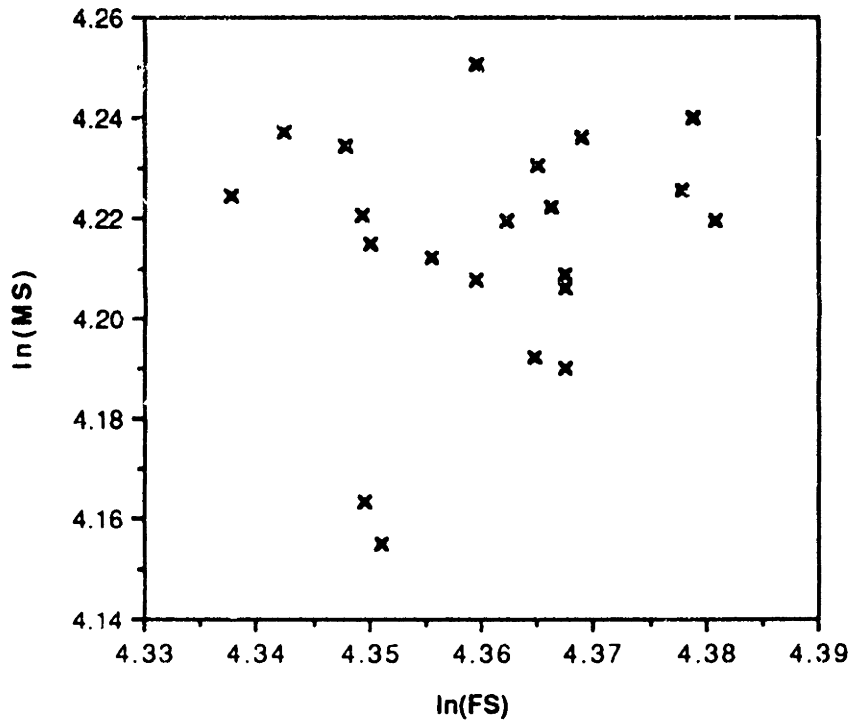


FIGURE 4.25
ln(FS) VS. ln(MS) IN OMAHA

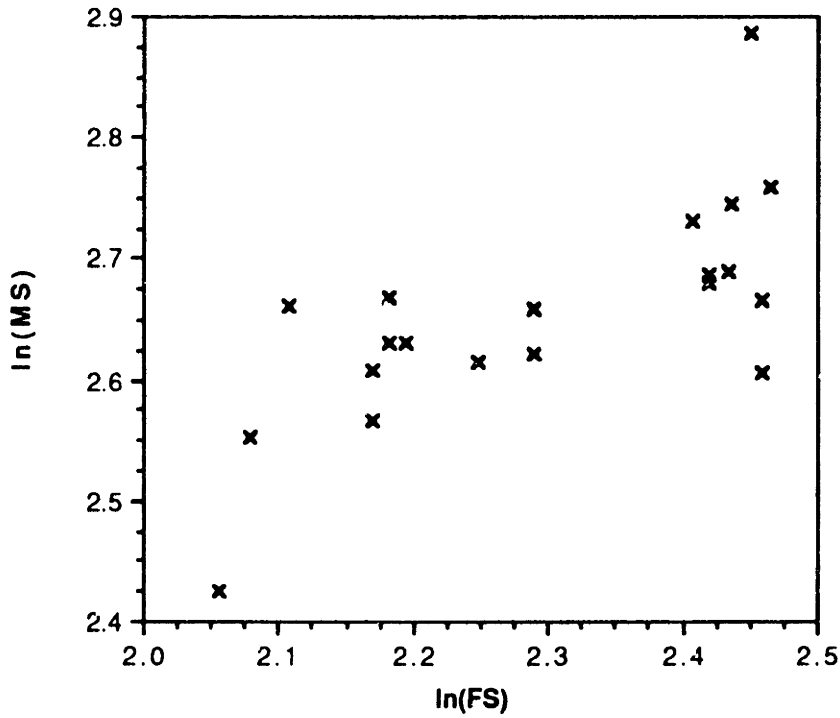


FIGURE 4.26
ln(FS) VS. ln(MS) IN ORLANDO

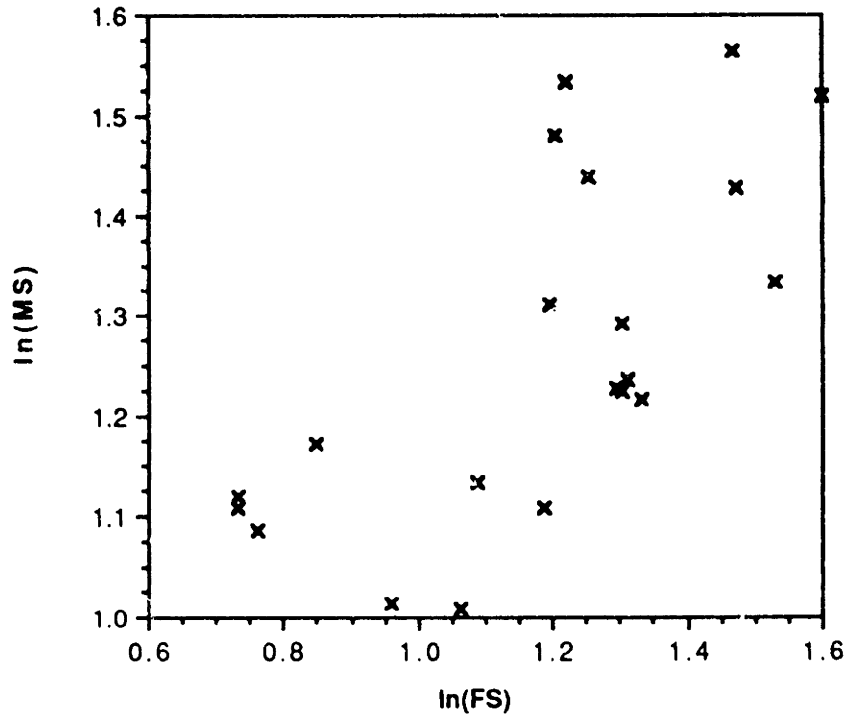


FIGURE 4.27
ln(FS) VS. ln(MS) IN ROCHESTER

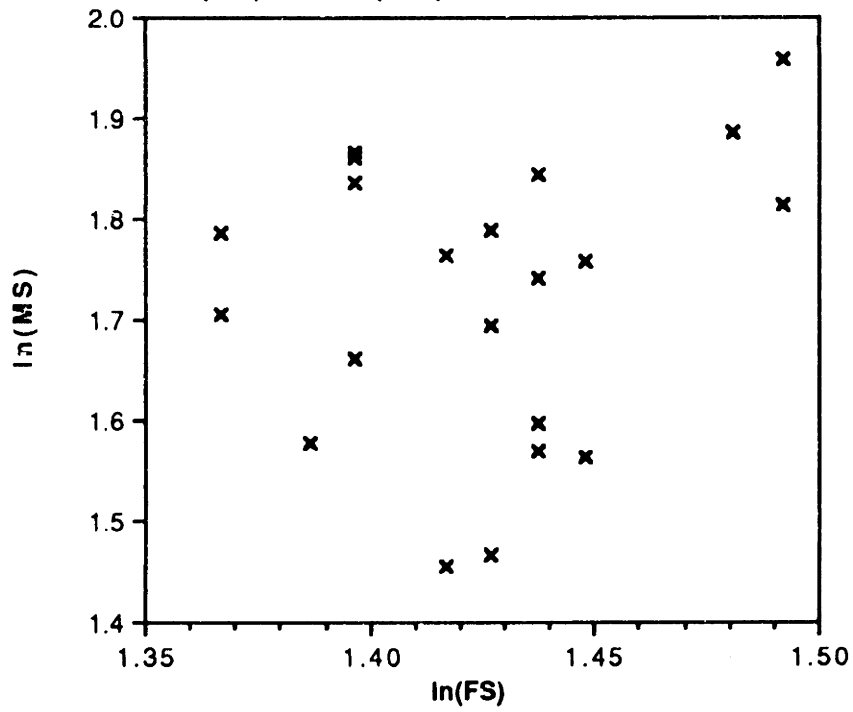


FIGURE 4.28
ln(FS) VS. ln(MS) IN SAGINAW

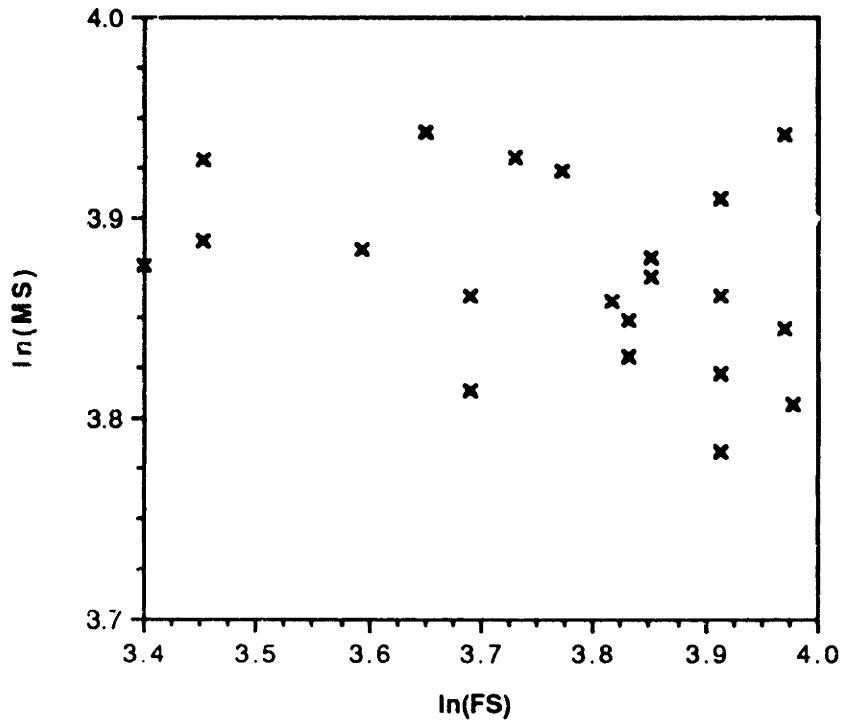


FIGURE 4.29
ln(FS) VS. ln(MS) IN SEATTLE

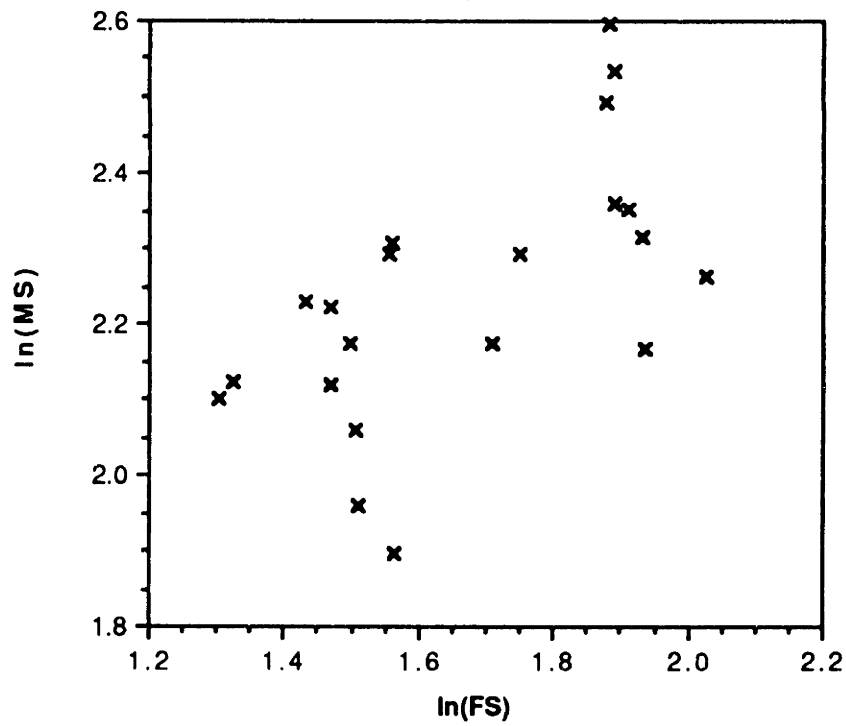
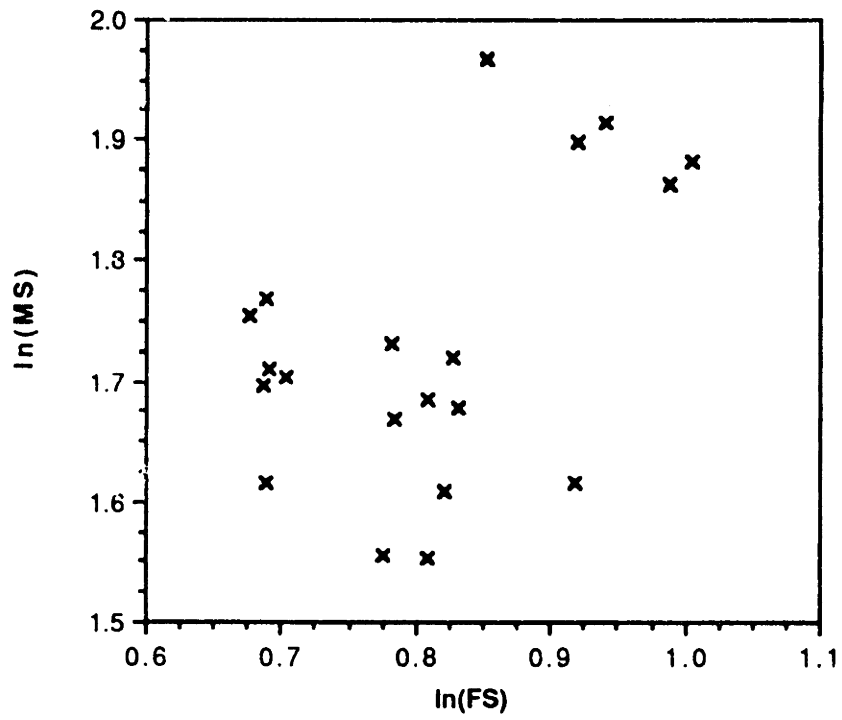


FIGURE 4.30
ln(FS) VS. ln(MS) IN ST. LOUIS



the weak relationships shown in these markets, the plots of $\ln(\text{MS})$ vs. $\ln(\text{FS})$ in Madison (Figure 4.21), Milwaukee (Figure 4.24), and Rochester (Figure 4.27) also show non-linear trends between these two variables. Based upon these figures, it is doubtful whether the inclusion of frequency share in Equation 4.3 will improve the overall accuracy of the regressive market share model in these markets.

Applying the method of linear least squares to Equation 4.13, the parameters α , β_1 , and β_2 are determined by minimizing

$$S(\alpha, \beta_1, \beta_2) = \sum_{i=1}^m [\ln(\text{MS}_i) - \ln(\alpha) - \beta_1 \times \ln(\text{SS}_i) - \beta_2 \times \ln(\text{FS}_i)]^2 \quad (4.14)$$

where, as before, $i=1, m$ is the number of months over which data is available. The minimization of Equation 4.14 is identical to that described in the previous section with the addition of a third partial derivative,

$$\frac{\partial S}{\partial \beta_2} = 0 \quad (4.15)$$

Equations 4.9, 4.10, and 4.15 form a set of linear equations that can be solved simultaneously to find the parameters α , β_1 , and β_2 .

Using the seat share, frequency share, and market share data described in Chapter Three, a multiple linear regression of Equation 4.13 was performed in each of the fifteen originating city markets. Table 4.2 shows the estimated parameters α , β_1 , and β_2 from the regression analysis along with the t-statistics (underneath the parameter and in parentheses) for each of the coefficients. Also

CITY	α	β_1	β_2	RMSE	R_C^2	\overline{MSP}
BOSTON	1.94	.4711 (1.78)	.3406 (1.09)	.95	.46	7.49
CHICAGO	2.48	.6897 (4.03)		.57	.41	5.32
DETROIT	85.34	.9226 (1.50)	-1.0514 (-1.84)	2.10	.17	45.63
HUNTSVILLE	3.16	.0970 (.40)	.6850	1.74	.72	24.35
LOS ANGELES	1.79	.7893 (2.32)	-.0925 (-.30)	.35	.63	4.21
MADISON	2.99	.6588 (3.54)		2.73	.37	35.94
MEMPHIS	1956		-.8509 (-1.64)	1.72	.08	44.61
MILWAUKEE	10.12	.3603 (1.47)	-.0167 (-.15)	1.68	.20	34.92
MINNEAPOLIS	7.90	-1.0267 (-1.56)	1.5123 (1.69)	1.56	.04	67.69
OMAHA	5.16		.4412 (6.87)	.98	.47	14.27
ORLANDO	1.88	-.3752 (-2.08)	.9688 (3.69)	.42	.53	3.57
ROCHESTER	2.11	1.4000 (3.89)	-.8733 (-1.09)	.57	.42	5.63
SAGINAW	95.07	-.1469 (-2.12)	.1025 (1.70)	2.10	.13	48.06
SEATTLE	1.80	.5887 (1.67)	.1962 (.92)	1.33	.40	9.45
ST. LOUIS	3.33		.6529 (2.77)	.59	.26	5.65

TABLE 4.2: RESULTS OF SS - FS REGRESSION ANALYSIS

shown in Table 4.2 are the RMSE; the adjusted coefficient of determination, R_c^2 ; and the average predicted market share, \overline{MSP} . As before, the RMSE should be compared with the standard deviation of the actual market share (σ_{MSA} , provided in Table 4.1) to assess the accuracy of the model.

In assessing the significance of each coefficient, t-statistics which are less than -2.0 and greater than 2.0 indicate an acceptable level of significance. The values of some coefficients were listed even if the t-statistic indicated a lack of confidence in the parameter estimate, except in some cases when the t-statistic indicated extremely low levels of confidence. Table 4.2 indicates that except for Orlando, there are no cities that show two significant values of β . However, the negative β_1 coefficient found in this city may not be of practical use.

Although this negative β_1 coefficient improves the overall "goodness of fit" for the regression model, the use of this model (with the negative coefficient) to forecast market share may lead to questionable results. Since there are many other forces that affect a carrier's market share, a negative elasticity does not necessarily imply a causal relationship (i.e. a loss in market share caused by an increase in seat share). Including this negative coefficient in a forecasting model would suggest to management that travel agency market share quotas should be reduced with a projected increase in seat share, which, of course, makes no sense.

The results in Table 4.2 show that in four of the cities: Huntsville, Orlando, Omaha, and St. Louis, frequency share is a more significant explanatory variable

than seat share. The lower RMSE statistic in each of these cities, compared with the RMSE of these same cities in Table 4.1, also indicates that frequency share is a better predictor variable than seat share. Except for the three hub cities (Detroit, Memphis, and Minneapolis) seat share was still the best predictor variable in the remaining cities. Although Table 4.2 shows that for some cities (e.g. Boston) neither β_1 nor β_2 was significant, Table 4.1 clearly indicates that a regression with seat share alone produced a significant β_1 . Thus, if the multiple regression of Equation 4.3 does not produce at least one significant β when both variables are included, then a step wise regression (discussed in the next section) can be used to identify which of the variables is significant.

Tables 4.1 and 4.2 indicate that neither seat share nor frequency share (or a combination of them) are very useful in explaining a carrier's market share in a hub city. The scatter plots of $\ln(\text{MS})$ vs. $\ln(\text{SS})$ and $\ln(\text{MS})$ vs. $\ln(\text{FS})$ in each of these markets did not reveal a strong linear relationship between these variables, and thus, this result was not unexpected. Perhaps the inclusion of other independent variables in Equation 4.3, conducted in the next section, will improve the results of the modeling in the hub cities.

Finally, although the R_c^2 statistic was lower in the four cities where frequency share was found to be a more useful variable, Table 4.2 suggests that further explanatory variables might still be included in Equation 4.3. In the next section of this chapter, two new variables, a carrier's share of the total non-stop markets served by all airlines and a carrier's share of the total possible markets served by all airlines, will be included in Equation 4.3. By calibrating this model

in the fifteen test markets, an assessment will then be made with respect to the usefulness of these new parameters.

4.4 Calibration of the Multivariable Multiplicative Regression Model

As demonstrated in the previous section, the structure of the simple multiplicative market share model (Equation 4.4) is extremely useful because it allows one to include additional independent variables in a manner that still permits calibration by linear least squares. In the previous section, frequency share was included in the original model and was found to be a more significant variable than seat share in four out of the fifteen test markets. However, the coefficient of determination, R_C^2 , in Table 4.2 indicated that there was still room for additional parameters to further explain the variation in a carrier's market share. In this section, two new independent variables will be added to Equation 4.3 in an attempt to further explain a carrier's market share in an originating city market.

The first variable to be included in Equation 4.3 is a carrier's share of the total non-stop markets served by all airlines out of an originating city market. Since air travelers generally prefer non-stop service over one or two-stop service, a carrier that offers a more than proportional share of the non-stop

flights should gain a more than proportional share of the originating market. If two carriers have identical seat shares and frequency shares, then the carrier serving more non-stop destinations should have a competitive advantage. Using data provided by Northwest Airlines (as tabulated from OAG tapes), its share of the total non-stop markets served by all airlines, given by,

$$\text{NSS} = \frac{\text{total \# Northwest non-stop markets}}{\text{total \# non-stop markets all carriers}} \quad (4.16)$$

was determined in the fifteen test markets over the twenty-two month sample period.⁴

The second variable to be included in Equation 4.3 is a carrier's share of the total possible markets (which includes through flights and all viable connections) served by all carriers out of an originating city market. Obviously, the more destinations that an airline offers the consumer, the better its chances are of meeting that consumer's demand. If two carriers operating in a market have identical seat shares and frequency shares, then the carrier who offers a greater number of destinations to the consumer should enjoy a competitive advantage. Using data provided by Northwest Airlines (again, as tabulated from OAG tapes), its share of the total possible markets served, given by,

$$\text{PMS} = \frac{\text{total \# Northwest markets served}}{\text{total \# markets served all carriers}} \quad (4.17)$$

was determined in the fifteen test markets over the twenty-two month sample period.

Including these two new variables into Equation 4.3 yields

$$MS = \alpha \times SS^{\beta_1} \times FS^{\beta_2} \times NSS^{\beta_3} \times PMS^{\beta_4} \quad (4.18)$$

where β_3 is the elasticity of market share with respect to non-stop share and β_4 is the elasticity of market share with respect to a carrier's share of possible markets served.

Similar to the previous models, Equation 4.18 must first be linearized before calibration by linear least squares is possible. Taking the natural log of both sides of Equation 4.18 yields

$$\ln(MS) = \ln(\alpha) + \beta_1 \times \ln(SS) + \beta_2 \times \ln(FS) + \beta_3 \times \ln(NSS) + \beta_4 \times \ln(PMS) \quad (4.19)$$

which simplifies to

$$\ln(MS) = \ln(\alpha) + \beta_1 \times \ln(SS) + \beta_2 \times \ln(FS) + \beta_3 \times \ln(NSS) + \beta_4 \times \ln(PMS) \quad (4.20)$$

This equation can now be calibrated using the technique of linear least squares. Again, it should be noted that Equation 4.20 assumes a linear relationship between $\ln(MS)$ and all the explanatory variables given in this equation.

Applying the method of linear least squares to Equation 4.14, the parameters α , β_1 , β_2 , β_3 , and β_4 are determined by minimizing

$$S(\alpha, \beta_1, \beta_2, \beta_3, \beta_4) = \sum_{i=1}^m [(\ln(MS_i) - \ln(\alpha) - \beta_1 \ln(SS_i) - \beta_2 \ln(FS_i) - \beta_3 \ln(NSS_i) - \beta_4 \ln(PMS_i))]^2 \quad (4.21)$$

over the $i=1, m$ months of data. The minimization of this equation is identical to that described in the previous sections with the addition of two partial derivatives,

$$\frac{\partial S}{\partial \beta_3} = 0 \quad (4.22)$$

and

$$\frac{\partial S}{\partial \beta_4} = 0 \quad (4.23)$$

Equations 4.8, 4.9, 4.15, 4.23, and 4.24 form a set of linear equations that can be solved simultaneously to find the parameters α , β_1 , β_2 , β_3 , and β_4 .

In calibrating Equation 4.20 in the fifteen originating city markets, a type of regression analysis referred to as "step-wise regression" was employed. A step-wise regression analysis will add the independent variables in Equation 4.21, one at a time, depending both on the contribution it makes to the model (i.e. whether the included variable reduces the RMSE) and on the value of its t-statistic. After the best independent variable is chosen, the other independent variables are

added to the equation, one at a time, and tested for significance.⁵ The procedure continues until no more variables can be added or deleted from the equation. The benefit of this type of regression analysis is that it insures that only the significant independent variables will be kept in the model.

Table 4.3 shows the results of this step-wise regression analysis of Equation 4.20 in the fifteen test markets. Values for α , β_1 , β_2 , β_3 , and β_4 are listed, with their respective t-statistics (underneath and in parentheses) if the variable was found to be significant. Also shown in Table 4.3 are the RMSE; the adjusted coefficient of determination, R_c^2 ; and the average predicted market share, \overline{MSP} . As before, the RMSE should be compared with the standard deviation of the actual market share data (σ_{MSA} , provided in Table 4.1) to assess the accuracy of the model.

The results of this analysis indicate that in fourteen of the fifteen test markets (except Memphis), at least one of the two new variables included in Equation 4.19 was found to be significant. In some of these markets, however, the coefficient was found to be negative, which, although statistically possible, is counter-intuitive. Table 4.3 also indicates that except for Saginaw and Seattle, at least one of the variables, either seat share or frequency share, was found to be significant. The R_c^2 statistic also showed improvement in markets where more than one explanatory variable was found to be significant.

A comparison of the RMSE statistic between Table 4.3 and Tables 4.1 and 4.2 clearly indicates that of the models tested thus far, this multivariable multiplicative regression model produces the best overall results. Since this

CITY	α	β_1	β_2	β_3	β_4	RMSE	R_c^2	\overline{MSP}
BOSTON	9.30	1.17 (7.44)	-.72 (-3.54)		.65 (7.71)	.44	.87	7.52
CHICAGO	33.10	.65 (2.00)	-.96 (-2.48)	.12 (1.45)	.47 (4.66)	.32	.79	5.33
DETROIT	.04	1.50 (4.05)		.71 (-4.90)	-.71 (-4.90)	1.87	.53	45.66
HUNTSVILLE	1.58		.77 (12.30)	-.33 (-6.57)		.99	.92	24.23
LOS ANGELES	.49	.96 (5.37)		-.30 (-1.85)		.34	.68	4.22
MADISON	34.81		.13 (1.50)		.27 (5.12)	2.28	.56	35.98
MEMPHIS	1959		-.85 (-1.64)			1.88	.08	44.51
MILWAUKEE	20.49		.23 (3.51)		.16 (2.63)	1.58	.35	34.92
MINNEAPOLIS	4.18	-.09 (-1.44)	.79 (3.68)		.08 (2.56)	1.65	.0	67.69
OMAHA	3.22		.63 (7.31)		.09 (1.69)	.90	.51	14.28
ORLANDO	4.62	-.23 (-1.60)	.52 (2.19)		.25 (3.63)	.32	.72	3.57
ROCHESTER	.002	2.31 (3.75)		-.99 (-1.60)	-.48 (-1.69)	.56	.50	5.64
SAGINAW	52.46				.12 (3.17)	2.02	.31	48.07
SEATTLE	77.48			.79 (5.01)	.07 (2.06)	1.23	.56	9.47
ST. LOUIS	101.5	1.19 (2.45)	-1.20 (-3.59)	.83 (6.55)		.36	.77	5.67

TABLE 4.3: RESULTS OF MULTIVARIABLE MULTIPLICATIVE REGRESSION ANALYSIS

model includes all of the independent variables tested in the previous sections, the step-wise regression of Equation 4.19 guarantees that, at a minimum, this model will be as good as the two previous models. The addition of the new independent variables, NSS and PMS, if found to be significant, can only improve upon the results of the two previous sections.

With these four independent variables: SS, FS, NSS, and PMS, a linear multiple regression model will be calibrated in the next section of this chapter. Having found variables that are significant in almost every city (except Memphis), this analysis will indicate whether the non-linearized structure of the models tested in the three previous sections can be altered so as to improve upon the accuracy of the model results obtained thus far.

4.5 Calibration of the Multivariable Linear Regression Model

In the last three sections of this chapter, four independent variables have been tested for their ability to explain the variation in a carrier's market share in an originating city market. The results have indicated that in thirteen of the fifteen test markets either seat share or frequency share was necessary to explain the variation in a carrier's market share. Of the two new variables included in the last section, it was determined that in thirteen of the fifteen test markets

either non-stop share or share of total possible markets served improved the overall accuracy of the model.

The three market share models presented thus far in this chapter have assumed a non-linear relationship between market share and the four independent variables. Recall that a natural logarithmic transformation was required to linearize the models before calibration by linear least squares could be carried out. Although the results in the last section have substantiated the use of such a model structure, there are other ways of combining these four variables which might lead to an improved market share model. In this section of Chapter Four, the linear multiple regression model

$$MS = A + B_1 SS + B_2 FS + B_3 NSS + B_4 PMS \quad (4.5)$$

where A , B_1 , B_2 , B_3 , and B_4 are constants, will be calibrated in the fifteen originating city markets. Having already calibrated a market share model using these same independent variables, the results of the calibration of Equation 4.5 will allow an assessment to be made concerning the general structure of these two models.

Since Equation 4.5 is already linear with respect to market share, the method of linear least squares can be applied to determine the constants A , B_1 , B_2 , B_3 , and B_4 . As before, these constants are determined by minimizing

$$S(A, B_1, B_2, B_3, B_4) = \sum_{i=1}^m (MS_i - A - B_1 SS_i - B_2 FS_i - B_3 NSS_i - B_4 PMS_i)^2 \quad (4.25)$$

where $i=1,m$ is the number of months over which data is available. The minimization of Equation 4.25 is identical to the procedure described in Section 4.4. Partial derivatives of S with respect to A , B_1 , B_2 , B_3 , and B_4 are set equal to zero to form five linear equations. These equations can then be solved simultaneously to find estimates of the five parameters.

The results of this model calibration are presented in Table 4.4. Estimates of the five parameters are listed with the respective t -statistics underneath and in parentheses, if the estimate was found to be significant. Also shown in Table 4.4 are the RMSE; the adjusted coefficient of determination, R_c^2 ; and the average predicted market share, \overline{MSP} . A comparison of the RMSE statistic with the standard deviation of the actual market share data (σ_{MSA} provided in Table 4.1) should be made to assess the model's performance.

Analysis of the results presented in Table 4.4 indicates that either the variable frequency share or the variable seat share still plays a dominant role in explaining a carrier's market share. At least one of these variables was determined to be significant in twelve of the fifteen cities. Although in five of the markets both of these variables were found to be statistically significant, one of the coefficients, B_1 or B_2 , was negative. As before, a negative coefficient implies a counter-intuitive relationship between market share and either seat share or frequency share.

CITY	A	B ₁	B ₂	B ₃	B ₄	RMSE	R _c ²	<u>MSP</u>
BOSTON	-.10	127.20 (8.23)	-163.25 (-4.33)		50.01 (8.83)	.42	.88	7.53
CHICAGO	3.98	117.90 (2.06)	-193.32 (-2.56)	32.43 (1.33)	24.84 (5.28)	.32	.81	5.34
DETROIT	-16.50	108.90 (3.41)		40.38 (1.50)	-141.29 (-4.10)	1.66	.43	45.68
HUNTSVILLE	13.75		142.95 (11.49)	-80.30 (-5.92)		.99	.91	24.24
LOS ANGELES	1.49	192.80 (3.70)	-82.49 (-1.17)	-39.48 (-1.93)	-7.74 (-1.65)	.32	.68	4.23
MADISON	22.46		20.73 (1.57)		35.57 (5.95)	2.10	.64	36.04
MEMPHIS	93.79		-60.93 (-2.10)		8.17 (1.44)	1.68	.12	44.64
MILWAUKEE	21.73		23.83 (4.00)		25.17 (3.01)	1.43	.43	34.95
MINNEAPOLIS	67.70					1.59	.0	67.70
OMAHA	7.14		93.08 (4.31)	-26.19 (-1.87)		.90	.51	14.30
ORLANDO	1.65	-20.42 (-2.13)	53.75 (2.42)		8.89 (4.31)	.29	.79	3.59
ROCHESTER	.48	252.60 (3.65)		-152.77 (-1.42)	-18.87 (-1.59)	.54	.52	5.66
SAGINAW	41.87				12.29 (3.33)	1.83	.34	48.10
SEATTLE	1.34			78.68 (4.00)	12.15 (1.97)	1.20	.50	9.53
ST. LOUIS	3.15	176.99 (1.27)	-262.20 (-2.60)	118.22 (2.98)	5.60 (1.38)	.35	.76	5.68

TABLE 4.4: RESULTS OF MULTIVARIABLE LINEAR REGRESSION ANALYSIS

The two new variables considered in the last section, non-stop share and total possible markets share, were also found to be significant in this model. As indicated in Table 4.4, at least one of these variables was found to be significant in fourteen of the fifteen test markets. However, in five of the markets where one of these variables was significant, either B_3 or B_4 was negative; this again, is counter-intuitive.

A comparison of the RMSE statistics in Tables 4.3 and 4.4 indicates that the multiple variable linear market share model (Equation 4.5) performed better than the multiplicative market share model (Equation 4.4) in eleven of the fifteen test markets. However, the inclusion of negative coefficients, which were found to be statistically significant in the linear market share model (Equation 4.5), can affect the calculation of this RMSE statistic. Analysis of Tables 4.3 and 4.4 indicates that in ten of the fifteen markets (ten from each table), negative coefficients were found to be significant. Since removal of the negative coefficients would affect the performance of the model, a comparison of these models based solely on the RMSE statistic may not be of great use.

4.6 Summary

Thus far, two broad categories of market share models have been calibrated in the fifteen originating city markets. In Chapter Three, the results of the

classical seat share - frequency share model calibrations indicated that these types of models were not extremely useful in explaining the variation in a carrier's market share out of an originating city. However, the results of Sections 4.2 and 4.3 indicated that these same two variables (seat share and frequency share) used in market share regression models, Equations 4.2 and 4.3, satisfactorily explain the variation in market share in the majority of the markets analyzed. Clearly, the structure of the models presented in this chapter was superior to the structure of the models calibrated in Chapter Three.

Four independent variables: a carrier's seat share, frequency share, share of the total non-stop markets, and share of the total possible markets have been identified and tested in this chapter. Each of these variables, to some extent, helped explain the variation in a carrier's market share. The results indicated that except for the three hub cities, a combination of these variables, although never more than two, could be found to satisfactorily explain a carrier's market share in an originating city market. A discussion of the model's poor performance in the hub cities will be included in the last chapter of this thesis.

To assess which of the four models presented in this chapter best explain a carrier's market share, it is really only necessary to compare the results obtained from the calibration of Equations 4.4 and 4.5. As stated earlier, the market share model in Equation 4.4 is really just an extension of the two previous models given in Equations 4.2 and 4.3. By using a step-wise linear regression to calibrate Equation 4.4, the results are guaranteed, at a minimum, to be as good as those obtained from Equations 4.2 and 4.3. Although a comparison of the RMSE statistics

from Tables 4.3 and 4.4 indicated that the market share model presented in Equation 4.5 was superior, the large number of negative coefficients obtained from the calibration of this model may not be desirable -- particularly if the objective is to forecast market share based on changes in the service variables. The final chapter of this thesis will discuss, in more detail, both the pros and cons of each of these models and the practical significance of the estimated β coefficients.

¹ The statistical software package "SOLO" was used in this analysis. SOLO is distributed by BMDP Statistical Software; Los Angeles, California.

² For a complete explanation and derivation of the t-statistic, see John A. Rice, Mathematical Statistics and Data Analysis, Wadsworth & Brooks, Pacific Grove, California, 1988, p. 461.

³ For a complete explanation and derivation of the R_c^2 statistic, see Rice, p. 494.

⁴ An additional month of data was obtained and used in the calibration of equation 4.18. This additional month worth of data was added to the data bases for the fifteen selected originating city markets.

⁵ For this analysis, the parameter was considered significant if its t-statistic was less than -1.5 or greater than 1.5.

CHAPTER 5

Summary and Conclusions

5.1 Overview

The major question addressed in this thesis concerns how an airline might predict its share of the passenger market out of an originating city market. As described in Chapter One, each year air carriers pay travel agents millions of dollars in commission fees for the sale of airline tickets. In addition to the standard commissions, travel agents are also given other monetary incentives (commission overrides), usually an additional fixed percentage of the total ticket price, when an agency's sales of that airline's tickets exceed a certain volume. How airlines structure these override programs can significantly affect their profit potential. Whether the override program is based on a volume of sales by the travel agent or a specific carrier share achieved by the agent, the setting of a baseline goal for a travel agent to meet can be of critical importance.

In an effort to predict or forecast a carrier's market share in an originating city, mathematical models which relate a carrier's market share to a set of exogenous variables describing an airline's operations, were employed. However, in order to make a confident forecast, these models first had to be tested and calibrated in various types of markets, using historical data to assess

their "goodness of fit". The purpose of this thesis was to apply previously developed market share models, and new derivatives of these models, in originating city markets in order to identify which variables are most important in predicting an airline's share of the originating passengers.

In Chapter Three of this thesis, classical seat share and frequency share models were calibrated and tested in fifteen originating city markets. Although these models were previously found to be very effective in modeling a carrier's market share in origin-destination (O-D) markets, their application in originating city markets was less encouraging. In Chapter Four, a completely different structure of a market share model, one which could be calibrated using linear regression techniques, was tested in the fifteen selected originating markets. The encouraging results obtained from this first model, which was also originally proposed for use in O-D markets, led to the development and testing of more extensive models in an attempt to further explain the variation in a carrier's market share. Both a linear and non-linear structure of a multivariable regression model were calibrated in the test markets and were able to pass standard statistical goodness of fit tests.

5.2 Findings and Conclusions

The results of the classical market share-frequency share model calibration (Equation 3.1) presented in Table 3.2 indicated that this model could not effectively explain the variation in a carrier's market share out of an originating city. A similar model (Equation 3.2), using seat share as the explanatory variable instead of frequency share, was also calibrated in the fifteen test markets. As indicated in Table 3.3, this classical market share-seat share model produced acceptable results in only six of the fifteen test markets. Although neither model was particularly useful in explaining the variation in a carrier's market share, a comparison of the results did reveal that seat share was more useful than frequency share in explaining a carrier's originating city market share.

The inability of these two models to correctly explain this variation in a carrier's market share was largely due to the lack of an S-shaped relationship between market share and frequency share and between market share and seat share. It was this S-shaped relationship, previously noted in O-D markets, that was the cornerstone upon which this model was built. Since Northwest's market share-seat share and market share-frequency share data clearly did not fit this S-shaped relationship, the model's inadequacy in originating markets was not surprising.

The structure of these classical market share models also posed several difficulties. Since neither model could satisfactorily explain the variation in market share, a model which included both of these variables would have been desirable. However, the structure of the models in Equations 3.1 and 3.2 could not easily accept more than one independent variable at a time -- at least in a manner that would have some theoretical basis. In addition to this difficulty, the exponent α in each of these models (in cities where the models were statistically sound) was difficult to interpret. Unlike the first three models of Chapter Four, the exponent was not a true elasticity of market share and, therefore, could not be used to assess the effects of changes in service variables (either frequency share or seat share) on market share.

The results of the seat share regression model in Section 4.2, indicated that this simple market share regression model was statistically significant in twelve of the fifteen selected markets (significant β coefficients were not found in the hub cities). By comparing this model, which used seat share as the only explanatory variable, with the classical seat share model calibrated in Chapter Three, it became clear that the market share regression model was superior. Since both models used only seat share in their formulations, this comparison allowed an assessment to be made concerning the model structure. Because the regression model was able to explain a carrier's market share in twice as many cities as the classical model, further testing of more extensive market share regression models was justified.

Analysis of the coefficient of determination, R_C^2 , from this seat share regression model (Table 4.1) indicated that other explanatory variables were needed to better explain the variation in a carrier's market share. In Section 4.3, frequency share was added to the non-linear regression model, and in Section 4.4, the variables non-stop share and share of the total possible markets served were also added. The resultant non-linear regression model in Section 4.4, which included four explanatory variables, was then calibrated using a step-wise regression technique.

The inclusion of these three additional variables in the basic seat share regression model significantly improved the explained variation in a carrier's market share, as indicated in Table 4.4. The results of this step-wise regression analysis also indicated that in almost every market tested, one of the variables, either frequency share or seat share (but not both), was found to be a significant explanatory variable. Since these two variables are highly correlated, it is no surprise that only one of them is needed in the model. However, including both of these variables in the initial model is still recommended because in some markets frequency share was found to be more significant, and in other markets seat share proved to be more significant. Also, by using the step-wise regression technique, variables that are not found to be significant are automatically discarded, and thus, the final model is not biased by the inclusion of insignificant variables.

It was also noted that of the two variables, non-stop share and share of the total possible markets, one of them, but rarely both, was found to be significant in

almost all of the test markets. By comparing the R_c^2 statistics in Tables 4.2 and 4.3, it is noted that in thirteen of the fifteen test markets, inclusion of at least one of these variables improves the amount of explained variation in a carrier's market share. As noted above, although one of the variables may not be significant in a given market, it is still prudent to include both variables in the model and let the step-wise regression analysis determine which of these two variables should be included in the final model specification.

Although the results of the non-linear regression analysis showed that this model was very useful in explaining the variation in a carrier's market share, another variation of this model, a linear combination of the four explanatory variables, was proposed and tested in the fifteen selected cities. A comparison of the RMSE statistics for the non-linear and linear regression models, given in Tables 4.3 and 4.4 respectively, indicates that both models contained approximately the same level of error (the RMSE for the linear model was only marginally better in eleven of the fifteen markets). Although the two models may be statistically equivalent, the non-linear regression model Equation 4.4 is preferred over the linear model for several reasons.

First, as its name indicates, the linear regression model assumes a linear relationship between the dependent variable (market share) and the four explanatory variables: seat share; frequency share; non-stop share; and share of the total possible markets. However, Figures 3.1-3.30, which show market share vs. frequency share and market share vs. seat share in the fifteen selected cities over the twenty-one month sample period, indicate that for the majority of the

markets, these relationships cannot be characterized as linear. Although the linear market share model was still effective in these test markets, attempting to explain a non-linear relationship with a linear model in other markets may not prove to be as successful.

The practical interpretation of the estimated parameters (β_1 , β_2 , β_3 , and β_4) in the non-linear model is a second reason why the non-linear model is preferred over the linear one. As described in Chapter Four, each of these parameters represents an elasticity of market share with respect to the four explanatory variables. These elasticities can be used by airline management to estimate the effect on market share with changes in the variable that they represent. As such, these elasticities can be used as an additional decision making tool for establishing the levels of service in the various markets. The coefficients b_1 , b_2 , b_3 , and b_4 , on the other hand, cannot be interpreted so easily. The coefficient b_1 , for example, represents the change in market share with a one unit change in seat share, with the three other variables held fixed (which is not a true elasticity). Although this may present management with some useful information, it is not as useful a tool as the true elasticities in helping management establish levels of service in the various markets.

As noted in Chapter Four, both the non-linear and linear regression models contained estimated coefficients that were found to be negative. This result was not surprising because the slope of the observed relationship between market share and seat share and between market share and frequency share in some of the fifteen selected cities was actually negative. Although these negative

coefficients certainly improved the overall goodness of fit for both models, using these negative coefficients in models used for forecasting could lead to questionable results. The counter-intuitive relationship observed in Saginaw, for example, does not necessarily imply that a decline in market share is caused by an increase in frequency share. Therefore, a model based on this counter-intuitive relationship may not be of great use.

Quite obviously, there are many other forces at work, such as competitor fares, advertising, and frequent flyer programs, which can affect a carrier's market share. With so many other factors affecting a carrier's share of the originating market, these counter-intuitive relationships may only occur in the short run, and therefore, forecasts based on these erroneous coefficients may be invalid. As discussed in the final section of this chapter, a regression analysis with these negative coefficients constrained to zero may be of more practical use to airline management.

5.3 Further Research

As shown in the last chapter, the structure of the non-linear regression model given in Equation 4.3 is extremely useful because it allows one to include further independent variables in a manner that still pennits calibration by a linear least squares method. Because Table 4.3 indicates that there is still more

room for additional explanatory variables, further research in this area, using this easily adaptable model, should attempt to define variables that could further reduce the amount of unexplained variation in a carrier's market share (especially in the hub cities). Using the step-wise regression technique, additional variables can easily be added to the model and tested for significance.

The market share data (described in Chapter Three) used in this analysis was defined as the percentage of tickets sold by travel agents for travel on Northwest in a given market. Within a specific market, however, there exist many different sizes of agencies, in terms of revenue generated, which could be broken down into three or four different subgroups. By tabulating the market share data as a function of agency size, and then performing a regression analysis for each of these subgroups, the influence of an agency's size (or market power) on a carrier's market share could be studied. If market share is found to be a function of agency size, management could establish different baseline market share goals as a function of this size.

Another known influence affecting the number of tickets sold on a particular airline by a travel agency is the Computer Reservation System (CRS) to which the agency subscribes. Due to the bias in these systems, travel agents may be prone to make a reservation on the airline that serves as their host CRS. Thus, travel agents using SABRE, owned by American Airlines, may be biased toward American when selling a ticket to a consumer. Extending the analysis in this thesis, the market share data presented in Chapter Three could be broken up into two groups: those travel agents that subscribe to PARS (owned jointly by

Northwest and TWA), and those agents that do not subscribe to PARS. A regression analysis using both groups of market share data would reveal the effect of the CRS host and would also allow for different market share quotas to be set for the two groups.

Aside from the inclusion of additional variables in the model and classifying the market share data in different ways, the effect of constraining significant parameters, that were found to be negative, to zero should be studied. Although it is clear that the constraint will increase the RMSE (if the parameter was significant) and decrease the amount of explained variation in a carrier's market share, the extent of these changes is unknown. Calibrating these market share regression models with this additional constraint will produce, as discussed in the last section, a more practical model, the results of which could be readily used for forecasting.

Finally, forecasts are only as good as the models used to produce them. Comparing the results of a forecast with the outcome of that period should give an indication of the model's validity. Also, as new data becomes available, the market share model can be re-calibrated so that forecasts for the next period reflect the carrier's most recent performance.

Selected References

- [1] Baily, E.E., Graham, D.R., and Kaplan, D.P. Deregulating the Airlines. Cambridge, Ma.: The MIT Press, 1985.
- [2] Bond, Richard O. "Flight Frequency Competition on Domestic Air Routes." Master of Science Thesis, MIT, Sloan School of Management, 1970.
- [3] Dorman, Gary J. Airline Competition: A Theoretical and Empirical Analysis. Berkeley, Ca.: Graduate Division of the University of California, Berkeley, 1976.
- [4] Fruhan, William E. Jr. The Fight for Competitive Advantage. Boston, Ma.: Graduate School of Business Administration, Harvard University, 1972.
- [5] Leet, James O. "Transportation in the Era of Deregulation." Transportation Quarterly, Volume XXXIX:2, April, 1985.
- [6] Levine, Michael E. "Airline Competition in Deregulated Markets: Theory, Firm Strategy, and Public Policy." Yale Journal of Regulation, Volume 4: 363, 1987.
- [7] Meyer, John R. and Oster, Clinton V. Deregulation and the Future of Intercity Passenger Travel. Cambridge, Ma.: The MIT Press, 1985.
- [8] Pan Am Corporation Annual Report. 1982.
- [9] Pindyck, Robert and Rubinfeld, Daniel L. Econometric Models & Economic Forecasts. 2nd ed. New York: McGraw-Hill Book Company, 1981.
- [10] Renard, Gilles. "Competition in Air Transportation, An Economic Approach." Master of Science Thesis, MIT, The Department of Aeronautics and Astronautics, 1970.
- [11] Rice, John A. Mathematical Statistics and Data Analysis. Pacific Grove, Ca.: Wadsworth & Brooks, 1988.
- [12] Simpson, Robert W. "A Market Share Model for U.S. Domestic Airline Competitive Markets." FTL Memorandum 70-5, The Department of Aeronautics and Astronautics, MIT, 1970.
- [13] Simpson, Robert W. "An Outline of Research Activities into Airline Market Share Models." FTL Memorandum 75-5, The Department of Aeronautics and Astronautics, MIT, 1975.

- [14] Simpson, Robert W. "The Demand for Air Transportation Services." Course Notes for Air Transportation Economics, The Department of Aeronautics and Astronautics, MIT, 1982.
- [15] Taneja, Nawal K. "Airline Competition Analysis." FTL Report R68-2, The Department of Aeronautics and Astronautics, MIT, 1968.
- [16] Taneja, Nawal K. The Commercial Airline Industry. Lexington, Ma.: D.C. Heath and Company, 1976.